Epigenetic and Transcriptional Regulation of Self Renewal in Acute Myeloid Leukemia

by

James P Ropa

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Molecular and Cellular Pathology)
in the University of Michigan
2019

Doctoral Committee:

Assistant Professor Andrew G. Muntean, Chair
Professor Eric R. Fearon
Professor Alexey I. Nesvizhskii
Professor Pavan Reddy
Assistant Professor Jean-François Rual

James P. Ropa

jropa@umich.edu

ORCID iD: 0000-0001-6120-1859

Dedication

To my grandfather, Robert Ropa, whose death motivated the start of my life as a cancer biochemist, and whose life still inspires me to strive every day to follow his example and be a good husband, a loving father, and a kind and compassionate person.

To my grandmother, Carol Doyle, who was always there for me with support, advice, or just a funny sarcastic remark to help get me through a tough day, and who was a shining example of how to reflect God's light in daily life.

I miss you both and I pray every day that I am making you proud.

Acknowledgements

The work presented in this dissertation epitomizes collaborative scientific research, and I am very thankful to the many people who have contributed to it both scientifically and in other ways. First, I would like to thank my mentor, Dr. Andrew Muntean. When I chose the University of Michigan as my new home and the place where I would develop as a scientist, I had no idea how much I would learn in such a short amount of time. His mentorship is the primary reason that I have grown more as a scientist, a thinker, an academic, a writer, and a communicator in the last five years than in any other time in my life. I would like to thank Andy for his willingness to allow me independence, for providing guidance when it was necessary, and for making the Muntean lab a truly enjoyable place to work. I also would like to thank him for his belief in me as a scientist, as that is what pushed me to really challenge myself in the next stage in my career. Further, I want to thank the members of the Muntean lab both past and present, who have all made important contributions to this work, including Zhiling (Lynn) Chen, Maria Mysliwski, Laura Adkins, Marisa Hildebrandt, Hsiangyu (David) Hu, and Blaine Teahan. Finally, I would like to especially thank Dr. Lili Chen, Dr. Nirmalya Saha, and Dr. Justin Serio, who were critical collaborators on this project, and who are great scientists and even better friends.

I would next like to thank my committee members for their invaluable guidance and support. Dr. Alexey Nesvizhskii has been a de facto second mentor to me through the last several years, particularly with regards to my pursuit of a Masters degree in Bioinformatics concurrent with my thesis work. His guidance and willingness to collaborate on proteomics studies made much of the work in this dissertation possible. Dr. Jean-Francois Rual has also been an important collaborator, providing reagents and instruments that were critical to the development of this story. Dr. Rual has provided important feedback and critique that helped to shape the studies I worked on throughout

my entire graduate school experience. Dr. Pavan Reddy has given invaluable insight and clinical perspective to this work, and I am very fortunate to have been given advice on a hematology project from a renowned hematologist. Dr. Eric Fearon has provided vital intellectual direction to this project. I am always amazed at his ability to really see what is missing from a study, and to understand the best path forward to add impact and wrap up a story. My thesis committee has helped make this project into an exciting story, and they have also helped me to grow as a scientist and a critical thinker.

I would be remiss if I did not thank the people who make graduate school at Michigan such an amazing experience, namely the faculty, staff, and students. Dr. Zaneta Nikolovska-Coleska is an incredible program director for MCP, a fantastic scientific mind, and a great friend. Laura Labut is the best program administrator at the University of Michigan, and probably anywhere else as well. I would also like to thank Dr. Margit Burmeister and Julia Eussen, the director and program administrator of the Bioinformatics graduate program, for their guidance. The Hematology Research Group has been an incredible resource, and I would like to give special thanks to Dr. Maria (Ken) Figueroa, Dr. Ivan Maillard, Dr. Mark Chiang, Dr. Yali Dou, Dr. Jolanta Grembecka, Dr. Tomek Cierpicki, and Dr. Russell Ryan for all their advice. Additionally, I would like to thank Dr. Robert Stahelin and Dr. Katherine Ward for their mentorship during my undergraduate research experience at the University of Notre Dame. Finally, I would like to thank the MCP student group, as well as “PiBS and Friends” (my program cohort), which are groups made up of some of the best scientists and people I know. I truly cherish the friendships I have made during my time at the University of Michigan.

Most importantly, I would like to thank my family. First, my extended family and my wife’s extended family have provided an amazing support system through all of my academic pursuits. In particular, my grandfather, Patrick Doyle, and my grandmother, Carolyn Ropa, have been incredibly supportive of all that I have done throughout my academic pursuits. Dave McKinney, Chris Surfus, Jared Swihart, and Caleb Eiler, who are the best friends anyone could wish for, have provided many laughs and many beers, both of which are important contributions to a successful graduate student experience. Jackie and John Widau, Allison and Brendon Widau, and Lauren and Chris Surfus are the best in-laws in the world, and I am so fortunate to have their support in

both good and difficult times. My sister and brother-in-law, Catharine and Dave Root, are my inspirations- I aspire every day to be as good of a person, parent, and colleague as they are. Cat has been my partner in crime for almost 29 years now, and I would not have had half of the success as I have had without her help and support. My mom and dad, Stephanie and Chris Ropa, have been my biggest cheerleaders, my biggest critics, my shoulders to cry on, and my best friends, depending on what the day called for. I am truly blessed to have such amazing parents and none of this work would have been possible without you challenging me and supporting me throughout my academic career. I would like to thank my daughter and son, Amelia and Noah, for giving me a reason to get up in the morning (metaphorically and literally), even in the toughest weeks. I would also like to thank my amazing wife, Dr. Rebecca Ropa, for being my rock, my best friend, and my partner in this crazy adventure for the last 12 years. I could write an entire dissertation about all that you mean to me but for the sake of space I will just remind you that nothing I do would be possible or worth it if I wasn't doing it with you by my side. Finally, I would like to thank God for this truly blessed life and for giving me the ability to pursue my passion for scientific discovery.

Table of Contents

List of Tables

List of Figures

List of Appendices

Abstract

Acute myeloid leukemia (AML) is diagnosed in >20,000 people/year in the United States alone and is associated with a poor prognosis. AML arises due to altered transcriptional programs resulting from mutations and chromosomal rearrangements. Frequently, this altered transcription is a consequence of epigenetic deregulation. Indeed, over 70% of AML patients harbor mutated epigenetic modifiers[1], which regulate chromatin accessibility and gene expression. Aberrant expression of the *HOXA* gene cluster, which can result from epigenetic deregulation, drives transformation of ~50% of AML, including those associated with poor prognosis[2–5]. One manner in which the *HOXA* gene cluster becomes aberrantly expressed is through 11q23 chromosomal translocations involving the *Mixed Lineage Leukemia 1* (*MLL1*) gene[6–10]. These events result in the formation of fusion genes encoding MLL fusion oncoproteins which transcriptionally activate oncogenes, including the *HOXA* cluster[5,7,11]. Our lab and others have demonstrated that the Polymerase Associated Factor complex (PAF1c), an epigenetic regulator complex, interacts directly with and recruits wildtype MLL1 and MLL-fusion oncoproteins to target loci like *HOXA9* and *MEIS1*[12,13]. The PAF1c-MLL interaction is required for leukemia cell proliferation, but dispensable for normal hematopoiesis[14]. Mutations and aberrant expression of subunits of the PAF1c are observed in various malignancies, suggesting that the PAF1c must be tightly regulated for proper cellular development[15–17]. However, the biochemical regulation of the PAF1c that allows for its dynamic regulation of gene expression in AML is not fully understood.

To better understand the regulation of the PAF1c, we use a proteomics approach to identify novel interaction partners of the PAF1c in AML cells. This study reveals a novel interaction between the PAF1c and the H3K9 methyltransferase SETDB1. The PAF1c-SETDB1 interaction represses the target genes *Hoxa9* and *Meis*1 in murine MLL-AF9 driven leukemic cells and human AML cell lines. SETDB1 mediated transcriptional

repression is correlated with an increase in promoter H3K9 trimethylation (H3K9me3). These data suggest that SETDB1 epigenetically represses pro-leukemic gene expression in AML. Therefore, we next explore the biological impact of SETDB1 expression and H3K9 methylation on AML. We note that expression of *SETDB1* in AML patient samples is significantly lower compared to normal hematopoietic cells. Further, higher *SETDB1* expression correlates with a significantly better overall survival in AML patients. These data are consistent with SETDB1 negatively regulating pro-leukemic genes and suggests that *SETDB1* expression and H3K9 methylation levels may be correlated with AML patient prognosis. We demonstrate that overexpression of *SETDB1* significantly delays MLL-AF9 mediated leukemogenesis *in vivo* by inducing differentiation of leukemic cells. We also explore how chemical inhibition of H3K9 methylation affects AML transformation. Treatment with H3K9 methyltransferase inhibitor UNC0638 is antagonistic to established AML cell growth. In contrast, UNC0638 preserves mouse hematopoietic stem and progenitor cells (HSPCs) in culture and increases the amenability of bone marrow cells to be transformed by the MLL-AF9 oncogene. Transcriptome analyses demonstrate that overexpression of SETDB1 downregulates *Hoxa* and pluripotency gene programs. ChIP-sequencing and ATAC-sequencing of AML cells show that overexpression of SETDB1 leads to the acquisition of a more compact, epigenetically silenced chromatin state at the promoters of genes that are critical for AML, including *Dock1* and MLL-AF9 target genes *Hoxa9* and *Six1*, and others. Together, these data reveal a previously unrecognized role for SETDB1 and H3K9 methylation in suppressing AML by epigenetically silencing pro-leukemic target genes and promoting differentiation.

Chapter 1 Introduction

1.1 Transcription and epigenetic regulation in normal and malignant hematopoiesis

Hematopoiesis is a tightly regulated process that is critical to both the establishment and the homeostasis of the blood system[18]. The process of hematopoiesis is largely governed by tightly regulated transcription of factors that cause cells to make specific lineage fate decisions. Furthermore, deregulation of these transcription processes leads to diseases and defects of the hematopoietic system[19]. Specifically, hematopoietic cancers are malignancies associated with transcriptional deregulation[19]. Epigenetic regulation of transcription is critical to normal hematopoietic development and epigenetic deregulation is common in hematopoietic malignancies. Therefore, epigenetic deregulation has become an area of great interest with regards to studying hematologic cancers[1,20,21].

Hematopoiesis

Hematopoiesis is the process by which blood cells develop and are regenerated throughout the lifespan of an organism[22]. During the process of adult hematopoietic differentiation, primitive cells with high self-renewal capacity and pluripotency, i.e. the ability to differentiate into different cell types[23], give rise to all primitive and mature blood cells in the myeloid, lymphoid, and erythroid lineages. Hematopoietic stem cells (HSCs) are the most primitive blood cells and are pluripotent, giving rise to multiple different hematopoietic lineages as well as other tissue subtypes[24]. Within the HSC population are two subpopulations of cells: long-term HSCs (LT-HSCs) and short-term HSCs (ST-HSCs). LT-HSCs have a high self-renewal capacity and are responsible for sustaining the HSC population[25]. LT-HSCs can differentiate into ST-HSCs, which are highly proliferative and can differentiate to give rise to more lineage committed progenitor cells[26,27]. ST-HSCs differentiate to become multipotent progenitor cells (MPPs), which in turn differentiate into progenitor cells that are fully committed to either a myeloid or

lymphoid lineage. These cells are known as Common Myeloid Progenitors (CMPs) or Common Lymphoid Progenitors (CLPs) [28]. In the lymphoid lineage, CLPs then make lineage decisions to give rise to the mature lymphoid cells, including Natural Killer (NK) cells, T cells, and B cells[28]. These mature lymphoid cells are primarily associated with immune system processes, with B cells and T cells functioning in adaptive immunity and NK cells functioning in the innate immune response[29,30]. In the myeloid lineage, CMPs exhibit the capacity to give rise to two additional types of progenitor cells: Megakaryocyte/ Erythroid Progenitors (MEPs) and Granulocyte/ Macrophage progenitors (GMPs) [28]. Interestingly, MEPs can also arise directly from MPPs or even HSCs[31]. MEPs differentiate to give rise to the terminally differentiated erythrocyte cell population, which are responsible for the transportation of oxygen and carbon dioxide throughout the body[32]. MEPs also differentiate to become megakaryocytes[28]. Megakaryocytes in turn can differentiate to produce platelets, which function in wound healing and other critical organismal processes[33]. GMPs differentiate to become basophils, neutrophils, eosinophils, and mast cells, which are all fully mature myeloid cells that function in a variety of critical processes, including infection response, inflammation, and vasodilation[34,35]. GMPs also give rise to monocytes, which in turn can differentiate into macrophages, cells that are involved in phagocytosis and immune response[35]. In summary, HSCs self-renew and differentiate in order to produce progenitor and terminally differentiated blood cells that encompass a wide variety of cell types that function in diverse processes. These lineage defining processes are the result of tightly regulated changes in specific gene programs. Aberrations in these processes can lead to hematopoietic deficiencies, immune disorders, and various malignancies. It is therefore critical to understand the regulatory processes involved in hematopoiesis.

As these differentiation processes take place, lineage specific gene programs must be turned on, while gene programs associated with self-renewal and potency are progressively turned off[2]. Hematopoietic stem and progenitor cells (HSPCs) have distinct gene expression profiles, which includes high expression of genes associated with self-renewal such as the *Homeobox A* (*HOXA*) cluster of genes and others[2,36,37]. As these cells differentiate, the expression of these self-renewal gene programs is

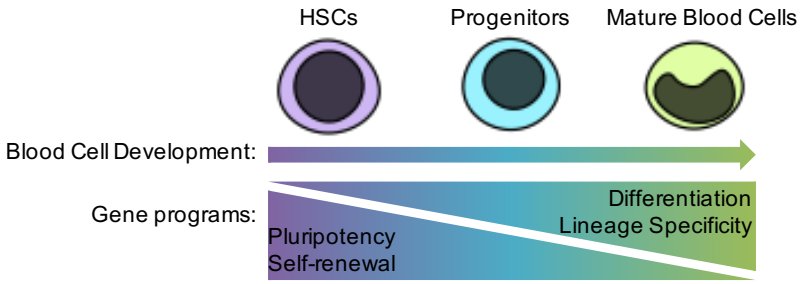


Figure 1-1: Changing gene programs during hematopoietic differentiation. Shown here is a model representing changes in specific gene expression programs as primitive HSCs and hematopoietic progenitor cells differentiation to become fully mature blood cells. As cells undergo myeloid differentiation, gene programs associated with pluripotency and self-renewal must be progressively downregulated while differentiation gene programs and lineage specific genes become activated.

progressively downregulated (Figure 1-1). Additionally, the expression of genes that are associated with lineage specific functions are upregulated. For instance, in mature blood cells we see high expression of *KLF3*, which plays an important role in erythroid cell maintenance[38]. There are several mechanisms by which these gene programs are regulated, including gene activation by lineage specific transcription factors and changes in the epigenome that lead to changes in the transcriptome.

Lineage Specific Transcription Factors

Transcription is the cellular process by which DNA template is used by protein machinery to produce RNA. RNA polymerase II (RNAPII) dependent transcription is associated with the production of mRNA, among other subtypes of RNA[39]. mRNA is the template by which proteins are produced through the process of translation, and proteins are one of the primary effectors of cellular changes and identity. Therefore, the process of RNAPII dependent transcription is a critical regulatory step in establishing gene expression programs that define various cellular states and cellular processes, including development[40,41]. Transcription is regulated in part by many different protein networks that consist of a wide variety of transcription factors and cofactors[40,41]. Frequently, the regulation of transcription consists of a dynamic interplay between factors and cofactors that are associated with active gene transcription and those that drive gene repression and silencing[42]. Deregulation of these transcriptional networks is associated with aberrant regulation of cellular processes and the development of disease states[43]. Transcription factors play a critical role in HSC maintenance as well as

the regulation of lineage fate determination throughout the process of hematopoietic differentiation[19,44].

HSCs require activated gene programs that allow for self-renewal and continuous repopulating ability[24]. Some of the transcription factor genes associated with self renewal of HSCs are Homeobox protein B4 (*HOXB4*), which is required for the proliferation of HSCs[45,46]; *IKAROS*, which is associated with maintaining the self-renewal capacity of LT-HSCs[25]; *GATA-2*, which is a pro-survival and pro-proliferative gene in HSCs[47–49]; and Stem cell leukemia/ T-cell acute lymphocytic leukemia protein 1 (*SCL/TAL1*), which is required for the long term potential of HSCs, and is also important for the development of more mature blood cells[18,50,51].

During the process of hematopoietic differentiation, there are many proteins that are required to drive specific fate decisions. These proteins are known as lineage specific transcription factors and are critical regulators of blood cell development[52]. Deregulation of the genes encoding these proteins is also associated with the disease states, including the development of malignancies[53]. For the sake of simplicity, some of the more well-studied transcription factors that govern myeloid lineage fate are reviewed here. Beginning with the earliest stage of myeloid differentiation, SCL/TAL1 has been established as an important regulator of not only HSCs, but also the differentiation of cells in the myeloid lineage[54], where it plays a critical role in regulating the cell cycle and proliferation capacity of monocytic progenitors[55]. Another transcription factor required relatively early in the myeloid differentiation process is PU.1. PU.1 is an E-26 transformation specific (ETS) domain containing protein and transcription factor that activates myeloid specific gene programs to drive MPP differentiation to GMPs and more terminally differentiated myeloid cells[56]. Interestingly, PU.1 is also required to sustain the quiescence of HSCs, indicating that it plays a role in both the maintenance of stem cells as well as driving hematopoietic differentiation[57]. It is important to note that some lineage decisions that are dependent on PU.1, like many other lineage fate decisions, also depend on other transcriptional regulators. Frequently, there is a balance of expression between different lineage factors that are all expressed in common progenitor cells[52]. For instance, PU.1 and the erythroid associated protein GATA binding factor 1 (GATA-1) act antagonistically to each other's' transcription factor

functions[58]. The differentiation of cells in the myeloid and erythroid lineages are therefore dependent on a balance of *PU.1* and *GATA-1* expression. *PU.1* high expression pushes cells to a monocytic lineage decision, whereas low expression of *PU.1* shifts the balance to GATA-1 driven transcription of genes associated with terminal erythroid differentiation[59]. Another protein that plays a crucial role in early progenitor cell fate decisions of myeloid differentiation is CCAT/enhancer binding protein α (C/EBPα). The lineage decisions that are impacted by the C/EBP family of genes are directly influenced by the expression of other, collaborating genes. C/EBPα drives myeloid differentiation[60], and is specifically expressed in progenitors that are committed to the granulocyte/ macrophage lineage[61]. However, it has been demonstrated that *GATA-2* co-expression with *C/EBPα* can push these cells toward eosinophil differentiation[52]. This again suggests that the lineage fate decisions in hematopoiesis are dependent on a complex balance of the expression of different factors. Interestingly C/EBPα has been reported to be required only through the CMP-GMP transition, and is not necessary for the terminal differentiation of these progenitors[62]. Further along the myeloid differentiation path are genes associated with regulating the differentiation of more mature myeloid cells, including terminal differentiation. For instance, Zinc finger protein GFI-1 (GFI-1) is required for neutrophil differentiation, and is expressed in lineage committed cells to cause them to commit to a final mature cell type[63–65]. These transcription factors make up just a small part of the total machinery required to define a fully functional hematopoietic system. In summary, these factors form a complex network in which a precise balance of regulation is required to sustain hematopoietic potential and drive the development of the various mature blood cell lineages. These cell fate decisions can also be affected by other biochemical regulators, such as epigenetic modifiers.

Epigenetics as a mode of transcriptional regulation

One manner in which lineage specific transcription factors and other transcriptional processes are regulated are through changes in the proximal or distal epigenetic landscape, as well as through interactions with the epigenetic modifying machinery[66–68]. Epigenetics refers to heritable changes in gene expression that are not

explicitly encoded by the genomic sequence, but instead are attributable to regulatory mechanisms such as changes in chromatin conformation and compaction. Chromatin is the scaffold of DNA, RNA, and proteins that makes up chromosomes, which is the structure that contains the genetic information for a cell. Cells contain ~ 3 meters of DNA that must be packaged to fit within a cell but must also be available when needed for various molecular processes, including transcription[69]. As a result of this balance, chromatin exists in a variety of states of compaction, from forming fully compact mitotic chromosomes, to open chromatin that is associated with active transcription, and a full spectrum of compaction and decompaction in between[69,70]. Cells effect changes in chromatin compaction or availability through chromatin remodeling proteins, which oftentimes are recruited based on the recognition of the epigenetic landscape at a given region of the genome[71]. Consequently, the epigenome is critical in mediating the availability and therefore the transcriptional output of the genome. This makes it another important mode of transcriptional regulation in hematopoiesis[20,21,72]. These changes in chromatin conformation play an important role in all biological processes, including cell development and the initiation, prognosis, and progression of malignancy[72]. Therefore, the epigenetic modifying proteins that induce these changes are particularly important to understand in the context of normal and malignant hematopoiesis.

There are a wide variety of epigenetic changes and mechanisms that affect the epigenome. Some of the most well studied types of epigenetic modifications involve RNA interference (RNAi), which involves various classes of RNA that can modulate chromatin availability or silence transcription or translation of specific targets[73]; chemical modifications on the DNA, such as methylation, which function by recruiting chromatin remodelers in order to effect changes in chromatin availability[74]; and histone modifications, which are covalent post translational modifications (PTMs) on histone proteins[75]. These different epigenetic changes frequently depend on each other to affect a functional change in the chromatin, often working in a stepwise fashion to establish a specific chromatin state or forming feedback loops to maintain the epigenomic state of the chromatin[73]. Epigenetic modifiers involved in the deposition, removal, and recognition of PTMs on histones are reviewed here. These modifiers play an important

role in development and are recurrently mutated, rearranged, or aberrantly expressed in hematopoietic malignancies[1].

Histone modifications

Histone proteins form a nucleosome octamer complex that is the basic unit of DNA packaging[75]. DNA is wound around this octamer, which can then be tightly packed with other histone octamers to make fully condensed, inaccessible chromatin, which is known as heterochromatin[76]. Heterochromatin makes up regions of the chromosome where little or no transcription takes place and are important for the structural integrity of the chromosome, genome stability, and packaging of chromatin[76]. It also marks genes that are associated with development that have been silenced because they are not necessary for or compatible with the given cell lineage[76]. Histone octamers can also be moved by chromatin remodeling proteins in order to create a permissive state of chromatin known as euchromatin. Euchromatin is the region of the genome where active transcription can occur[76].

The histone proteins that make up the nucleosome octamer contain disordered N-terminal “tail” regions that can be chemically modified in a variety of ways, and these modifications can lead to changes in chromatin conformation or effect changes in transcriptional state[75]. These tails have been shown to be phosphorylated, acetylated, methylated, glycosylated, sumoylated, and ubiquitinated, among others[77]. Importantly, these modifications have profound effects on the chromatin state, by recruiting chromatin remodelers to either compact or de-compact chromatin, depending on the specific combination of modifications that are present at a given region of the genome[69,75,78]. There are several different types of proteins involved in establishing this “histone code”[79] and in effecting the chromatin structural changes that the modifications dictate. These proteins are known as writers, readers, and erasers. They function by depositing, recognizing and binding to, or removing PTMs on histone tails[75].

Epigenetic writers, readers, and erasers

Proteins that contain catalytic domains that function by depositing covalent modifications on histone tails are known as epigenetic writers. There are many families

of proteins responsible for modifying histone tails, including lysine methyltransferases (KMTases), arginine methyltransferases (PRMTs), histone acetyl transferases (HATs) ubiquitin ligases, kinases, and sumoylases, among others[80]. In contrast, the proteins that have function by removing PTMs from histone tails are known as epigenetic erasers. Some families of erasers are lysine demethylases (KMDases), arginine demethylases, histone deacetylases (HDACs), deubiquitinating enzymes (DUBs), and phosphatases[80]. Proteins that recognize and bind histone modifications are known as epigenetic readers. Examples of epigenetic readers and a few of their known binding sites include bromodomain proteins, which can bind to acetylated histones; PHD domain containing proteins, which can bind to acetylated and methylated histones; and tudor domain containing proteins, which can bind to methylated lysine residues[71]. These domains may also bind to other modified histones and have functions independent of their interactions with histone tails. Frequently, epigenetic writers and erasers contain reader domains, that aid in recruiting writers or erasers[71]. Chromatin remodelers such as proteins in the Remodel the structure of chromatin (RSC) complex are also frequently epigenetic readers, allowing for them to be recruited to regions of the genome that have been marked for compaction or decompaction[71]. Another class of reader are transcription factors that recognize and bind to specific epigenetic modifications[66]. For example, General transcription factor IID (TFIID) is recruited by and engages a histone PTM, Histone 3 Lysine 4 trimethylation (H3K4me3) through the Plant homeodomain (PHD) fingers of its TBP associated factor 3 (TAF3) subunit. TFIID is one of the first protein complexes to bind promoter regions to initiate transcription and H3K4me3 is a modification associated with active promoters, suggesting that TAF3's affinity for H3K4me3 allows for it to be recruited to promoters marked for active transcription. The dynamic interplay of these proteins and the modifications they modulate has important effects on chromatin availability and transcriptional output.

Epigeneticists have identified a large number of histone modifications that play important roles in the availability of chromatin and suspect that even more exist that have yet to be fully studied. Some of the most common modifications that play a role in modulating transcription are lysine and arginine methylation of histone 3 (H3), acetylation of H3, and ubiquitination of histone 2B (H2B)[43,69,75,81]. Some of the H3 lysine

(H3K) residues that are methylated that are associated with affecting transcriptional output are H3K4, H3K9, H3K27, H3K36, and H3K79[82]. These residues can be mono, di, or trimethylated (me1/me2/me3), and frequently the different degrees of methylation have different functional effects on the chromosome by recruiting different proteins[82]. Some of the writers associated with depositing lysine methylation include SET1A/B and MLL1/2/3/4, which deposit H3K4 methylation[83]; the SUV39H family, which deposits H3K9 methylation[77]; EZH1/2 which is associated with H3K27 methylation[84]; SETD2/3, ASH1L, and NSD1/2/3 which deposit H3K36 methylation[85,86]; and DOT1L, which is the only known mammalian H3K79 methyltransferase[8]. Erasers of lysine methylation include the JARID1 family of proteins, which removes H3K4 methylation; the JMJD1 family of H3K9 demethylases; the JMJD3 family which demethylates H3K27; and the JHDM1 family of proteins, which is associated with H3K36 methylation[87]. Interestingly, there are also several families of lysine demethylases with less specific substrate specificity, including the JMJD2 family, which can demethylate H3K9 and H3K36[88]; the PHF family of H3K9/K27 demethylases[89]; and LSD1, which demethylates both H3K4 and H3K9[88]. To date, only one study has identified a bona fide H3K79 demethylase, JMJD1B[90], so further studies may be required to understand the full extent of the dynamics surrounding this modification.

Many of the same lysine residues that can be methylated also undergo acetylation (ac), which is a mutually exclusive mark from methylation and is frequently antagonistic to the chromatin state that is induced by methylation on the same residue[91]. Some of the residues that are acetylated include H3K4, H3K9, H3K14, H3K27, and H3K36, though the relationship between H3K4ac and H3K36ac and transcription is less well known[92]. These acetylation marks are deposited by histone acetyltransferases (HATs), and GEN5 and p/CAF are the two HATs responsible for depositing the majority of H3Kac modifications[81]. Another pair of HATs that are well studied in the context of development and disease are E1A binding protein p300 (p300) and its paralog Creb binding protein (CBP). P300/CBP can acetylate H2,H3, and H4, are involved in a variety of signaling pathways, and have important implications in malignancy, especially solid tumors[93]. There are two families of histone deacetylases (HDACs) including the HDAC family and the Sirtuin (SIRT) family of HDACs, however

the substrate specificity of these enzymes is not well understood[81]. Ubiquitination of lysine 120 on H2B (H2Bub) is another critical modification that is crucial for active the regulation of transcription[94]. This ubiquitination is deposited by the E3 ubiquitin ligase complex RNF20/40[95], in coordination with the E2 ubiquitin ligase RAD6[94]. H2Bub is removed by deubiquitinating enzymes such as USP22, which is a member of the Spt-Ada-Gcn5-acetyltransferase (SAGA) complex that acts a transcriptional co-activator[96]. Deubiquitination of H2B is likely important for nucleosome turnover following RNAPII dependent transcription[97,98]. In summary, there are many different proteins associated with modulating a wide array of PTMs on histones, and it is important to understand the functions of these modifications in the context of cellular development and malignancy.

Several epigenetic modifications are associated with transcriptional activation and elongation. H3K4me3, H2Bub, and H3K9ac are modifications that frequently accumulate at the promoter regions of genes that are actively transcribed by RNAPII[99]. H3K4me1 alone is a modification associated with poised enhancers, which are enhancers that are predetermined to be activated, such as lineage specific enhancers, but are not yet active[100]. H3K4me1 together with H3K27ac are found to mark active gene enhancers[101]. H3K79me2 is a mark associated with transcriptional elongation and is often seen marking the gene bodies of actively transcribed genes[99]. H3K36me2 is thought to be important for the initiation of RNAPII dependent transcription[85]. Together, these combinations of histone modifications are associated with active transcription.

In contrast, many histone modifications are also associated with repressed or silenced regions of the chromatin. H3K9me2/3 are associated with transcriptional repression and silencing at euchromatic genes[76,102]. These modifications also accumulate in heterochromatic regions of the genome, so they are important for both genome stability as well as transcriptional regulation. H3K27me3 is a repressive mark that is found at enhancer regions of the genome[101]. H3K36me3 is associated with posttranscriptional silencing of a gene and marks the gene body after RNAPII transcribes it[85]. H3K36me3 has also been shown to be present on heterochromatin, suggesting that it too plays a role in maintaining genomic stability and chromosomal structure[85]. With respect to transcription, epigenetic modifiers and modifications effect changes in chromatin states that range from permissive for active transcription to

silenced heterochromatin (Figure 1-2). Because transcription is such an essential aspect of regulating blood cell development, these various histone modifications and the proteins associated with modulating them are interesting both in the sense of understanding normal hematopoiesis as well as examining deregulation that results in malignancy.

The proteins that write, read, or erase histone modifications play critical functions in regulating transcription. In blood cell development, there is a tight transcriptional control that is required to maintain blood homeostasis, but there are also dynamic changes in gene programs that must take place for cells to make lineage decisions and differentiate. Deregulation of these dynamic processes can directly affect blood cell development and contributes to malignancy[21,66].

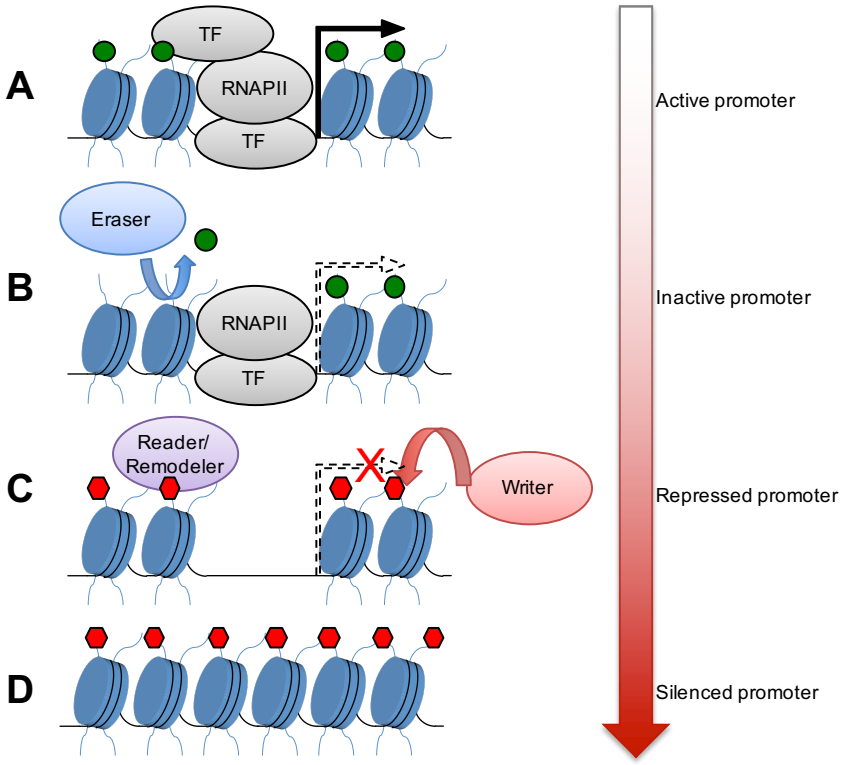


Figure 1-2: Epigenetic writers, readers, and erasers modulate chromatin states. A-D) Different states of chromatin permissibility to transcription are shown. Moving from A to D shows how an active gene might become silenced by epigenetic changes**.** A) An active RNAPII-bound gene promoter is shown. Two transcription factors (TF) that are important for active transcription are shown. B) An epigenetic eraser protein has removed an epigenetic modification that recruits one of the critical TFs to the promoter region. RNAPII cannot actively transcribe without this TF. C) An epigenetic writer protein is depositing a modification associated with repressing transcription. D) An epigenetic reader that functions by remodeling chromatin has been recruited by the repressive mark and has compacted the chromatin, silencing the promoter.

Polycomb group and trithorax group proteins: an example of dynamic epigenetic regulation

A well-studied example of dynamic epigenetic regulation of transcription involves the antagonistic relationship between Polycomb group (PcG) proteins and Trithorax group (TrxG) proteins. PcG proteins were identified as essential repressors of Homeobox (Hox) genes in *Drosophila*. *Hox* genes are essential regulators of proper anterior-posterior development in *Drosophila*, and they play critical roles in human development and malignancy[103,104]. PcG mutations lead to inappropriate expression of Hox genes, which lead to developmental defects in anterior-posterior segmentation of the flies[105,106]. Similarly, TrxG proteins were discovered when mutations in the genes encoding them were found to phenocopy a loss of *Hox* gene expression, and also lead to anterior-posterior defects[106,107]. These contrasting effects of perturbing PcG and TrxG proteins lead to the hypothesis that these proteins act antagonistically to each other.

These proteins are highly conserved from *Drosophila* to mammals, with an important example found in MLL1[108]. MLL1 was discovered primarily due to its tendency to take part in chromosomal translocations that are recurrent in leukemia, and it has a high homology to *Drosophila* TrxG proteins[108]. Similarly, Enhancer of zeste homolog 2 (EZH2) and its associated complex, Polycomb repressive complex 2 (PRC2) were found to function similarly to the *Drosophila* PcG proteins in regulating gene expression[106,109]. Indeed, these proteins act antagonistically to each other to regulate the transcription of genes, such as *HOX* genes, by deposition of histone modifications associated with different chromatin states[110]. MLL1 is an H3K4 trimethyltransferase, depositing a modification associated with active transcription[83]. Conversely, EZH2 deposits H3K27me3, a modification associated with repression of transcription[84]. Mutations or deregulation in either can lead to aberrant upregulation or downregulation of their regulatory targets. For instance, translocations in *MLL1*, which cause the hyperactivation of a subset of MLL1 targets, lead to aberrant upregulation of the *HOXA* cluster of genes in leukemias[7,111]. Similarly, low expression of *EZH2* in myelodysplastic syndrome (MDS) is associated with higher expression of the *HOXA-D* clusters of genes due to lower *HOX*-associated H3K27me3[112]. These data demonstrate a critical balance

between these protein groups is necessary for cell homeostasis, and that deregulation of this balance is associated with disease states. This is a good example of deregulated transcriptional and epigenetic regulation that is commonly associated with hematopoietic malignancy.

Transcriptional and epigenetic deregulation in acute myeloid leukemia

Genetic aberrations such as mutations or chromosomal aberrations in hematopoietic cells can lead to the initiation of hematologic malignancies[113,114]. These malignancies include myeloproliferative neoplasm (MPN), MDS, and acute or chronic leukemias, including acute lymphoblastic leukemia (ALL), chronic lymphoblastic leukemia (CLL), chronic myeloid leukemia (CML) and acute myeloid leukemia (AML)[113,114]. Specifically, many hematologic malignancies arise from the deregulation of gene programs in HSPCs[113]. Leukemia is a cancer that is initiated in blood cells and is associated with the deregulation of hematopoietic differentiation[115]. Leukemic initiation can result from a block in the expression of gene programs that are associated with driving differentiation and lineage specific functions of mature blood cells. Another aspect of leukemic initiation involves aberrant overexpression of gene programs that must be progressively downregulated as HSPCs differentiate into mature blood cells, such as pro-proliferative gene programs and genes associated with promoting self-renewal[116,117]. These genetic aberrations accumulate until there is the initiation of a leukemia that has a high capacity for self-renewal and is less capable of differentiating to mature blood cells[116,117].

AML develops when this deregulation of the hematopoietic differentiation process occurs during myeloid cell development due to a deregulation of specific gene programs[3,118]. Therefore, AML is a disease associated with transcriptional deregulation, including epigenetic deregulation. In fact, a recent whole genome sequencing study revealed that over 70% of AML patients harbor mutations in genes encoding epigenetic modifiers[1]. AML is an aggressive malignancy, with an overall five-year survival rate of only 26%, and is associated with a poor prognosis both children and adults[119]. Standard treatments for AML involve broadly cytotoxic chemotherapies, which affect the body's normal cells as well as the cancerous cells[119]. These therapies have severe,

debilitating, and even deadly side effects. Therefore, there is a clear unmet need for targeted therapies that allow for a more precise treatment of leukemia cells while reducing dangerous side effects associated with cytotoxicity. This relies on a strong understanding of the molecular mechanisms that lead to deregulated transcription in AML.

Transcription factor deregulation in AML

Myeloid-specific transcription factors are mutated in over 20% of AML[1], and aberrant expression of other transcription factors, like *HOXA* genes, occur in more than 50% of AMLs[2,37]. A few of the most commonly deregulated transcription factors in AML are reviewed here.

RUNX1

Runt related transcription factor 1 (RUNX1) is a highly conserved transcription factor that has been shown to be required for hematopoietic development by maintaining HSPC self-renewal capacity[120]. *RUNX1* exhibits loss of function mutations in AML, and mutated *RUNX1* is associated with poor prognosis in patients[120]. However, other groups have shown that RUNX1 promotes survival of murine AML cells as well[121]. These data suggest its function may be more dynamic than just suppressing cell growth

CEBPα

CEBPα is a myeloid specific transcription factor that is necessary for specific lineage fate decisions, such as the differentiation of an MPP to a GMP. CEBPα is also frequently mutated in AML. Interestingly *CEBPα* mutation is considered to be an independent marker for favorable prognosis in AM[122]. *CEBPα* has loss of function mutations in AML, but it is also required for leukemic transformation[123]. This suggests that the level of *CEBPα* is critically regulated in normal myeloid development and is also important for the establishment of myeloid malignancies[123].

Hoxa9/Meis1

HOXA9 is a member of the homeobox (HOX) family of proteins. These proteins are transcription factors that are highly conserved and are associated with cell development in many different tissues[103]. There are four orthologous clusters of *HOX* genes in the mammalian genome, labelled *HOXA/B/C/D*[37]. The HOXA/B proteins have been shown to be critical for normal hematopoietic development. For instance, HOXB4 is a regulator of HSC self-renewal capacity, and forced expression of it has been shown to expand HSC colony forming capacity *ex vivo*[45]. Similarly, *HOXA9* is highly expressed in HSPCs and skews commitment of embryonic stem cells to a hematopoietic lineage[36]. HOXA9 and its cofactor MEIS1 are associated with promoting self-renewal and are progressively downregulated during hematopoietic differentiation (Figure 1-3A) [36,37,124,125]. However, aberrant expression of *HOXA9* and *MEIS1* prolongs the self-renewal capacity of hematopoietic cells and blocks differentiation, resulting in

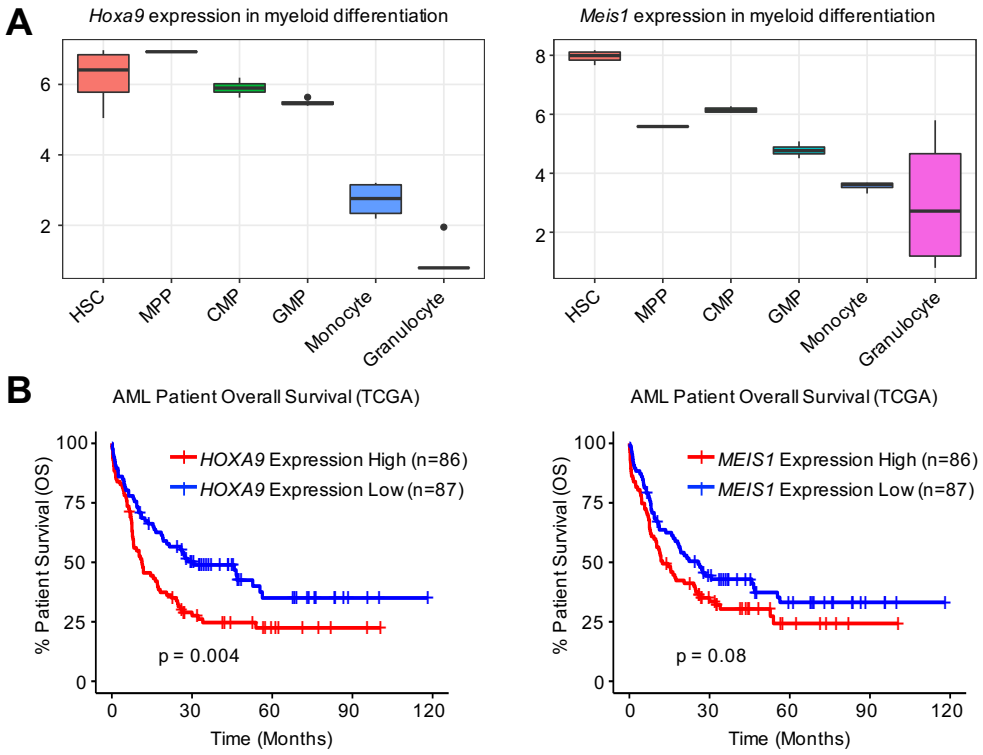


Figure 1-3: HOXA9 and MEIS1 are important regulators of self-renewal and are aberrantly upregulated in AML. A) RNA-seq expression values of *Hoxa9* and *Meis1* in different hematopoietic cell types. *Hoxa9* and *Meis1* are progressively downregulated as cells become more differentiated. B) TCGA AML patient survival data showing overall survival stratified by gene expression of HOXA9 or MEIS1. The patient samples are divided by those with expression above the median or below the median (n=173).

malignancy[36,124,125]. HOXA9 is of particular interest in the context of AML because it is aberrantly upregulated in over 50% of AMLs[2,37]. Furthermore, high expression of *HOXA9* and its cofactor *MEIS1* are considered to be poor prognostic markers in AML (Figure 1-3B)[3,4,11].

HOXA9 and MEIS1 are transcription factors that directly interact[4]. These proteins also bind to PBX3, which is an additional cofactor in HOXA9/MEIS1 driven transcription[126]. Importantly, they have been shown to share binding sites with myeloid specific transcription factors like PU.1[127] and CEBPα[128], which are also frequently deregulated in AML. HOXA9/MEIS1 transcriptionally activates targets, such as IGF-1, which has oncogenic functions in AML[129]. They are also shown to play a role in repression of target genes, like *ARF* and *INK4*, which are tumor suppressors in AML[130]. This demonstrates a two different molecular functions for HOXA9/MEIS1 that both drive leukemogenesis. Therefore, therapeutically targeting factors that control *HOXA9* and *MEIS1* transcription is of biomedical interest.

HOXA9/MEIS1 can be aberrantly upregulated through a variety of mechanisms. One common mechanism of upregulation is mutation of *Nucleophosmin* 1 (*NPM1*). These mutations lead to the cytoplasmic accumulation of NPM1, which drives *HOXA9* and *MEIS1* expression, though the mechanism not clearly understood[131]. *HOXA9* is also upregulated due to loss of function mutations in the epigenetic repressive *EZH2* gene, a PcG protein known to repress *HOX* gene expression[112]. Another example of this is found in inactivating mutations of *ASXL1*, the encoded protein of which has been shown to associate with the EZH2 containing complex PRC2 and drive repression of *HOXA9*[132]. Finally, various chromosomal rearrangements have been shown to be associated with driving high *HOXA9* and *MEIS1* expression. These rearrangement events include NUP98 translocations, rearrangements resulting in expression of the CALM-AF10 fusion, and translocations of the TrxG homolog *MLL1*[2,37]. Due to their frequency and associated poor prognosis, MLL rearrangements are of particular interest in the development of AML and will be further reviewed here. The aberrant expression of *HOX* genes in malignancy is a good example of how cancer associated deregulation of transcription factors is frequently tied to the deregulation of epigenetic modifiers.

Epigenetic deregulation in AML

Epigenetic deregulation has emerged as an important contributor to oncogenesis and disease progression in a variety of malignancies, including leukemia[69,133–135]. One class of epigenetic regulation implicated in blood malignancies involves the post translational modifications of histones[134,136–138]. Histone modifying proteins are commonly mutated, rearranged, aberrantly expressed and required in leukemias, providing potential therapeutic targets. Several of the most commonly deregulated epigenetic modifiers in AML are reviewed here.

KDM6A

KMD6A (UTX) exhibits the highest mutation rate of any demethylase in hematopoietic malignancies[139,140]. KDM6A is an H3K27 demethylase[73]. KDM6A exhibits loss of function mutations in myeloid malignancies. Normally, KDM6A functions by repressing gene programs driven by the transcription factor ETS domain containing proteins, which are pioneering transcription factors that include the myeloid specific factor PU.1[141]. These factors can be oncogenic when aberrantly expressed[141]. Further, KDM6A upregulates tumor suppressor genes, such as GATA protein driven genes[141]. Additionally, low expression of KDM6A are associated with relapse of AML after chemotherapy treatment[142]. These data implicate KDM6A as a tumor suppressor in myeloid malignancies.

EZH2

EZH2 is a PRC2 associated H3K27 methyltransferase that is associated with repressing gene targets[143]. EZH2 exhibits loss of function mutations in MDS and AML, suggesting that it has tumor suppressive functions[84]. As discussed previously, EZH2 normally functions by repressing HOX gene expression. In myeloid malignancies EZH2 is shown to repress expression of oncogenes such as HOXA9[37,144]. However, EZH2 is also a potential therapeutic target in MLL-fusion driven AML, where it is required for cell growth[145]. Interestingly, depletion of EZH2 does not affect HOXA9 levels in MLL-rearranged leukemia and instead induces differentiation by relieving repression of differentiation associated genes, such as G-

CSF[145,146]. These data suggest that EZH2 plays an important but context dependent role in myeloid malignancies.

MLL translocations

MLL1 is a conserved histone methyltransferase of the SET domain superfamily of methyltransferases. The Su[var]3-9, enhancer of zeste, trihorax (SET) domain found on the C-terminal domain of MLL1 is a catalytic domain that is responsible for the trimethylation of H3K4. MLL1 also contains a CxxC domain that is critical for its DNA binding function[147–149] and four plant homeodomain (PHD) fingers that are important for protein-protein interactions, localization, and recognition/ binding to H3K4me3[12,150,151]. Yu and colleagues demonstrated that MLL1 is required for embryonic development, and loss of *MLL1* leads to defects in the skeletal and hematopoietic systems, among others[152]. Interestingly, the authors show that MLL1 is required for maintained expression of target genes such as Hoxa7, but is not required for their activation, suggesting it plays a role in transcriptional memory/maintenance. MLL1 has been shown to be critical for adult hematopoiesis, where MLL1 null bone marrow cells are incapable of reconstituting lethally irradiated bone marrow, suggesting that MLL1 is absolutely required for HSC maintenance[153]. Furthermore, MLL1 is shown to function as a critical regulator of *Hox* genes in normal hematopoietic development[154]. As previously mentioned, *Hox* genes are crucial genes in the development of many tissues, including the blood system, and play an important role in the self-renewal capacity of HSPCs[104,153,154]. Remarkably, the self-renewal capacity of HSPCs can be rescued in MLL null cells by re-expression various *Hox* genes individually, such as *Hoxa9*[153]. This suggests that the regulation of *Hox* gene expression is a primary function of MLL1 in HSPCs. Mechanistically, MLL1 is one side of an antagonistic regulation duo that includes trithorax group proteins, such as the MLL1 containing complex COMPASS, and polycomb group proteins, such as the EZH2 containing complex PRC2[42]. While COMPASS is responsible for the deposition of H3K4me3 to effect transcriptional activation, PRC2 deposits H3K27me3, a mark associated with repression of target genes[42]. Interestingly, recent work has suggested that the

methyltransferase activity of MLL1 is not required for hematopoiesis, but its critical function is the recruitment of MOF, an H4K16 acetyltransferase, to *Hox* genes to maintain active transcription[155]. Regardless, MLL1 and COMPASS appear to be requisite in the maintenance of HSPC self-renewal capacity. The expression of *Hox* genes, therefore, is dependent on the balance or lack thereof of the trithorax group protein or polycomb group protein levels and capacities to localize to target genes. In summary, MLL1 is a dynamic regulator of *HOX* genes and its regulation is crucial to maintaining the appropriate expression of *HOX* genes in HSPCs. Therefore, deregulation of MLL1 has dramatic consequences in hematopoiesis.

MLL1 is of particular interest in the field of hematology due to its high rate of translocation in leukemia. MLL rearrangements are present in approximately 10% of leukemias overall, with a disproportionate number of infant leukemias and therapy related leukemias represented[156,157]. Between 70-80% of infant leukemias and 70% of therapy related leukemias harbor an MLL translocation[156]. Leukemias harboring these translocations are associated with a particularly poor prognosis, and little progress has been made in developing successful therapies to combat MLL-fusion driven leukemia[158]. MLL1 contains a region of high genomic instability that leads to frequent DNA double strand breaks that can then lead to rearrangements with other chromosomes, which are known as 11q23 chromosomal rearrangements[159]. This rearrangement leads to MLL-fusion transcripts, which in turn are translated to oncogenic MLL-fusion proteins which consist of the N-terminal domain of MLL and one of over eighty different fusion partners[160]. Some of the most common fusion partners of MLL include AF4, AF9, AF10, ENL, and ELL[160]. Another common mechanism of MLL rearrangement is partial tandem duplications (PTD), wherein exons 5-11/12 of MLL1 are duplicated and re-inserted into the MLL1 locus[161]. These rearrangement events can lead to the development of an AML, ALL or mixed lineage leukemia.

MLL rearrangements have been studied extensively in the context of AML. Though there are many fusion partners for MLL translocations, the mechanism by which these rearrangements drive malignancy is consistently observed as an upregulation of the *HOXA* cluster of genes, *MEIS1*, and other genes associated

with self-renewal[13,104,162]. The manner in which these genes are upregulated relies on the protein-protein interactions formed with the MLL-fusion proteins. Given the variability of fusion partners that are found in MLL rearrangements, there is likely not just one protein complex formed with MLL-fusions, but rather many different complexes that are context dependent[156,157,160]. Some of the known protein interactors that play a critical role in MLL rearranged leukemogenesis include Menin (MEN1) and Lens epithelial derived growth factor (LEDGF), with MEN1 serving as an interaction that connects MLL to the chromatin binding LEDGF protein[163,164]. Yokoyama and colleagues demonstrated that both of these interactions were required for MLL-ENL driven leukemia, and that the MLL-fusion-MEN1-LEDGF interaction is conserved regardless of fusion partner[163,164]. Additionally, some of the most common fusion partners of MLL are a part of transcriptional regulating super complex of proteins. The bona-fide members of this complex are disputed, though many groups have reported common proteins that exist in this complex. This complex is known by different groups as the elongation assisting proteins (EAP) compex[165], the super elongation complex (SEC)[166] or the AF4 family/ENL family/P-TEFb (AEP) complex[167]. Some of the proposed members of the complex include ENL, AF4 and AF5, all of which are translocation partners of MLL. Additional proteins that associated with this complex include the Positive transcription elongation factor (P-TEFb)[167]; DOT1-like (DOT1L), which is an H3K79 methyltransferase associated with active transcription[165]; the PAF1c, which is an epigenetic co-modifying complex that is associated with transcriptional regulation[13]; and Bromodomain containting protein 4 (BRD4), which is a chromatin binding protein that is connected to MLL through P-TEFb and the PAF1c. Notably, BRD4 has been demonstrated as a valuable therapeutic target to treat MLL-fusion driven leukemias[168,169]. These protein-protein interactions are interesting in the context of therapeutic development and understanding the dynamic interactions that regulate MLL-fusion driven leukemia. Of particular interest is the PAF1c, because it has been recently shown that the PAF1c-MLL interaction is critical for MLL-driven leukemias but is dispensable for normal hematopoietic cell growth[14]. The PAF1c will be more thoroughly reviewed here.

1.2 The Polymerase Associated Factor Complex in Development and Disease

The polymerase associated factor complex (PAF1c) is an epigenetic co-modifying complex and transcription co-factor that plays an important role in the regulation of a large subset of genes in both yeast and mammals[16,170–172]. The PAF1c, since its discovery in yeast and subsequent discovery in mammals, has been associated with regulating both activation and repression of transcription in development, and has been shown to play a critical role in a variety of malignancies[99].

The PAF1c is a transcriptional co-regulating complex

The PAF1c is a highly conserved, six-subunit protein complex that is involved in a variety of transcriptional processes. The PAF1c was first discovered in yeast as a RNAPII associated complex. The proteins that make up the PAF1c include Polymerase Associated Factor 1 (PAF1), Cell Division Cycle 73 (CDC73), Left Open Reading Frame 1 (LEO1), and Cin Three Requiring 9 (CTR9)[173–177]. Restore TBF Function 1 (RTF1) is a core component of the yeast PAF1c, but has been found to be a more transient interacting partner of the complex that has independent functions in mammals[16,177,178]. Additionally, there is a mammalian specific PAF1c component that is not found in the yeast complex, WD Repeat Domain 61 (WDR61/SKI8)[179] (Figure 1-4).

The PAF1c binds directly to both the phosphorylated and unphosphorylated tail of RNAPII[177,180–182]. Cryogenic electron microscopy (CryoEM) on yeast PAF1c proteins demonstrates that this interaction consists of multiple contacts between the PAF1c and RNAPII. First, the heterodimer formed between Paf1 and Leo1 binds to the RNAPII subunit Rpb2[181]. A second set of interactions consists of an interaction between Cdc73 and the RNAPII subunits Rpb3/11[181]. These two separate interaction domains are bridged by Ctr9, so that the PAF1c covers a large surface of the external RNAPII machinery[181]. We can speculate that this structure may present a good platform for recruitment of the epigenetic modifiers that are known to play a critical role in PAF1c mediated regulation of transcription[99]. The PAF1c is also known to bind to a variety of RNAPII associated transcription factors, such as the DRB sensitivity inducing factor (DSIF) complex, which contains SPT4/5 proteins and is a critical regulator of all RNAPII

dependent transcription[181–183]; pTEFB, which is a positive regulator of transcriptional elongation[184]; the Facilitates chromatin transcription (FACT) complex, which is a chromatin remodeling complex[185,186]; and others. While the order in which these various transcriptional components are recruited to sites of transcription is not well understood, it is clear through these interactions that the PAF1c is a central component of the transcriptional machinery.

While the subunits of the PAF1c have no known catalytic activity, this complex plays a critical role in the deposition of several histone modifications impacting RNAPII-dependent transcription. The PAF1c acts as a platform to recruit epigenetic modifiers to transcription start sites and allowing these enzymes to deposit histone modifications in order to affect the histone code and chromatin conformation at the given loci. The PAF1c has been shown to be important for the deposition of several histone

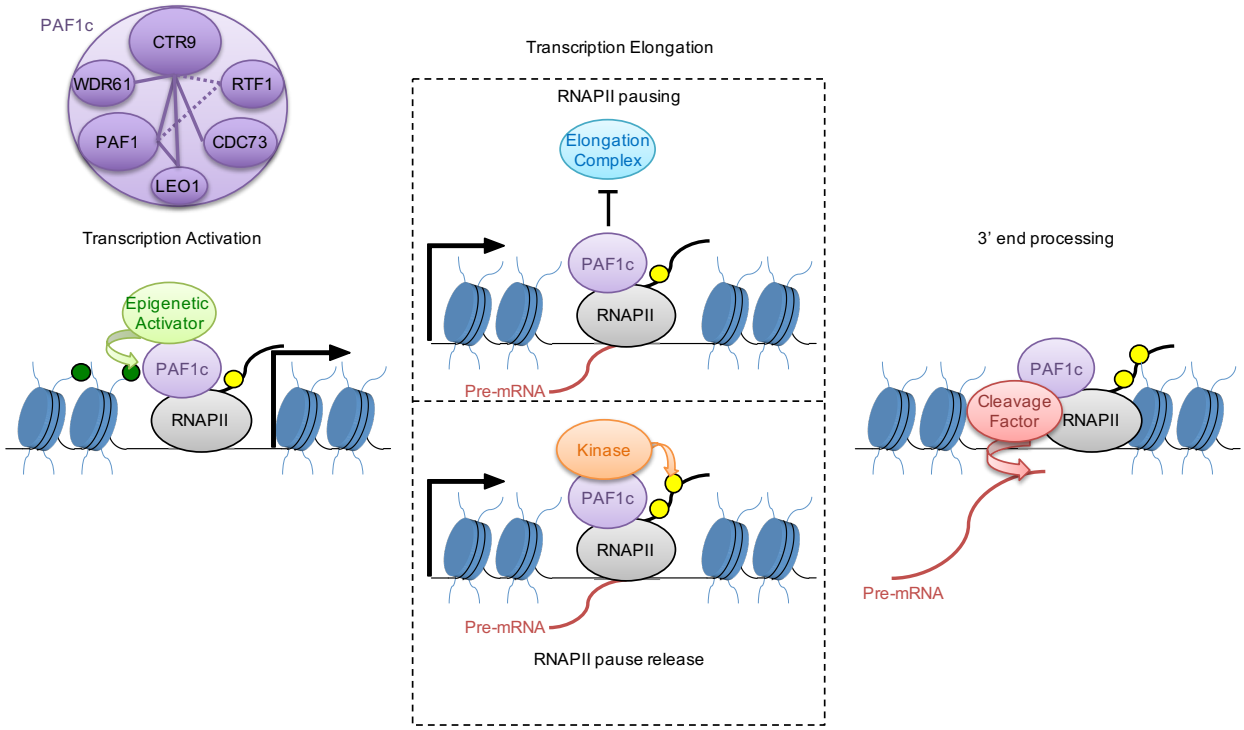


Figure 1-4: The PAF1c regulates transcription. Left, top) The PAF1c is a six-subunit complex in mammals. WDR61 is mammalian specific. RTF1 is a more transient interaction partner with the PAF1c in mammals (dotted lines). Left, bottom) The PAF1c can activate transcription by recruiting epigenetic modifiers to promoters, where they deposit a modification that leads to RNAPII-dependent transcription activation. Middle, top) The PAF1c can stabilize RNAPII pausing by inhibiting an interaction between RNAPII and elongation machinery required for elongation. Middle, bottom) The PAF1c can release RNAPII pausing by recruiting proteins that phosphorylate the C-terminal tail of RNAPII (black curved line), which is required for transcriptional elongation. Right) The PAF1c plays a role in mediating 3' end processing of pre-mRNA (red curved line) by recruiting proteins required for cleavage and polyadenylation.

modifications associated with transcriptional activation. For example, the PAF1c interacts with the H3K4 methyltransferases Set1 (yeast) or MLL1 (mammals) and the associated COMPASS complexes to recruit promoter H3K4me2/3, which is a modification correlated to active transcription[12,13,187]. Both yeast and mammalian PAF1c have also been shown to be required for Dot1/ DOT1L mediated H3K79me2, which marks the gene bodies of actively transcribed genes[188]. Additionally, the Rtf1/ RTF1 subunit makes direct contact with the E2/E3 ubiquitin ligases Rad6 and Bre1 or their human homologs RNF20/40 to direct H2B K123 (or K120 in mammals) ubiquitination (H2Bub) of target gene loci[94,95,187,189–192]. This H2Bub modification is broadly required for active transcription. There has also been evidence linking the PAF1c to Mof-mediated H4K16ac and Set2-mediated H3K36me in yeast, both of which are also associated with an active transcription state[16] (Figure 1-4).

Recently, the PAF1c has also been linked to promoter proximal RNAPII pausing[170–172]. In mammals, pausing is a process in which RNAPII transcription is activated, but the polymerase and machinery only process between 20-60 nucleotides downstream of the transcription start site before pausing[170]. While the importance and function of pausing still needs further study, there are several pieces of evidence that point to this as an additional mechanism of regulation. Pausing has been suggested to play a role in synchronizing expression of genes, keeping the chromatin around transcription start sites accessible to transcriptional machinery. This serves as a checkpoint separate from transcriptional activation and elongation during which the transcriptional output of the gene can be modulated based on cues from the cell[193]. The PAF1c has been proposed to play a role in both releasing RNAPII pausing[170] and stabilizing RNAPII pausing[171,172]. This is interesting because it suggests that the PAF1c is involved in both promoting and repressing transcription, which suggests a dynamic regulation of PAF1c regulated genes (Figure 1-4).

The PAF1c also plays a role in 3’ end processing, which includes a series of key posttranscriptional processes that are crucial for proper regulation of mRNA products. The PAF1c was identified in yeast as being critical to the 3’ end processing of mRNA products (Figure 1-4). Deletion of the *PAF1* subunit in yeast revealed a dysregulation in polyadenylation (polyA) site usage or a shortening of polyA tails, which leads to

unstable transcripts that were prone to nonsense mediated decay[194,195]. polyA tails are deregulated upon loss of Paf1 in part due to a loss of recruitment of the cleavage factor and polyadenylation factor Pcf11[195]. Furthermore, the CDC73 subunit of the PAF1c was found to be a part of the posttranscriptional machinery complex in human cells[196]. CDC73 was additionally shown to directly interact with CSTF and CPSF, which are two complexes associated with cleavage and polyadenylation of mRNA transcripts[197]. Depletion of CDC73 leads to decreased mRNA product through decreased association of these posttranscriptional processing proteins with mRNA[197]. Importantly, the PAF1c is not a general transcription factor, as only a subset of genes have perturbed expression after depletion of the PAF1c[99,170,198]. In summary, the PAF1c plays many roles in regulating transcriptional processes.

The PAF1c in malignancy

The subunits of the PAF1c and the complex itself have been demonstrated to play critical roles in a variety of cancers. Interestingly, the role of the complex in disease states appears to be context dependent, functioning as both a tumor suppressor or an oncogene depending on the tissue.

Solid tissue malignancies

Perhaps the most well studies subunit of the PAF1c in the context of disease is the *CDC73* gene, which encodes Parafibromin. Over 70% of patients with hyperparathyroidism jaw tumor (HPT-JT) exhibit mutations in the subunit *CDC73* which are usually a loss of heterozygosity mutation[199]. *CDC73* is also exhibits loss of function mutations in other endocrine cancers as well, including familial isolated hyperparathyroidism[200]. Mechanistically, CDC73 acts as a tumor suppressor by mediating the repression of Cyclin D1 (*CCND1*) and *MYC* through recruitment of the H3K9 methyltransferase SUV39H1[201–203]. Loss of *CDC73* allows for aberrant upregulation of *CCND1* and *MYC*, which are critical oncogenes in many subtypes of cancer, including HPT-JT[201–203]. Further, CDC73 is a tumor suppressor in Wilms Tumor, a kidney cancer that is highly linked to HPT-JT. In these tumors, *CDC73* also exhibits loss of heterozygosity mutations, one of which was functionally shown to cause a loss of

CDC73 mediated *CCND1* repression[204]. CDC73 has also been shown to act as a tumor suppressor in oral squamous cell carcinoma, where it is repressed by miRNA-155 and its overexpression leads to reduced cell growth[205]. It acts as a tumor suppressor in ovarian cancer, where overexpression of CDC73 leads to reduced cancer cell growth and reduced metastasis associated phenotypes[206]. CDC73 protein levels have also been shown to inversely correlate with tumor size in breast cancer[207]. These data suggest a role for CDC73 as a tumor suppressor in a wide variety of cancers. Recently, however, CDC73 has been shown to act as an oncogene in gastric adenocarcinoma cells[208,209]. When dephosphorylated by SHP2 phosphatase, CDC73 displays enhanced binding to β-catenin[210], Notch Intracellular Containing Domain (NICD) and GLI1. This drives increased Wnt, Notch, or Hedgehog signaling, respectively[208,209]. Aberrant Wnt, Notch, and Hedgehog signaling are strongly associated with oncogenesis in several different types of cancer. The implication of these studies is that CDC73, and the PAF1c, are dynamically regulated.

The PAF1 subunit also plays a critical role in malignancy and has been most well-studied in the context of pancreatic cancer. *PAF1* is depleted in early neoplastic

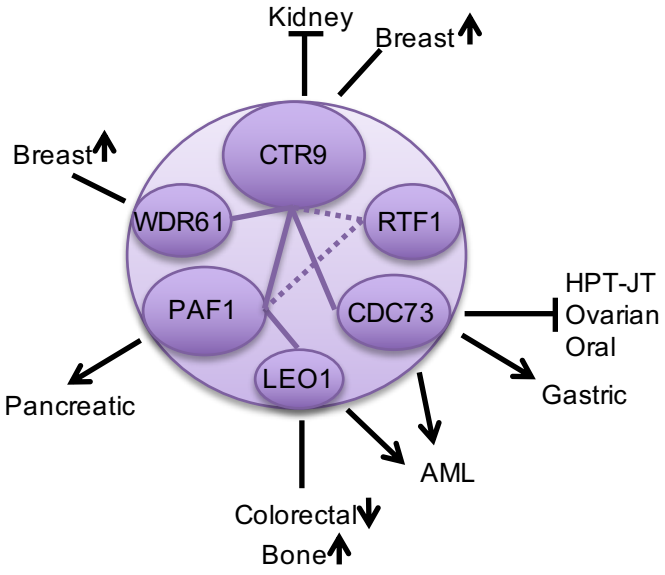


Figure 1-5: The PAF1c in malignancy. Here is shown a summary of the known roles or implications of the subunits of the PAF1c in different types of cancers. An arrow pointed to a subtype indicates that the subunit exhibits oncogenic functions in that cancer. An inhibitory pointer indicates that there are potential tumor suppressor functions for that subunit in that disease subtype. A small arrow pointing up or down indicates that the subunit is amplified/upregulated or deleted/downregulated recurrently in that malignancy subtype.

cells during the initiation of pancreatic malignancy[211]. However, PAF1 was initially connected to pancreatic malignancy because it is located on chromosome 19q13, which is a region that is recurrently amplified in pancreatic adenocarcinoma[212]. Functionally, the overexpression of PAF1 in NIH3T3 fibroblast cells enhances cellular proliferation *ex vivo* and tumor growth *in vivo,* demonstrating that PAF1 can drive malignant transformation of cells[212]. This initial reduction in expression followed by amplification of expression during the progression of pancreatic malignancy suggests that PAF1 may regulate gene programs that initiate tumorigenesis and genes that can further drive malignant progression[211]. Further, there is high expression of PAF1 in pancreatic ductal adenocarcinoma (PDAC), which drives tumor growth and metastasis by upregulating oncogenes like *CCND1* and *MYC*[213]. Similarly, PAF1 drives *MYC* expression in non-small cell lung cancer, where high expression of PAF1 correlates with poor prognosis in patients[214]. Interestingly, *MYC* and *CCND1* are repressed by CDC73, suggesting that there may be a more dynamic PAF1c-centric regulation of these genes that can be dysregulated in different ways depending on the cellular context. *PAF1* is also highly expressed in pancreatic and ovarian cancer stem cells (CSCs), where it drives self-renewal by interacting with POU Class 5 Homeobox 1 (POU5F1), a transcription factor that is responsible for the expression of gene programs associated with self-renewal[215,216]. In summary, the PAF1 subunit appears to primarily play a role as an oncogene in solid tissue cancers. However, the dynamic interplay between the oncogenic functions of PAF1 and the tumor suppressor roles of other PAF1c subunits has not been well studied.

While CDC73 and PAF1 are the two most well-studied subunits of the PAF1c, there have also been links between malignancy and the other subunits of the PAF1c. CTR9 is an important regulator of ER positive breast cancer by recruiting ERα and RNAPII to target genes in response to estrogen[217,218]. In contrast, a recent patient screen identified loss function mutations in *CTR9* in Wilms Tumor, suggesting inactivating mutations in *CTR9* predispose patients to developing Wilms Tumor[219]. In colorectal cancers, the chromosomal location for *LEO1* is frequently lost while in malignant bone histiocytoma this chromosomal region is frequently amplified[220,221]. A small patient sample study identified increased WDR61 protein in breast cancer

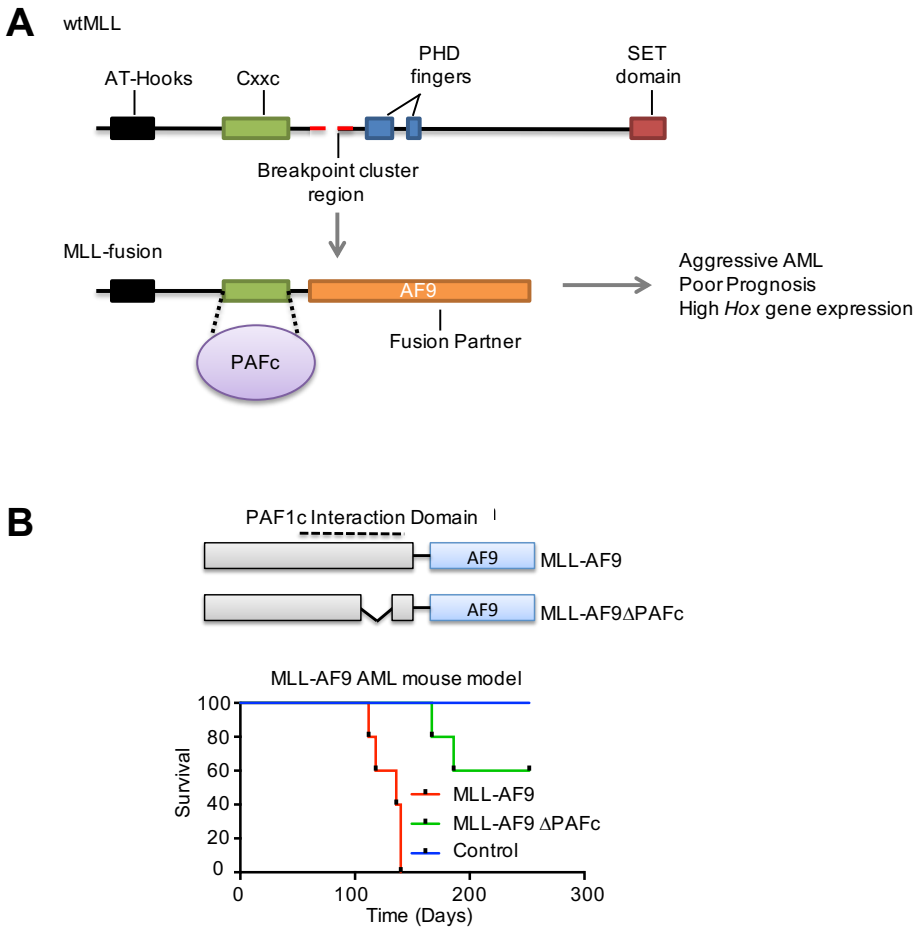


Figure 1-6: The PAF1c interacts with MLL-fusion partners and is critical for MLL-fusion driven leukemia. A) Here is shown a schematic of wildype MLL and MLL-fusion proteins. MLL has a breakpoint cluster region that can undergo double strand DNA breaks and fuse to another chromosome, forming an MLL-fusion gene that translates to an MLL-fusion protein. These proteins drive aggressive leukemias. A recent proteomics study by our lab found that the PAF1c binds to MLL and MLL-fusion proteins. B) Introduction of an MLL-AF9 transcript that lacks the PAF1c interaction domain delays AML disease latency in a mouse model of AML compared to full length MLL-AF9.

compared to its normal tissue counterpart[222]. RTF1 regulates the Notch signaling pathway, which is deregulated in cancer[223]. Taken together, these studies show that the PAF1c is an important regulatory complex in a variety of solid tissue cancers and can be deregulated in different ways depending on cellular context.

Hematopoietic malignancies

There have been several studies that establish a critical role for the PAF1c in hematopoietic cancers. In the context of MLL-fusion protein driven AML, there is a direct physical interaction between the PAF1c and MLL1 protein, which regulates the AML associated *HOXA* cluster of genes in hematopoiesis[12,13]. Not only does the PAF1c interact with wildtype MLL1, but it also binds to MLL-fusion proteins[12,13].

Specifically, the PAF1c is required for the deposition of H3K4me2/3, H3K79me2, and H2Bub at the *Hoxa9* or *Meis1* loci in MLL-AF9 driven leukemia. Interestingly, Milne and colleagues also demonstrated that recruitment of wildtype MLL1 protein by the PAF1c was required for the displacement of the H3K9 methyltransferase SETDB1 from the *Hoxa9* promoter, allowing for the recruitment of MLL-fusion proteins to the *Hoxa9* promoter by the PAF1c[12]. Furthermore, the PAF1c was shown to be required for AML, as deletion of the *Cdc73* subunit led to a dramatic loss of murine AML cell growth[14,224]. Additionally, disruption of the PAF1c-MLL interaction leads to reduced *Hox* gene activation and inhibition of MLL-AF9 driven AML growth[14]. This data shows that the PAF1c plays a critical role in MLL-fusion-mediated regulation of *Hox* genes in AML.

In a separate model of MLL-fusion driven AML, Hetzner, et al. identified an interaction between PAF1 and ENL, which is a YEATS domain containing protein that is also a common fusion partner in MLL-translocations[225]. This study showed that PAF1 makes physical contacts with both MLL and ENL in the context of an MLL-ENL fusion protein, leading to hyperactivation of target genes, such as *Hoxa9* and *Meis1*. This provides a novel mechanism in which the PAF1c is important for AML growth by regulation of the SEC in the context of MLL-ENL driven leukemia[225].

The PAF1c plays an important role in a different subtype of AML where overexpression of Phosphatase of regenerating liver-3 (PRL-3) drives oncogenesis. A series of studies show that LEO1 is a critical downstream target of PRL-3, an oncogene that is overexpressed in more than 50% of AML[226]. PRL-3 mediates the de-repression of *Leo1* in murine AML cells by stabilizing JMJD2C, an H3K9 demethylase, at the promoter region of *Leo1*. Loss of *Leo1* destabilizes the PAF1c and leads to reduced AML growth. Additionally, PRL-3 dephosphorylates LEO1, which promotes its interaction with β-catenin and helps drive aberrant Wnt signaling in AML cells[227]. This is especially interesting given interaction between dephosphorylated CDC73 and β-catenin in other cellular contexts[208,209]. In summary, the PAF1c is critical to a variety of AML subtypes and regulates many different oncogenic gene programs.

The PAF1c as a potential therapeutic target

Because of the roles described above, the PAF1c is an important target for the possible therapeutic development. There are several different ways in which the PAF1c may be targeted, many of which require further elucidation of the structure-function relationship of the subunits and the complex as a whole. One manner in which the complex may be targeted is by interrupting the recruitment to DNA. CDC73 has a Ras-like domain that is important for the association of the PAF1c with chromatin[180], and PAF1 as well as CDC73 are critical for the PAF1c interaction with the C-terminal tail of RNAPII[180,181,228]. Either of these interaction domains may be a potential therapeutic target. Another possible target lies in the interactions of the complex itself. Destabilization of the PAF1c complex inhibits malignant cell growth. This particular strategy must be pursued with caution, as the PAF1c is important for normal development in yeast and mammals[16], so a specific therapeutic window would have to be elucidated. Another manner in which to approach targeting the PAF1c is to disrupt the protein-protein interactions with the PAF1c instead of the complex itself. One example of this has already been discussed, as targeted disruption of the PAF1c-MLL interaction is effective in inhibiting AML cell growth. Importantly, disruption of this interaction did not have an effect on normal hematopoietic reconstitution, suggesting that there may be less of a requirement for the PAF1c-MLL interaction in normal HSPCs relative to AML cells. This provides a potential therapeutic window allowing for intervention that has a stronger effect on suppressing cancer cell growth as opposed to normal cell development. Given the many other protein-protein interactions that have been discovered with the PAF1c, this may provide the groundwork for different ways to target the regulatory functions of the PAF1c that are specifically deregulated in malignancy.

In summary, the PAF1c is an important regulatory complex in the context of both normal transcription and in cancer. Aberrant expression of the PAF1c is correlated with various types of cancer, making it an interesting complex to study in the context of disease. However, there are still aspects of its regulation in various disease states, including AML, that are not fully understood. Specifically, though the PAF1c is known to act as a platform by recruiting other proteins to promoters, its full protein-protein

interactome in the context of disease remains relatively understudied. Protein-protein interactions that are either disrupted or stabilized could have a dramatic effect on PAF1c transcriptional targets and may lead to malignancy.

One particular aspect of the PAF1c that will be interesting to explore is its interactions with proteins associated with transcriptional repression. In this study, we will provide new evidence that the PAF1c interacts with several different H3K9 methyltransferases, providing strong support linking the PAF1c to H3K9 methylation. While CDC73 has been shown to interact with the H3K9 methyltransferase SUV39H1[201], the role this interaction plays in the many malignancies that have implied roles for the PAF1c are not fully understood and should be further elucidated. Further, it is interesting to consider potential interactions with the PAF1c and epigenetic repressor proteins, as the epigenetic modifiers that interact with the PAF1c are largely thought to be epigenetic activators. Additionally, H3K9 methylation is a modification found at promoters of genes associated with development[229], as is the PAF1c. Further investigation into H3K9 methylation in the context of the PAF1c could provide important insight into how the PAF1c regulates its target genes in normal development, including those that are associated with self-renewal and must be downregulated during hematopoietic differentiation, such as *Hoxa9* and *Meis1*. Furthermore, this could provide further insight into the mechanism of deregulation of these PAF1c target genes in the context of disease.

1.3 H3K9 methylation in normal and malignant hematopoiesis

H3K9 methylation is an epigenetic modification that is generally associated with gene repression and chromatin condensation[76,230]. H3K9 methylation also plays a role in the regulation of various cellular processes. A primary cellular process affected by H3K9 methylation is the formation of heterochromatin. H3K9 methylation and the proteins associated with depositing H3K9 methylation are known to accumulate on heterochromatic regions of the genome[231,232]. H3K9 methylation is a unique histone modification because it marks large regions of the genome, such as pericentric heterochromatin, but it also has been shown to play a functional role in the dynamic repression of genes in euchromatic regions of the genome as well[75,102,230,233–236].

Proteins associated with deposition of H3K9 methylation come from two different families of proteins: the SUV39 family and the PRDM family. The SUV39 family of H3K9 methyltransferases consists of SUV39H1/2[237–239], G9a/GLP (*EHMT1/2*)[240–243], and SETDB1/2[244,245]. The H3K9 methyltransferases in the PRDM family are PRDM2[246], PRDM3[247], PRDM8[248], and PRDM16[247]. Proteins associated with removing the H3K9 methylation modification include the JMJD1/2 families of proteins, as well as PHF2/8 proteins. H3K9 methylation is critical for genome stability, chromatin conformation, and regulation of gene expression. Therefore, it is a critical mode of regulation in both normal and malignant development. The most well-known enzymes associated with H3K9 methylation are reviewed here, with a particular focus on H3K9 methyltransferases, due to the potential importance of their interactions with the PAF1c in malignancy.

H3K9 methylation in solid tissue malignancies

Many of the proteins associated with modulating H3K9 methylation have been shown to be aberrantly regulated in solid tissue malignancies. For example, SETDB1 has been implicated as an oncogene in a variety of malignancies, including melanoma, breast cancer, liver cancer, and lung cancer[249–253]. G9a/GLP(EHMT2/1) have also been proposed to promote oncogenesis in ovarian, breast, and lung cancers[77,254–256]. However, G9a has also been proposed to activate TP53 tumor suppressor protein in lung cancer, indicative that its function may be highly dependent on cellular context[257]. Similarly, SUV39H1 has been shown to help drive melanoma progression[258]. Interestingly, SUV39H1 has also been proposed to be a tumor suppressor through its interaction with the PAF1c[201,203]. The H3K9 lysine demethylases also have implications in solid tumor malignancies, though the roles are more ambiguous. For instance, PHF8 has been implicated as an oncogene in breast cancer[259], while PHF2 has been shown to be a inhibit breast cancer cell proliferation[260]. The JMJD2 family of proteins has been associated with oncogenic functions in breast cancer and lung cancer[261,262], whereas the JMJD1 family has tumor suppressor functions in germ cell tumors[263]. In summary, there is no clearly defined role for H3K9 methylation in solid tissue malignancy, and therefore each individual disease state must be looked at separately to begin to

elucidate the overall importance of H3K9 methylation in the initiation and progression of disease. This is also the case in hematological malignancies. To begin to gain a better picture of the different roles of H3K9 methylation in normal and malignant hematopoiesis, the known functions of the different H3K9 methylation associated enzymes are reviewed here.

H3K9me2/3 in normal HSPC self-renewal

Regulation of H3K9me2/3 plays a critical role in the maintenance of normal HSPC growth. For instance, mouse hematopoietic cells redistribute H3K9me3 during differentiation. While the overall amount of H3K9me3 does not change, the modification shifts to a radial nuclear distribution in more mature blood cells. In mice, normal HSPCs defined by Lin-, Kit+, and Ska1+ (LSK) that are treated with a G9a inhibitor (UNC0638) maintain a higher number of LSKs in culture compared to vehicle treated cells. Similarly, primitive CD34+ human hematopoietic cells that are treated with UNC0638 maintain a higher percentage of CD34+ cells and have a higher engraftment capacity in NOD-SCID mice. Taken together, these data suggest a role for G9a-dependent H3K9 methylation in repressing genes associated with self-renewal in normal HSPCs.

H3K9 methyltransferases in normal and malignant hematopoiesis

SETDB1/2

Recently, SETDB1 has been linked to critical functions in normal and malignant hematopoiesis. First, Koide and colleagues demonstrated that *Setdb1* is required for both hematopoietic cell growth as well as MLL-AF9 driven AML in murine cells[264]. Cells with genetically deleted *Setdb1* are incapable of reconstituting the bone marrow of lethally irradiated mice. Additionally, deletion of *Setdb1* leads to extended disease latency in an AML mouse model. Interestingly, this study found that *Setdb1* depletion induced apoptosis in hematopoietic cells by dysregulation of non-hematopoietic specific genes, such as metabolism-associated genes like *Fbp1/2*[264]. SETDB1 also directly interacts with PML[265]. PML is a protein that is associated with tumor suppressor functions in promyelocytic leukemia (PML). PML localizes to PML-nuclear bodies (PML-

NBs) where SETDB1 is required for their stability and function, including repression of PML target genes[265]. Furthermore, Pasquarella et al. has shown that SETDB1 is essential for the silencing of retroviruses in pro-B cells and loss of Setdb1 in the B-cell lineage leads to apoptosis of pro-B cells[266]. Finally, Cuellar et al. found SETDB1 silences retroviral elements in human AML cell lines [267]. Loss of *SETDB1* led to a loss of silencing of retroviral elements, which induced a cytotoxic interferon response in AML cells[267]. The paralog of SETDB1, SETDB2, has less clearly defined roles in malignancy. However, it was initially characterized in human cells as a gene that is frequently deleted in B-cell CLL[268], suggesting it may have tumor suppressor functions. In summary, SETDB1 displays both oncogenic and tumor suppressor functions in heme malignancies, which stands in contrast its oncogenic role in most solid tissue tumors that it has been studied in. Further, the mechanistic reasons behind these different functions are not clearly understood. Therefore, the function of SETDB1 and its regulation in hematopoietic cancers should be further elucidated.

G9a and GLP

G9a (encoded by the gene *EHMT2*) and GLP (encoded by *EHMT1*) play critical roles in hematopoiesis and leukemia. As mentioned above, G9a/GLP play an important role in differentiation associated changes in H3K9 methylation patterning. Interestingly, however, G9a is not required for normal HSC reconstitution of lethally irradiated mice[269].Functionally, G9a is critical for AML progression and initiation. A recent study demonstrated that G9a is required for AML growth in in vivo mouse models where loss of G9a extends disease latency. Further, G9a binds *Hoxa9* and is necessary for *Hoxa9* transcription factor function in murine AML cells[269]. These authors proposed that this presents a valuable therapeutic window in blood cells where G9a/GLP may be targeted selectively in AML cells without affecting normal HSC function[269]. In keeping with the theme of therapeutic targeting, in cellular studies performed by Loh et al. and Savickiene et al., the authors observe that inhibition of G9a/GLP sensitizes CML cells to interferon treatment and PML cells to retinoic acid (ATRA)[270,271]. Furthermore, in vivo xenograft studies have revealed that inhibition of G9a/GLP leads to reduced leukemia growth in vivo[272]. ChIP-sequencing of K562 AML cells has also revealed that there is an

accumulation of H3K9me2 across large genomic regions of the genome in AML cells[273]. These regions are larger than 100Kb in size and are enriched for regions of the genome that are commonly deregulated in AML, either by mutation, chromosomal rearrangement, or downregulation. The accumulation in these regions seem to be dependent on G9a and its binding partner GLP because inhibition of G9a/GLP's methyltransferase activity leads to a reduction in H3K9me2 accumulation[273]. Finally, GLP itself has recently been demonstrated to be a prognostic marker in CLL. Alves-Silva et al. showed that high expression of GLP is associated with a poor prognosis in CLL and that inhibition of GLP/G9a induces cell death in a CLL cell line[274]. These studies all point to an oncogenic function of G9a/GLP. In contrast, Son and colleagues have recently demonstrated that ATRA-induced differentiation of PML cell lines requires G9a[275]. In this study, the authors observe that G9a represses *JAK2* expression via promoter H3K9me2 after cells are treated with ATRA. Consistent with this, the authors see a stabilization of G9a protein after treatment with ATRA[275]. Taken together, these data suggest that G9a and GLP are important factors in blood malignancies and that their role is highly context dependent.

SUV39H1/H2

SUV39H1/H2 also have important functions in normal HSPCs and malignant blood cancers. A study of germline deletion of *Suv39h1/h2* in mice revealed that while double null mice are largely non-viable, loss of either *Suv39h1* or *Suv39h2* resulted in viable, though developmentally impaired mice[232]. Interestingly, heterozygous or homozygous null *Suv39h1* mice develop B-cell lymphomas at a rate of about 28%, which would be indicative of some tumor suppressive functions for SUV39H1[232]. *Suv39h2* heterozygous or homozygous null mice develop B-cell lymphomas at a much lower rate (<5%)[232]. Further, Djeghloul and colleagues have also demonstrated that age-related loss of SUV39H1 leads to impaired B-cell development and that overexpression of SUV39H1 leads to a stronger production of B-cells in old HSCs[276].

In malignant hematopoiesis, it has been observed that low levels of *SUV39H1* and high levels of *SUV39H2* are associated with increased acquisitions of karyotypic abnormalities in CLL[277]. In contrast, depletion of *SUV39H1* in AML cells leads to a lower

rate of chromosomal abnormalities in Myc-driven malignancy[278]. Taken together, these data imply that deregulation of H3K9 methylation-mediated chromatin conformation could contribute to a higher mutational burden and onset of heme malignancies. Further, SUV39H1 has been described as an oncogene in acute promyelocytic leukemia (APL), where it functions by physically binding with the PML-RAR fusion oncoprotein and silencing its target genes, driving malignancy[279]. It is noteworthy that a small molecule inhibitor of SUV39H1 (chaetocin), delays cell growth and induces differentiation of AML cells, both alone and in combination with other chemotherapeutics[280–282]. Mechanistically, treating AML cells with this chemical leads to re-expression of tumor suppressor genes that are repressed by SUV39H1 in AML

Table 1-1: A summary of H3K9 methyltransferases known associations with hematopoietic malignancies

	H3K9 modification	Hematologic Disease Association
GLP (*EHMT1*)	me1/me2	High expression in CLL is a poor prognostic marker
G9a (*EHMT2*)	me1/me2	Required for AML by mediating Hoxa9 driven transcription; Chemical inhibition sensitizes AML, CML, PML cells to differentiation and death; mediates accumulations of H3K9me2 across large genomic regions in AML; Necessary for ATRA-induced differentiation of PML cells
SETDB1	me1/me2/me3	Required for MLL-AF9 driven AML; Represses oncogenic self-renewal genes in AML; Represses retroviral elements in AML; Interacts with PML tumor suppressor
SETDB2	me3	Recurrently deleted in CLL
SUV39H1	me2/me3	Germline deletion leads to B-cell lymphomas (high penetrance); Low expression associated with karyotypic abnormalities in CLL; Depletion leads to lower rate of chromosome abnormalities in AML; oncogene in APL by silencing PML-RAR
SUV39H2	me2/me3	Germline deletion leads to B-cell lymphomas (low penetrance); Upregulated in ALL;
PRDM2 (RIZ1)	me1/me2/me3	Downregulated in CML and ALL; further downregulated during CML progression; Stabilized expression leads to MM cell death
PRDM8	me2	Downregulated in E2A-PBX1 driven ALL

cells[283]. However, recent studies have suggested that chaetocin is a non-selective inhibitor of many different lysine methyltransferases, and that its mechanism of action may vary drastically from its reported targeting of SUV39H1[284–286]. This data urges caution when using chaetocin as a proposed inhibitor to SUV39H1.

The SUV39H1 paralog, *SUV39H2*, is aberrantly upregulated in ALL, which is surprising given that it is thought to be a testes specific gene[238,277,287]. Muntonga et al. have shown that genetic knockdown of SUV39H2 in both B-cell ALL and T-cell ALL leads to reduced leukemic cell viability[287]. Furthermore, high expression of *SUV39H2* is shown to make ALL cells more resistant to cytarabine[277]. Mechanistically, SUV39H2 methylates LSD1, a known oncogene in ALL, which protects LSD1 from ubiqutination and subsequent degradation[288]. In summary, SUV39H1 and SUV39H2 can play a variety of context dependent roles in hematological malignancies.

The PRDM family of H3K9 methyltransferases

The role of the PRDM family of H3K9 methyltransferases hematologic malignancies is less clearly understood. However, there are several studies demonstrating PRDM2 (RIZ1) as a tumor suppressor gene in leukemia and MDS. First, *PRDM2* exhibits reduced expression in CML cell lines and patient samples, and its expression is further reduced during CML disease progression[289,290]. Furthermore, overexpression of PRDM2 in CML cells leads to differentiation and reduced cell growth[289,290]. Additionally, the *PRDM2* promoter is methylated in T-cell and B-cell ALL, AML and MDS resulting is decreased PRDM2 expression. In MDS low PRDM2 expression correlates with higher risk[291–293]. Mechanistically, PRDM2 represses IGF-1 via deposition of promoter H3K9 methylation[290]. Finally, PRDM2 has been shown to be a potential therapeutic target in multiple myeloma (MM). Treatment of MM cells with triptolide leads to increased *PRDM2* expression and apoptosis[294] suggesting tumor suppressor function for PRDM2 in blood malignancies.

The roles of the other members of the PRDM family of H3K9 methyltransferases in blood malignancies are not well understood. PRDM8 is a bona-fide H3K9 methyltransferase that is downregulated in ALL cell lines driven by E2A-PBX1 fusion oncoproteins, though no functional studies have demonstrated a mechanism for or

effect of this downregulation[295]. PRDM3 (MDS1-EVI1), which is encoded by MECOM, has only recently been described to have H3K9 methyltransferase activity[247]. PRDM3 has been reported to have oncogenic functions in AML but it is not clear whether this is a result of its H3K9 methyltransferase functions or its protein-protein interaction network[296–299]. Additionally, other studies have suggested that PRDM3 has no H3K9 methyltransferase activity[300]. Similarly, PRDM16 has only recently been reported as an H3 lysine methyltransferase, and there have been discrepancies over whether it shows substrate specificity for H3K4 or H3K9[247,301]. *PRDM16* is required for hematopoiesis by maintaining the self-renewal capacity of LT-HSCs. It has also been implicated as both a tumor suppressor gene and an oncogene in AML, though in neither case was the H3K9 methyltransferase function explored as the mechanism of action[301,302]. Whether PRDM3 and PRDM16 are bona-fide H3K9 methyltransferases and what the role of this methyltransferase activity is in disease still needs to be elucidated.

An H3K9 methyltransferase complex

H3K9 methyltransferases may not always function as independent units. Members of the SUV39 family of H3K9 methyltransferases form a heterotetrameric complex that functions to repress euchromatic gene expression. SETDB1, SUV39H1, G9a and GLP physically associate and localize to known gene targets of G9a[234]. Furthermore, several of these complex members have been proposed to have similar roles in various malignancies, suggesting the whole complex may be dysregulated and not just one individual H3K9 methyltransferase. For instance, G9a SUV39H1 interact with both EVI-1 and the longer MDS1-EVI-1 isoform (PRDM3) proteins[299,300,303]. EVI-1 functions as an oncoprotein and high expression of EVI-1 is associated with a poor prognosis in AML patient samples. One mechanism of action by which EVI-1 drives malignancy is by repressing TGF-β signaling. Goyama and colleagues propose that G9a and SUV39H1 are required for this repression by deposition of H3K9me2/3[300]. Importantly, these authors also demonstrate a lack of H3K9 methyltransferase activity for PRDM3, but the recruitment of SUV39H1 and G9a may account for the correlation between PRDM3 and H3K9 methylation that has been previously reported[247]. Another study also proposed a shared oncogenic function of G9a and SUV39H1, wherein both

of these proteins are recruited to epigenetically repress the tumor suppressor gene SOCS1 in AML[304]. In summary, H3K9 methyltransferases may be important to development and disease as single units or may act in concert as a complex. If the complex is playing a major role in disease, the biochemical regulation of the complex would need to be elucidated to fully understand how it functions.

H3K9 demethylases in normal and malignant hematopoiesis

JMJD1A/B/C (JHDM2A/B/C;KDM3A/B/C)

The Jumanji domain containing protein 1 (JMJD1) family plays important roles in blood cell development and are required for AML cell growth. Li and colleagues demonstrated that JMJD1B demethylates H3K9me2 (as well as H4R3me2) at promoter regions of genes that are critical for hematopoietic development, including GATA3, STAT3, NOTCH1, and SOX2/4[305]. As a consequence of this regulation, deletion of JMJD1B leads to anemia, high white blood cell count, and a bias towards neutrophil differentiation[305]. Furthermore, *JMJD1B* is downregulated as cells differentiate, suggesting a role in regulating primitive hematopoietic cells. In the context of malignancy, JMJD1A/B/C have oncogenic properties in different subtypes of hematopoietic diseases. For instance, JMJD1A is a critical regulator of oncogene expression in multiple myeloma (MM) [306], including hypoxia-induced genes that are required for the survival of a subset of MMs[307]. Further, *JMJD1B* is upregulated in ALL patient samples and drives expression of the ALL oncogene Lmo2[308]. Conversely, *JMJD1B* is downregulated in AML patient cells and its overexpression induces AML cell differentiation[309]. Finally, *JMJD1C* is required for the growth of MLL-AF9 cells, and its depletion leads to AML cell differentiation and apoptosis in MLL-AF9 and AML-ETO9a translocation driven cell leukemias[310,311,312]. These data suggest an importance for the JMJD1 family of proteins in both normal and malignant hematopoiesis.

JMJD2A/B/C (JHDM3A/B/C;KDM4A/B/C/D)

The JMJD2 family have essential roles in malignant cell growth but appear to be more dispensable for normal hematopoietic development. Though the precise role of

the JMJD2 family in normal blood cell development is unclear, concurrent depletion of Jmjd2a/b/c does not affect the ability of cells to reconstitute the bone marrow a lethally irradiated mice, suggesting these proteins are not required for HSPC growth[313]. In contrast, loss of Jmjd2a/b/c in MLL-AF9 driven AML cells induces cell differentiation and apoptosis due to a downregulation of *Interleukin 3 receptor* α (*Il3ra*), which is important for survival of AML cells[313]. Interestingly, a separate study demonstrated that depletion of *Jmjd2c* itself is in fact enough to induce differentiation and apoptosis of AML cells, in this case through downregulation of *Hoxa9*[314]. Further, targeted inhibition of JMJD2C in MLL-fusion or MOZ-TIF2 fusion driven AML leads to decreased cell growth and induced differentiation and apoptosis[314]. Similarly, in lymphomas JMDJ2C acts synergistically with JAK2 to mediate the JAK/STAT signaling pathway and drive cell growth, specifically by upregulating MYC and IL-4[315,316]. These data show that the JMJD2 family of K9 demethylases plays an important role in the maintenance of AML, as well as other blood malignancies.

Potential of therapeutically targeting H3K9 methyltransferases/ demethylases

H3K9 methyltransferases and demethylases play an important role in blood cell development and can have either suppressive or oncogenic functions in different cellular contexts. Therefore, there is a strong interest in the development of molecules that can inhibit their function in order to gain a deeper understanding of their possible utility as therapeutic targets. There have been several studies that have shown the promise of targeting proteins that modify H3K9 methyltransferases and demethylases in hematopoietic malignancies. In fact, several small molecule inhibitors exist which can be used to selectively target these classes of proteins. One example of this is a series of small molecule inhibitors that can target G9a/GLP, including BIX01294[317], UNC0638[318], and A-366[272]. Small molecules also target the JMJD1 family including QC6352[319]. All of these molecules have been used to successfully delay the growth or induce differentiation of malignant cells. However, as has been previously discussed in this section, the proteins that modulate H3K9 methylation do not always function in the same manner regardless of cellular context. In some cases, they are required for cell growth, while in other cases they suppress cell growth. Furthermore, given the many

different biochemical roles of H3K9 methylation in regulating cellular processes, it is unlikely that inhibition of these proteins will simply disrupt one specific pathway. Inhibition of these proteins may disrupt one pathway in a manner that induces cell death, while simultaneously disrupting another axis of regulation that leads to an increase in genes associated with aggressive malignancy. In fact, a group recently demonstrated that while inhibition of G9a using small molecules can delay tumor growth in skin cancer, the tumors that did develop were significantly more aggressive, suggesting that off-target effects can induce a larger pool of cancer progenitor cells[320]. Therefore, the functions and regulation of these proteins and the effects of inhibiting them should be fully elucidated in a disease specific context.

1.4 Summary and Goals

Here we have reviewed a large body of evidence that emphasize a few important points. First, transcriptional regulation, including epigenetic regulation, are key processes in blood cell development. Deregulation of these processes is associated with inappropriate expression of genes associated with self renewal and a block in the expression of differentiation specific genes, and this can result malignant cell growth. Next, the PAF1c is an epigenetic co-modifying complex that is a key regulator of gene expression in normal development and in malignancy. It has either tumor suppressor or oncogenic functions in different solid tissue tumors. Further, the PAF1c is required for AML cell growth, at least partially through its interaction with MLL-fusion proteins and its regulation of *Hox* gene programs. An interesting but underexplored function of the PAF1c involves its role in epigenetically repressing genes, potentially through interactions with H3K9 methyltransferases. H3K9 methylation is a critical histone modification that plays a central role in normal cell functions and in cancer cell growth. H3K9 methyltransferases are largely seen as oncogenic proteins in solid tissue tumors, with some evidence indicating a more cell context dependent role. Similarly, changes in H3K9 methylation can function by both suppressing cell growth or promoting cell growth in hematopoietic cancers. Further study of the proteins involved in modulating this mark may help to elucidate novel modalities of regulation and uncover new ways to precisely target malignant cells in the treatment of cancer.

To that end, there are several outstanding questions related to the PAF1c and H3K9 methyltransferases and their role in hematopoietic malignancies like AML. First, while we know that the PAF1c is important in AML, a thorough study of its targets in AML and normal hematopoietic cells has not yet been performed. Further, given its role in transcriptional regulation as a platform complex that recruits other proteins, the PAF1c interactome should also be studied in the context of AML. Here we will show that the PAF1c interacts with H3K9 methyltransferases, which leads to new questions. First, it is important to understand the roles of H3K9 methyltransferases and the PAF1c-H3K9methyltransferase interaction in regulating transcription in AML. Next, the role of many H3K9 methyltransferases have only been studied in the context of AML using genetic deletion systems. However, it is important to understand the consequences of increased expression in AML as well, given the ambiguous roles of H3K9 methylation as tumor suppressive or oncogenic. Finally, because the inhibition of H3K9 methyltransferases has been proposed as a promising therapeutic target, it is critical to fully elucidate the effects this inhibition might have both on AML cells as well as on normal HSPCs. Our studies seek to answer these questions so that new insight will be provided into the role of epigenetic and transcriptional regulation in AML.

Chapter 2 The PAF1c-dependent transcriptome in AML and HSPCs

2.1 Introduction

Epigenetic modifiers are mutated in greater than 70% of AMLs and are frequently dysregulated in many other cancers, and therefore are promising potential therapeutic targets. The Polymerase Associated Factor complex (PAF1c) is an example of an epigenetic modifying complex that is implicated in endocrine, gastric, breast, pancreatic, and bone cancers. Our lab and others have recently demonstrated the PAF1c importance in leukemia. The PAF1c is a protein complex composed of six subunits: CDC73, PAF1, CTR9, LEO1, WDR61, and RTF1. The PAF1c interacts with RNA Polymerase II and recruits epigenetic modifiers to target genes. The PAF1c has been shown to play a critical role in the chemical modification of histones such as H3K4 methylation, H3K79 methylation, H3K36 methylation and H2b ubiquitination, all of which are epigenetic modifications associated with active transcription. Recent work by our lab and others has shown that the PAF1c binds and recruits MLL fusion proteins to target oncogenes and that this interaction is required for leukemic maintenance but is less critical for normal hematopoiesis. Thus, the PAF1c may be a promising therapeutic target. However, little is known about PAF1c mediated transcription in leukemia.

Despite the differences in PAF1c function in AML relative to normal hematopoietic cells, a full transcriptome analysis of the targets of the PAF1c in MLL-AF9 driven leukemia compared to the PAF1c regulated transcriptome in normal HSPCs has not been reported. Here we describe a series of transcriptomic studies in HSPCs and AML cells that demonstrate a role for the PAF1c in the regulation of several different gene programs in AML by exploring perturbations to the transcriptome in MLL-AF9 driven AML cells after genetic depletion of the PAF1c subunit CDC73. Additionally, these studies demonstrate that the PAF1c is responsible for the regulation of unique and overlapping gene programs in AML compared to normal HSPCs.

2.2 Methods

Cell proliferation assays

Cell lines were generated as previously described from lineage negative (lin-) mouse bone marrow: either wild type C57Bl/6 (Taconic Farms), *Cdc73fl/fl*-CreER^{T2} (*Cdc73fl/fl*-CreER)[321], or CreER^{T2} (CreER)[14]. Briefly, Platinum-E (Plat-E) viral packaging cells were transfected with retroviral vectors MSCVneo-MLL-AF9 (MA9), MSCVpuro-CreER (CreER), or MSCVpuro (EV). Viral supernatants were collected and bone marrow or AML cells were spun with the virus and 5ug/mL polybrene (Millipore). For AML studies, *Cdc73fl/fl*-CreER or CreER cells were spinfected with MA9. For HSPC studies, *Cdc73fl/fl* cells were spinfected with CreER or EV. All bone marrow-derived MLL-AF9 cells were cultured in IMDM supplemented with 15% Stem Cell FBS (Millipore), 1%pen/strep, and 10ng/mL IL-3 (R&D). Normal HSPCs were also supplemented with 100ng/mL SCF (R&D). Cells transduced with MA9 were selected with 1mg/mL neomycin. Cells transduced with CreER or EV were selected with 1ug/mL puromycin.

For all proliferation assays, $5x10^4$ cells were seeded in 2mL of normal growth media containing either 2.5nM 4-hydroxytamoxifen (4-OHT) for AML studies, 7.5nM for HSPC studies, or an equivalent percentage ethanol vehicle control (vehicle). Viable cell number was counted each day for 3 days (AML) or 11 days using Trypan Blue (Invitrogen). Every 2 days, the cells were supplemented with fresh media, IL-3 (+SCF for HSPCs), and 4-OHT or vehicle. Statistics were done by pairing treated samples with the control samples from the same biological replicate and comparing all biological replicates using generalize linear modeling followed by ANOVA (n=2-5).

In vivo mouse modeling

For secondary AML transplantation studies, primary AMLs were generated by tail vein injecting lethally irradiated mice (950 rads) with freshly spinfected MLL-AF9 cells on a *Cdc73fl/fl*-MxCre background. Mice were monitored, and moribund mice were euthanized and their bone marrow harvested. Bone marrow from leukemic mice were then tail vein injected into n=10 sublethally irradiated mice (650 rads), engraftment was allowed to proceed for 5 days, and mice were treated with 50ug Polyinosinic-

polycytidylic acid (poly(I:C)) (GE lifesciences) by intraperitoneal (IP) injection. Mice were monitored, euthanized when moribund, and leukemia was confirmed by splenomegaly and histology (data not shown). For normal adult hematopoiesis excision of *Cdc73* studies, Mx1Cre+/0 (MxCre), *Cdc73fl/+*-MxCre, or *Cdc73fl/fl*-MxCre mice (n=11,11,13 respectively) at 8 to12 weeks old were treated with 5 doses of 50ug/dose poly(I:C) every other day intraperitonially. Mice were monitored and euthanized when moribund. Genetic excision of *Cdc73* was confirmed (data not shown). Survival statistics were performed using Log-Rank tests.

RNAseq analysis

For AML studies, RNA was harvested from MA9*Cdc73fl/fl*-CreER or MA9-CreER cells 48 hours after treatment with 7.5nM 4-OHT to induce genetic excision of *Cdc73*. For Ckit+ HSPC studies, RNA was harvested from Ckit+ bone marrow. Ckit+ cells were isolated from *Cdc73fl/fl*-MxCre or MxCre mice 48 hours after 50ug IP injection of poly(I:C) by staining with anti-Ckit-APC antibody (BioLegend) followed by flow cytometry/ sorting for APC+ cells at the University of Michigan Flow Cytometry Core. Library preparation and single-end 50bp sequencing on a HiSeq4000 (Illumina) were performed by the University of Michigan Sequencing Core. Sequencing reads were obtained from the University of Michigan Sequencing Core. Separately RNA-seq reads from THP-1 cells with and without siRNA mediated PAF1 KD were obtained from the GEO database (GSE62171[170,322]) as fastq files. Following is a brief description of the optimized RNA-sequencing analysis pipeline. Tool version, authors, and any parameters used that are not defaults of the software/ tools used are listed in Table 2.

Quality of sequencing data was analyzed using FastQC. For Ckit+ HSPC studies, due to high sequence duplication levels, reads were analyzed for rRNA content using Bowtie2 alignment to the rRNA “chromosome” (GRCm38). 20-30% of each sequencing run was made up of rRNA reads, but analysis of raw counts with and without rRNA reads revealed that gene counts were not significantly affected by the presence of rRNA reads, so rRNA reads were kept for accurate library size determination. Reads were trimmed for quality using Cutadapt. Indices for alignment were generated using the GENCODE M20 release of the GRCm38 mouse genome assembly or the version 29

release of GRCh38 with STAR. Reads were aligned to the genome using STAR. Aligned reads were counted by gene using HTSeq, which only keeps uniquely aligned reads by default. These counts were analyzed for differential gene expression using two separate programs implemented in R: DESeq2 and the edgeR-voom-eBayes pipeline (edgeR). DESeq2 and edgeR both perform computational normalizations based on effective library size, with the base assumption that most genes are not differentially expressed. Both analyses yielded similar results with DESeq2 yielding a more stringent same to same comparison for genes in treated versus control samples, so the significance of differential expression of individual genes was based on DESeq2 analysis. edgeR is more suited to piping into gene set analysis (GSA) using CAMERA, so it was used for downstream GSA and generation of barcode plots. The following analyses were all implemented in R using the listed packages: The MA plot was generated using ggplot2. MA plot is a visualization application to represent differential expression of genes in RNAseq data computed based on log-intensity ratios of treated relative to control (M) versus log-intensity averages (A) of all samples in the comparison. Gene ontology analysis was performed using GOseq or GAGE. Venn diagrams were produced using VennDiagram. CAMERA and ROAST were used for gene set analyses using curated MSigDB[323] groups of gene sets or individual gene sets, respectively. Barcode plots were produced by modifying code from the function barcodeplot (limma). All analyses were performed in R version 3.5.1.

Chromatin Immunoprecipitation sequencing (ChIP-seq) analysis

ChIP-seq data for ChIP performed in THP-1 cells using antibodies against PAF1, LEO1, CDC73, RNAPII, H3K4me3, and H3K79me2 were obtained from the GEO database (GSE62171[170,322]) as fastq files. Following is a brief description of the optimized ChIP-seq analysis pipeline. Tool version, authors, and any parameters used that are not defaults of the software/ tools used are listed in Table 2.

Quality control was performed using FastQC. Adapters were trimmed and reads were trimmed for quality using Cutadapt. Reads were aligned to the GRCh38 genome assembly using Bowtie2. SAMtools was used to sort the aligned reads, filter out mitochondrial reads, and keep only reads that were confidently uniquely mapping (0.995

probability of mapping uniquely). DeepTools was used to generate sequencing tracks normalized by total library size. Integrated Genome Browser (IGB) was used to visualize sequencing tracks.

Table 2-1: List of software and tools for analysis of Next Generation Sequencing

Software/Tool	Version	Citation	Use	Parameters
SRAtoolkit	2.9.4		Download sequencing data from GEO	
FastQC	0.11.7	Babraham Bioinformatics[324]	Quality statistics	
Cutadapt	1.18	Martin, 2011[325]	Quality filtering of fastq files for RNA-seq/ ATAC-seq; adapter trimming	-q 20,20 --minimum-length 30
AfterQC	0.9.6	Chen, et al. 2017[326]	Quality filtering of fastq files for ChIP-seq; adapter trimming	-q 20 -u 19 -p 19 -s 25
STAR	2.6.0c	Dobin, et al. 2013[327]	Alignment of RNA-seq data	genomeGenerate: --sjdbOverhang 49 Alignment: --outSAMtype BAM SortedByCoordinate
Bowtie2	2.3.4.3	Langmead & Salzberg 2012[328]	Alignment of ChIP-seq data	--very-sensitive -X 2000
SAMtools	1.9	Li, et al 2009[329]	Sorting, indexing, quality filtering of aligned BAM files	view -b -q 25 -f 0x2
Picard	2.18.19	Broad Institute[330]	Removing PCR duplicates from ATAC-seq/ChIP-seq	
MACS2	2.1.2	Zhang, et al. 2008[331]	Calling peaks for ATAC-seq/ ChIP-seq	Narrow ChIP-seq peaks (H3K9ac): -t <input> -f BAMPE -g mm --keep-dup all Broad ChIP-seq peaks (H3H9me3, H3K79me2): -t <input> -f BAMPE -g mm --scale-to-large --broad --keep-dup all ATAC-seq peaks: -f BAM -g mm --nomodel --shift -100 --extsize 200 --keep-dup all
deepTools (bamCoverage)	3.1.3	Ramirez, et al. 2014[332]	Generation of normalized bigwig files for sequencing tracks	RNA-seq: --binSize 10 --scaleFactor <sizeFactor from DESeq2> --normalizeUsing None

				ChIP-seq/ATAC-seq: --binSize 10 --normalizeUsing RPGC --effectiveGenomeSize 2150570000 --extendReads
HTSeq	0.11.0	Anders, et al. 2014[333]	Counting reads assigned to features (exons) for RNA-seq	-s reverse -r pos
Integrated Genome Browser	9.0.2	Nicol, et al 2009[334]	Visualization of sequencing tracks	
DESeq2	3.8	Love, et al. 2014[335]	Differential gene expression analysis	alpha = 0.05; lfcThreshold = log2(1.5)
edgeR	3.8	Robinson, et al. 2010[336] McCarthy, et al. 2012[337]	Differential gene expression analysis	
limma	3.8	Ritchie, et al 2015[338]	Barcode plots; library normalization quality control	
ChIPpeakAnno	3.8	Zhu, et al. 2010[339,340]	Establishing consensus peak sets and gained/lost peaks in ChIP-seq/ ATAC-seq analyses	Peaks were annotated to promoters (-5000, +2000 from TSS) for ATAC and H3K9ac; Peaks were annotated to any overlapping gene feature for H3K79me2 and H3K9me3
DiffBind	3.8	Stark, et al 2011[341,342]	Analyzing differential peak signals for ChIP-seq/ATAC-seq	method = DBA_DESEQ2; bFullLibrarySize=TRUE
CAMERA		Wu and Smyth 2012[343]	Differential gene expression analysis to generate statistics for barcode plots	inter.gene.cor=0.01
ROAST		Wu, et al. 2010[344]	(see CAMERA, for stand alone gene sets)	
ggplot2	3.1.0	Wickham, et al	Generation of graphs	
VennDiagram	1.6.2	Boutros, Paul	Generation of venn diagrams	

2.3 Results

These data were generated as part of a completed study and an ongoing study in our lab. The project exploring the role of the PAF1c in AML was done in collaboration with Justin Serio[224]; the bioinformatic analyses for this study were performed by Jingya Wang and myself. The project exploring the role of the PAF1c in normal HSPCs was done in collaboration with Nirmalya Saha (Saha, et al. *In revision*); I was responsible for all bioinformatic analyses in this study. Here the implications of the transcriptome studies are the primary focus.

Genetic excision of Cdc73 suppresses growth of AML cells and normal HSPCs

It has recently been demonstrated that the PAF1c is required for the cellular proliferation of MLL-AF9 and E2A-HLF driven AML cell lines using a Cre-lox system. This system allows us to induce genetic excision of the PAF1c subunit *Cdc73* in AML cells by treatment with 4-hydroxytamoxifen (4-OHT). We first sought to extend the analysis of the requirement for Cdc73 in different AML subtypes. Cell lines were generated on the homozygous Cdc73floxed and hemizygous CreER^{T2} background (*Cdc73fl/fl*-CreER) or on a Cdc73 wildtype and hemizygous CreER^{T2} background (CreER) by retrovirally transducing different oncogenes into lineage negative (lin-) mouse bone marrow cells. Excision of Cdc73 was confirmed and proliferation assays reveal that loss of Cdc73 leads to a dramatic loss of AML cell growth in MLL-AF9 (MA9) and E2A-HLF (EHF) AML cells (Figure 2-1A, top 2 panels). We next explored whether Cdc73 was required for AML cells driven by overexpression of Hoxa9 and Meis1 (H/M), which are oncogenes that are downstream of the MLL-AF9 oncogene, and AML-ETO9a (AE9), which drives a separate oncogenic program from MLL-AF9 or E2A-HLF. Treatment of H/M*Cdc73fl/fl*-CreER and AE9*Cdc73fl/fl*-CreER cells with 4-OHT revealed a strong proliferative defect in cells after induced excision of *Cdc73* (Figure 2-1A, bottom 2 panels). The next important questions was whether Cdc73 was required for normal hematopoietic cell growth. Lin- bone marrow was isolated from *Cdc73fl/fl* cells, retrovirally infected with CreER^{T1}, and the cells were selected for three days. The cells were then treated with 4-OHT and monitored proliferation. In cells that were transduced with a control empty vector and treated with 4-OHT, a proliferative burst that plateaued

after 8 days was observed. In cells expressing CreER, treatment with 4-OHT resulted in a loss of the proliferative burst (Figure 2-1B). Next, we sought to ask whether there is a requirement for Cdc73 in AML and HSPC cell growth *in vivo*. To this end, we performed secondary AML transplants tail-vein injecting sublethally irradiated mice with primary MLL-AF9 cells that are on a Cdc73fl/fl or wildtype background that are also hemizygous for the hematopoietic specific inducible Mx1-Cre (MxCre). These MA9*Cdc73fl/fl*-MxCre cells can be induced to excise *Cdc73 in vivo* by injecting the mice with Polyinosinic-polycytidylic acid (poly(I:C)). After injection with poly(I:C), we observed a significant increase in AML disease latency in mice that were injected with MA9 *Cdc73fl/fl*-MxCre

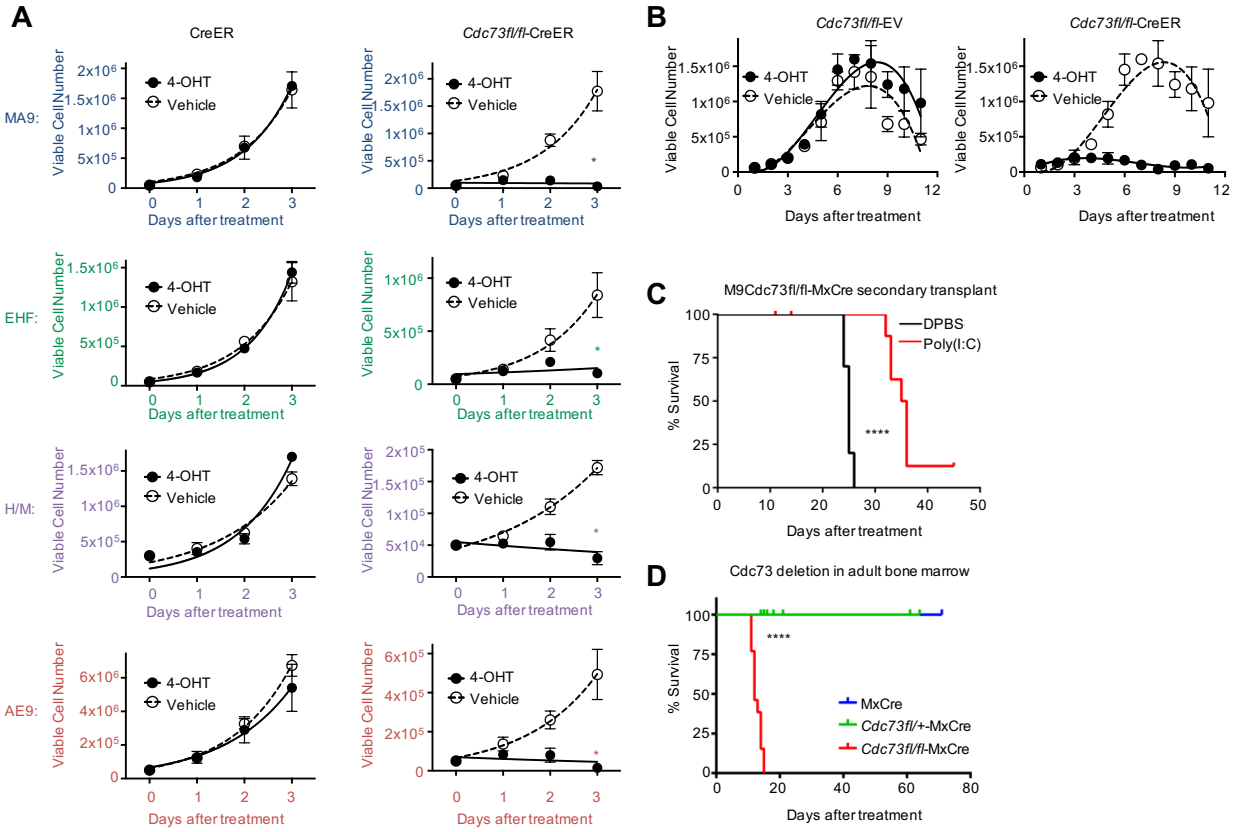


Figure 2-1: PAF1c subunit Cdc73 is required for AML cell and HSPC growth. A) Cell lines transformed with the indicated oncogenes (MA9=MLL-AF9; EHF=E2A-HLF; H/M=Hoxa9/Meis1; AE9=AML-ETO9a) on a *CDC73fl/fl*-CreER^{T2} or CreER ^{T2} background were treated with 2.5nM 4-OHT or EtOH vehicle control and proliferation was monitored. Shown are representative experiments of n=2-5. B) Lin- bone marrow from a Cdc73fl/fl mouse was isolated and transduced with CreER^{T1} or empty vector control. Cells were selected for 3 days then treated with 7.5nM 4-OHT or EtOH vehicle and proliferation was monitored. C) Primary MA9Cdc73fl/fl-MxCre cells were injected into sublethally irradiated mice. After 5 days of engraftment, mice were treated with poly(I:C) or PBS control to induce excision of Cdc73. Kaplan-Meier curve of survival. D) Mice of the indicated background were treated with poly(I:C) to induce hematopoietic specific excision of Cdc73 and survival was monitored. * $p < 0.05$; EV= Empty Vector

cells relative to those injected with MA9-MxCre cells (Figure 2-1C). Importantly, the mice that did eventually get leukemia in the poly(I:C) treated group were genotyped and we determined that the leukemias arose from clones that had escaped excision of Cdc73, indicating a strong selective pressure to maintain PAF1c expression in AML. We next asked whether Cdc73 depletion affects hematopoietic reconstitution *in vivo*. *Cdc73fl*/+-MxCre, Cdc73fl/fl-MxCre, or MxCre mice were treated with poly(I:C) to induce excision of Cdc73 specifically in the bone marrow. Strikingly, all mice from the *Cdc73fl/fl*-MxCre group died within 20 days of the poly(I:C) injections from bone marrow failure, while no mice succumbed to hematopoietic failure in either the control or heterozygous excision group (Figure 2-1D). Together, these data demonstrate an essential role for Cdc73 in AML and normal HSPC cell growth.

Depletion of Cdc73 in AML leads to downregulation of oncogenic gene programs

We next sought to characterize the PAF1c regulated transcriptome in AML cells. To this end we performed RNA-sequencing (RNA-seq) experiments. RNA was harvested from MA9*Cdc73fl/fl*-CreER or MA9-CreER cells after 48 hours of treatment with 4-OHT, at which time point Cdc73 protein levels are ablated[14,224]. After sequencing, differential gene expression analysis using DESeq2 reveals that 1896 genes are upregulated and 1329 genes are downregulated after excision of Cdc73 with an alpha = 0.05 and a fold-change threshold of > 1.5 (Figure 2-2A). Gene ontology studies were performed for both upregulated and downregulated gene set/ pathway analysis using generally applicable gene set enrichment for pathway analysis (GAGE). These analyses revealed that genes associated with apoptosis and differentiation are upregulated upon loss of Cdc73 and that gene programs associated with methyltransferase activity are downregulated (Figure 2-2B). Next gene set analysis using correlation adjusted mean rank gene set test (CAMERA) was performed on the MsigDB curated gene sets list. Gene programs associated with differentiation are highly upregulated upon loss of Cdc73 (Figure 2-2C). Further, oncogenic gene programs like the genes that are upregulated by Hoxa9/Meis1 are significantly downregulated upon loss of Cdc73 (Figure 2-2D). Finally, gene programs associated with histone methyltransferase activity were also downregulated upon loss of Cdc73 (Figure 2-2E). This methyltransferase

program became the focus of a now published study demonstrating that Protein arginine methyltransferase 5 (PRMT5) is a critical downstream target of the PAF1c and that PRMT5 itself is critical to supporting AML cell growth[224]. Finally, overlap analysis between genes that are significantly changed in after excision of Cdc73 in MLL-AF9 cells and genes that are direct binding targets of either MLL-AF9 or MLL-ENL show a significant overlap. 35 genes were found in this overlap, including targets critical for AML cell growth such as Myc, Eya1, and Cdkn1b. Taken together, these data show that the PAF1c regulates gene programs that are essential for blocking differentiation and promoting self-renewal in AML cells.

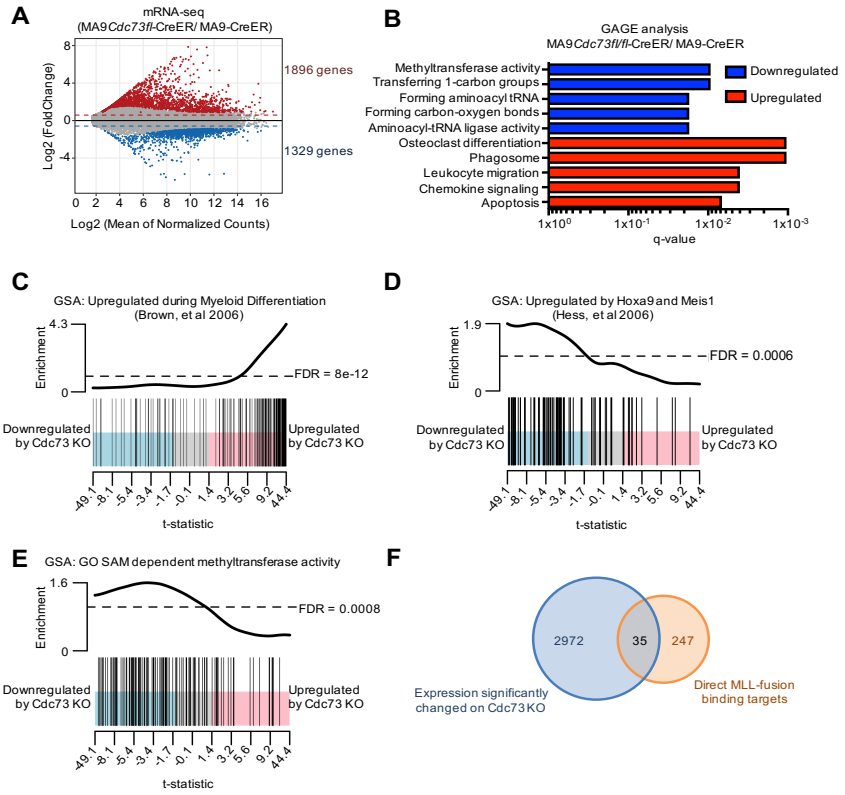


Figure 2-2: Deletion of Cdc73 in MLL-AF9 cells leads to downregulation of oncogenic gene programs. A-E) MA9*CDC73fl/fl*-CreER or MA9-CreER cells were treated with 4-OHT for 48 hours. RNA was harvested and sequenced. A) MA plot showing fold change of gene expression in MA9Cdc74fl/fl-CreER cells compared to MA9-CreER cells after 4-OHT treatment versus average expression for all samples. Red highlighted genes are significantly upregulated and blue highlighted genes are significantly downregulated (alpha = 0.05, fold change > 1.5). B) GAGE analysis showing gene ontology enrichments for up and downregulated genes after deletion of Cdc73. C-E) Gene set analysis performed with CAMERA for the indicated gene sets. F) Overlap analysis performed between all genes that are significantly changed upon loss of Cdc73 in AML cells compared to known direct binding targets of MLL-fusion proteins (MLL-AF9: Bernt, et al. 2011; MLL-ENL: Garcia-Cuellar, et al. 2016.

The PAF1c has conserved targets in mouse AML and human AML

I next sought to validate these results in the context of human AML. Yu and colleagues have recently performed RNA-seq experiments on human THP-1 AML cells with and without siRNA-mediated knockdown (KD) of PAF1[170]. These data were used to demonstrate that the PAF1c plays a critical role in RNAPII pause release but have not been fully analyzed and published in the context of evaluating the PAF1c transcriptome

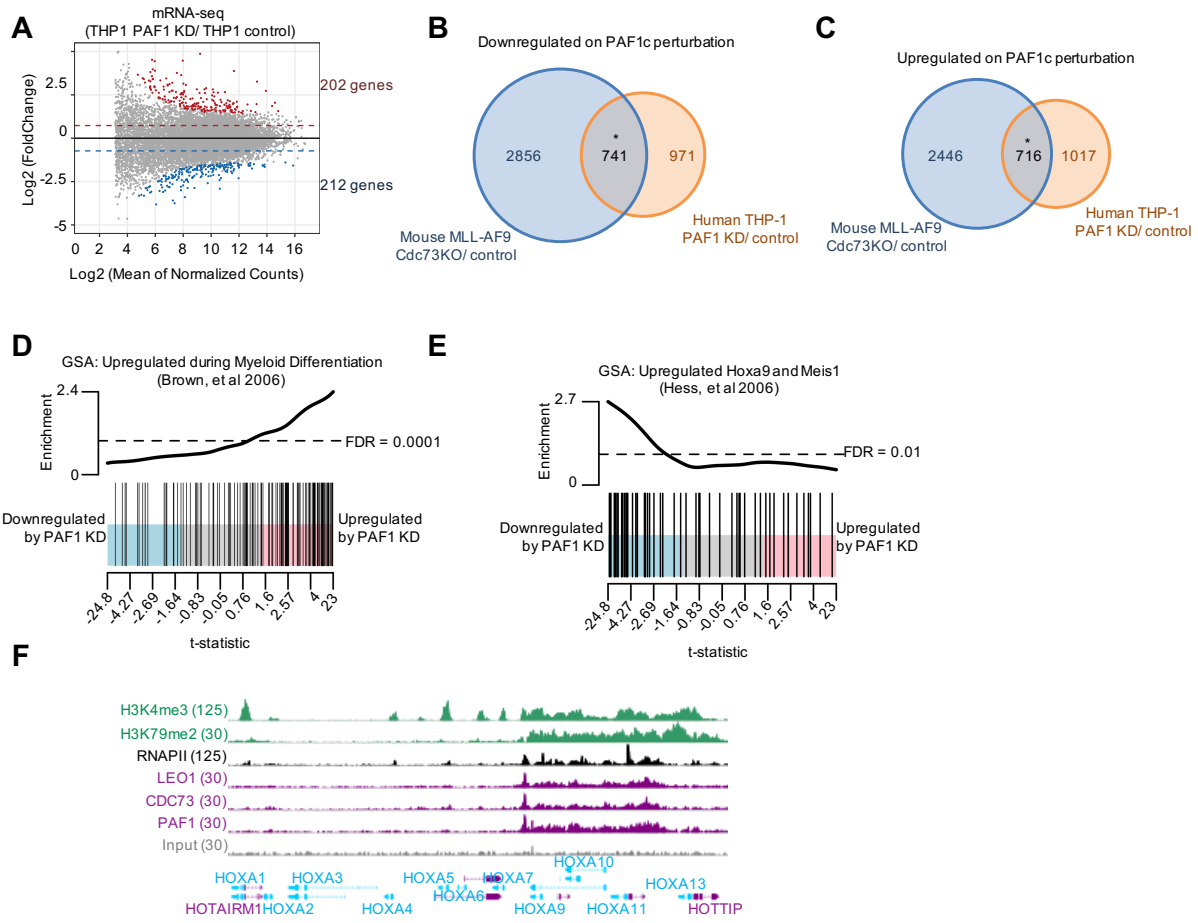


Figure 2-3: KD of PAF1 in THP-1 cells leads to downregulation of oncogenic gene programs. A-E) RNA-seq data for siRNA mediated PAF1 knockdown (KD) was downloaded from GEO (SRP048744) A) MA plot showing fold change of gene expression in PAF1 KD THP-1 cells compared to THP-1 control cells. Red highlighted genes are significantly upregulated and blue highlighted genes are significantly downregulated (alpha = 0.05, fold change > 1.5). B-C) Overlap analysis between genes that are significantly upregulated (B)/ downregulated (C) in MA9*CDC73fl/fl*-CreER and PAF1 KD THP-1 cells. Significance was relaxed to alpha = 0.05. D-E) Gene set analysis performed with CAMERA for the indicated gene sets. F) ChIP-seq tracks for pulldown using antibodies against PAF1c components (LEO1, CDC73, PAF1), RNAPII, and H3K4me3/H3K79me2 showing the HOXA locus in THP-1 cells. Raw data was downloaded from GEO (SRP048744) and re-analyzed.

in AML. This is also a good dataset for comparison to our studies because this cell line harbors an MLL-AF9 fusion. Differential gene expression analysis using DESeq2 reveals that genes are upregulated, and genes are downregulated after knockdown of PAF1 with an alpha = 0.05 and a fold-change threshold of > 1.5 (Figure 2-3A). Next, an overlap analysis of gene expression changes in THP-1 cells after PAF1 KD and expression changes in MLL-AF9 cells after excision of Cdc73 was performed. Because of the difficulty of overlapping human and mouse gene set analyses and due to the low number of significantly changed genes with stringent thresholds (fold-change > 1.5) in the PAF1 KD THP-1 cell experiment, the parameters were relaxed so that all genes with padj < 0.05 were considered significant in both gene sets, using a total gene universe of all genes that were expressed in either cell type that had an orthologue in the compared cell type. This analysis revealed that a significant amount of the genes regulated by the PAF1c is conserved between the mouse MLL-AF9 cells and the human THP-1 cells (Figure 2-3B). Further, this group of overlapping genes included common gene programs, such as an upregulation of genes associated with myeloid differentiation (Figure 2-3C) and a downregulation of genes regulated by Hoxa9/Meis1 (Figure 2-3D). Finally ChIP-seq data from the same study by Yu et al. reveals that several components of the PAF1c bind to the *Hoxa* cluster of genes in THP-1 cells, suggesting the PAF1c regulation of this gene cluster is a direct regulation (Figure 2-3E).

Depletion of Cdc73 in HSPCs results in cell cycle defects

We next asked what effects the PAF1c has on the transcriptome in normal HSPCs following depletion of Cdc73. RNA was harvested from Ckit+ bone marrow from mice that were on a Cdc73fl/fl-MxCre background or MxCre control mice. The Ckit+ cell population is made up of primitive hematopoietic stem and progenitor cells. Bone marrow from these mice were harvested 48 hours after 1 injection of poly(I:C) to induce hematopoietic specific excision of Cdc73. After sequencing, differential gene expression analysis using DESeq2 reveals that 399 genes are upregulated, and 433 genes are downregulated after knockdown of PAF1 with an alpha = 0.05 and a fold-change threshold of > 1.5 (Figure 2-4A). GOseq analysis to determine biological processes or molecular functions gene sets that are significantly changed upon depletion of Cdc73

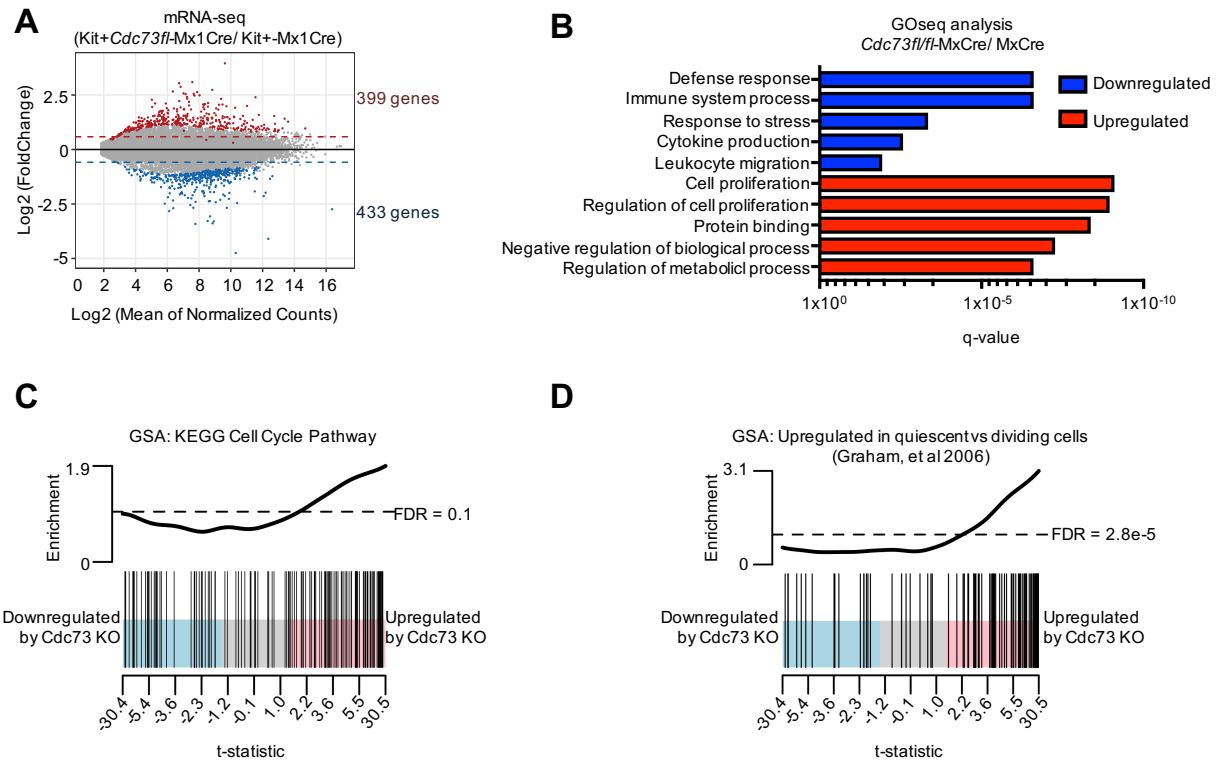


Figure 2-4: Deletion of Cdc73 in hematopoeitic cells leads to deregulation of cell cycle gene programs. A-E) *CDC73fl/fl*-MxCre or MxCre cells were harvested from mice 48 hours after treatment with poly(I:C) and Ckit+ cells were isolated. RNA was harvested and submitted to sequencing. A) MA plot showing fold change of gene expression in *CDC73fl/fl*-MxCre cells compared to MxCre cells after poly(I:C) treatment versus average expression for all samples. Red highlighted genes are significantly upregulated and blue highlighted genes are significantly downregulated (alpha = 0.05, fold change > 1.5). B) GOseq analysis showing gene ontology enrichments for up and downregulated genes after deletion of Cdc73. C-D) Gene set analysis performed with CAMERA for the indicated gene sets.

were performed. These analyses reveal that immune response gene programs are downregulated while gene programs associated with regulating cell proliferation are upregulated after loss of Cdc73 (Figure 2-4B). Next, gene set analysis using CAMERA revealed that gene sets associated with cell cycle regulation are significantly upregulated following perturbation of the PAF1c in HSPCs (Figure 2-4C). Further, a gene set specifically associated with quiescence was upregulated upon loss of Cdc73 (Figure 2-4D). These data suggest that the PAF1c is a critical regulator of cell cycle in normal HSPCs.

The PAF1c regulates unique and overlapping gene programs in AML and HSPCs

Finally, we explored whether the PAF1c dependent transcriptome is unique or distinct between AML cells and normal HSPCs. To do this, we first performed an overlap analysis looking at genes that are significantly changed in AML cells or HSPCs after loss of Cdc73. This analysis revealed both overlapping and unique genes that are regulated by the PAF1c in these cell types (Figure 2-5A), with an enrichment in gene programs that are differentially regulated in AML and HSPCs. To determine whether the PAF1c is regulating the same gene programs in AML and HSPCs, gene set analysis was performed using CAMERA on both RNA-seq experiments. Interestingly, there were 37 gene sets that were enriched in both RNA-seq experiments out of 238 significant gene sets in the AML study and 112 significant gene sets in the HSPC study (data not shown). More importantly, we explored whether gene sets that could account for the

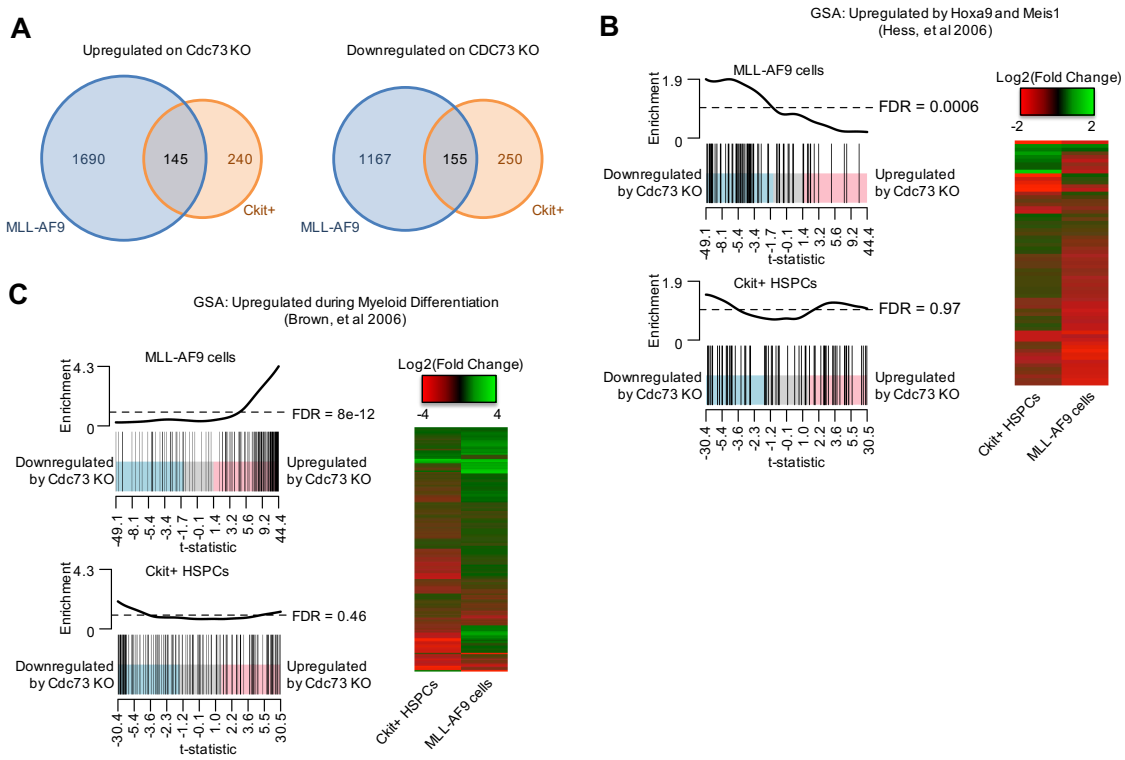


Figure 2-5: Deletion of Cdc73 leads to changes in different gene programs in AML and normal HSPCs. A) Overlap analysis of genes that are up/downregulated in *CDC73fl/fl*-MxCre (Ckit+) or MA9*CDC73fl/fl*-CreER (AML) relative to their respective controls after induced excision of Cdc73. B-E) Comparative analysis of the indicated gene sets that are regulated by the PAF1c in AML cells. The top left panel shows the CAMERA gene set analysis for MA9*CDC73fl/fl*-CreER cells; the bottom panel shows the analysis for *CDC73fl/fl*-MxCre; the right panel is a heatmap showing fold-changes of all genes included in the gene set used for the gene set analysis in both cell types after deletion of Cdc73.

strong phenotypic changes observed in AML cells upon loss of Cdc73 were regulated by the PAF1c in normal HSPCs as well. Strikingly, HSPCs did not exhibit changes in gene sets associated with myeloid differentiation upon loss of Cdc73, while they constitute the top gene set changes in AML cells after loss of Cdc73 (Figure 2-5B,C). Further, gene programs regulated by Hoxa9 and Meis1, while changing significantly in AML cells upon loss of Cdc73, do not exhibit a unidirectional change in HSPCs upon loss of Cdc73 (Figure 2-5D,E). These data show that the PAF1c regulates common and unique genes in AML cells and HSPCs. These data also point to the PAF1c regulating gene clusters uniquely between AML cell and HSPCs and strongly suggests the PAF1c has overlapping and unique functions associated with self-renewal and differentiation.

2.4 Discussion

Here we described a series of transcriptomic studies to determine the PAF1c transcriptional targets in AML cells and HSPCs. We first showed that the PAF1c subunit Cdc73 is required for both AML cell and HSPC growth (Figure 2-1). Previously, it has been shown by our lab and others that the PAF1c has a direct physical interaction with both wildtype MLL1 and MLL-fusion protein[12,13]. Interestingly, the interaction with MLL is required for AML cell growth, but disruption of the interaction does not significantly affect the ability of HSPCs to populate the bone marrow[14]. In AML, the PAF1c functions at least in part by regulating oncogenic gene programs including direct targets of MLL-fusion proteins. This includes critical downstream targets of MLL-AF9 such as Eya1 and the Hoxa9/Meis1 gene program. However, the interaction with wildype MLL does not play a critical role in normal hematopoietic growth. Despite this lack of requirement for the regulation of MLL targets, the PAF1c subunit Cdc73 is still required for HSPC growth. This suggests that the PAF1c may regulate different targets in HSPCs. To test this, we performed RNA-seq experiments in several different cell systems to determine the AML specific and HSPC specific PAF1c-dependent transcriptome.

I demonstrated that the PAF1c regulates critical oncogenic programs in both mouse and human AML (Figure 2-2,2-3). These oncogenic programs include genes regulated by Hoxa9 and Meis1. Hoxa9 and Meis1 are critical oncogenes in MLL-fusion driven leukemia[4,7,11,124,162]. Further, they are upregulated in over 50% of AML[2,37]. This

validates the PAF1c as a critical upstream regulator of oncogenic programs that are necessary for the initiation, progression, and maintenance of a high percentage of AML. Additionally, this study revealed a novel target of the PAF1c, the PRMT family of methyltransferases. More specifically, PRMT5 was found to be significantly downregulated by loss of Cdc73. A deep exploration of this axis of regulation revealed that PRMT5 is a critical downstream regulator of PAF1c-dependent transcription and that loss of either leads to AML differentiation and growth arrest[224]. Further, depletion of Cdc73 resulted in an upregulation of gene programs associated with differentiation. This is suggestive that loss of the PAF1c leads to a relief of the AML differentiation block by a loss of activation of genes associated with self-renewal. This provides evidence that the PAF1c may be a valuable therapeutic target in the treatment of AML, but we first need to elucidate its function in HSPCs.

To that end, we performed RNA-seq on Ckit+ HSPCs after inducing loss of Cdc73 (Figure 2-4). We revealed novel regulatory roles for the PAF1c in cell cycle processes. Of particular interest was the enrichment of a gene program after loss of Cdc73 indicating that genes associated with quiescence are upregulated when the PAF1c is depleted. A more thorough analysis of the role of the PAF1c in normal hematopoietic cells reveals that depletion of the PAF1c in the hematopoietic system leads to the depletion of HSPCs, and this phenotype is attributable to loss of cell cycle regulation, an increase in apoptosis, and an arrest of cell growth (Saha, et al. *In revision*). Importantly, we sought to determine whether there are differences in the PAF1c regulated transcriptome in AML or HSPCs (Figure 2-5). Strikingly, we see that the gene programs that are most changed and are most attributable to the phenotype in AML cells after loss of Cdc73 are not significantly changed after loss of Cdc73 in HSPCs. This suggests that the PAF1c regulates these self-renewal pathways specifically in AML and may account for the difference in requirement of the PAF1c-MLL interaction in hematopoietic cells compared to AML. While the PAF1c is clearly required for the growth of hematopoietic cells, the different requirements for protein-protein interactions and the difference in gene programs that are regulated by the PAF1c in these different cellular contexts suggest that there may be a therapeutic window in which the PAF1c can be targeted without severely disrupting normal hematopoietic

growth. However, to fully explore this possibility, a thorough study of the AML specific regulation of the PAF1c is necessary.

Chapter 3 A novel SETDB1-PAF1c interaction regulates *Hoxa9* and *Meis1* in AML

3.1 Introduction

The PAF1c is a highly conserved complex that was first identified in yeast as a transcriptional regulating complex that co-purified with RNA polymerase II (RNAPII)[173,176]. The PAF1c is composed of several subunits: PAF1, CDC73, CTR9, LEO1, RTF1, and the mammalian specific subunit WDR61[174,175,177]. While lacking any known catalytic activity itself, the PAF1c plays a critical role in the dynamic regulation of epigenetic landscapes at gene loci. The complex modulates epigenetic landscapes via protein-protein interactions with epigenetic modifying proteins[16,99]. For example, the PAF1c has been shown to be important for histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 79 dimethylation (H3K79me2) modifications through its interaction with Mixed Lineage Leukemia (MLL) histone methyltransferase and the Super Elongation Complex (SEC)[12,13,184]. Additionally, the PAF1c is necessary for H2B monubiquitination (H2Bub) through its recruitment of ringer finger proteins RNF20/40 and the ubiquitin ligase RAD6[94,95,185,187–192,345]. The PAF1c and these epigenetic modifications that it modulates are critical for transcriptional elongation at a subset of genes in yeast and mammals[170,181,183,184,188,224,345,346].

Importantly, there is a requisite role for the PAF1c complex in leukemias harboring a *MLL (KMT2a)* translocation[12–14]. MLL is a histone methyltransferase that deposits the H3K4me3 modification associated with promoter regions of actively transcribed genes. *MLL* is involved in chromosomal translocations with a variety of gene fusion partners that result in oncogenic MLL fusion proteins that drive transcription of genes critical for leukemogenesis, such as *MEIS1* and *HOXA9*[11,111,162]. In a series of studies to understand the regulation of these leukemogenic target genes, it was revealed that there is a direct physical interaction between the PAF1c and MLL or MLL-fusion proteins[12,13]. Importantly, disruption of the PAF1c-MLL interaction selectively inhibits the growth of MLL leukemias but is tolerated by normal hematopoietic cells

pointing to cancer specific functions for the PAF1c[14]. Despite these differences, the biochemical regulation of the PAF1c that allows for the dynamic regulation of target genes, such as *Hoxa9* and *Meis1*, remains poorly understood.

While much of the work on the PAF1c has demonstrated a role in active transcription elongation, there is also evidence that subunits of the PAF1c are involved in transcriptional repression. For instance, the PAF1 subunit is necessary for proper promoter-proximal pausing of RNAPII[172]. Further, the hyperactivation of a subset of transcriptional enhancers is restrained by PAF1 illustrating a role in transcriptional repression via enhancer regulation[171]. Furthermore, the PAF1c subunit CDC73 can transcriptionally repress oncogenic targets, such as *MYC* and *CCND1*[203,347]*.* Additionally, CDC73 has been shown to promote *CCND1* promoter H3K9me3 by recruitment of the H3K9 methyltransferases SUV39H1 or G9a[201]. This may contribute to tumor suppressor activity attributed to CDC73, which is mutated in hyperparathyroidism-jaw tumor syndrome (HPT-JT), exhibiting a loss of function mutation in more than 80% of these malignancies[348,349]. These studies suggest that the complex may be dynamically regulated to function as a transcriptional co-activator or co-repressor in different cellular contexts.

Here, we explore the biochemical regulation of the PAF1c in AML through an Affinity Purification-Mass Spectrometry (AP-MS) approach using CDC73 as bait. We identified both known and novel protein-protein interactions with the PAF1c. These interacting partners included a group of H3K9 methyltransferases, including a novel interacting partner SETDB1. H3K9 methyltransferases are epigenetic modifying proteins associated with transcriptional repression and heterochromatin formation[77]. There is also emerging evidence that SETDB1 plays a role in mediating H3K9me3 at dynamically regulated gene loci, such as the ID2 promoter[234,251,264,265,350,351]. The developmental *HoxA* gene cluster is also regulated by H3K9me3 in embryonic stem cells and melanoma cells[229,251]. Despite the importance of the *HoxA* gene cluster and co-factor *Meis1* in AML, we do not understand the role of H3K9me3 in regulating these genes in leukemic cells. In this study, we identified SETDB1 as a novel PAF1c interacting protein and explored the role of this interaction in modulating transcription of the known PAF1c pro-leukemic target genes *Hoxa9* and *Meis1*.

3.2 Methods

Plasmid cloning and mutagenesis

The human SETDB1 isoform 1 expression construct was a gift from Dr. Jianyong Shou and pcDNA3.1-HA-SETDB1 isoform 3 was a gift from Dr. Jean-Francois Rual. HA-SETDB1 isoform 1 was cloned into MSCVpuro retroviral vector and confirmed by sequencing. Human FLAG-CDC73 and HA-CDC73 were cloned into MigR1 retroviral vector and confirmed by sequencing. MigR1-HA-CDC73_3YF, MigR1-FLAG-CDC73_3YF, and MSCVpuro-HA-SETDB1_C1226A were generated using site-directed mutagenesis using the QuikChangeXL kit according to the manufacturer's protocol (Agilent). CMV-MYC-FLAG-PAF1, CMV-MYC-FLAG-CTR9, and CMV-MYC-FLAG-LEO1 were purchases from Origene.

Cell Line Generation and Cell Culture Conditions

Cell lines were generated as described in Chapter 2. After establishment of MLL-AF9 cell lines on a *Cdc73fl/fl*CreER or a wildtype background, AML cells were spinfected with MigR1(EV), MigR1-CDC73, or MigR1-CDC73_3YF; or MSCVpuro-HA-SETDB1, MSCVpuro-HA-SETDB1_CD, or MSCVpuro(EV), respectively. Separately, M1 murine leukemia cells were spinfected with MigR1(EV), MigR1-CDC73, or MigR1-CDC73_3YF. Cells transduced with MLL-AF9 were selected with 1mg/mL neomycin. Cells transduced with MigR1, CDC73, or CDC73_3YF were sorted for GFP positivity. Cells transduced with MSCVpuro, SETDB1, or SETDB1_CD were selected with 1ug/mL puromycin. All bone marrow-derived MLL-AF9 cells were cultured in IMDM supplemented with 15% Stem Cell FBS (Millipore), 1%pen/strep, and 10ng/mL IL-3 (R&D). M1 murine AML cells were cultured in RPMI supplemented with 10% FBS and 1% pen/strep. HEK293T cells are grown in DMEM supplemented with 10% FBS.

Proliferation Assays and Luciferase Assays

MLL-AF9-Cdc73fl/fl-CreER^{T2} cells expressing MigR1, MigR1-HA-CDC73, or MigR1-HA-CDC73_3YF were seeded at 5x10^4 cells in 2mL of normal growth media containing either 2.5nM 4-OHT or vehicle. Viable cell number was counted each day for 4 days using Trypan Blue. On day 2, the cells were supplemented with fresh media, IL-

3, and tamoxifen or vehicle. Luciferase assays were performed using the Dual Luciferase assay kit and a GloMax 20/20 Luminometer (Promega) as previously described [13].

Colony Formation Assays

MLL-AF9-Cdc73fl/fl-CreER^{T2} cells expressing MigR1, MigR1-HA-CDC73, or MigR1-HA-CDC73_3YF grown in normal growth media were pretreated with 2.5nM 4-OHT or vehicle for 24 hours. They were then seeded at a density of 1x10^3 cells in 2mL semi-solid methylcellulose medium for mouse cells (STEMCELL M3234) containing 10ng/mL IL-3. Colonies were counted and 2x images of the 5-phenyl tetrazolium chloride (INT) stained dishes were taken after 7 days of growth.

Immunoprecipitations (IP) and antibodies

For HEK293T transient transfection experiments, cells are transfected following the Fugene®6 (Promega). Cells were collected, lysed and IP'd with anti-HA high affinity beads (Roche) or M2 anti-FLAG magnetic beads (Sigma). Antibodies used for western blotting were anti-HA (Abcam9110), anti-FLAG (Sigma F7425), anti-SETDB1 (Abcam107225, Bethyl A300-121), from Bethyl: anti-EHMT2 (A301-642), anti-PAF1 (A300-172), anti-CTR9 (A301-395), anti-LEO1 (A300-174), or anti-WDR61 (A305-191). Antibodies used for western blotting and ChIP were anti-H3 (Abcam1791, western blot 1:4000), anti-H3K9Me3(Abcam 8898), anti-rabbitIgG (Millipore 12-370). All antibodies were generated in rabbit and were used at a dilution of 1:1000 for western blot and 4ug/reaction for ChIP unless otherwise noted.

Affinity Purification-Mass Spectrometry

1x10^9 cells M1 murine AML cells that stably express FLAG-CDC73 (n=2) or FLAG-CDC73_3YF (n=3) were harvested and lysed in 300mM KCl lysis buffer containing protease inhibitors and IGEPAL CA-630. Lysates were incubated with M2 FLAG magnetic beads (Sigma). Beads were washed 6 times with 0.3M-1M KCl and eluted with 15ug 3X FLAG peptide. Proteins were denatured in 8M urea. Cysteines were reduced with 10 mM DTT and alkylated using 50 mM chloroacetamide. Proteins

were digested with 500 ng of sequencing grade, modified trypsin (Promega). Reaction was terminated by acidification with trifluoroacetic acid (0.1% v/v) and peptides were purified using SepPak C18 cartridge following manufacturer's protocol (Waters Corp) and dried. Peptides were reconstituted in HPLC loading buffer and resolved on a nano-capillary reverse phase column (Acclaim PepMap C18, 2 micron, 50 cm, ThermoScientific) using 0.1% formic acid/acetonitrile gradient at 300 nl/min (2-25% acetonitrile in 105 min; 25-40% acetonitrile in 20 min followed by a 90% acetonitrile wash for 10 min and a further 30 min re-equilibration with 2% acetonitrile) and directly introduced in to *Orbitrap Fusion Tribrid* mass spectrometer (Thermo Scientific, San Jose CA). MS1 scans were acquired at 120K resolution (AGC target=$2e^5$, max IT=50ms). Data-dependent high-energy C-trap dissociation MS/MS spectra were acquired for the most abundant ions for 3 seconds following each MS1 scan (15K resolution; AGC target=$5e^4$; relative CE ~32%). Proteins were identified by searching the data against *Mus musculus* (Swissprot, v2016-04-13) using SEQUEST-HT (Proteome Discoverer v2.1, Thermo Scientific). Search parameters included MS1 mass tolerance of 10 ppm and fragment tolerance of 0.05 Da; two missed cleavages were allowed; carbamidimethylation of cysteine was considered fixed modification and oxidation of methionine, deamidation of asparagine and glutamine, phosphorylation of serine, threonine and tyrosine were considered as potential modifications. False discovery rate (FDR) was determined using Percolator and proteins/peptides with an FDR of ≤1% were retained for further analysis. Raw data files were uploaded to PRIDE data repository.

Scoring of protein-protein interactions

Interactions with CDC73 and CDC73_3YF were scored using MS2 spectral counting (PSM counts). Using PSM counts as measures of protein abundance in each sample, SAINT probabilities for each interaction were calculated using the CRAPome online resource [352–354]. Three FLAG-IP replicates of cells expressing only the empty vector MigR1 were used as controls. SAINT calculates the probability that an interaction is a true positive using a model where true-positive and false-positive interactions for each bait have distinct Poisson distributions. A value of 1 indicates a high probability of

a bona-fide interaction [352]. SAINT parameters used were: average=best 2 replicates; virtual controls=10; iter(2000,4000); normalization=1. SAINT probabilities for all identified proteins are found in Supplemental Table 1. Proteins with a SAINT probability of >=0.7 for either or both bait proteins were kept as potential interacting proteins. We began our filtering with this relatively low threshold due to the low IP efficiency of CDC73_3YF relative to CDC73 and the possibility that our phenotypic effects were due to a more transient interaction. For the discovery of protein interactions involved in transcriptional repression that were potentially bound specifically to CDC73_3YF, all prey proteins with a SAINT probability >=0.7 were analyzed using GeneMANIA [355]. The network of interactions derived using GeneMANIA was filtered to include only those proteins with a published physical protein-protein interaction, including those contained in the BioGRID interaction database [356]. This protein-protein interaction network was then investigated to observe biologically interesting interaction subnetworks. When an interesting sub-network contained at least three proteins with SAINT probability > 0.7, that sub-network was used as a new search node for GeneMANIA. Protein-protein interaction networks shown in Fig. S2 were generated with GeneMANIA. Networks shown in Fig. 1 were adapted from GeneMANIA using Microsoft PowerPoint 2011.

Quantitative PCR (qPCR) for gene expression

RNA was harvested from MLL-AF9 + MSCVpuro control cells, MLL-AF9 + SETDB1, or MLL-AF9 + SETDB1_CD overexpression cells; or from MLL-AF9-Cdc73fl/fl-CreER^{T2} MigR1, CDC73, or CDC73_3YF cells treated with 2.5-5nM tamoxifen or vehicle control. 1-5x10^6 cells were harvested and RNA was extracted using the Qiagen RNeasy mini plus kit. cDNA synthesis was performed using oligo-dT priming and the SuperScript III kit (Invitrogen). qPCR was performed using the fast SYBR-green mastermix protocol (Thermo Fisher). Primer sets used were: Meis1_F- 5'ATCAGAGCGCCAGGACCTAT3'; Meis1_R- 5'CTTCCCCCTGGCTTTCGATT3'; Hoxa9_F- 5'GAATGAGAGCGGCGGAGAC3'; Hoxa9_R- 5'GAGCGAGCATGTAGCCAGTTG3'; β-Actin_F- 5'GCCCTGAGGCTCTTTTCCAG3'; β-Actin_R- 5'TGCCACAGGATTCCATACCC3'. Statistical analysis was performed using two sample t-tests.

Chromatin immunoprecipitation-qPCR (ChIP-qPCR)

ChIP experiments were performed as previously described in MLL-AF9 + MSCVpuro control cells, MLL-AF9 + SETDB1, or MLL-AF9 + SETDB1_CD cells; or in MLL-AF9-Cdc73fl/fl-CreER^{T2} MigR1, CDC73, or CDC73_3YF cells treated with 2.5-5nM tamoxifen or vehicle control[357]. Briefly, 3x10^7 AML cells were crosslinked with 1% formaldehyde, lysed with 1% SDS and sonicated on a Bioruptor Pico sonication device (Diagenode). Cleared lysates were immunoprecipitated with anti-Histone H3 or anti-Histone H3 (trimethylated K9) using Protein G dynabeads. The IPs were washed with a low salt buffer, a high salt buffer, and a stringent lithium chloride wash buffer. Protein-DNA complexes were eluted in 1% SDS, were decrosslinked in high salt and treated with RNaseA and ProteinaseK. DNA was purified with a Qiagen PCR purification kit. qPCR was performed using the fast SYBR-green mastermix protocol. Primer sets used were: Meis1_promoter_F-5'TCAAAGTGACAAAATGCAAGCA3'; Meis1_promoter_R-5'CCCCCCGCTGTCAGAAG3'; Hoxa9_promoter_F-5'TGACCCCTCAGCAAGACAAAC 3'; Hoxa9_promoter_R- 5'TCCCGCTCCCCAGACTG 3'. ChIP-qPCR data were analyzed for statistical significance using ANOVA on fitted linear models. ANOVA models in SAS (PROC GLM) or R (lm) were used to compare treated and control groups. The model includes group (treated/control), cell lines and the interaction between group and cell lines. If the interaction is not significant, we report the p-value for the main effect of group. Significance is determined if $p<0.05$. All analyses were conducted using SAS (version 9.4, SAS Institute, Cary, NC) or R. All graphs were generated in R or Prism (version 7.0c). All technical replicate (3 per biological replicate) values were included for statistical analyses.

3.3 Results

Point mutations in CDC73 disrupt AML cell growth without affecting PAF1c complex integrity

Recent studies have linked the PAF1c subunit Cdc73 with the WNT, Hedgehog and Notch signaling pathways through protein interactions with β-catenin, Gli1 and Notch intracellular domain [209,210]. A tyrosine mutational analysis has identified a trio of

tyrosine residues on CDC73 that regulate its interaction with β-catenin. Mutation of these residues to phenylalanine stabilizes the CDC73 interaction with β-catenin and enhances WNT signaling in gastric carcinoma cells [208,209]. A critical role for β-catenin in MLL rearranged leukemias prompted us to investigate this triple tyrosine to phenylalanine mutant, CDC73-Y290/293/315F (CDC73_3YF) (Fig. 3-1A) AML cells [358–

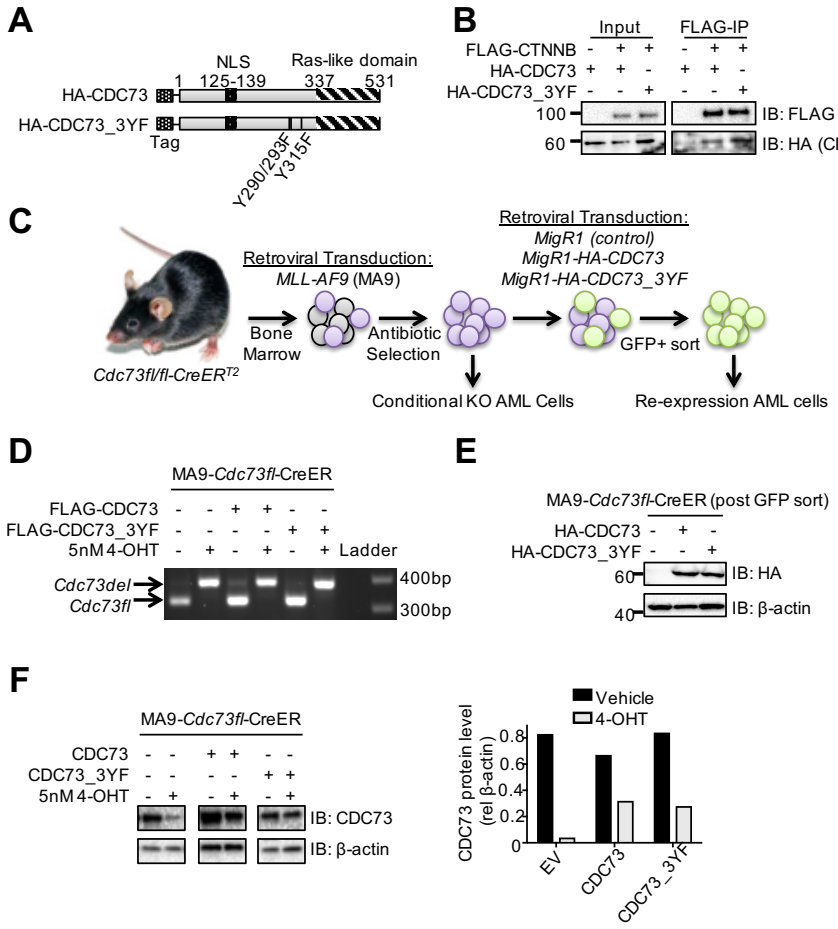


Figure 3-1: CDC73_3YF is a CDC73 mutant that has an enhanced interaction with β-catenin. A) Schematic of HA-CDC73 showing known domains of CDC73 and the location of the three tyrosine mutations introduced to make CDC73_3YF. B) HEK293T cells were transiently co-transfected with either HA-CDC73 or HA-CDC73_3YF and FLAG- β-catenin. FLAG-IPs were performed and immunoblotted with the indicated antibodies. **C)** Diagram representing the workflow for the generation of MLL-AF9 transformed *CDC73fl/fl*-CreERT2(MA9-*Cdc73fl*-CreER^{T2}) CDC73 or CDC73_3YF re-expression cells, as well as MigR1 control cells (EV Control). Mouse image from Taconic Biosciences, Inc.[392] D) Genotyping of MLL-AF9-*Cdc73fl/fl*-CreER^{T2} cells treated with 4-OHT or vehicle control to confirm genetic excision of the *Cdc73* allele. E) Western blot of whole cell lysates from MLL-AF9 transformed *CDC73fl/fl*-CreER^{T2} CDC73 or CDC73_3YF re-expression cells and immunoblotted with the indicated antibodies demonstrating expression of HA-CDC73 and HA-CDC73_3YF. F) Whole cell lysates from MLL-AF9-*Cdc73fl/fl*-CreER^{T2} cells treated with 4-OHT or vehicle control immunoblotted with anti-CDC73 antibody (left). Densitometry was performed on CDC73 protein bands and values were plotted relative to β-actin protein band signal (right).

[360]. As reported, CDC73_3YF displayed enhanced interaction with β-catenin compared to wild type CDC73 following transient transfection of HEK293T cells (Fig. 3-1B). To explore the biological impact of CDC73_3YF in AML cells, we transduced *Cdc73fl/fl*-CreER^{T2} mouse bone marrow with MLL-AF9 (MA9) packaged retrovirus to generate stable AML cell lines that can be induced to genetically excise *Cdc73* by treatment with 4-hydroxytamoxifen (tamoxifen) (Fig. 3-1C, 3-1D) [321]. Tamoxifen treatment results in almost complete loss of the CDC73 protein by 48 hours (Fig. 3-1E)[14,224]. We used

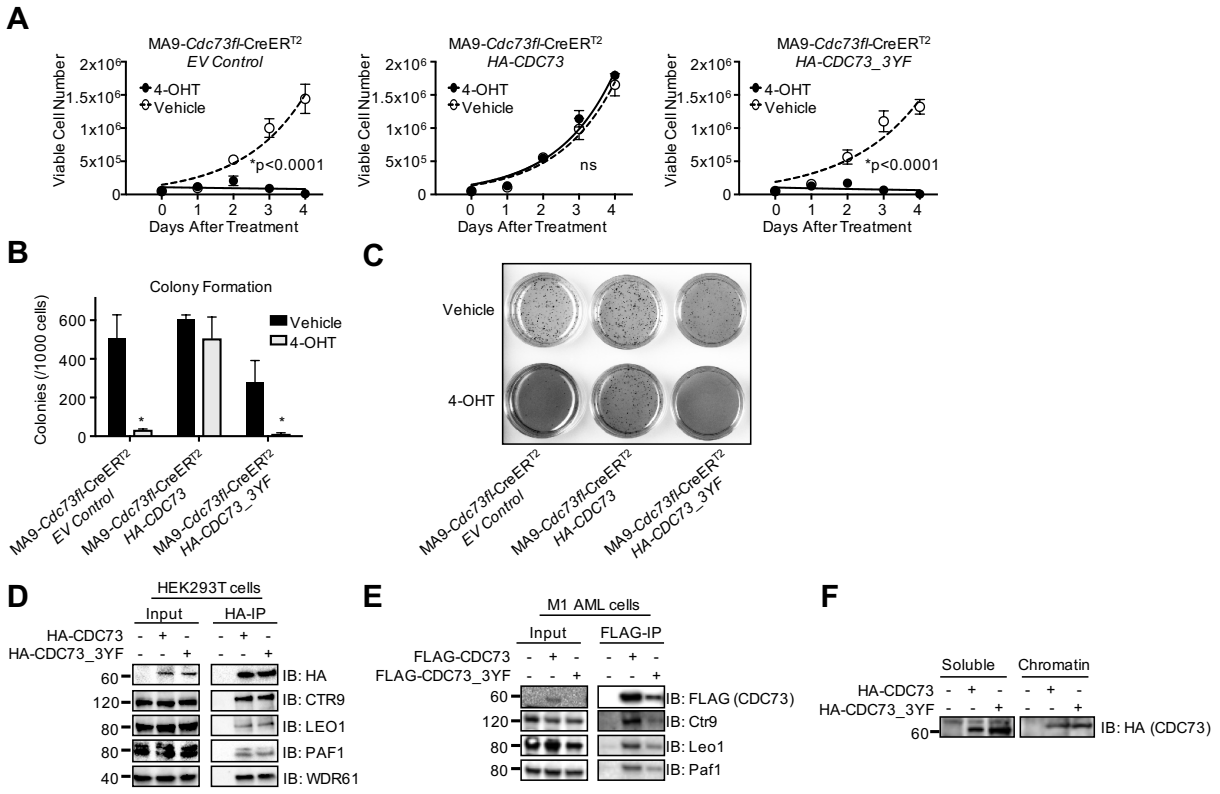


Figure 3-2: A CDC73 tyrosine mutant does not support AML cell growth. A) The indicated MLL-AF9 transformed *CDC73fl/fl*-CreER^{T2} CDC73 re-expression cells were plated on day 0 in the presence of 4-OHT or vehicle control. Viable cells were counted every day for four days. Shown is one representative assay of n>5 biological replicates. B) The indicated MLL-AF9 transformed *CDC73fl/fl*-CreER^{T2} CDC73 re-expression cells were pretreated with 4-OHT or vehicle for one day prior to being plated in semisolid methylcellulose. Colonies were counted after 7 days (n=2 biological replicates). C) Representative 1x magnification images of INT stained colonies that are quantified in E. *p<0.05. D) Wild type and CDC73_3YF were immunoprecipitated following transient transfection of HEK293T cells. HA-IPs were performed followed by immunoblots with the indicated antibodies for PAF1c components. E) M1 mouse AML cells that express stable retroviral FLAG-CDC73 or FLAG-CDC73_3YF are subjected to FLAG-IPs. PAF1c proteins are detected using western blotting with the indicated antibodies. (F) HEK293T were transiently transfected with CDC73 or CDC73_3YF. Whole cell lysates were collected and run as soluble fraction. Cells were lysed, the insoluble chromatin fraction was spun down and washed three times, then immunoblotted with HA antibody as the chromatin fraction.

retroviral transduction to stably express *CDC73* or *CDC73_3YF* in *Cdc73fl/fl*-CreER^{T2} AML cells to test the rescue capacity of CDC73_3YF upon loss of *Cdc73* (Fig. 3-1C). We confirmed expression of the tagged CDC73 and CDC73_3YF (Fig. 3-1E). Upon deletion of *Cdc73*, we confirmed that CDC73 protein was reduced to <5% of vehicle treated cells (Fig. 3-1F). Expression of wild type CDC73 or CDC73_3YF following tamoxifen treatment resulted in protein levels at about 50% that of endogenous Cdc73 observed in vehicle treated cells expressing an empty vector control (MigR1) (3-1F). Proliferation assays demonstrate that, following excision of *Cdc73,* MLL-AF9 cells expressing an empty vector (MigR1) exhibit a significant reduction in cell proliferation, whereas re-expression of CDC73 fully rescued proliferation similar to that of cells treated with vehicle control (Fig. 3-2A). Surprisingly, despite stabilized interaction with β-catenin, CDC73_3YF failed to rescue cellular proliferation similar to MigR1 control cells (Fig. 3-2A). We further explored the effects of re-expression of CDC73_3YF on the colony forming unit ability of AML cells. Following excision of *Cdc73,* MLL-AF9 cells expressing an empty vector showed a significant reduction in colony forming unit capacity. This phenotype was rescued by re-expression of CDC73 but not CDC73_3YF suggesting this mutant does not support leukemic colony forming unit potential (Fig. 3-2B, C). To confirm that CDC73_3YF is capable of binding to the other components of the PAF1c, we transiently transfected HEK293T cells with HA-CDC73 or HA-CDC73_3YF and performed immunoprecipiation (IP)-western blots. We observed co-precipitation of CTR9, LEO1, PAF1, and WDR61 with both CDC73 and CDC73_3YF (Fig. 3-2D). We also confirmed PAF1c co-purification with CDC73_3YF in M1 mouse AML cells that stably express retroviral FLAG-CDC73_3YF (Fig. 3-2E). We further observed that HA-CDC73 and HA-CDC73_3YF are present in the chromatin fraction of transiently transfected HEK293T cells (Fig. 3-2F). These data indicate that the CDC73_3YF mutant assembles into the PAF1c and stabilizes interaction with β-catenin, but does not support proliferation and colony formation capacity of AML cells.

CDC73_3YF stabilizes interaction with H3K9 methyltransferases SETDB1 and G9

To determine whether the proliferation deficient phenotype associated with CDC73_3YF was due to differences in protein-protein interactions, we performed

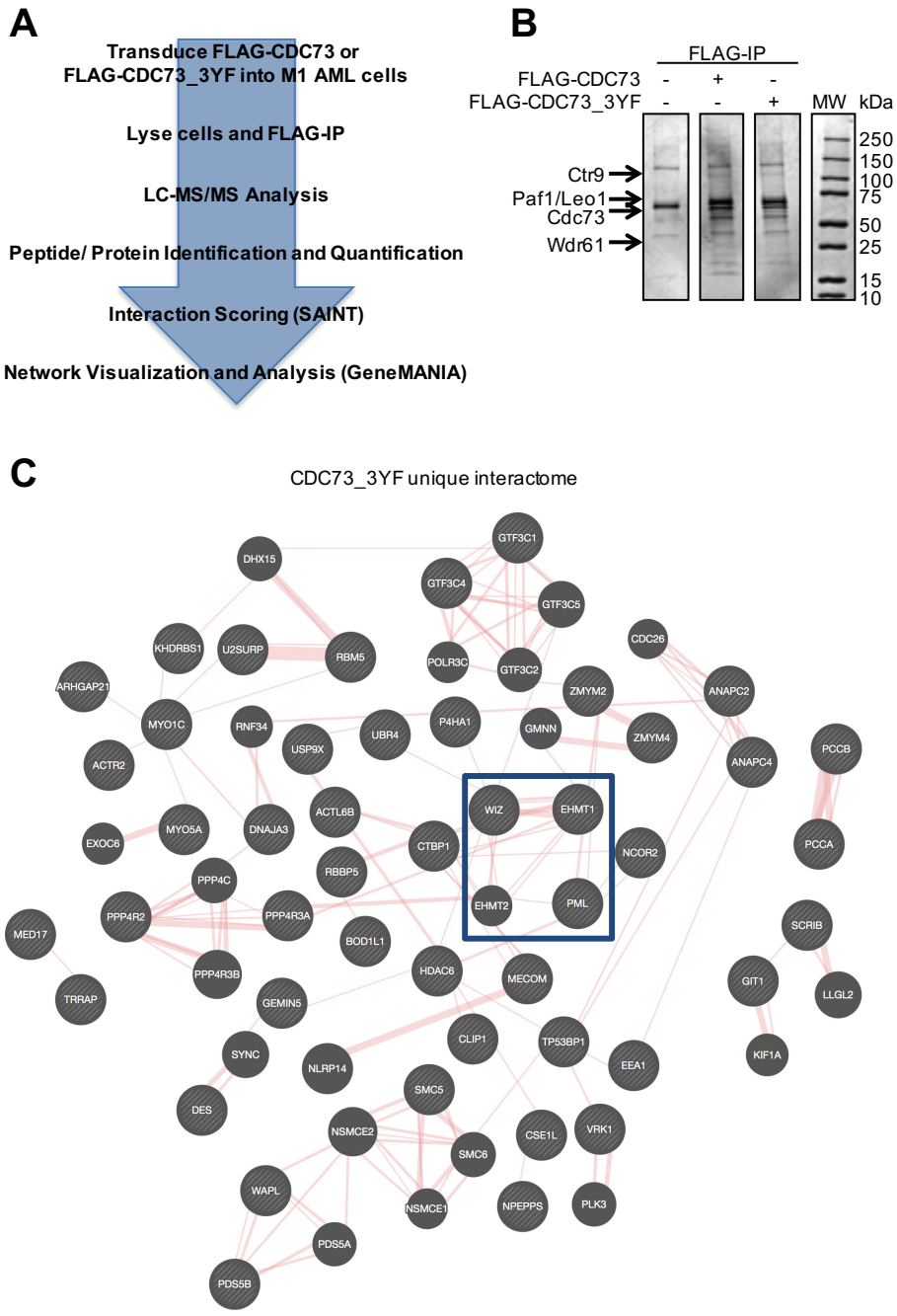
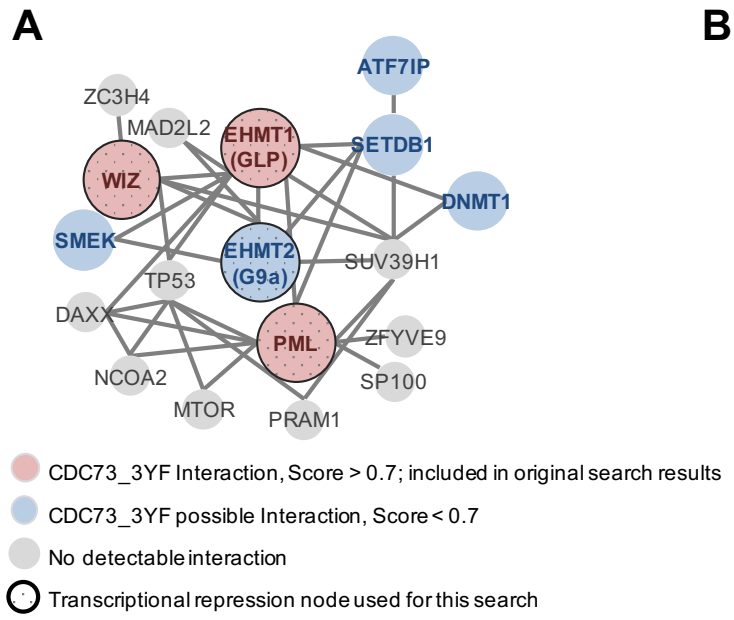


Figure 3-3: CDC73 and CDC73_3YF have overlapping and distinct interactomes. A) Schematic demonstrating the approach for the AP-MS experiment to determine interaction partners of CDC73 and CDC73_3YF in M1 mouse leukemia cells. B) FLAG-IPs of HEK293T cells transiently transfected with CDC73 or CDC73_3YF were eluted and run on a gel that was stained with Coomassie Blue. Arrows indicate the appropriate size for protein bands for the given PAF1c components. C) Interaction network output by GeneMANIA for proteins that co-purified with CDC73_3YF in the AP- MS experiment. The interactions were filtered to include only those described as physical interactions. Circles containing diagonal lines indicate proteins included in the search node. The blue rectangle indicates the sub-network of proteins that was found to be associated with transcriptional repression. The proteins contained in this rectangle were used in the subsequent targeted GeneMANIA search.

Affinity Purification-Mass Spectrometry (AP-MS) to find the interactome of CDC73 and

CDC73_3YF (Fig. 3-3A). First, HEK293T cells were transiently transfected with FLAG tagged CDC73_3YF or wildtype CDC73 and subjected to FLAG IP. Bait and co-purifying proteins were observed by Coomassie Blue staining, and differential banding patterns suggested different interactomes for CDC73 and CDC73_3YF (Fig. 3-3B). To identify CDC73 interactions that are specifically relevant to AML, M1 murine AML cells were stably transduced with FLAG-CDC73 or FLAG-CDC73_3YF. Co-IPs from M1 cells were analyzed using liquid chromatography tandem mass spectrometry (LC-MS/MS). Using the Contaminant Repository for Affinity Purification online resource (CRAPome), protein-protein interactions were scored using Significance Analysis of Interactome (SAINT) probabilistic scoring (Fig. 3-3A). Interestingly, there was a distinct subset of 87 potential CDC73_3YF interacting proteins that were not found in the interactome of CDC73. Analysis of this interactome using the GeneMANIA database to search for previously reported protein-protein interactions uncovered a group of interacting proteins in this group associated with transcriptional repression (Fig. 3-3C). We studied this group of proteins as a new search node in GeneMANIA and cross-referenced this



B

Protein	CDC73 PSMs	CDC73_3YF PSMs	Control PSMs
PAF1	438	477	23
LEO1	205	253	6
CTR9	1580	1053	30
CDC73	2029	870	28
WDR61	246	176	3
RTF1	0	0	0
WIZ	0	12	0
PML	0	13	1
ATF7IP	0	12	4
EHMT1	0	9	0
G9a	0	3	0
SETDB1	0	2	0

Known CDC73 interactors (PAF1c components)
H3K9Methyltransferases and interactors

Figure 3-4: CDC73_3YF binds to a network of proteins associated with transcriptional repression. A) Interaction network for proteins that co-purified with CDC73_3YF adapted from an output from GeneMANIA using the proteins marked by solid borders as a search node. Solid lines connecting proteins indicate a reported physical interaction between two proteins. Circle sizes are proportional to the degree of connectivity to the proteins in the search node. B) Total peptide spectrum matches (PSMs) for the PAF1c or proteins associated with H3K9 methyltransferase components generated from the AP-MS experiment for either CDC73 (sum of 2 replicates), CDC73_3YF (sum of 3 replicates), or MigR1 control (sum of 3 replicates).

protein interaction network with our AP-MS data. This identified a group of epigenetic modifying proteins that catalyze histone H3 lysine 9 methylation (H3K9me) that preferentially associated with CDC73_3YF (Fig. 3-4A, B). These include: EHMT1 (GLP), EHMT2 (G9a) and SETDB1. In addition, several proteins associated with these H3K9 methyltransferases co-purified with CDC73_3YF, including WIZ, which associates with GLP and G9a; PML, which associates with GLP and SETDB1; and ATF7IP which associates with SETDB1 [236,265,361]. (Fig. 3-4A, B)

CDC73 has previously been reported to interact with H3K9 methyltransferases G9a and SUV39H1. Yang *et al*. demonstrated that CDC73 recruits these methyltransferases to the *CCND1* promoter in HeLa cells and promotes H3K9 di- and tri- methylation (H3K9me2/3) [201]. For this study, we focused on H3K9 methyltransferases that were found specifically in our AP-MS data. Despite having the highest SAINT probability score, we were unable to validate an interaction between CDC73 and GLP, possibly related to antibody efficiency (data not shown). However, IP-western blots demonstrated that HA-CDC73 and HA-CDC73_3YF both pulled down endogenous SETDB1 and G9a in transiently transfected HEK293T cells. Consistent with the AP-MS data, there was a stabilized interaction between CDC73_3YF and SETDB1 or G9a compared to the interactions with CDC73 (Fig. 3-5A, B). We confirmed this interaction between endogenous proteins in human AML THP-1 cells by subjecting cells to a CDC73-IP. Immunoblotting revealed an interaction between endogenous CDC73 and endogenous SETDB1 and G9a (Fig. 3-5C, D). We were also interested in whether this interaction was a PAF1c dependent interaction or an independent function of CDC73. We therefore performed IPs on FLAG-tagged PAF1c components CTR9 and LEO1 in transiently transfected HEK293T cells. We found that CDC73, CTR9, and to a lesser degree LEO1 co-immunoprecipitate endogenous SETDB1 and G9a suggesting the interactions occur with the PAF1c (Fig. 3-5E, F). Due to the novelty of the SETDB1 interaction, we focused our studies on SETDB1. To further confirm the stabilized interaction with CDC73_3YF and to determine the SETDB1 domains that are involved in the CDC73 interaction, we utilized a natural isoform of SETDB1 (isoform 3; NCBI reference sequence NP_001157114.1) that lacks the methyl binding domain (MBD) and

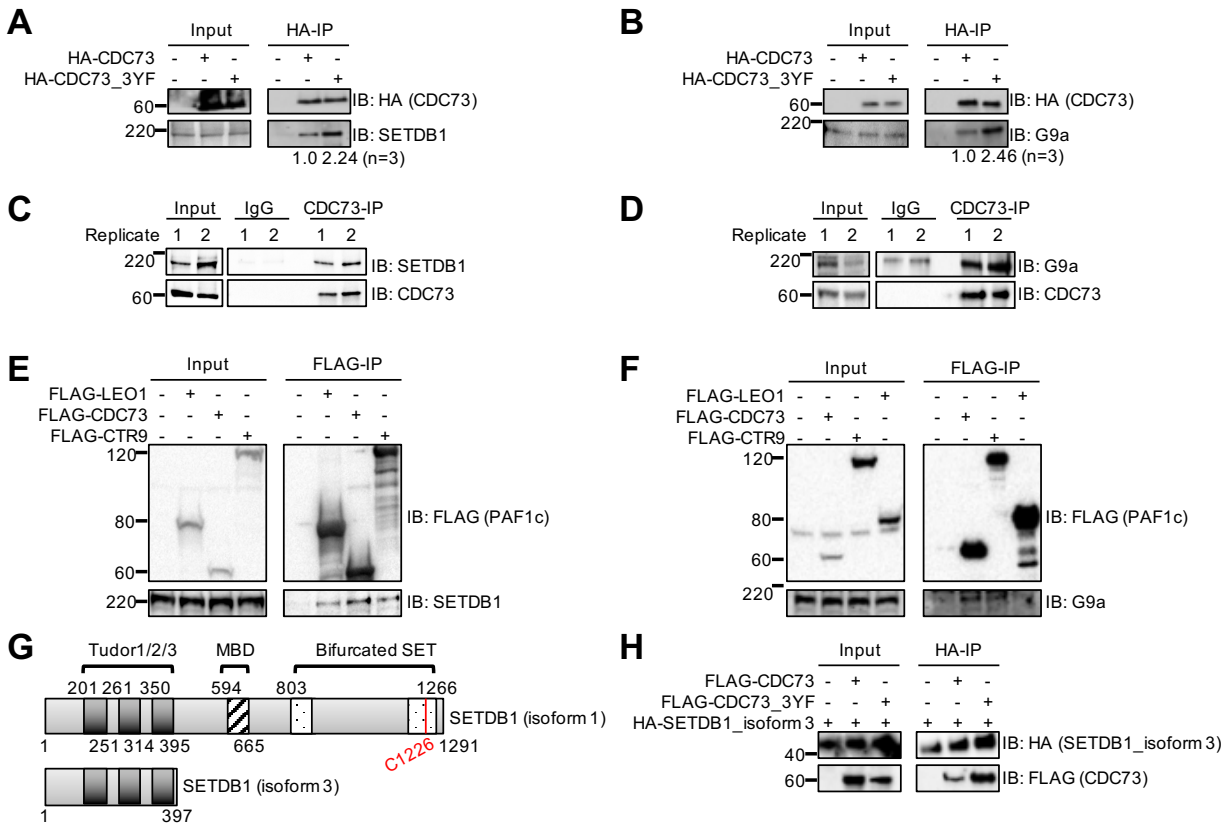


Figure 3-5: The PAF1c interacts with H3K9methyltransferases SETDB1 and G9a. A-B) HEK293T cells were transiently transfected with HA tagged CDC73 or CDC73_3YF. HA- IPs were performed and co-purifying H3K9 methyltransferase proteins were detected by western blotting with indicated antibodies. Densitometry was performed on H3K9methyltransferase bands normalized to the HA-IP bait bands. Shown is the average of n=3 quantifications. C-D) CDC73-IPs were performed in THP-1 human AML cells and co-purifying H3K9 methyltransferase proteins were detected by western blotting with indicated antibodies. Two biological replicates are shown. E-F) Flag tagged PAF1c components were transiently expressed in HEK293T cells. FLAG-IPs were performed and co-purifying proteins were detected with the indicated antibodies for H3K9 methyltransferases. Shown are representative blots of n=2-3 biological replicates. G) Schematic showing the domains that are conserved between SETDB1 isoform 1 (top) and SETDB1 isoform 3 (bottom) as reported by Jurkowska, et al. and annotated in UniProt Q15047. MBD = methyl CpG binding domain. H) HA-tagged isoform 3 of SETDB1 and FLAG-CDC73 or FLAG-CDC73_3YF were transiently transfected in HEK293T cells. HA-IPs were performed and the CDC73 constructs were detected with a FLAG immunoblot. Shown is a representative blot of n=4 biological replicates. EV=Empty Vector control (MigR1).

the catalytic bifurcated SET domain (Fig. 3-5G) [362,363]. HEK293T cells were transiently transfected with HA-hSETDB1_isoform 3 and FLAG-CDC73 or FLAG-CDC73_3YF. The cells were then subjected to HA-IPs and western blots. We observed that CDC73 copurified with the shorter isoform of SETDB1, and CDC73_3YF demonstrated a stronger co-purification (Fig. 3-5H). Together, these data show that CDC73 interacts with at least SETDB1 and G9a and that CDC73_3YF stabilizes interaction with these H3K9 methyltransferases.

SETDB1 modulates expression of PAF1c oncogenic target genes

We hypothesized that the function of the PAF1c may be regulated, in part, via its interactions with SETDB1. To test this, we investigated transcription of PAF1c targets in AML cells expressing CDC73_3YF. We collected RNA from MLL-AF9-*Cdc73*fl/fl-CreER^{T2} cells expressing either CDC73 or CDC73_3YF treated with tamoxifen or vehicle. Interestingly, while CDC73 completely rescued expression of known PAF1c targets *Meis1* and *Hoxa9* upon tamoxifen induced deletion of *Cdc73*, CDC73_3YF was incapable of rescuing expression (Fig. 3-6A, 4B). To more directly assess the effects of CDC73_3YF on transcription, we utilized a luciferase reporter construct. We used a *Hoxa9* luciferase reporter, which is a known MLL-AF9 target gene that is dependent on the PAF1c for full expression [13]. As expected, overexpressing CDC73 significantly augmented MLL-AF9 mediated transactivation of the *Hoxa9* promoter in a dose dependent manner (Fig. 3-6C). Interestingly, overexpression of CDC73_3YF displayed no transcriptional synergy with MLL-AF9 in activating the *Hoxa9*-luciferase reporter (Fig. 3-6C), consistent with *Hoxa9* transcript levels observed in CDC73_3YF expressing MLL-AF9 cells (Fig. 3-6B). Due to the role of H3K9 methyltransferases in repressing gene transcription, we hypothesized that the transcriptional phenotype associated with *Cdc73*-/- cells re-expressing CDC73_3YF was due, in part, to H3K9 methyltransferase mediated transcriptional repression of known PAF1c target genes. Thus, we generated MLL-AF9 AML cell lines that overexpress human MSCV-HA-SETDB1 (referred to hereafter as SETDB1) and MLL-AF9-MSCV control cell lines by retroviral transduction and collected RNA for gene expression analysis. qPCR experiments demonstrate reduced expression of *Meis1* and *Hoxa9* in MLL-AF9 cells overexpressing SETDB1 compared to control MLL-AF9 cells (Fig. 3-6D). To determine whether the methyltransferase activity of SETDB1 is necessary for the reduction of *Meis1* and *Hoxa9* expression, we generated MLL-AF9 AML cell lines expressing the catalytic dead SETDB1_C1226A (SETDB1_CD) (Fig. 3-5G) [364]. qPCR demonstrates that overexpression of this mutant did not lead to a significant change in *Meis1* or *Hoxa9* expression relative to MLL-AF9 control cells, suggesting an important role for the methyltransferase activity of SETDB1 in regulating the transcription of these genes (Fig.

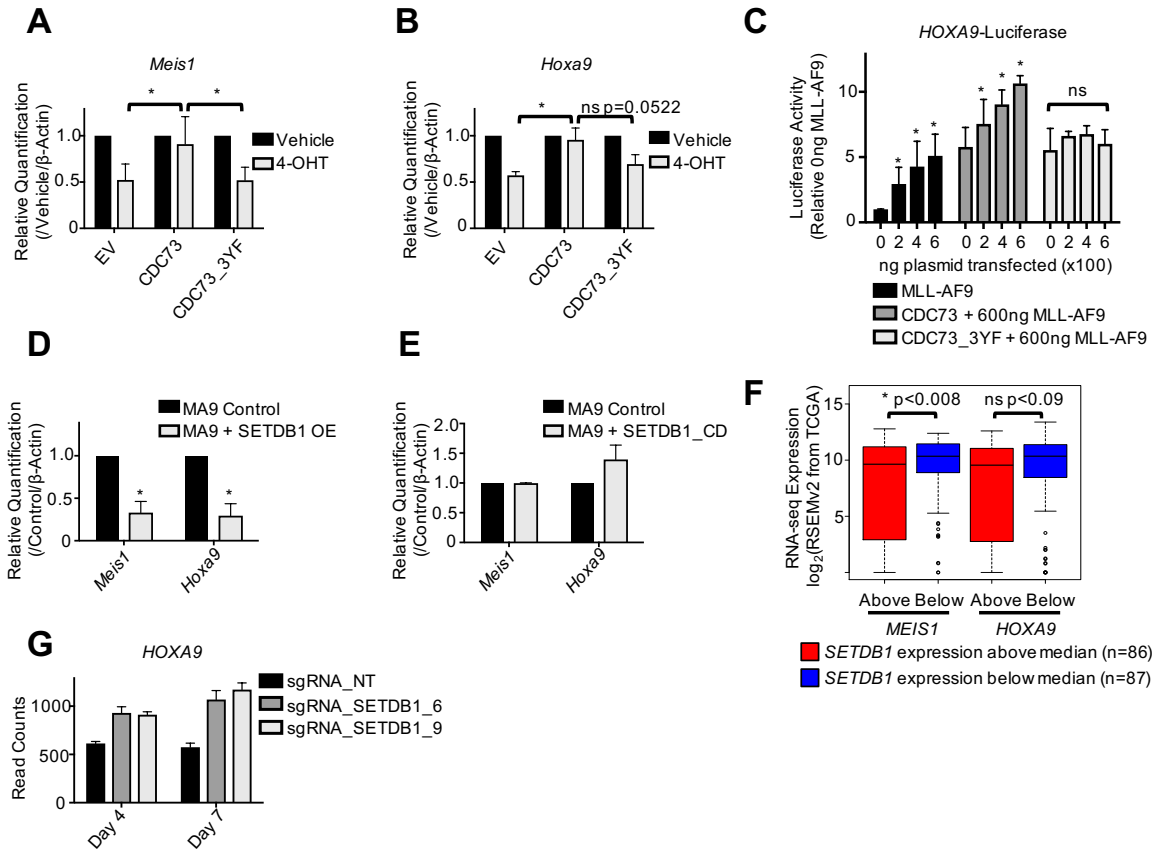


Figure 3-6: SETDB1 and CDC73_3YF mediated repression of *Hoxa9* and *Meis1* transcription. A-B) MLL-AF9 transformed *CDC73fl/fl*-CreER^{T2} CDC73 re-expression cells were plated in the presence or absence of 4-OHT and RNA was collected after 48 hours. qPCR was performed to detect expression of *Meis1* or *Hoxa9*. 4-OHT treated group was normalized to vehicle treated group for each cell type. (biological replicates n=5, n=3, respectively). C) HEK293T cells were transiently transfected with the indicated plasmids, Hoxa9 luciferase, and Firefly-Renilla. Luminescence readings were taken at 48 hours and plotted relative to empty vector control transfections. Statistics were calculated using 2-way ANOVA with post-hoc Dunnett's testing (biological replicates n=5). D-E) qPCR detection of *Meis1* and *Hoxa9* expression in MLL-AF9 control cells or MLL-AF9 co-transduced with (D) SETDB1 or (E) SETDB1_CD overexpression vector. (biological replicates n=2-3). F) RNA-seq data, downloaded through the cBioPortal analytical tool, was deposited to the TCGA by Ley et al. Data was mined for expression correlation between *SETDB1* and *MEIS1* or *HOXA9*. Patient samples were divided into *SETDB1* expression higher than the median and those that were lower, and the gene of interest expression was plotted on the y-axis. G) RNA-seq data, downloaded through the Gene Expression Omnibus (GEO), was deposited by Cuellar et al. Data was mined for HOXA9 expression in human THP1 cells expressing Cas9 and one of two small guide RNAs (sgRNA_SETDB1_6 and sgRNA_SETDB1_9) designed to interfere with SETDB1 expression. Control cells expressed Cas9 and a non-targeting sgRNA (sgRNA_NT). RNA was collected and sequenced at two different time points- 4 days after introduction of the sgRNA and 7 days after introduction of the sgRNA (n=3). $^{*}p<0.05$; EV = Empty Vector control (MigR1); OE = Overexpression.

3-6E). To determine if increased expression of *SETDB1* correlates with the reduced expression of *MEIS1* and *HOXA9* in human AML samples, we mined RNA-seq data from 173 AML patients deposited in The Cancer Genome Atlas (TCGA) [1]. Consistent

with our data, patient samples with *SETDB1* expression that was greater than the median expression of all samples in the set (n=86) had a significantly lower expression of *MEIS1* (Fig. 3-6F). This trend held true for *HOXA9* expression, though the difference was not significant (Fig. 3-6F). To explore the effect of SETDB1 loss of function on AML cells, we mined data from a recent study that utilized CRISPR-Cas9 targeting of SETDB1 in the human MLL-AF9 driven AML cell line, THP-1 [267]. Loss of SETDB1 protein in THP-1 cells leads to increased *HOXA9* expression at both four days and seven days after introduction of small guide RNAs (sgRNAs) targeting Cas9 to *SETDB1*, consistent with a repressive role for SETDB1 at the *HOXA9* locus (Fig. 3-6G). These data suggest that SETDB1 overexpression leads to the reduced transcription of at least a subgroup of PAF1c target genes.

Stabilization of the SETDB1-PAF1c interaction increases promoter H3K9me3

We next asked how the SETDB1-PAF1c interaction alters the epigenetic regulation of PAF1c target genes. We asked whether CDC73_3YF or overexpression of SETDB1 affected H3K9 methylation at the *Meis1* and *Hoxa9* gene promoters. To explore this we employed Chromatin Immunoprecipitation (ChIP) assays followed by qPCR on MLL-AF9 *Cdc73*-/- cells re-expressing CDC73 or CDC73_3YF. Cells expressing a control empty vector or re-expressing CDC73 did not display a difference in H3K9me3 at the *Meis1* promoter. However, cells re-expressing CDC73_3YF exhibited a significant increase in H3K9me3 at this locus (Fig. 3-7A). Similarly, MLL-AF9 cells overexpressing SETDB1 showed an increase in H3K9me3 at the *Meis1* promoter compared to MLL-AF9 control cells (Fig. 3-7C). We also performed ChIP-qPCR at the *Hoxa9* promoter. Here, loss of *Cdc73* in cells expressing the control vector exhibited little change in H3K9me3, whereas expression of CDC73_3YF again resulted in increased H3K9me3 (Fig. 3-7B). We also observed that AML cells overexpressing SETDB1 had increased H3K9me3 at the *Hoxa9* promoter (Fig. 3-7C). To determine whether the increase in H3K9me3 requires SETDB1 catalytic activity, we performed

CHIP-qPCR in cells overexpressing SETDB1_CD. Surprisingly, these cells exhibited a significant loss of H3K9me3 at the *Meis1* and *Hoxa9* promoter regions (Fig. 3-7D). This suggests that SETDB1 is responsible for the increase in H3K9me3 at these promoters in cells overexpressing SETDB1 and that SETDB1 may be responsible for depositing a basal level of H3K9me3 at the *Meis1* and *Hoxa9* promoter genes in resting state MLL-AF9 AML cells. We also tested for global changes in H3K9me3 using acid extracted histones. Western blots demonstrate that AML cells overexpressing SETDB1 have a marked increase in global H3K9me3 (Fig. 3-7E), however, neither loss of *Cdc73* nor re-expression of CDC73_3YF affected the global levels of H3K9me3 (Fig. 3-7F). These data may point to a more locus specific phenotype in CDC73_3YF cells compared to SETDB1 overexpressing AML cells. Together, this suggests the PAF1c may be

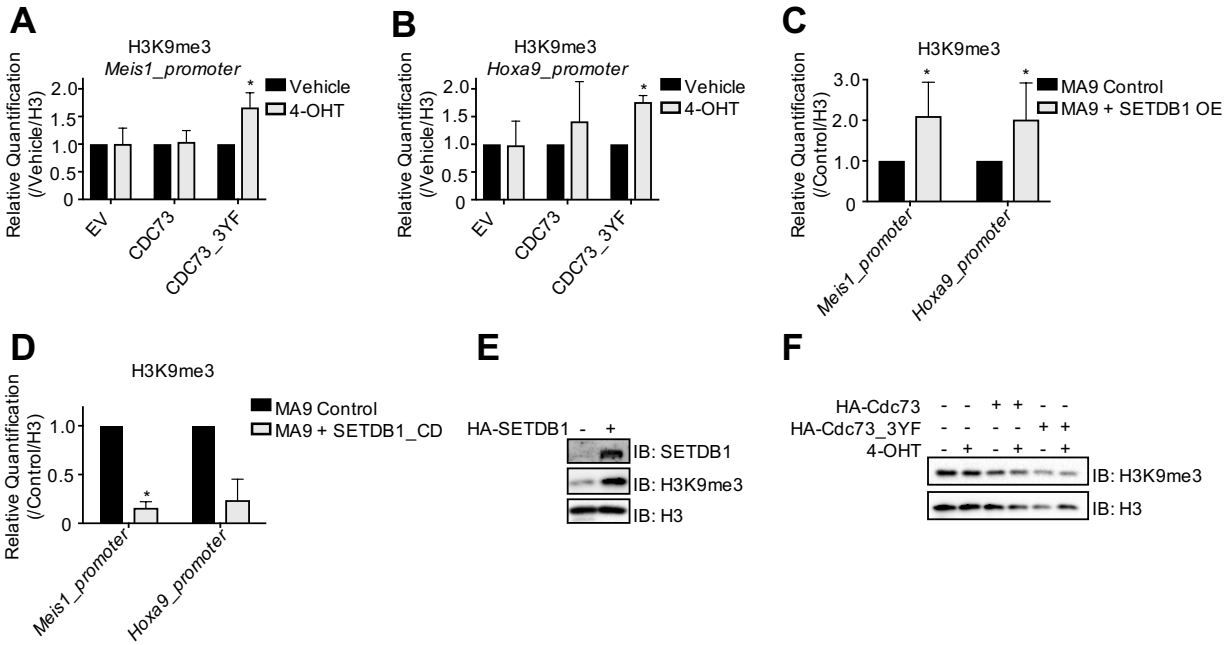


Figure 3-7: SETDB1 or CDC73_3YF contribute to epigenetic remodeling at the Meis1 and Hoxa9 promoter. A-B) ChIP-qPCR experiments for H3K9me3 at the (A) *Meis1* and (B) *Hoxa9* promoter were performed in MLL-AF9 transformed *CDC73fl/fl*-CreER^{T2} CDC73 re-expression cells treated with 4-OHT or vehicle control (biological replicates n=3). **C-D)** ChIP-qPCR experiments for H3K9me3 at the *Meis1* and *Hoxa9* promoters were performed in MLL-AF9 control cells with and without transduced overexpressed (C) SETDB1 or (D) SETDB1_CD (biological replicates n=2-3). **E)** Western blot of acid extracted histones from MLL-AF9 control cells or MLL-AF9 cells co-transduced with SETDB1 overexpression vector. Western blots were blotted with the indicated antibody. The top panel is a whole cell lysate demonstrating overexpression of SETDB1. **F)** Western blots of acid extracted histones from MLL-AF9 transformed *CDC73fl/fl*-CreER^{T2} CDC73 re-expression cells. Western blots were blotted with the indicated antibody. $^{*}p<0.05$; EV = Empty Vector control (MigR1); OE = Overexpression.

regulated by interactions with SETDB1 that modulate H3K9me3 at target genes critical for leukemogenesis (Fig. 3-8).

3.4 Discussion/Conclusions

Here we describe a proteomics approach to identify interactions with the PAF1c subunit CDC73 that may play a critical role in the regulation of PAF1c functions. We found that a CDC73_3YF mutant, which displays enhanced interaction with β-catenin, surprisingly does not support AML cell growth or colony formation capacity, in contrast to its function in gastric carcinoma [208]. Using proteomic and bioinformatics analyses, we found that CDC73 interacts with G9a and SETDB1; an interaction that is stabilized by the CDC73_3YF mutant. These interactions with H3K9 methyltransferases elucidate a mechanism of PAF1c functional regulation that may explain recent reports of transcriptional repression associated with subunits of the PAF1c [171,172]. Indeed, we demonstrate that interaction with SETDB1 modifies PAF1c mediated epigenetic and transcriptional regulation of leukemic target genes *Hoxa9* and *Meis1* in a manner dependent on SETDB1 catalytic activity (3-8).

To determine the role of SETDB1 and the effects of enhancing the interaction between CDC73 and SETDB1 on modulating transcription of PAF1c target genes, we utilized two systems. First, we genetically deleted *Cdc73* and re-expressed wild type CDC73 or CDC73_3YF in MLL-AF9 AML cells. In the second, we stably overexpressed

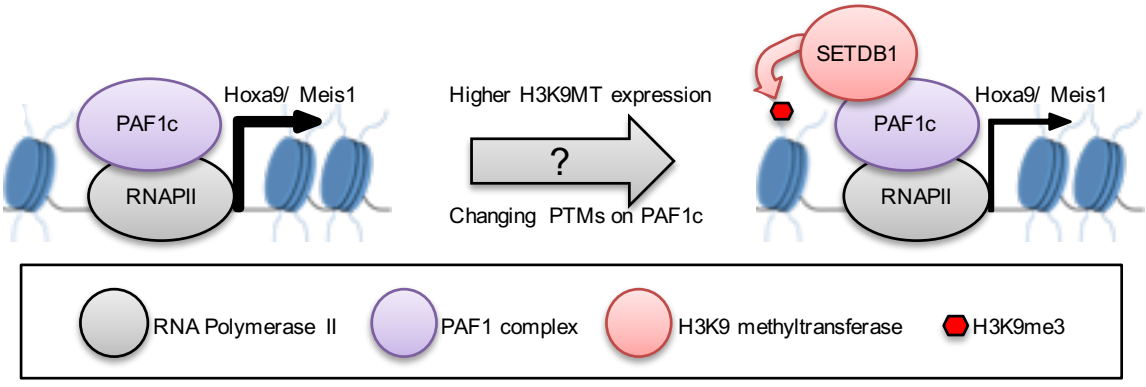


Figure 3-8: Working model for the role of the PAF1c-SETDB1 interaction in transcriptional regulation of Hoxa9/Meis1. Increased expression of H3K9 methyltransferases, such as SETDB1, or stabilized interaction between H3K9 methyltransferases and the PAF1c, possibly mediated by post-translational modifications (PTMs), leads to promoter H3K9 methylation and reduced expression of *Hox* genes.

SETDB1 in mouse MLL-AF9 AML cells. We note that both CDC73_3YF re-expression and overexpression of SETDB1 leads to the reduced transcription of *Hoxa9* and *Meis1*, while the catalytically inactive SETDB1_CD does not affect transcription. This data is remarkably consistent with RNA-seq data mined from the TCGA, which shows that higher expression of *SETDB1* correlates to significantly lower expression of *MEIS1* and a trend downward in the expression of *HOXA9* [1]. Importantly, we also observe that promoter H3K9me3 is increased at the *Meis1* and *Hoxa9* loci in cells re-expressing CDC73_3YF or overexpressing SETDB1, while the SETDB1_CD overexpression leads to a reduction in H3K9me3 (Fig. 3-7D). These data illustrate a novel interaction with SETDB1 that functionally modulates PAF1c activity through the SETDB1-mediated deposition of H3K9me3 and transcriptional repression.

Previous reports have demonstrated that phosphorylation of CDC73 at key tyrosine residues can act as a molecular switch to regulate interactions with β-catenin. In solid tissue tumors de-phosphorylation (mimicked using CDC73_3YF) enhanced the oncogenic activation of WNT targets by CDC73 [208,209]. This prompted us to explore phosphorylation of CDC73 in AML cells. Interestingly, we were unable to observe phosphorylation on CDC73 in AML cells by IP western blots or mass spectrometry searches including phosphorylated peptides (data not shown). Therefore, it appears that phosphorylation of CDC73 may be dependent on cellular context. This raises interesting questions regarding the stabilization of the SETDB1 and G9a interaction with CDC73_3YF. Further studies will be necessary to determine what, if any, post-translational modifications are affected by mutating these three tyrosine residues in AML cells and, mechanistically, how this affects interactions with G9a and SETDB1 (Fig. 3-8).

CDC73 has previously been reported to interact with SUV39H1 and G9a proteins, though the authors were unable to validate the interaction between endogenous G9a and CDC73 [201]. Here we confirmed the interaction with G9a and uncovered a novel interaction with SETDB1 while also confirming a transcriptional consequence of these interactions at leukemogenic target loci. Additionally, we found that the interactions with SETDB1 and G9a occur in the context of the PAF1c and are not specific to the CDC73 subunit. Our AP-MS data also provides evidence that CDC73

may interact with GLP, though further studies are needed to validate this interaction. Taken together, these data provide varying degrees of evidence that the PAF1c binds to four different H3K9 methyltransferases. Thus, it is possible that the PAF1c binds to a previously reported multimeric H3K9 methyltransferase complex consisting of G9a, GLP, Suv39H1 and SETDB1 [234]. Further investigation is necessary to determine whether the PAF1c interacts with this H3K9 methyltransferase complex or individual H3K9 methyltransferases at specific loci. Given that CDC73 differs in its post translational modifications in a cell context manner, it will also be necessary to evaluate the nature of the PAF1c interaction with H3K9 methyltransferases in different cell types (Fig. 3-8). We must further investigate how this interaction affects different cellular phenotypes such as cell cycle and differentiation. Indeed, global increases in heterochromatin formation and H3K9 methylation are observed following differentiation of both embryonic stem cells and hematopoietic stem and progenitor cells [276,365–367]. The current study examined the role of SETBD1 in modulating the transcription of PAF1c target genes *Hoxa9* and *Meis1*. Due to the important role of Hoxa9 and Meis1 in development, normal hematopoiesis, and hematologic malignancies like AML, the interaction between SETDB1 and the PAF1c has major implications on our understanding of the regulation of these genes and the potential to therapeutically target these epigenetically regulated pathways.

Chapter 4 SETDB1 mediated H3K9 methylation suppresses MLL-fusion target expression and leukemic transformation

4.1 Introduction

Epigenetic deregulation has emerged as an important contributor to oncogenesis and disease progression in a variety of malignancies, including leukemia[69,133–135]. Deep sequencing has revealed that genes encoding epigenetic modifying proteins are mutated in over 70% of AML patients[1]. H3K9me1/2/3 marks large regions of condensed transcriptionally inactive chromatin, such as pericentric heterochromatin. H3K9me2/3 also plays a functional role in the dynamic repression of genes in euchromatic regions of the genome[75,102,230,233–236]. Two families of proteins are associated with deposition of H3K9 methylation: the SUV39 family and the PRDM family. The SUV39 family of H3K9 methyltransferases consists of SUV39H1/2[237–239], EHMT1/2[240–243], and SETDB1/2[244,245]. Our lab and others have previously demonstrated that members of the SUV39 family of H3K9 methyltransferases bind to the Polymerase Associated Factor complex (PAF1c)[201,368]. SETDB1, G9a (*EHMT2*), and GLP (*EHMT1*), were identified in a proteomics study exploring the interactome of the PAF1c in AML[368,201,368]. The PAF1c is an epigenetic regulator complex that physically associates with RNA polymerase II (RNAPII) and both positively and negatively regulates gene transcription[15,16,99,170–173,177,185,188,189,191]. In AML, the PAF1c is critical for the regulation of a pro-leukemic *HOXA* gene program in AML cells through the recruitment of MLL and MLL-fusion proteins to the *Hoxa* locus via direct physical interactions [12–14,224]. *HoxA9* and its co-factor *Meis1* are upregulated in about 50% of AML and are associated with a poor patient prognosis[2]. Given our recent data linking H3K9 methyltransferases with *Hoxa9* and *Meis1* repression along with altered H3K9me3 in AML patients compared to CD34+ cells[369], it is important to understand the epigenetic and biological impact of H3K9 methyltransferases on AML.

SETDB1 is a H3K9 mono/di/tri-methyltransferase involved in heterochromatin regulation[236] and euchromatic gene silencing[364]. SETDB1 has been shown to bind to gene loci associated with development in mouse embryonic stem (ES) cells, such as the *Hoxd* cluster of genes[229]. SETDB1 has been implicated as an oncogene in melanoma, breast cancer, liver cancer, and lung cancer[249–253]. Importantly, Ceol, et al. reported amplification of SETDB1 in melanoma results in aberrant binding and regulation of the *HOXA* locus[251]. In contrast to the oncogenic roles for SETDB1 described above, Avgustinova and colleagues report that depletion of the H3K9 methyltransferase G9a in squamous tumors leads to a delayed, but more aggressive phenotype due to expanded cancer progenitor pools with increased genomic instability[320]. In the hematopoietic system, the methyltransferase activity of G9a is required for leukemogenesis due to a physical interaction with Hoxa9. Importantly, loss of G9a has no effect on hematopoietic stem cells[264,269]. *Setdb1*, on the other hand, is required for both HSPC maintenance and leukemic stem cells[264]. Further, Cuellar and colleagues show that SETDB1 mediated silencing of endogenous retroviral elements is required for the growth of AML cell lines[267]. Together, these studies suggest that therapeutic targeting of SETDB1 may benefit AML patients. However, we recently demonstrated that SETDB1 negatively regulates the expression of the pro-leukemic *HoxA9* and *Meis1* genes in MLL-AF9 transformed AML cells through association with the PAF1c, which localizes to *HoxA* and *Meis1* loci. The PAF1c-SETDB1 interaction mediates promoter H3K9me3 and repression of *HoxA9* and *Meis1* expression[368]. Further, *SETDB1* expression is inversely correlated with *HOXA9* and *MEIS1* expression in AML patient samples[368]. These data imply a more complex role for H3K9 methylation in AML similar to skin tumors whereby H3K9 methyltransferases display both oncogenic and suppressive roles[251,320]. Thus, further investigation into the role of H3K9 methyltransferases in AML is required.

Here we show that AML patients with higher expression of *SETDB1* display a better prognosis, consistent with repression of *HOXA9* and *MEIS1* expression. SETDB1 overexpression induces cellular differentiation and delays disease onset in a mouse model of AML, recapitulating AML patient survival. We also investigated the effect of inhibiting H3K9 methyltransferases in AML cells and HSPCs, demonstrating that

inhibition of H3K9 methylation in HSPCs leads to retention of self-renewal capacity in HSPCs and more efficient transformation by the MLL-AF9 fusion protein. Finally, we show that SETDB1 regulates gene expression by inducing changes in the epigenetic landscape and chromatin accessibility at gene targets critical to leukemogenesis.

4.2 Methods

Patient sample data

Data for gene expression in patient samples relative to normal hematopoietic cells were mined from the the BloodPool group of datasets on BloodSpot database[370], where AML patient samples' transcriptomic profiles are assigned to a closest normal hematopoietic counterpart and are then compared to obtain a fold-change relative to normal hematopoietic cells[371]. AML patient gene expression and survival data were mined from The Cancer Genome Atlas (TCGA)[1] using cBioPortal[372,373]. Survival data was analyzed using survival and survminer packages in R 3.5.1.

Cell line generation and proliferation assays

Cell lines were generated from wild type C57Bl/6 (Taconic Farms) mouse bone marrow or from SETDB1fl/fl[374] mouse bone marrow as described in Chapter 2. Cells were spinfected with the indicated combination of MSCVneo-FLAG-MLL-AF9 (MA9), MSCVneo-FLAG-E2A-HLF, MSCVhygro-FLAG-EHMT2 (Kai Ge lab; Addgene plasmid #41721)[375], MSCVpuro-HA-SETDB1 or empty vector (EV) controls. Cells were selected with 1mg/mL G418 (Invitrogen) and 1ug/mL puromycin (Invitrogen) or 200ug/mL hygromycin (Invitrogen). All cells were cultured in IMDM supplemented with 15% Stem Cell FBS (Millipore), 1%pen/strep (Invitrogen), 10ng/mL IL-3 and 100ng/mL SCF (R&D). Statistical analysis was generalized linear modeling followed by ANOVA.

Leukemia colony formation assay

Cells were retrovirally transduced as described above. They were then seeded at a density of $1x10^3$ cells in 2mL semi-solid methylcellulose medium for mouse cells (STEMCELL M3234) containing selection antibiotics (1mg/mL G418 and 1ug/mL puromycin or 100ug/mL hygromycin), 10 ng/mL IL-3, GM-CSF, and IL-6 and 100ng/uL

SCF (R&D). Colonies were counted, pooled, and replated in the same way for two additional rounds of colony formation and counting. 1x images of the 5-phenyl tetrazolium chloride (INT) stained dishes were taken after round 2. Statistical analysis was generalized linear modeling followed by ANOVA.

HSPC colony formation

Mouse bone marrow was harvested and lineage depletion was performed using the Hematopoeitic Progenitor Isolation Kit (STEMCELL) or CD34+ cells were isolate from mobilized blood. One group of mouse cells were transduced with shRNA against SETDB1 as previously described. These cells were pretreated for 4 days with puromycin before being plated in colony formation assays. The other group of mouse cells and the human cells were treated with the indicated dose of UNC0638 or DMSO vehicle control in liquid culture containing cytokines: 100ng/mL SCF and 10ng/mL IL-3. Cells were incubated for two days, retreated and given fresh media and cytokines, and incubated for another two days. For normal HSPC colony formation assays, 10,000 cells were plated in semi-solid methylcellulose in the presence of cytokines: 100ng/mL SCF and 10ng/mL IL-3, IL-6, and GM-CSF. Colonies were allowed to form for 10 days (mouse) or 14 days (human) and were then counted. Cells were pooled from colonies and 20,000 cells were replated in methylcellulose. Colonies were counted again after 10 days (mouse) or 14 days (human). Each mouse harvest was considered a biological replicate: n=3 for shRNA, n=4 for UNC0638. Each patient sample was considered a biological replicate for CD34+ cell assays: n=2. Statistical analysis was generalized linear modeling followed by ANOVA.

In vivo mouse modelling

Primary MLL-AF9 mouse leukemia cells were retrovirally transduced with MSCVpuro-HA-SETDB1 or EV. Cells were selected in 2ug/mL puromycin for 4 days and 100,000 cells were injected via tail vein into sublethally irradiated (650 rads) C57Bl/6 mice. Mice were monitored for survival and moribund mice were euthanized and bone marrow, spleen, and liver were harvested. AML was confirmed by splenomegaly and histology.

Quantitative PCR (qPCR) for gene expression

RNA was harvested from the indicated cells using the Qiagen RNeasy mini plus kit. cDNA synthesis was performed using random hexamer priming and the SuperScript III kit (Invitrogen). qPCR was performed using the fast SYBR-green mastermix protocol (Thermo Fisher). Statistical analysis was two sample t-tests.

RNA-seq

RNA was harvested from MLL-AF9+EV control cells or MLL-AF9+SETDB1 cells (n=3 each) using the Qiagen RNeasy mini plus kit. Quality control, library preparation using TruSeq stranded mRNA library prep kit (Illumina), and single-end 50bp sequencing on the HiSeq4000 (Illumina) were performed by the University of Michigan sequencing core. Separately, raw count files for THP-1 cells subjected to Crispr-Cas9 KD of SETDB1 were downloaded from GEO (GSE103409)[267,322]. RNA-seq analysis was performed as described in Chapter 2 and Table 2, using GRCm38 genome assembly for mouse samples and GRCh38 for human samples.

ChIP-seq

ChIP experiments were performed as previously described[357] with slight modifications. Briefly, $3x10^7$ MA9 + EV control cells or MA9 + SETDB1 overexpression cells (n=2 each) were crosslinked with 1% formaldehyde, quenched with 125mM Glycine, lysed with 1% SDS and sonicated on a Bioruptor Pico sonicator (Diagenode). Cleared lysates were immunoprecipitated with 4ug antibodies using Protein G dynabeads (Invitrogen). All antibodies were validated for specificity using Histone peptide arrays from EpiCypher (Antibodies used: anti-H3K9me3: Active Motif; anti-H3K9ac: EpiCypher; anti-H3K79me2: Abcam). The IPs were washed with a low salt buffer, a high salt buffer, and a lithium chloride wash buffer. Protein-DNA complexes were eluted in 1% SDS, were decrosslinked in high salt and treated with RNaseA and ProteinaseK (Invitrogen). DNA was purified with a Qiagen PCR purification kit. Libraries were prepared using the iDeal Library Preparation kit (Diagenode) according manufactuerer's recommendations. Library amplification was optimized by monitoring

amplification cycles using SYBR I fluorophore and qPCR. Paired-end 38bp sequencing was performed on the NextSeq500 (Illumina) by the University of Michigan Sequencing Core.

ChIP-seq analysis was performed up to alignment as described in Chapter 2 and Table 2, using GRCm38 genome. For the following analyses, any parameters used that are not defaults are also listed in Table 2. After reads were aligned, peaks were called using MACS2. Peak analysis in two separate ways. First, consensus peaks were determined for MA9+EV or MA9+SETDB1. Only peaks that were called in both replicates for a given sample were kept for this analysis. Next, overlap analysis was performed using ChIPpeakAnno to determine peaks that are found in both samples, or in only one sample. Finally, these peaks were annotated to promoter regions (-5000bp, +2000bp from transcription start site) for H3K9ac ChIP or for any peak overlapping a gene locus for H3K9me3 or H3K79me2 ChIP. The second analysis used DiffBind to analyze differential signals for the different experiments. DiffBind keeps any peaks that are found in at least two samples in the entire experiment as consensus peaks, and then re-counts the reads for all samples at those peak regions. After obtaining differential binding results, the peaks were annotated as described above. Overlap analyses were performed with this differential binding analysis to determine peaks that were significantly changed in the different ChIP conditions, as well as RNA-seq and ATAC-seq. Signal tracks were generated using DeepTools and normalizing by reads per genomic content.

Assay for Transposase-Accessible Chromatin (ATAC-seq)

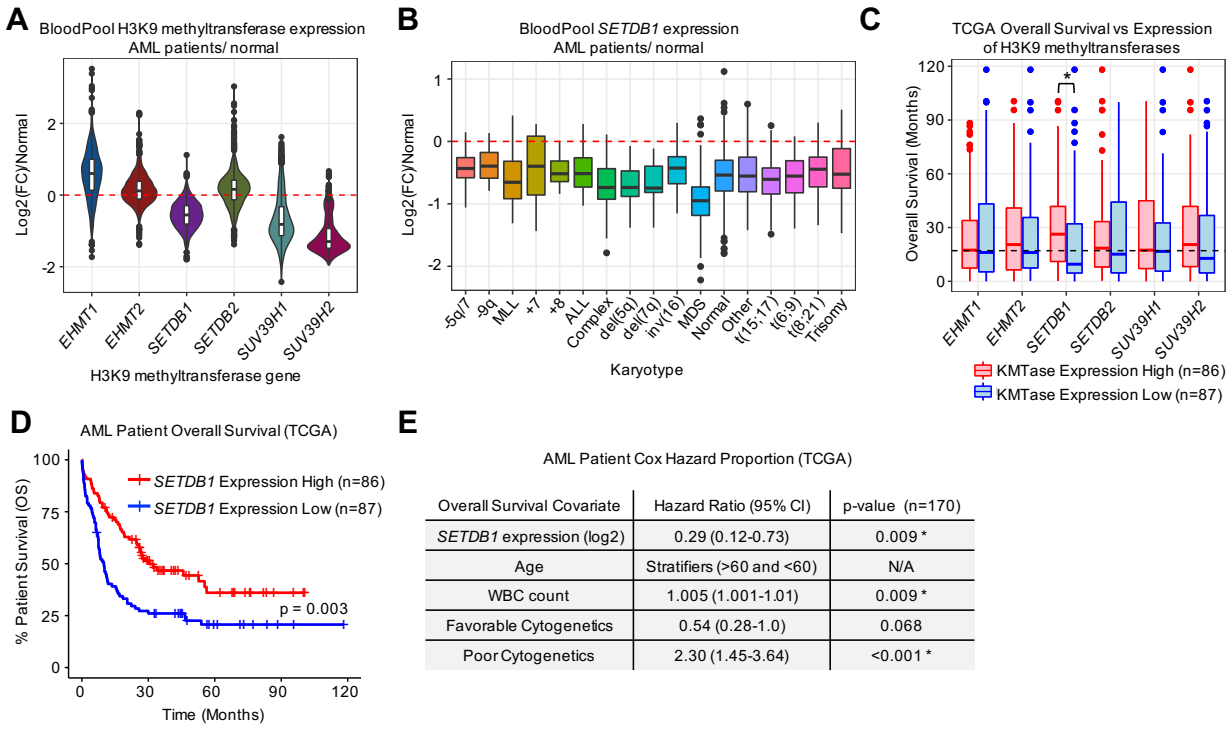
ATAC-seq was performed as described in Buenrostro, et al.[376] Briefly, MA9+EV or MA9+SETDB1 established cell lines (n=2 each) were treated with 200units/mL DNAse I. Cells were then slow frozen. Nuclei were isolated, treated with transposase, and purified by Active Motif. Active Motif prepared libraries and performed 42bp paired-end sequencing. Fastq files were obtained from Active Motif. Analysis was performed the same as ChIP-seq with two modifications. First, MACS2 parameters were amended to center reads on the site of transposition (Table 2). Second, signal tracks were

generated from MACS2 (using deepTools to convert pileup files to bigwig) in order to preserve this read shifting.

4.3 Results

SETDB1 expression is correlated with AML patient prognosis

Given our data linking H3K9 methyltransferases with *HOXA9* and *MEIS1* expression, we investigated their expression in AML patient sample data. Microarray data demonstrates that *SETDB1*, *SUV39H1*, and *SUV39H2* exhibit lower expression in AML patient samples when compared with normal hematopoietic cells, with median



Overall Survival Covariate	Hazard Ratio (95% CI)	p-value (n=170)
SETDB1 expression (log2)	0.29 (0.12-0.73)	0.009 *
Age	Stratifiers (>60 and <60)	N/A
WBC count	1.005 (1.001-1.01)	0.009 *
Favorable Cytogenetics	0.54 (0.28-1.0)	0.068
Poor Cytogenetics	2.30 (1.45-3.64)	<0.001 *

Figure 4-1: SETDB1 expression is low in AML and is correlated with AML patient prognosis. A) BloodPool data mined from Bloodspot shows AML patient gene expression relative to the nearest normal hematopoietic counterpart for the indicated genes encoding H3K9 methyltransferases. B) BloodPool data divided by karyotype for SETDB1 in AML samples occurs regardless of genetics driving the disease. C) TCGA patient sample data is divided by "high" expression (above median) and "low" expression (below median) for the indicated gene encoding an H3K9 methyltransferase. Boxplot shows overall survival for each stratified gene. D) Kaplan Meier curve showing overall survival of AML patients stratified by SETDB1 expression above ("high") or below ("low) median. E) Multivariate analysis using Cox Hazard Proportion analysis to assess the hazard ratio associated with changing levels of SETDB1 expression. SETDB1 expression is expressed in log2(RSEM), so the hazard ratio is associated with a 2-fold change in SETDB1 expression. Age did not satisfy the model and was stratified by patients < 60 years old and patients > 60 years old. Statistics: * p/padj < 0.05; one-sample

expression levels that are 66%, 56%, and 41% relative to their nearest normal hematopoietic counterparts, respectively (Figure 4-1A)[370,371]. The downregulation of these genes, specifically *SETDB1*, in AML samples was consistent regardless of AML karyotype (Figure 4-1B)[370,371]. We next tested whether H3K9 methyltransferase gene expression significantly correlated with patient survival. Dividing patient gene expression data based on median expression, we found that *SETDB1* is the only gene whose expression significantly correlated with patient survival (Figure 4-1C). Median

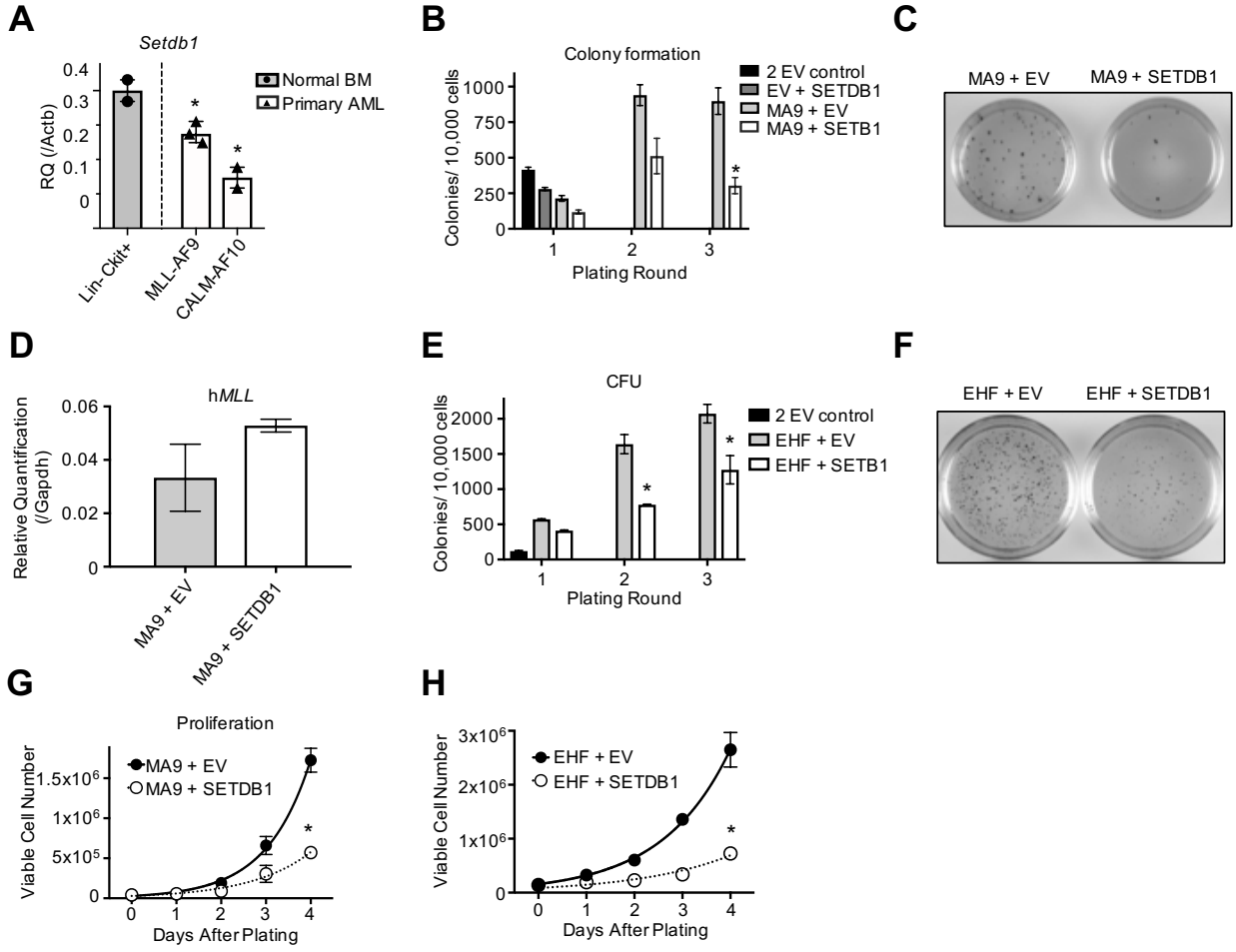


Figure 4-2: Overexpression of SETDB1 delays AML growth. A) qPCR using primers for m*Setdb1* to determine expression levels in primary MLL-AF9 (n=3) or CALM-AF10 (n=2) AML cells compared to lin-Ckit+ mouse bone marrow HSPCs (n=2 pools of 5 mice each). B/E) Mouse lin- bone marrow was retrovirally transduced with the indicated plasmid vectors and plated in methylcellulose. Colonies were counted after 7 days and re-plated, for a total of three rounds. Shown is one representative experiment of n=4 (B) or n=2 (E). C/F) Representative INT staining of colony assay plates for MA9 cells with or without SETDB1 overexpression. D) qPCR using primers for h*MLL* to determine expression levels of MLL-AF9 in MA9+EV or MA9+SETDB1 cells (n=2). G/H) Lin- bone marrow cells retrovirally transduced with MA9 in the presence or absence of SETDB1 overexpression, selected for 2 weeks, then proliferation was monitored by viable cell count daily. Shown is one representative experiment of n=4 (G) or n=2 (H). EV=empty vector control. * $p < 0.05$.

survival for patients with *SETDB1* expression above the median was 26.3 months and 9.5 months in patients with *SETDB1* expression below the median (Figure 4-1D)[1]. Univariate and multivariate analyses reveal that higher expression of SETDB1 is associated with a higher overall survival rate with a $p < 0.003$ and a lower expected hazard ratio of 0.29 per 2-fold change of SETDB1 expression (Figure 4-1D, 4-1E)[1].

SETDB1 or G9A expression reduces AML cell growth and colony forming capacity

We next explored the biological effects of SETDB1 expression on cellular transformation and the growth of AML cells. First, we explored whether there is a difference in *Setdb1* expression in mouse AML relative to normal HSPCs by isolating lineage negative (lin^-) $cKit^+$ cells from mouse bone marrow. qPCR demonstrates that *Setdb1* expression is significantly reduced in primary MLL-AF9 and CALM-AF10 AML cells compared to normal HSPCs, consistent with the patient sample data (Figure 4-2A). We next performed colony replating assays where lin^- mouse bone marrow cells were

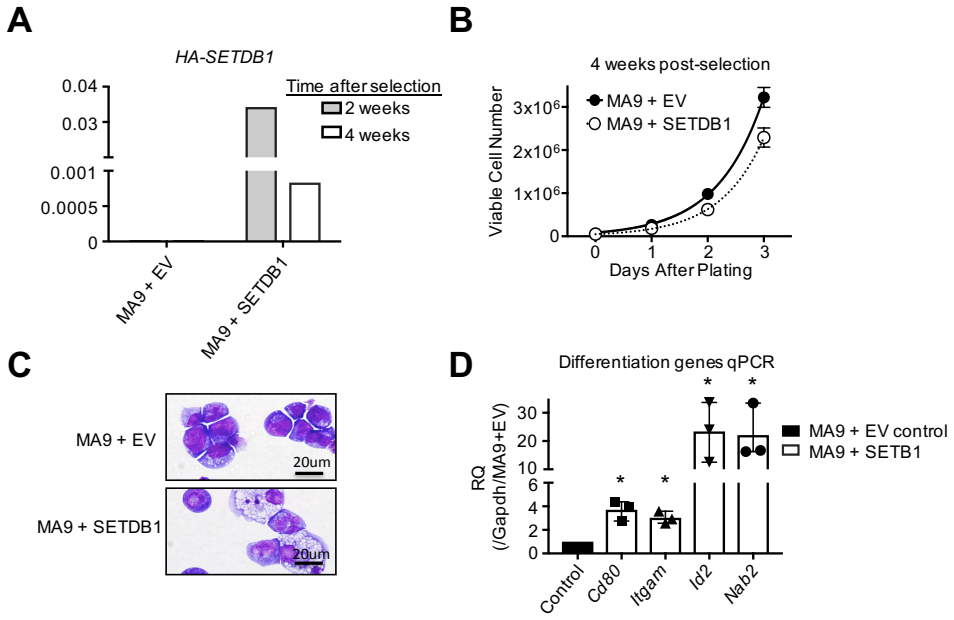


Figure 4-3: Overexpression of SETDB1 induces AML differentiation. A) qPCR using primers for HA-SETDB1to determine expression of exogenous SETDB1 in MA9 cells with or without SETDB1 overexpression 2 and 4 weeks after selection media is withdrawn. B) Proliferation assay of MA9 cells 4 weeks after selection media is withdrawn. C) Cytospin and Hema3 staining of MA9 cells with or without SETDB1 overexpression. D). qPCR to measure gene expression of genes associated with differentiation in MA9 cells in the presence or absence of SETDB1 overexpression (n=3). * $p < 0.05$.

retrovirally co-transduced with *MLL-AF9* fusion (MA9) oncogene with and without *SETDB1* and plated in semi-solid methylcellulose. Overexpression of *SETDB1* significantly reduced MLL-AF9 mediated colony formation (Figure 4-2B, C). Overexpression of *SETDB1* did not change expression of the *MLL-AF9* retrovirus as determined by qPCR (Figure 4-2D)[267]. Colony formation driven by a separate leukemic fusion protein, E2A-HLF (EHF), was also reduced in the presence of SETDB1, suggesting a general effect on AML transformation (Figure 4-2E, F). *Ex vivo* proliferation assays demonstrate that overexpression of *SETDB1* in either MLL-AF9 or E2A-HLF transformed AML cells leads to a significant reduction in cellular proliferation (Figure 4-2G,H). Interestingly, we observe a strong selective pressure to reduce exogenous *SETDB1* expression in cultured MLL-AF9+SETDB1 cells (Figure 4-3A). This results in a rescue of MLL-AF9 cellular proliferation, consistent with a selective pressure against high *SETDB1* expression (Figure 4-3B). Cytospins revealed that MLL-AF9 cells that overexpress *SETDB1* undergo morphological changes consistent with differentiation (Figure 4-3C). Genes associated with hematopoietic differentiation,

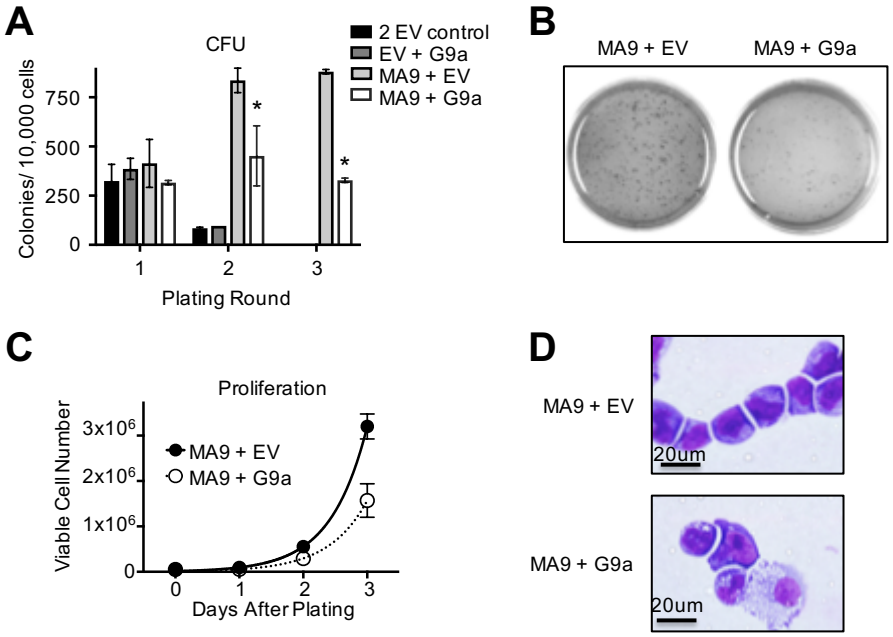


Figure 4-4: Overexpression of G9a delays AML growth. A) Mouse lin- bone marrow was retrovirally transduced with the indicated plasmid vectors and plated in methylcellulose. Colonies were counted after 7 days and re-plated, for a total of three rounds. Shown is one representative experiment of n=3. B) Representative INT staining of colony assays from A. C) Proliferation assay of MA9 cells with or without G9a overexpression. D) Cytospin and Hema3 staining of MA9 cells with or without G9a overexpression.

including Id2, Cd80, Nab2, and Itgam, have increased expression upon overexpression of SETDB1 (Figure 4-3D). We tested another H3K9 methyltransferase and found that G9a (EHMT2) overexpression also reduces colony formation and cellular proliferation of MLL-AF9 cells (Figure 4-4A-C) and induces morphological changes consistent with differentiation (Figure 4-4D). These data demonstrate that expression of H3K9 methyltransferases reduces AML cell proliferation and colony forming potential and induces hematopoietic differentiation.

SETDB1 expression delays MLL-AF9 mediated AML

To examine the effects of SETDB1 *in vivo*, we transplanted primary mouse MLL-AF9 AML cells retrovirally transduced with or without SETDB1 into sublethally irradiated syngeneic recipient mice and monitored survival. Consistent with AML patient data, overexpression of SETDB1 significantly delays MLL-AF9 mediated leukemogenesis *in*

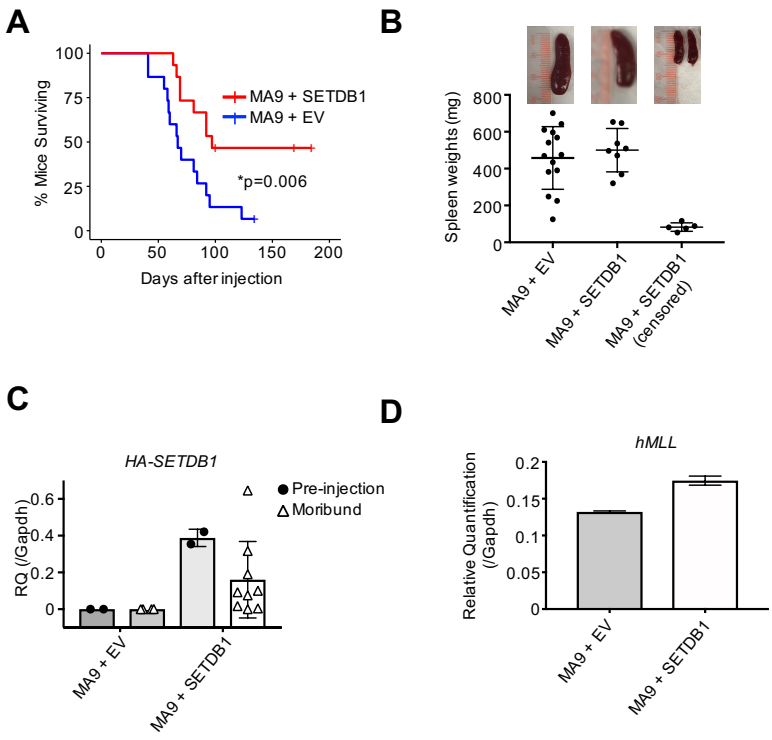


Figure 4-5: SETDB1 overexpression delays AML cell growth in vivo. A) Kaplan Meier survival curve of mice that were transplanted with primary MLL-AF9 cells that were transduced to overexpress SETDB1 or empty vector (EV) control. B) Spleen weights of moribund or censored mice. C) Expression of exogenous SETDB1 in moribund mouse bone marrow was determined using qPCR and primers specific for the HA-SETDB1 construct. D) MLL-AF9 expression was confirmed using primers for human MLL

vivo (Figure 4-5A). All moribund mice from both the control MLL-AF9 group and MLL-AF9+SETDB1 group exhibited splenomegaly (Figure 4-5B). We confirmed similar MLL-AF9 expression levels in MLL-AF9 and MLL-AF9+SETDB1 leukemic cells (Figure 4-5D). We measured expression of exogenous SETDB1 in MLL-AF9 leukemic cells before injection and after bone marrow harvest from moribund mice and observed a decrease in exogenous SETDB1 expression in 7/9 mice (Figure 4-5C). These data demonstrate that *SETDB1* expression suppresses MLL-AF9 mediated leukemic progression *in vivo*.

H3K9 methylation impairs HSPC colony formation

We next examined how inhibition of SETDB1 affects AML transformation and progression. In contrast with our observation that increased expression of SETDB1 impairs AML, Koide et al. reported that Setdb1 is required for hematopoietic stem cells and MLL-AF9 driven leukemogenesis[264]. Using inducible CreER-mediated knock out of *Setdb1*, we confirmed that Setdb1 is required for MLL-AF9 cell growth and demonstrated that Setdb1 is also required for the growth of E2A-HLF transformed leukemic cells (Figure 4-6A, B). Taken together, these observations suggest that there may be a window of *SETDB1* expression that allows for AML initiation or progression. Previous studies that have shown that *Setdb1* is required for AML growth have proposed chemical inhibition of H3K9 methyltransferases for the treatment of AML. However, our data suggests that inhibition might have the unintended effects of de-repressing genes associated with driving self-renewal and AML initiation and progression, such as *HOXA9*. To test this hypothesis, we sought to treat both AML cells and normal HSPCs with a small molecule inhibitor targeting H3K9 methyltransferases. Without a selective SETDB1 small molecule inhibitor, we explored the utility of inhibiting H3K9 methylation in AML using UNC0638, a selective inhibitor for G9a. Treatment of AML cells with UNC0638 leads to a reduction in both H3K9me2 and H3K9me3 (Figure 4-6C, D)[365,377]. Consistent with previous studies, treatment with UNC0638 results in

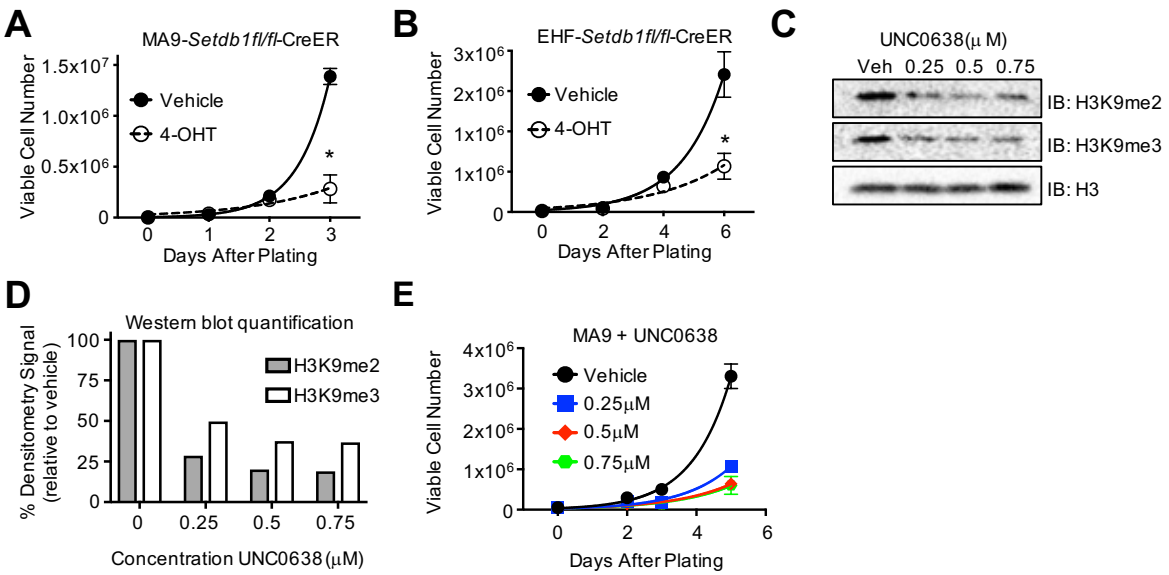


Figure 4-6: Validation of UNC0638 as a tool to reduce H3K9me2/3 levels in cells. A/B) Proliferation assays of MA9 or EHF cells following treatment with 5nM 4-OHT to induce genetic excision of Setdb1. 1 representative experiment of n=2. C) Western blot of whole cell extracts from MA9 cells treated with the indicated doses of UNC0638 for 5 days. IB=immunoblot with the indicated antibody. D) Quantification of C. Densitometry signals were normalized to H3 and vehicle (Veh/0uM) control. E) Proliferation assay monitoring MA9 cell growth after treatment with UNC0638. * $p < 0.05$.

reduced cellular proliferation of MLL-AF9 cells (Figure 4-6E)[269]. Given our data demonstrating self-renewal genes like *Hoxa9* and *Meis1* are repressed by H3K9 methylation[368], we tested the effect of UNC0638 on normal HSPCs self-renewal capacity. Previous reports have demonstrated that mouse Lin-Ska1+Ckit+ cells (LSKs) are preserved in culture following treatment with UNC0638[365]. To explore this further, Lin- cells were treated twice in liquid culture with increasing doses of UNC0638 for five days prior to plating in semi-solid methylcellulose in the presence of SCF and IL3. Interestingly, treatment with UNC0638 significantly increased colony formation capacity of lin- cells in a dose-dependent manner (Figure 4-7A, B). Further, human CD34+ cells isolated from mobilized peripheral blood and treated with increasing doses of UNC0638 demonstrate increased colony formation re-plating capacity (Figure 4-7C). This is consistent with reports that CD34+ cells that are treated with UNC0638 retain a higher percentage of CD34+ cells[366]. To address the possibility of off target effects, we performed shRNA mediated knock down experiments targeting Setdb1. Our data demonstrate a trend in higher capacity for colony formation following knock down of *Setdb1* relative to non-silencing controls (Figure 4-7D). Thus, chemical or genetic

inhibition of H3K9 methyltransferases preserves self-renewal capacity of bone marrow cells. Further, the level of H3K9 methyltransferase activity appears to be essential to its

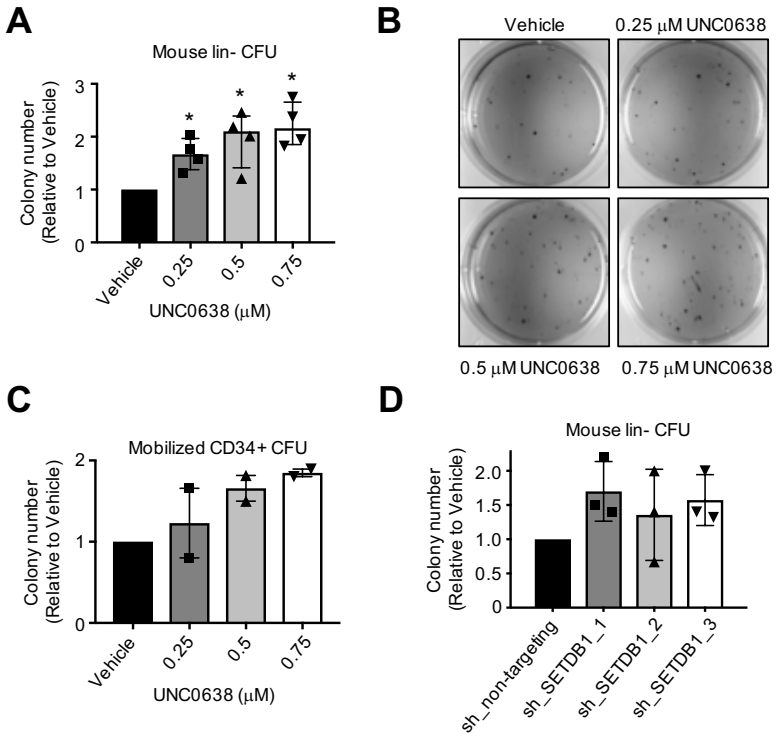


Figure 4-7: H3K9 methyltransferase inhibitor UNC0638 enhances HSPC colony formation capacity. A) Lin- mouse bone marrow was isolated and treated in culture for 4 days with the indicated dose of UNC0638. Cells were plated in methylcellulose and colonies were counted after 7 days, n=4. B) INT stained representative colonies from A. C) Isolated human CD34+ cells were treated for 4 days with the indicated doses of UNC0638 and plated in methylcellulose. Cololnies were counted after 14 days, n=2. D) Lin- mouse bone marrow was isolated and spinfected with shRNA targeting Setdb1. Cells were pre-selected with puromycin for 7 days, then plated in methylcellulose. Colonies were counted after 7 days. A/C/D: Colony numbers are shown relative to the vehicle/ non-silencing control for each replicate.

biological function in AML and normal HSPCs. Complete deletion of *Setdb1*/*Ehmt2* leads to AML cell death, whereas partial inhibition of H3K9 methyltransferases can activate self-renewal genes. Additionally, overexpression of H3K9 methyltransferases leads to impaired AML cell growth by inducing differentiation.

H3K9 methylation suppresses leukemic transformation of bone marrow cells

Because inhibition of H3K9 methyltransferases leads to an increased self-renewal capacity of HSPCs, we next asked whether UNC0638 alters MLL-AF9 mediated bone marrow transformation. First, we found that UNC0638 treatment of lin-

bone marrow cells preserves more primitive Ckit+ and Cd11b- populations (Figure 4-8A-D). Lin- cells also display increased *Hoxa9* expression in response to UNC0638

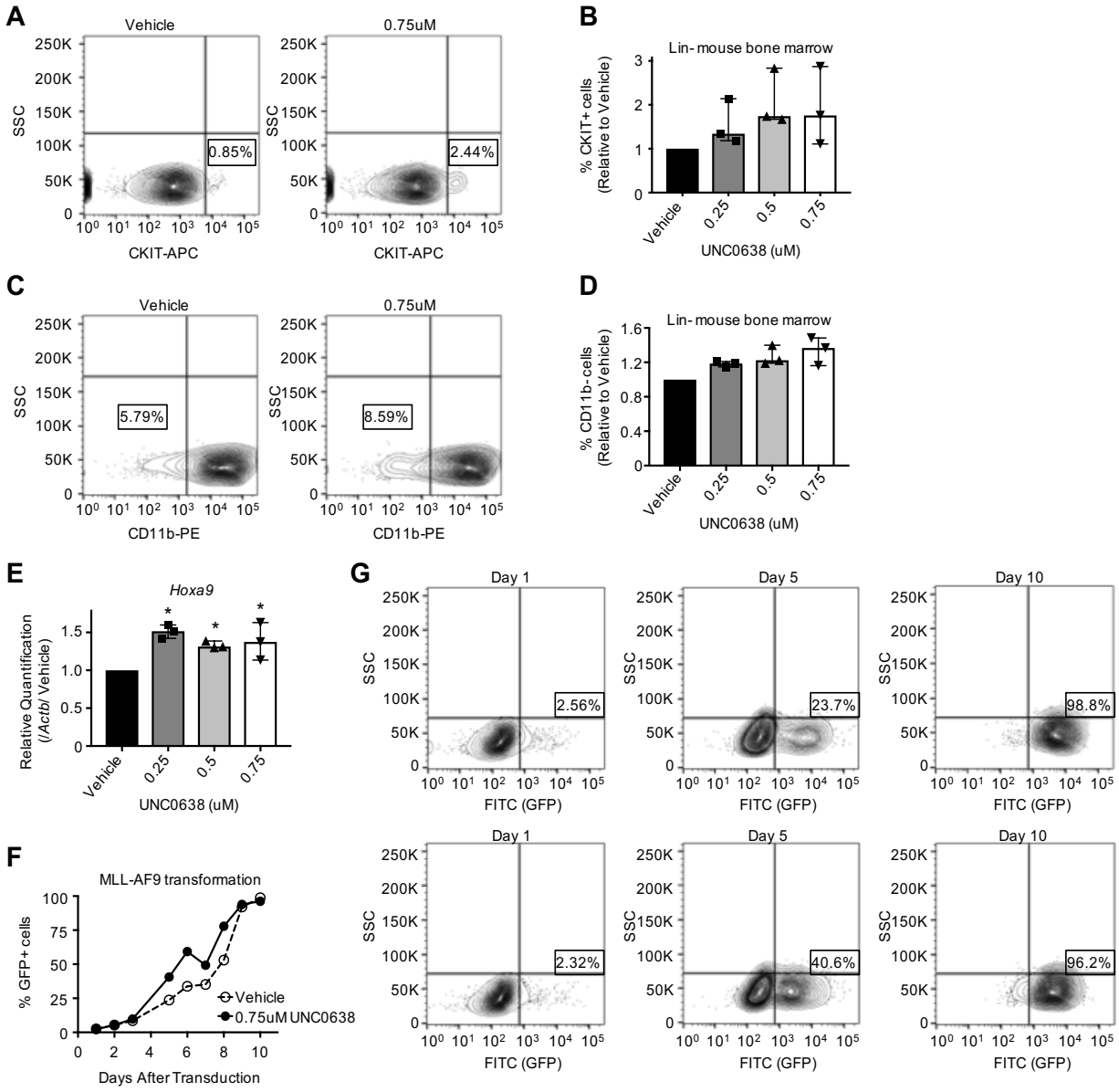


Figure 4-8: H3K9 methyltransferase inhibitor UNC0638 preserves pool of cells amenable to MLL-AF9 transformation. A-G) Lin- mouse bone marrow was treated with the indicated doses of UNC0638 for 4 days. A-D)Cells were stained with anti-Ckit conjugated to APC fluorophore (A/B) or anti-Cd11b conjugated to PE fluorophore (C/D). Flow cytometry was performed to analyze Ckit+ or Cd11b- populations. A/C are representative flow plots for 0.75μM UNC0638 treatments. B/D show the biological replicates for Ckit+ populations (B) or Cd11b- populations (D) of treated cells relative to vehicle (n=3). E) RNA was harvested after 4 days of treatment with UNC0638 and qPCR was used to determine changes in *Hoxa9* expression. F) After treatment with UNC0638, cells were spinfected with MigR1-MLL-AF9, which also expresses a GFP reporter. Cells were monitored for GFP expression by flow cytometry until 100% GFP was achieved. G) Representative flow plots from different time points during the GFP monitoring experiment. For F/G: shown is 1 representative experiment of n=4.

treatment (Figure 4-8E). Lin-kit+ and related primitive hematopoietic cell populations are more amenable to transformation than other more differentiated hematopoietic cellular subtypes[378,379]. We therefore hypothesized that UNC0638 treatment of bone marrow may result in greater transformation efficiency by the MLL-AF9 fusion protein. To explore this, we pretreated lin- bone marrow cells with UNC0638 for four days followed by retroviral transduction with MigR1-MLL-AF9 containing a GFP reporter, which we monitored by flow cytometry. We observed a more rapid expansion of GFP+ MLL-AF9 cells following treatment with UNC0638 compared to vehicle (Figure 4-8F). The percentage of GFP cells increased 1.4-fold to 1.7-fold for UNC0638 treated cells compared to vehicle before the population achieved 100% GFP+ MLL-AF9 cells (Figure 4-8F, G). One of four replicates demonstrated a more rapid expansion of vehicle treated cells, however, GFP+ cells were increased 2-fold in vehicle treated cells at day 1 suggesting this is attributable to differences in transduction rates. Our combined data suggests that several H3K9 methyltransferases can suppress leukemic transformation. Taken together with data from Koide, et al.[264] and Lehnertz, et al.[269] that shows that these H3K9 methyltransferase genes are required for AML inititaiton and progression, this points to a narrow window of H3K9 methylation that is optimal for leukemic transformation and aggressive cell growth.

SETDB1 regulates oncogenic gene programs in AML

To explore the gene programs regulated by SETDB1 in AML, we performed RNA-sequencing experiments on MLL-AF9 cells overexpressing SETDB1. These studies revealed changes in a subset of the transcriptome, with 2285 genes being upregulated and 1771 genes being downregulated by SETDB1 overexpression with a fold-change of 1.5 or more and an FDR of < 0.05 (Figure 4-9A). Notably, many genes found in the HoxA cluster were significantly downregulated, including *Hoxa3*, *Hoxa5*, *Hoxa6*, *Hoxa9*, and the *Hoxa9* cofactor *Meis1*; while only one of the Hoxa cluster of genes was upregulated, the long non-coding RNA Hoxa11os (Figure 4-9A, highlighted genes). In fact, gene programs that are upregulated by forced expression of Hoxa9 and Meis1 in mouse cells are shown to be significantly downregulated by SETDB1 overexpression using gene set analysis[323,380,381] (Figure 4-9B). We also saw a

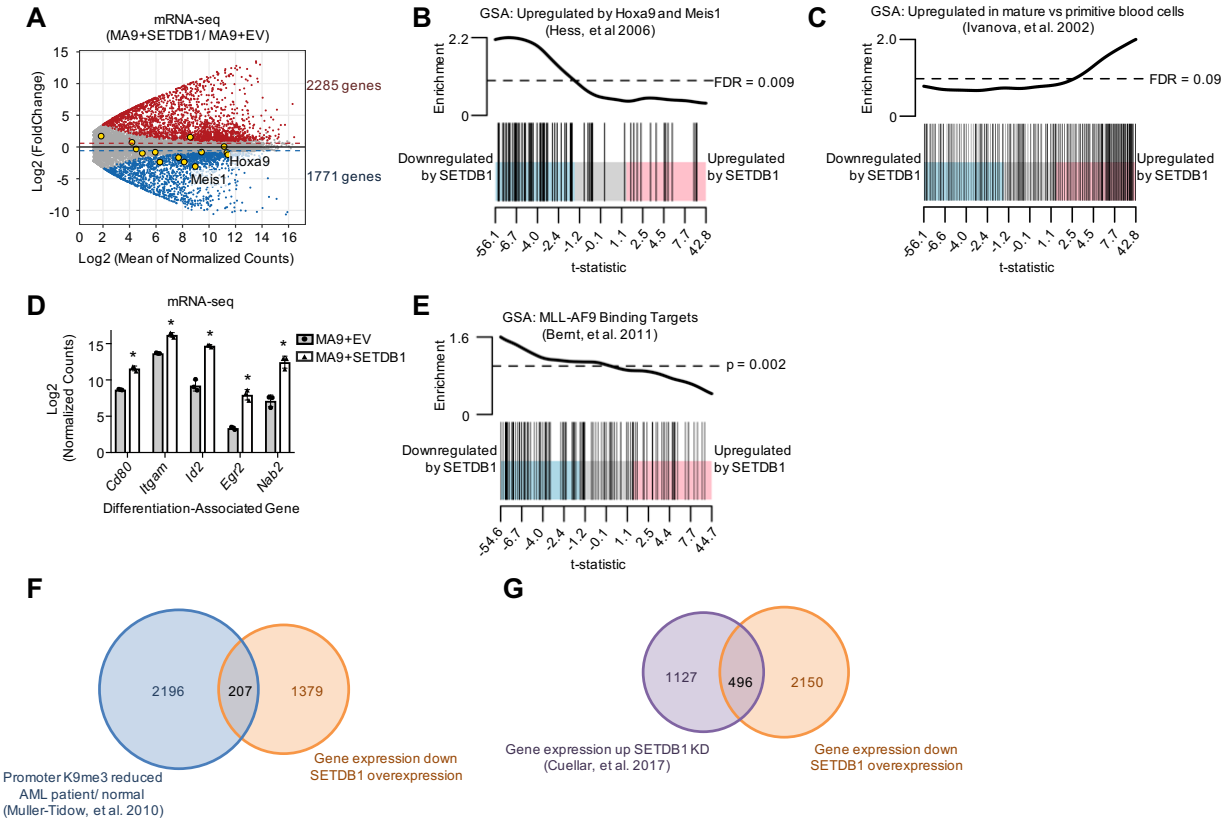


Figure 4-9: SETDB1 overexpression downregulates oncogenic gene programs and upregulates differentiation gene programs in AML. A) MA plot shows fold changes of genes in MA9+SETDB1/ MA9+EV versus the average expression of those genes in all samples. Highlighted genes are the *Hoxa* cluster of genes and the *Hoxa9* cofactor *Meis1*. B) GSA analysis using CAMERA shows that genes that are upregulated by Hoxa9 and Meis1 are downregulated by SETDB1. C) GSA analysis using CAMERA shows that genes that are upregulated in mature blood cells relative to primitive blood cells are upregulated by SETDB1. D) DESeq2-normalized RNA-seq counts of genes associated with differentiation. E) GSA analysis using ROAST shows that genes that are directly bound by MLL-AF9 are downregulated by SETDB1. F) Overlap analysis of genes that are downregulated by SETDB1 in mouse MA9 cells and genes that have reduced promoter H3K9me3 in AML patient samples compared to normal human CD34+ cells. G) Overlap analysis of genes that are downregulated by SETDB1 in mouse MA9 cells and genes that are upregulated upon SETDB1 knockdown by Crispr-Cas9 in human THP-1 cells. * padj < 0.05

significant upregulation of genes that exhibit increased expression in mature blood cells compared to HSCs and primitive progenitor cells, consistent with the cellular differentiation phenotypes observed upon SETDB1 overexpression[26,64,323,381,382] (Figure 4-9C, D). Because SETDB1 binds the PAF1c, which is required for localization of MLL fusion proteins[368], we asked how SETDB1 expression affects direct targets of MLL-AF9. Interestingly, we observed that genes directly bound by MLL-AF9 were significantly downregulated upon SETDB1 overexpression[8], suggesting H3K9me3 regulation of an MLL-AF9 gene program in leukemic cells (Figure 4-9E). 207 genes downregulated by

SETDB1 overexpression are reported to have reduced promoter H3K9 methylation in AML relative to normal CD34+ cells[369], suggesting SETDB1 may be responsible for regulating a subset of these genes, including *Kit*, *Cbl*, *Ptpn11*, *Six1,* and other genes that are important in AML (Figure 4-9F). There is also a significant overlap between data showing genes downregulated by SETDB1 and genes upregulated by Crispr/Cas9 mediated knockdown of SETDB1 in human THP-1 leukemic cells harboring an MLL-AF9 fusion protein (Figure 4-9G). This suggests SETDB1 regulates conserved pro-leukemic gene programs in leukemic cells including direct MLL fusion targets.

SETDB1 regulates the epigenome to affect changes in chromatin availability and gene expression

To understand the epigenomic changes induced by SETDB1 in AML cells, we performed ChIP-sequencing studies for H3K9me3 in MLL-AF9 and MLL-AF9+SETDB1 leukemic cells. We also performed ATAC-seq to assess changes in chromatin condensation in response to SETDB1. We first explored differences at the *HOXA* locus due to its importance in a large subset of AML, including MLL leukemias[4,368,5]. First, we see a consistent reduction in ATAC-seq signal across consensus peaks, suggesting a general decrease in chromatin availability at observed peaks. We next sought to characterize loci that are known to be regulated by SETDB1 and are important for AML growth. Overexpression of SETDB1 reduced chromatin accessibility and increased H3K9me3 at posterior *Hoxa* genes, which results in reduced transcription of *Hoxa9* (Figure 4-10A). We sought to define whole genome epigenetic regulation mediated by SETDB1 but observed only 552 consensus H3K9me3 peaks (Figure 4-10B), with the exception of regions enriched for H3K9me3, such as the zinc finger protein cluster on chromosome 7 (Figure 4-10C). It is possible sequencing of H3K9me3 is challenging because it is enriched at heterochromatic regions associated with sequence repeats, which consume sequencing depth and are difficult to map with stringency. We overcame this by performing ChIP-Seq for H3K9ac, which is mutually exclusive and acts antagonistically with H3K9me3. We performed H3K9ac ChIP-seq on MLL-AF9 and MLL-AF9+SETDB1 cells and saw changes at over 6,000 promoter regions, including both increased and decreased H3K9ac signal (Figure 4-10D). Gene ontology analysis

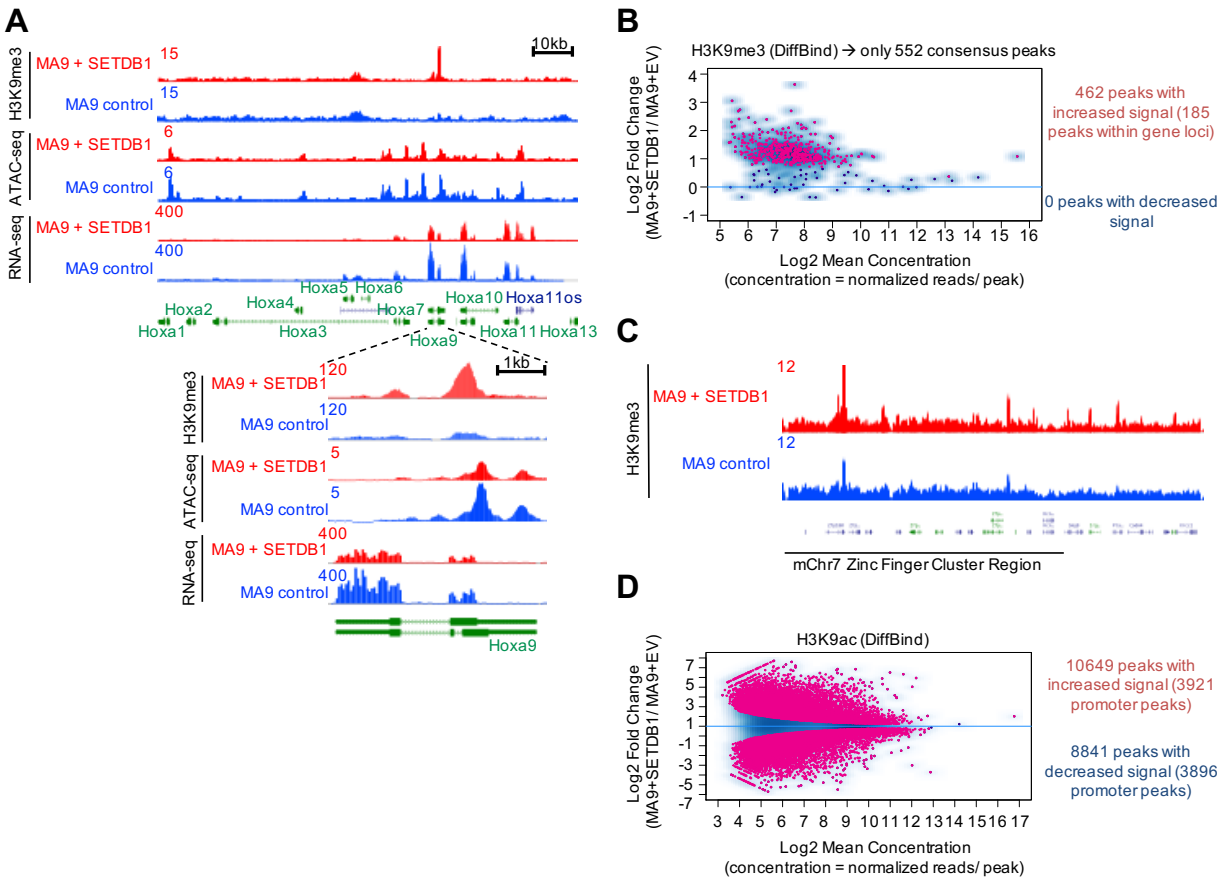
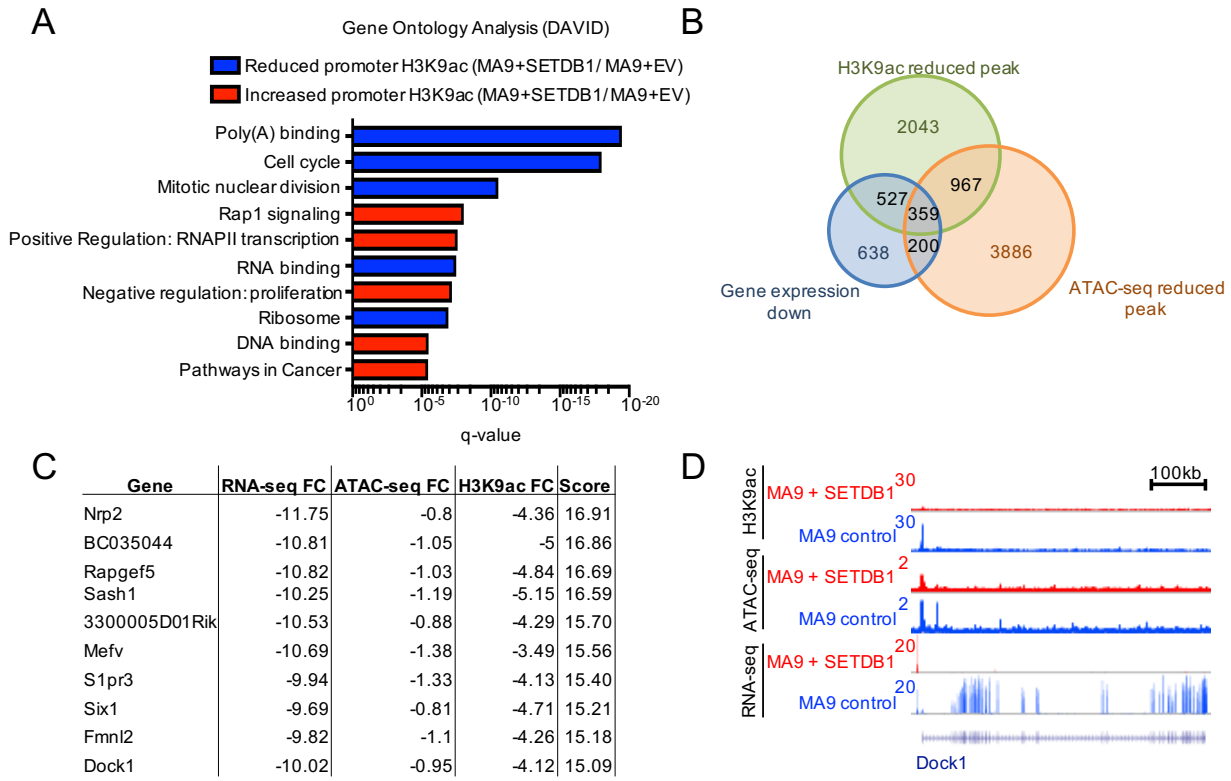


Figure 4-10: SETDB1 regulates Hoxa9 H3K9me3, chromatin availability, and expression. A) Sequencing tracks showing H3K9me3 ChIP-seq (top), ATAC-seq (middle), and RNA-seq signals for MA9+EV (blue) or MA9+SETDB1 (red) cells. Shown here is the entire *Hoxa* cluster of genes and a closer view of *Hoxa9* specifically. B) DiffBind analysis showing H3K9me3 ChIP-seq peaks that demonstrate differential binding. Only 552 consensus peaks were found in the analysis. C) H3K9me3 ChIP-seq track of the mouse Chromosome 7 Zinc Finger Cluster, which is known to be enriched for H3K9me3. D) DiffBind analysis showing H3K9ac ChIP-seq peaks that demonstrate differential binding.

reveals that the genes with decreased promoter H3K9ac in MLL-AF9+SETDB1 cells were associated with cell cycle and RNA binding, whereas genes with increased promoter H3K9ac were associated with signaling pathways and negative regulation of apoptosis (Figure 4-11A). Next, we overlapped genes that are downregulated with genes that increase chromatin compaction and have reduced promoter H3K9ac upon SETDB1 overexpression (Figure 4-11B). Scoring the genes found in this overlap by their overall fold changes across all three sequencing experiments results in several interesting targets including Nrp2, Six1 and Mefv, which are implicated as biomarkers in AML patients[383,384] (Figure 4-11C). We also observed significant loss of promoter H3K9ac, chromatin compaction and reduced gene expression of Dock1 (Figure 4-11D),



Gene	RNA-seq FC	ATAC-seq FC	H3K9ac FC	Score
Nrp2	-11.75	-0.8	-4.36	16.91
BC035044	-10.81	-1.05	-5	16.86
Rapgef5	-10.82	-1.03	-4.84	16.69
Sash1	-10.25	-1.19	-5.15	16.59
3300005D01Rik	-10.53	-0.88	-4.29	15.70
Mefv	-10.69	-1.38	-3.49	15.56
S1pr3	-9.94	-1.33	-4.13	15.40
Six1	-9.69	-0.81	-4.71	15.21
Fmnl2	-9.82	-1.1	-4.26	15.18
Dock1	-10.02	-0.95	-4.12	15.09

Figure 4-11: SETDB1 regulates promoter H3K9ac, chromatin availability, and expression of AML biomarkers. A) Gene ontology analysis done in DAVID to look for biological process, molecular function, and Kegg pathway gene sets that are overrepresented in the following groups: H3K9ac ChIP-seq peaks that have significantly reduced (blue) or increased (red) signal intensity in MA9+SETDB1 compared to MA9+EV. B) Overlap analysis of genes that are downregulated by SETDB1, genes where SETDB1 drives reduced promoter H3K9ac, and genes where SETDB1 drives reduced promoter ATAC-seq signal. C) Top ten scoring genes when fold changes for the three datasets in B are summed. D) Sequencing tracks showing H3K9ac ChIP-seq (top), ATAC-seq (middle), and RNA-seq signals for MA9+EV (blue) or MA9+SETDB1 (red) cells. Shown here is the locus for *Dock1*.

which is a prognostic marker of AML that displays changes in DNA methylation in AML patient samples relative to normal HSPCs[212,385,386]. We identified several direct binding targets of MLL-AF9 that underwent epigenetic remodeling and expression changes with overexpression of SETDB1. Thus, we asked whether SETDB1 impacts H3K79me2, which is deposited by DOT1L and associated with MLL-fusion proteins[6,8,387]. We performed H3K79me2 ChIP-seq in MLL-AF9 and MLL-AF9+SETDB1 cells to identify genes coregulated by H3K79me and H3K9me. We found a marked decrease of H3K9ac and ATAC-seq signal at H3K79me2 peaks in MLL-AF9 cells suggesting a role for H3K9 modifications in regulating genes marked with H3K79me2 (Figure 4-12A, B). To further explore the role of SETDB1 in regulating MLL-AF9 targets, we performed an

overlap analysis for genes that, upon SETDB1 overexpression, are downregulated, lose promoter H3K9ac, compact chromatin, lose gene body H3K79me2 signal, and are included in the published list of known binding targets of MLL-AF9[8]. This helped us to define a narrow target list of genes that may be coregulated by SETDB1 and MLL-AF9. Included in this group is Gfi1, which has been shown to affect AML cell growth [388]; Rap1gds1, a guanine nucleotide exchange factor; Arid1b, a member of the SWI/SNF complex; and Six1, which has been shown to promote formation of leukemic stem cells[7]

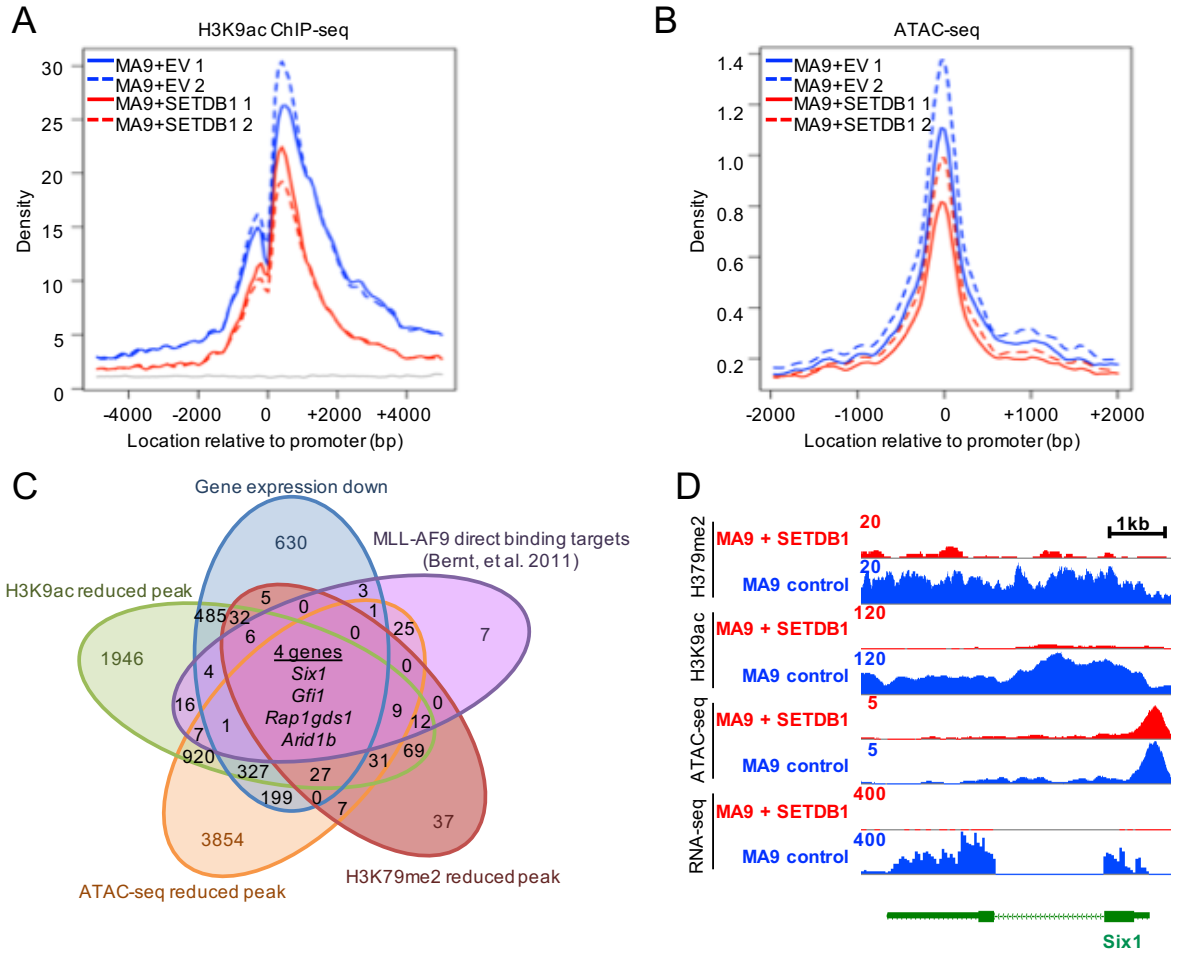


Figure 4-12: SETDB1 regulates the epigenetic landscape of a subset of MLL-AF9 target genes. A) Signal track for H3K9ac ChIP-seq centered around promoters of genes that have reduced H3K79me2 overlapping with the gene's locus in MA9+SETDB1 relative to MA9+EV cells. B) Signal track for ATAC-seq centered around promoters of genes that have reduced H3K79me2 overlapping with the gene's locus in MA9+SETDB1 relative to MA9+EV cells. C) Overlap analysis of genes that are downregulated by SETDB1 in AML, have reduced H3K9ac ChIP-seq, ATAC-seq, or H3K79me2 ChIP-seq upon SETDB1 overexpression in AML, or are a part of the published dataset of MLL-AF9 direct binding targets. D) Sequencing tracks showing H3K79me2 ChIP-seq (top), H3K9ac ChIP-seq (second from top), ATAC-seq (third from top), and RNA-seq (bottom) signals for MA9+EV (blue) or MA9+SETDB1 (red) cells. Shown here is the *Six1* gene locus.

(Figure 4-12C). Six1 was of particular interest given its role in promoting leukemogenesis and the striking reductions that we observed in H3K9ac, H3K79me2, ATAC-seq signal, and gene expression (Figure 4-12D). These data demonstrate profound effects of SETDB1 on the epigenome and transcriptome that includes genes critical for AML and are consistent with a role for SETDB1 in suppressing leukemic transformation.

4.4 Discussion

Here we demonstrate that gain of SETDB1 and H3K9 methylation suppresses AML disease progression *in vivo* through the repression of pro-leukemic genes including direct MLL-fusion protein targets. We found that AML patient samples exhibit lower *SETDB1* expression compared to normal hematopoietic cells and that higher *SETDB1* expression correlates strongly with better overall AML patient survival. We recapitulated these findings in mice where forced expression of *SETDB1* in MLL-AF9 driven AML induces differentiation of AML cells and increases disease latency. These data suggest SETDB1 suppresses AML cell growth and self-renewal by relieving the block in differentiation.

We attribute the phenotypes in AML cells to altered H3K9 methylation. We altered H3K9 methylation levels by manipulating the function of G9a both genetically and through small molecule inhibition. Similar to our results with SETDB1, manipulation of G9a suggests H3K9 methylation can suppress AML progression by promoting differentiation. Thus, H3K9 methylation may have a more general effect on AML initiation and progression. Interestingly, Lehnertz and colleagues reported a pro-leukemic function for G9a whereby overexpression accelerated a *HoxA9*/*Meis1* overexpression leukemia model. Mechanistically, G9a binds to Hoxa9, which is necessary for target gene expression[269]. These phenotypic differences may be explained by experimental models. MLL-AF9 directly drives endogenous *Hoxa9* and *Meis1* expression leading to transformation, which is repressed by SETDB1 in our model. A Hoxa9/Meis1 overexpression model may overcome this regulation. Further, mouse and human G9a possess unique functions in malignant cells including the ability to stimulate p53 transcriptional activity[257,269]. Consistent with a role for H3K9

methylation in suppressing hematopoietic transformation, deletion of the H3K9 methyltransferase, SUV39H1 (and to a lesser degree SUV39H2), leads to the development of B-cell lymphomas in mice[232]. Additionally, SETDB2 resides in a region of chromosome 13 that is commonly deleted in chronic lymphocytic leukemia (CLL)[245,268]. Thus, H3K9 methylation is likely exquisitely regulated in hematopoietic cells and performs context dependent functions that require further investigation to fully understand its role in AML.

Mechanistically, we found that SETDB1 is linked with altered H3K9 methylation and acetylation, decreased chromatin accessibility and transcriptional repression of critical AML oncogenes. These genes included several that have been implicated in MDS and AML[383,384]. We show SETDB1 regulates *Dock1* expression, which is correlated with leukemic stem cell gene signatures and a poor prognosis in AML patients[385,386]. We also observed that SETDB1 represses genes associated with AML, such as *Hoxa9* and *Six1*, which are direct targets of MLL-AF9[2,7,11]. These data point to SETDB1 negatively regulating a pro-leukemic gene program, many of which are potential therapeutic targets. Thus, understanding the mechanisms regulating *SETDB1* at the transcriptional and post-translational level may be a valuable therapeutic

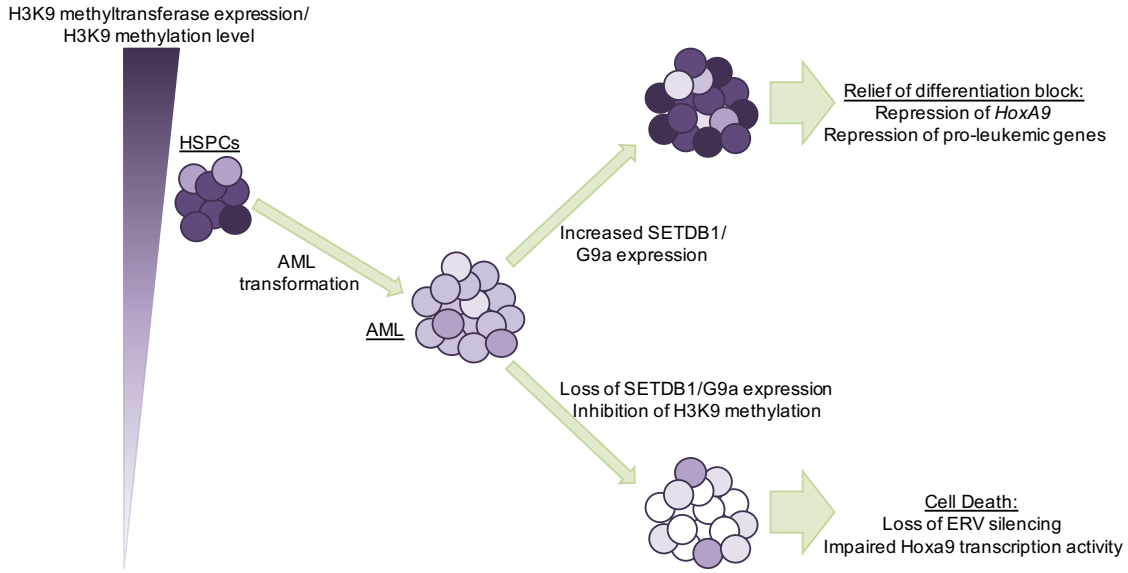


Figure 4-13: Working model for the proposed role of SETDB1/H3K9 methylation in AML initiation and maintenance. AML is initiated after an attenuation of H3K9 methyltransferase function, expression, or H3K9 methylation by other mechanism. Alternatively, AML arises from HSPCs that already have low levels of H3K9 methylation/ methyltransferases. After establishment of an AML, inhibition of H3K9 methyltransferases leads to loss of retroviral silencing and cell death. Stabilization of SETDB1 or G9a leads to repressed Hox gene expression and a relief of the differentiation block.

approach for AML. For example, miRNA29 is a critical mediator of *SETDB1* expression [249]. Another potential mediator of H3K9 methylation is the PAF1c. We showed SETDB1 binds to the PAF1c and mediates promoter H3K9me3 of the *Hoxa9* and *Meis1* loci[368]. Further, we and others have also identified G9a and SUV39H1 as interacting partners of the PAF1c[201 368]. Interestingly, SETDB1, G9a, GLP and SUV39H1 can exist in a complex that directs H3K9 methylation to euchromatic gene promoters[234]. Thus, the PAF1c may recruit H3K9 methyltransferases to specific targets to mediate gene repression. The PAF1c is a critical regulator of transcription in AML cells through direct physical interaction with wild type MLL and MLL-fusion proteins[12,13]. Target genes include *Hoxa9*, *Meis1, Prmt5* and others critical for AML cell growth[13,14,224]. It will be interesting to consider the biochemical interplay between H3K9 methyltransferases and MLL-fusion proteins with the PAF1c.

Previous studies have demonstrated that SETDB1 and G9a are required for AML initiation and progression[264,267,269]. In light of our current data demonstrating that SETDB1 suppresses AML growth, AML cells may maintain a narrow *SETDB1* expression level. We show that increased SETDB1 expression induces differentiation of AML cells through H3K9me3 and repression of self-renewal genes. Conversely, loss of SETDB1 is detrimental to leukemic cells due to derepression of endogenous retroviral elements (ERVs) and inhibition of Hoxa9 transcriptional activity[267,269] (Figure 4-13). Given the requirement for SETDB1 in leukemia, small molecule inhibition of H3K9 methyltransferases has been proposed as a therapeutic option[264,269]. However, a recent study shows that depletion of G9a increased a cancer progenitor cell population that initiates a delayed but more aggressive disease state[320]. Our data is consistent with a role for H3K9 methyltransferases in suppressing MLL-AF9 leukemia. Thus, it is critical to fully understand the effects of chemically inhibiting of H3K9 methylation as a treatment for AML. Further investigation into the roles of SETDB1, G9a and more generally H3K9 methylation levels will likely shed light on the precise role of these methyltransferases in normal and malignant hematopoiesis and determine the value of these epigenetic modifiers as therapeutic targets.

Chapter 5 Conclusions and future directions

5.1 Summary and conclusions

As the field of precision medicine continues to explore the value of targeting proteins that modulate the epigenetic landscape to treat malignancy, it will remain important to fully understand the role and regulation of these epigenetic modifiers in different disease states. Here, we sought to define novel regulatory functions for a specific epigenetic commodifying complex, the PAF1c, in the context of acute myeloid leukemia. To that end, we demonstrated that the PAF1c is critical for both AML cell growth and normal HSPC cell growth (Chapter 2). Interestingly, the mechanism by which the PAF1c supports the growth and self-renewal of these cells is through regulation of distinct gene subsets of the transcriptome. In AML, the PAF1c regulates oncogenic programs associated with self renewal, such as gene programs regulated by Hoxa9/Meis1, direct binding targets of MLL-fusion proteins, and a group of histone methyltransferases known as the PRMT family. In HSPCs, the PAF1c is responsible for maintaining proper cell cycle regulation, as evidence by the deregulation of cell cycle and quiescence gene programs upon depletion of Cdc73. Finally, we show that there are overlapping and unique transcriptional targets of the PAF1c in AML cells and HSPCs, but that many of the gene programs affected by perturbation of the PAF1c that are most likely to account for the phenotypic changes that are seen are distinct from each other. This discovery and the knowledge that the PAF1c-MLL interaction seems preferentially required in AML relative to normal hematopoietic cells lead us to ask what mechanisms of regulation are responsible for the leukemia-specific functions of the PAF1c.

To explore the biochemical regulation of the PAF1c in AML, we defined the PAF1c interactome using the CDC73 subunit in an affinity purification mass spectrometry experiment (Chapter 3). This proteomics study revealed both known and novel binding partner candidates of the PAF1c. One particularly interesting group of

proteins that were candidate interacting partners of the PAF1c were three H3K9 methyltransferases of the SUV39 family of proteins: SETDB1, G9a, and GLP. The interaction between SETDB1 and CDC73 and the interaction between G9a and CDC73 were validated. Interestingly, a triple point mutant of CDC73, CDC73_3YF, has a stabilized interaction with these proteins and does not support AML cell growth. Overexpression of SETDB1 in AML cells and utilized the CDC73_3YF mutant to examine the effects of the stabilized interaction between the PAF1c and SETDB1. This provided two ways to determine the functional role of SETDB1 in PAF1c regulated transcription. SETDB1 regulates deposition of promoter H3K9me3 to repress the expression of both *Hoxa9* and *Meis1*. Further, expression of the mutant to generate a stabilized PAF1c-SETDB1 interaction mimicked these results. Finally, *SETDB1* expression inversely correlates with *HOXA9* and *MEIS1* in patient samples, suggesting the findings of these studies are relevant to human AML.

Given the known oncogenic role of HOXA9 and MEIS1 in AML and our newly discovered correlation between SETDB1 expression and these oncogenes, we sought to determine the biological role of SETDB1 in AML (Chapter 4). We first noted that SETDB1 expression correlates with overall survival and that SETDB1 expression is lower in AML patient samples than in normal hematopoietic cells. Next self-renewal and proliferation assays demonstrated that increased expression of SETDB1 suppresses AML cell growth *in vivo*. Mechanistically, SETDB1 represses the expression of oncogenic programs, including genes regulated by *Hoxa9*/*Meis*, genes that are directly targeted by MLL-fusion proteins, and other known prognostic biomarkers of AML. SETDB1 represses these genes by inducing changes in the epigenome, including increased gene body/promoter H3K9me3, decreased promoter H3K9ac, decreased gene body H3K79me2, and decreased promoter chromatin availability. Finally, we examined the effects of inhibiting H3K9 methyltransferases in the context of AML and normal hematopoietic cell growth. While inhibition of H3K9 methylation is toxic to established AML cells, in HSPCs it relieves suppression of genes associated with malignancy, increases the self-renewal capacity of the cells, and creates a pool of cells that are more amenable to transformation by an oncogene.

These data reveal novel roles for H3K9 methylation, SETDB1/G9a/H3K9 methyltransferases, and the PAF1c in AML. We propose a role for PAF1c mediated H3K9 methylation as a mechanism by which PAF1c target genes, such as Hoxa9 and Meis1, are regulated in normal hematopoietic cells. When the H3K9 methylation axis of regulation becomes disrupted, it can create cells that are more likely to become malignant, depending on the manner in which and degree to which it is disrupted. After establishment of AML cells, stabilization of H3K9 methyltransferase expression or stabilization of the PAF1c-H3K9 methyltransferase interaction leads to a repression of PAF1c target genes and relief of the AML differentiation block. Conversely, loss of H3K9 methylation results in loss of normal cellular processes, such as the silencing of endogenous retroviral elements, and leads to cell death through apoptosis. This suggests that there is a window of expression of H3K9 methyltransferases and a level of H3K9 methylation that is allowable for transformation and maintenance of an AML, and that changes outside of this window in either direction can be used to delay the growth of AML (Figure 5-1).

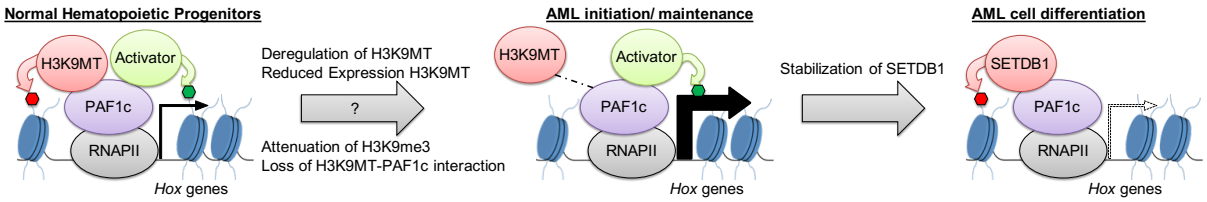


Figure 5-1: Working model and hypotheses for the proposed role of the PAF1c-SETDB1 interaction in AML initiation and maintenance. The PAF1c binds to epigenetic modifiers associated with both activation and repression, including H3K9 methyltransferases (H3K9MT). An attenuation of H3K9 methylation through various mechanisms, including downregulation of the H3K9MT or loss of the H3K9MT-PAF1c interaction, could lead to a loss of repression for genes associated with self-renewal. However, stabilization of this interaction or of H3K9MT like SETDB1 leads to repression of self-renewal genes and relief of the differentiation block.

5.2 Future directions

The PAF1c/SETDB1 regulation of the transcriptome/ epigenome in AML/ HSPCs

In chapters 2 and 4, an important role for the PAF1c and for H3K9 methyltransferases in the regulation of the transcriptome and epigenome in AML cells was shown. However, there are many lingering questions that should be addressed in the future. First, it will be important to fully understand the PAF1c-H3K9

methyltransferase interaction as it pertains to modulating changes in the epigenome. For instance, are all PAF1c targets also targets of SETDB1? Is the opposite case true? Or do they coregulate just a subset of genes? Transcription factor ChIP-seq analysis will be required to fully explore the overlap of regulation. Furthermore, to this point we have only approached this from the perspective that H3K9 methylation plays a role in regulating known PAF1c target genes. However, it is possible that the PAF1c is also playing an as yet undiscovered role in regulating known targets of H3K9 methyltransferases. For instance, we might speculate that the PAF1c is also bound to and playing an important role in regulating the expression of additional euchromatic targets of SETDB1/G9a apart from *Hoxa9* and *Meis1*, which are targets of H3K9 methylation as revealed in this study. Additionally, it is possible that the PAF1c may have a part to play in the regulation of heterochromatin formation that is mediated by SETDB1 and other H3K9 methyltransferases. In yeast, it has been shown that the PAF1c prevents the spreading of heterochromatin and converts the boundary regions of heterochromatin into permissive chromatin states for transcription[389,390]. However, this has not been explored in mammals and there are no studies to our knowledge that explore whether the PAF1c regulation of heterochromatin is related to interactions with H3K9 methyltransferases or if its regulation of this process may play a role in disease. This could provide an interesting new angle to pursue with regards to the PAF1c deregulation in disease.

The PAF1c interactome in AML

In chapter 3, we demonstrated the utility of using a proteomics-based approach to explore novel and known interactors of the PAF1c in leukemia cells. By using an affinity purification mass spectrometry (AP-MS) approach and network analyses, we were able to identify with confidence an interactome in AML cells for both wildtype CDC73 and CDC73_3YF. However, there remains a large amount of information from these experiments that has not yet been thoroughly explored or interpreted.

First, there were additional candidate interactions that were identified as being preferentially bound to CDC73_3YF that may also be playing a role in the PAF1c-H3K9 methylation axis of regulation. A noteworthy example of this is MECOM/EVI-1 (PRDM3),

which is interesting both because of its oncogenic roles in AML as well as the H3K9 methyltransferase activity of the longer protein isoform. Further, several known interacting partners of H3K9 methyltransferases that are critical for their functions were also identified by AP-MS, including WIZ, which interacts with G9a/GLP; ATF7IP, which is critical for the deposition of SETDB1-mediated H3K9me3; and PML, which is a tumor suppressor in leukemia and interacts with G9a/GLP/SETDB1/SUV39H1. Validation and elucidation of these interaction networks may provide important insight into the PAF1c regulation of H3K9me3.

It will also be important to further explore the interaction network of the PAF1c. Thus far, we have only examined interactions that are enhanced by introducing artificial mutations to CDC73. While these interactions also occur with wildtype CDC73 and are therefore important to study, it will also be interesting to pursue strong interactions that occur with the wildtype PAF1c. This may include interactions that are disrupted by the CDC73_3YF, which would provide more evidence of the dynamic interplay between the protein-protein interactions that drive transcription and those that repress transcription mediated by the PAF1c. Finally, a full proteomics study looking at the interactions that occur with all of the PAF1c subunits could provide a clearer picture of the protein-protein interactions that are most important for the functions of the PAF1c in leukemia.

The PAF1c-H3K9 methylation interaction network

In chapter 3, we specifically explored the interactions between the PAF1c and the H3K9 methyltransferase SETDB1. However an interaction between the PAF1c and G9a was also validated. Further, the AP-MS study provided evidence that GLP, the heterodimeric binding partner of G9a, may also interact with the PAF1c, though this interaction is unvalidated. Finally, previous reports have demonstrated an interaction between CDC73 and SUV39H1, though the PAF1c dependency of this interaction has not been show. This provides varying degrees of evidence that all four H3K9 methyltransferases that act together in a euchromatic gene repressing complex bind to the PAF1c. This provides us with the interesting question of whether the PAF1c binds to these H3K9 methyltransferases individually, to the whole H3K9 methyltransferase complex, or to different subcomplexes of these methyltransferases depending on

context. Further, different interacting partners of these H3K9 methyltransferases may also bind to the PAF1c as revealed by our proteomics data. In light of this, we may also ask whether these proteins are also a part of this H3K9 methyltransferase complex and whether they too bind to the PAF1c in this context. We might speculate that, much as the PAF1c plays an integral role in the function of large complexes of proteins that drive transcription, it may also play a role in the function of large complexes of proteins that repress transcription. To elucidate these mechanisms, we can perform further proteomics analyses that involve using the members of the PAF1c as well as the H3K9 methyltransferases and associated proteins as bait proteins in AP-MS to add clarity to the protein-protein interaction networks involving the PAF1c. We can also biochemically isolate the different complexes that the PAF1c is bound in to probe the interaction partners of complexes of varying sizes.

If we were to find that these different complexes are forming, it could have important implications into the dynamic regulation of and context-dependent roles of the PAF1c. Elucidating the mechanisms that would cause one type of complex to form over another could give us insight into how the PAF1c is regulated during development and deregulated in malignancy.

The PAF1c-H3K9 methylation regulation axis in development

In chapter 4, we showed that SETDB1 and H3K9 methylation repress genes associated with self-renewal and oncogenesis, such as *Hoxa9*. Further, this repression of genes and suppression of self-renewal occurs in both AML and normal HSPCs. Given this data and a wide range of evidence that Hoxa9 and genes associated with self-renewal must be progressively downregulated during hematopoietic differentiation, H3K9 methylation and the proteins that modulate this modification may play an important role in this downregulation. While it has been shown by many different groups that H3K9 methylation itself is dramatically changed during differentiation, the mechanisms by which these changes occur are not clear. There could exist a wide array of mechanisms by which cells induce the downregulation of genes associated with self-renewal through changes in H3K9 methylation. Some of these could include

changes in expression, protein levels, or catalytic function of H3K9 methylation modulating proteins as cells undergo differentiation.

Another interesting potential mechanism of regulation could involve changes in protein-protein interactions, potentially with the PAF1c. In chapter 4 we showed that *Hoxa9* and *Meis1* repression could be induced not just by overexpression of SETDB1, but also by stabilizing the interaction with the PAF1c by expressing CDC73_3YF. This mutant may be representative of a change in interaction stability that occurs in nature. For instance, we can speculate that the PAF1c-H3K9 methyltransferase interaction may be a weak and transient interaction in HSCs and primitive cells, but the interaction is enhanced as cells differentiate in order to drive repression of PAF1c target genes as opposed to activation. If deregulation of the PAF1c or H3K9 methyltransferase lead to a loss of this stabilized interaction, it could lead to aberrant expression of genes associated with self-renewal and make cells more amenable to malignant transformation.

The stabilization of this interaction could occur through a number of mechanisms, including changes in PTMs on the PAF1c, which could be represented by CDC73_3YF. It is noteworthy that CDC73_3YF has been reported to be a phosphodeficient mutant of CDC73, though our group has been unable to demonstrate that CDC73 is phosphorylated in AML. It is possible that the level of phosphorylation is below the limit of detection in AML, that we are disrupting a different PTM, or that something else altogether is causing the enhanced interaction with SETDB1. Further work determining what, if any, PTMs on CDC73 exist and are changed during hematopoiesis and AML initiation would provide insight into a possible mechanism for stabilizing this interaction. Also, a thorough study should be performed exploring changes in expression of H3K9 methyltransferases and the PAF1c, as well as changes in protein-protein interactions in primitive hematopoietic cells relative to more mature cells. This could reveal a novel PAF1c-centric mechanism by which self renewal genes are regulated during blood cell development, again providing insight into how they might become aberrantly expressed in disease contexts.

The PAF1c-SETDB1-MLL interaction

In chapter 4, we demonstrated that the PAF1c and SETDB1 interact in AML cells. Furthermore, SETDB1 regulates targets of both the PAF1c and MLL-fusion proteins in AML cells. Because we know that the PAF1c also interacts with MLL-fusion proteins in AML cells, it would be interesting to determine the dynamics of this relationship between these three proteins. Milne and colleagues addressed a separate but related question when they demonstrated that SETDB1 and H3K9me3 have a higher binding signal at the Hoxa9 locus when *MLL1* is deleted, which are reduced when *MLL1* is re-expressed. However, they did not explore what the cause of this was mechanistically and they did not examine whether this chromatin binding exclusivity remained true in the context of MLL-fusion proteins. Based on our data showing coregulation of a subset of MLL-fusion targets by SETDB1, it is possible that the reason for the loss of SETDB1 and H3K9me3 signal that they observe could be due to a PAF1c-MLL interaction that excludes SETDB1. If there is a mutually exclusive interaction between the PAF1c and SETDB1/MLL-fusions, it would be interesting to explore what causes the PAF1c-MLL-fusion interaction to be favored in AML to drive Hoxa9 transcription.

Furthermore, in light of Milne et al.'s data, it would be important to examine whether there is a shift in PAF1c binding affinity for MLL1 relative to SETDB1 in primitive cells that require high expression of MLL targets such as Hoxa9 compared to differentiated cells that exhibit downregulation of these genes. As we previously speculated, there might be a change in the binding affinity of the PAF1c for different interaction partners as cells differentiate. With that in mind, it is possible that the downregulation of *Hox* genes would occur not just through the transcriptional repression mediated by H3K9 methylation, but also by a loss of the H3K4me3 activating modification. This could provide another mechanism by which the PAF1c and SETDB1 function to regulate *Hox* gene expression during development.

Post translational modifications on the PAF1c in leukemia

We have speculated that the PAF1c may be differentially post-translationally modified in different hematopoietic or leukemic cell contexts. As mentioned, CDC73 has

been shown to be phosphorylated in gastric cells, and when it is dephosphorylated, it functions as an oncogenic driver. Further, LEO1 has been shown to drive cellular proliferation when dephosphorylated in AML. This shows that the PAF1c function can change a result of PTMs. Though our data doesn't support the phosphorylation of CDC73 in the context of AML, preliminary mass spectrometry data identified a novel methylation site on CDC73 in AML. An exploration of whether this methylation plays a functional role and what proteins regulate the deposition of this PTM could provide a novel mechanism by which the PAF1c is regulated in blood cells that may be deregulated in AML. For example, it is possible CDC73 must be demethylated in order to increase the affinity of the PAF1c for protein interactions that repress genes associated with self-renewal. These studies would reveal how that process is regulated. Furthermore, it would be interesting to apply a mass spectrometry approach to identifying PTMs on the PAF1c. This could be done using recombinant proteins as a starting point to identify novel PTM sites. However, the really interesting experiments would be to compare the PTMs on the PAF1c in AML cells relative to normal cells, as well as in more primitive hematopoietic cells compared to mature cells. These studies might show us new ways that the PAF1c is regulated.

Therapeutic targeting of H3K9 methyltransferases in AML

Finally, in chapter 5 we demonstrated that H3K9 methylation and H3K9 methyltransferases play an important role in the repression of self-renewal genes and the suppression of leukemic cell growth. This data is a critical piece of information as we explore the possibility of therapeutically targeting epigenetic modifiers in AML. Frequently, it is tempting to suggest that epigenetic modifiers are attractive therapeutic targets simply because they are required for the survival of AML cells. However, in this study we show that even though SETDB1/G9a are required for AML cell growth, inhibition of these proteins may create a pool of cells that are more amenable to leukemic transformation. This suggests that while treatment with H3K9 methyltransferase inhibitors may be toxic for a large majority of AML cells, it is possible that any surviving AML cells will in fact be a more aggressive clone, expressing higher levels of genes associated with aggressive disease states, such as *HOXA9*. Further, it

is important to consider off-target effects. Our data suggests that normal hematopoietic progenitors that are present when an AML is treated with an H3K9 methyltransferase inhibitor will survive, will have a higher self-renewal capacity, and will be more likely to be transformed to an AML by an oncogenic event.

To further understand the utility in targeting H3K9 methyltransferases, a more complete study should be performed to elucidate the full effects of treating an AML with an H3K9 inhibitor. For instance, we might imagine using a mouse model of AML where we transplant primary AML cells that have been treated with an H3K9 inhibitor. The likely result (and indeed reported result[269]) of this initial experiment would be that the AML disease progression would be delayed. However, if we then perform tertiary transplants, do we see the development of a more aggressive disease state? In a separate but related experiment, we could simply treat non-malignant bone marrow with the inhibitor, transplant these cells into mice, and determine whether there is a higher penetrance of malignant hematopoiesis in the treated mice compared to control. This study would mimic the possible off-target effects of treating patients with H3K9 methyltransferase inhibitors. In summary, these data suggests that caution should be used when thinking of H3K9 methyltransferases as therapeutic targets, and a more thorough study of the effects of using these H3K9 inhibition in mouse models would help to understand the full picture of how HSPCs and AML cells are affected by treatment with these chemicals.

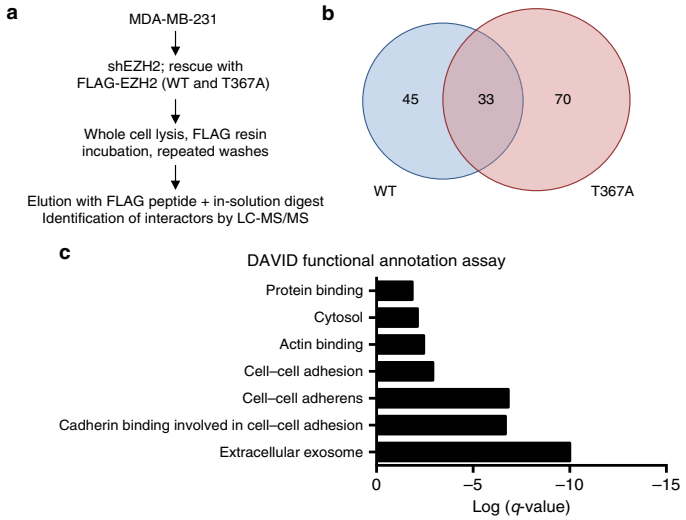
5.3 Final thoughts

Here, we sought to explore the role and regulation of the PAF1c and its interaction partners in AML and normal HSPC cell growth. While this study was primarily focused on the role of PAF1c-mediated H3K9 methylation, and more specifically on the H3K9 methyltransferases SETDB1 and G9a in AML, it is probable that these findings have implications with regards to other epigenetic modifying proteins. Further, it is likely that these findings have implications in other malignancies. The dynamic nature of transcriptional and epigenetic regulation makes it unlikely that any one epigenetic modifier or epigenetic modifying complex has one specific role in development. It is therefore unlikely that their contribution to disease is as simple as

being involved in the deregulation of one specific process. While a genetic knockout of these modifiers may be toxic to cancer cells, this could often simply be the result of losing an essential regulator of cellular homeostasis. For instance, depletion of SETDB1 was shown by Cuellar and colleagues to lead to a toxic interferon response due to a loss of retroviral element silencing. However, inhibition of these modifiers may prove to be more complex. If the inhibitors are not as efficient at attenuating the function of these modifiers as depletion is, we might simultaneously be targeting the desired function of the modifier as well as an undesired function. In our example, while we may be inducing cell death, we are also de-repressing genes associated with self-renewal. It is therefore important to consider the full body of effects that are induced by targeting epigenetic modifiers, as opposed to simply determining whether they have essential functions in cancer cells or not. These more nuanced studies will be critical as the field moves forward and continues to explore the exciting possibility of treating malignancy with chemicals targeting epigenetic modifiers.

Appendix 1 Proteomics analyses reveals phosphorylation dependent interactome of EZH2 in metastatic breast cancer

This study was a collaboration with Talha Anwar and Celina Kleer. In this study, Anwar, et al. demonstrated that EZH2 is phosphorylated during breast cancer progression. They demonstrate that this phosphorylation event (which is on residue T367 of EZH2) drives EZH2 localization to the cytoplasm, where it is required to be



d

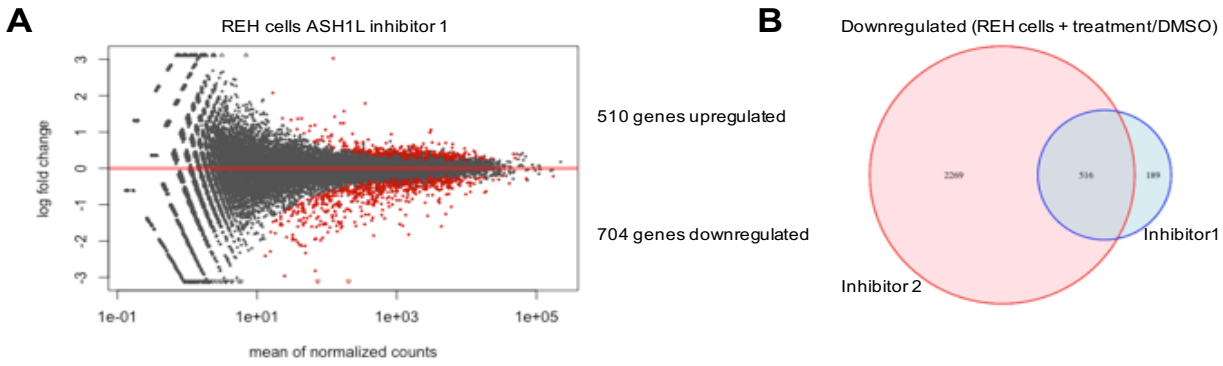
GENE_ID	WT_FC_A	T367A_FC_A	AVERAGE_WT_SC	AVERAGE_T367A_SC	AVERAGE_control_SC	WT_SP	T367A_SP
SYNE2	22.9	4.8	27	4	0	1	0.7
EPS8	7.2	0.6	22	0.3	2	0.9	0
MLPH	5.3	1.1	16	2	1.7	1	0.2
MYO1F	5.14	4.97	16	13.3	2	0.94	0.67
VCL	4.2	3.1	26.3	16.7	4	1	0.8
DBNL	3.9	3	3.7	2.3	0	1	0.7
EPS8L2	3.5	1.7	10.3	4	1.3	1	0.5

Figure A-1: The interactome of pEZH2(T367) reveals new cytoplasmic binding proteins. A) Schematic of mass spectrometry experiment to identify binding partners of EZH2 in MDA-MB-231 cells. Experiment was performed in triplicate. B) Venn diagram displaying interaction partner candidates overlap in proteins co-immunoprecipitating with WT- or T367A-EZH2 from the three biological replicates analyzed. C) DAVID functional annotation analysis of processes enriched in WT-EZH2 over T367A-EZH2. D) Top interaction candidates identified from a published actin-binding protein set with fold-change (FC) scores and normalized FC scores (relative to total bait protein pulldown). Average WT and T367A spectral counts (SC) and SAINT probabilities (SP) are also tabulated. (Anwar, et al 2018).

localized for its function in driving metastatic breast cancer. They also sought to define the phosphorylation dependent interactome of EZH2 in breast cancer cells. To do this, they utilized a phosphorylation deficient mutant of EZH2 (T367A-EZH2) to eliminate phosphorylation dependent interactions from the interactome. My role in this study was to analyze data generated from an affinity-purification mass spectrometry experiment in order to provide a list of candidate interacting partners for EZH2 that might contribute to its function in driving cancer progression. This analysis revealed that wildtype EZH2 binds to cytoplasmic and exosomal proteins to a higher degree than T67A-EZH2 and revealed vinculin as a candidate protein-protein interaction that is specific for phosphorylated EZH2. Anwar, et al. went on to demonstrate that this interaction played a role in localizing vinculin to focal adhesions, which is important for cell spreading. They propose that this interaction may be a part of the mechanism by which phosphorylated EZH2 plays a role in breast cancer progression[391].

Appendix 2 Transcriptomic effects of treating cancer cell lines with epigenetic inhibitors

Here I contributed to a series of studies in collaboration with Jolanta Grembecka and Tomek Cierpicki. The goal of the first study was to observe changes in the transcriptome when treating different leukemia cell lines with a small molecule inhibitor that targets ASH1L, which is an H3K36 methyltransferase. Three different human leukemia cell lines were treated with the inhibitor for 6-7 days and RNA was harvested and submitted for mRNA-sequencing. The second study involved treatment of two



C

Gene_Symbol	log2FoldChange_REH+Inhibitor1	padj_REH+Inhibitor1	log2FoldChange_REH+Inhibitor2	padj_REH+Inhibitor2
INHBE	-2.817895533	1.05E-17	-4.5354116	5.11E-12
GPR146	-2.330779848	2.44E-09	-3.270120485	2.48E-27
SPNS3	-1.562393688	1.90E-05	-3.207335705	1.24E-16
TSC22D3	-1.822349851	3.40E-43	-3.008286892	0
AK4	-1.626876834	1.22E-06	-2.957628492	7.81E-42
SNX20	-1.446135711	0.040082837	-2.64837463	0.003122858
CD9	-1.849706003	2.95E-44	-2.61356339	9.19E-146
IL21R	-3.646765469	1.45E-24	-2.612847617	8.02E-18
DDIT4	-1.533153244	3.13E-42	-2.575327233	1.73E-207
ALDOC	-1.433075655	9.61E-21	-2.515859656	6.08E-82

Figure A-2: Sample plots and tables for the differential gene expression analysis of human cancer cell lines treated with epigenetic inhibitors. A) MA plot showing fold-changes of human leukemia cells treated with an ASH1L inhibitor relative to control treated cells plotted against average expression of the genes across all samples. B) Overlap analysis of transcriptomic effects of two separate molecules targeting ASH1L in a human leukemia cell line. Shown are the genes that are significantly downregulated upon treatment with either molecule C) List of top ten genes that are downregulated upon treatment with either of the two ASH1L inhibitors.

different human breast cancer cell lines with a small molecule inhibitor targeting NSD1, a separate epigenetic modifier that also methylates H3K36. The final study sought to examine the effect of a separate molecule targeting ASH1L on a human leukemia cell line. This molecule is more potent and the cell line used was one of those used in study 1, so an overlap analysis comparing the more potent molecule to the weaker was the primary objective. I obtained the sequencing files and ran them through my analysis pipeline to determine differentially expressed genes. These data revealed significant changes in the genome upon treatment with either of the ASH1L targeting molecules. However, the gene expression changes did not exhibit a striking overlap across the different cell lines. In contrast, the two different molecules targeting ASH1L showed a remarkable overlap in gene expression changes, which is supportive of their selectivity. Further, the more potent molecule induced downregulation of a larger number of genes, as would be expected. Finally, there were also significant changes in transcriptomes when breast cancer cells were treated with the NSD1 inhibitor, but the changes were again variable across the two different cell lines. I also analyzed the gene programs that were significantly changed after treatment with these molecules using CAMERA gene set analysis and discovered that there are dramatic changes in the gene programs after treatment with these molecules in all cell lines. These data will be used to identify targets and critical gene programs that are affected by these molecules to gain further insight into the mechanism by which they affect cancer cell growth. representative plots and tables that were generated in this differential gene expression analysis.

Bibliography

1. Cancer Genome Atlas Research Network, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* 2013;368(22):2059–2074.

2. Collins CT, Hess JL. Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets. *Oncogene*. 2016;35(9):1090–1098.

3. Jung N, Dai B, Gentles AJ, Majeti R, Feinberg AP. An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis. *Nat. Commun.* 2015;6:8489.

4. Thorsteinsdottir U, Kroon E, Jerome L, Blasi F, Sauvageau G. Defining roles for HOX and MEIS1 genes in induction of acute myeloid leukemia. *Mol. Cell. Biol.* 2001;21(1):224–234.

5. Skvarova Kramarzova K, Fiser K, Mejstrikova E, et al. Homeobox gene expression in acute myeloid leukemia is linked to typical underlying molecular aberrations. *J Hematol Oncol*. 2014;7:94.

6. Slany RK. The molecular biology of mixed lineage leukemia. *Haematologica*. 2009;94(7):984–993.

7. Wang Q-F, Wu G, Mi S, et al. MLL fusion proteins preferentially regulate a subset of wild-type MLL target genes in the leukemic genome. *Blood*. 2011;117(25):6895–6905.

8. Bernt KM, Zhu N, Sinha AU, et al. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer Cell*. 2011;20(1):66–78.

9. Pui C-H, Gaynon PS, Boyett JM, et al. Outcome of treatment in childhood acute lymphoblastic leukaemia with rearrangements of the 11q23 chromosomal region. *Lancet*. 2002;359(9321):1909–1915.

10. Krivtsov AV, Armstrong SA. MLL translocations, histone modifications and leukaemia stem-cell development. *Nat. Rev. Cancer*. 2007;7(11):823–833.

11. Faber J, Krivtsov AV, Stubbs MC, et al. HOXA9 is required for survival in human MLL-rearranged acute leukemias. *Blood*. 2009;113(11):2375–2385.

12. Milne TA, Kim J, Wang GG, et al. Multiple interactions recruit MLL1 and MLL1

fusion proteins to the HOXA9 locus in leukemogenesis. *Mol. Cell*. 2010;38(6):853–863.

13. Muntean AG, Tan J, Sitwala K, et al. The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. *Cancer Cell*. 2010;17(6):609–621.

14. Muntean AG, Chen W, Jones M, et al. MLL fusion protein-driven AML is selectively inhibited by targeted disruption of the MLL-PAFc interaction. *Blood*. 2013;122(11):1914–1922.

15. Tomson BN, Arndt KM. The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. *Biochim. Biophys. Acta*. 2013;1829(1):116–126.

16. Van Oss SB, Cucinotta CE, Arndt KM. Emerging Insights into the Roles of the Paf1 Complex in Gene Regulation. *Trends Biochem. Sci.* 2017;42(10):788–798.

17. Karmakar S, Dey P, Vaz AP, et al. PD2/PAF1 at the crossroads of the cancer network. *Cancer Res.* 2018;78(2):313–319.

18. Vagapova ER, Spirin PV, Lebedev TD, Prassolov VS. The role of TAL1 in hematopoiesis and leukemogenesis. *Acta Naturae*. 2018;10(1):15–23.

19. Nakajima H. Role of transcription factors in differentiation and reprogramming of hematopoietic cells. *Keio J Med*. 2011;60(2):47–55.

20. Sharma S, Gurudutta G. Epigenetic regulation of hematopoietic stem cells. *Int J Stem Cells*. 2016;9(1):36–43.

21. Kosan C, Godmann M. Genetic and epigenetic mechanisms that maintain hematopoietic stem cell function. *Stem Cells Int.* 2016;2016:5178965.

22. Ng AP, Alexander WS. Haematopoietic stem cells: past, present and future. *Cell Death Discov.* 2017;3:17002.

23. Cell Potency: Totipotent vs Pluripotent vs Multipotent Stem Cells | Technology Networks.

24. Ogawa M, LaRue AC, Mehrotra M. Hematopoietic stem cells are pluripotent and not just “hematopoietic”. *Blood Cells Mol. Dis.* 2013;51(1):3–8.

25. Papathanasiou P, Attema JL, Karsunky H, et al. Self-renewal of the long-term reconstituting subset of hematopoietic stem cells is regulated by Ikaros. *Stem Cells*. 2009;27(12):3082–3092.

26. Ivanova NB, Dimos JT, Schaniel C, et al. A stem cell molecular signature. *Science*. 2002;298(5593):601–604.

27. Challen GA, Boles N, Lin KK-Y, Goodell MA. Mouse hematopoietic stem cell identification and analysis. *Cytometry A*. 2009;75(1):14–24.

28. Graf T. Differentiation plasticity of hematopoietic cells. *Blood*. 2002;99(9):3089–3101.

29. Vivier E, Tomasello E, Baratin M, Walzer T, Ugolini S. Functions of natural killer cells. *Nat. Immunol.* 2008;9(5):503–510.

30. Husby S, Grønbæk K. Mature lymphoid malignancies: origin, stem cells, and chronicity. *Blood Advances*. 2017;1(25):2444–2455.

31. Woolthuis CM, Park CY. Hematopoietic stem/progenitor cell commitment to the megakaryocyte lineage. *Blood*. 2016;127(10):1242–1248.

32. Kim J, Lee H, Shin S. Advances in the measurement of red blood cell deformability: A brief review. *JCB*. 2015;1(1):63–79.

33. Machlus KR, Italiano JE. The incredible journey: From megakaryocyte development to platelet formation. *J. Cell Biol.* 2013;201(6):785–796.

34. Weiskopf K, Schnorr PJ, Pang WW, et al. Myeloid cell origins, differentiation, and clinical implications. *Microbiol. Spectr.* 2016;4(5):.

35. Engblom C, Pfirschke C, Pittet MJ. The role of myeloid cells in cancer therapies. *Nat. Rev. Cancer*. 2016;16(7):447–462.

36. Ramos-Mejía V, Navarro-Montero O, Ayllón V, et al. HOXA9 promotes hematopoietic commitment of human embryonic stem cells. *Blood*. 2014;124(20):3065–3075.

37. Collins CT, Hess JL. Deregulation of the HOXA9/MEIS1 axis in acute leukemia. *Curr Opin Hematol*. 2016;23(4):354–361.

38. Pearson RCM, Funnell APW, Crossley M. The mammalian zinc finger transcription factor Krüppel-like factor 3 (KLF3/BKLF). *IUBMB Life*. 2011;63(2):86–93.

39. Buratowski S. Progression through the RNA polymerase II CTD cycle. *Mol. Cell*. 2009;36(4):541–546.

40. Chen FX, Smith ER, Shilatifard A. Born to run: control of transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* 2018;19(7):464–478.

41. Liu X, Bushnell DA, Kornberg RD. RNA polymerase II transcription: structure and mechanism. *Biochim. Biophys. Acta*. 2013;1829(1):2–8.

42. Schuettengruber B, Martinez A-M, Iovino N, Cavalli G. Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.*

2011;12(12):799–814.

43. Bradner JE, Hnisz D, Young RA. Transcriptional addiction in cancer. *Cell*. 2017;168(4):629–643.

44. Zhu J, Emerson SG. Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene*. 2002;21(21):3295–3313.

45. Antonchuk J, Sauvageau G, Humphries RK. HOXB4-induced expansion of adult hematopoietic stem cells ex vivo. *Cell*. 2002;109(1):39–45.

46. Björnsson JM, Larsson N, Brun ACM, et al. Reduced proliferative capacity of hematopoietic stem cells deficient in Hoxb3 and Hoxb4. *Mol. Cell. Biol.* 2003;23(11):3872–3883.

47. Rodrigues NP, Tipping AJ, Wang Z, Enver T. GATA-2 mediated regulation of normal hematopoietic stem/progenitor cell function, myelodysplasia and myeloid leukemia. *Int. J. Biochem. Cell Biol.* 2012;44(3):457–460.

48. Tsai FY, Orkin SH. Transcription factor GATA-2 is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood*. 1997;89(10):3636–3643.

49. Tsai FY, Keller G, Kuo FC, et al. An early haematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371(6494):221–226.

50. Lacombe J, Herblot S, Rojas-Sutterlin S, et al. Scl regulates the quiescence and the long-term competence of hematopoietic stem cells. *Blood*. 2010;115(4):792–803.

51. Mead PE, Deconinck AE, Huber TL, Orkin SH, Zon LI. Primitive erythropoiesis in the Xenopus embryo: the synergistic role of LMO-2, SCL and GATA-binding proteins. *Development*. 2001;128(12):2301–2308.

52. Iwasaki H, Mizuno S, Arinobu Y, et al. The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev.* 2006;20(21):3010–3021.

53. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. *Cell*. 2013;152(6):1237–1251.

54. Porcher C, Chagraoui H, Kristiansen MS. SCL/TAL1: a multifaceted regulator from blood development to disease. *Blood*. 2017;129(15):2051–2060.

55. Dey S, Curtis DJ, Jane SM, Brandt SJ. The TAL1/SCL transcription factor regulates cell cycle progression and proliferation in differentiating murine bone marrow monocyte precursors. *Mol. Cell. Biol.* 2010;30(9):2181–2192.

56. Nerlov C, Graf T. PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. *Genes Dev.* 1998;12(15):2403–2412.

57. Staber PB, Zhang P, Ye M, et al. Sustained PU.1 levels balance cell-cycle regulators to prevent exhaustion of adult hematopoietic stem cells. *Mol. Cell*. 2013;49(5):934–946.

58. Stopka T, Amanatullah DF, Papetti M, Skoultchi AI. PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. *EMBO J.* 2005;24(21):3712–3723.

59. Seshasayee D, Gaines P, Wojchowski DM. GATA-1 dominantly activates a program of erythroid gene expression in factor-dependent myeloid FDCW2 cells. *Mol. Cell. Biol.* 1998;18(6):3278–3288.

60. Friedman AD. C/EBPα in normal and malignant myelopoiesis. *Int J Hematol*. 2015;101(4):330–341.

61. Nerlov C. The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends Cell Biol.* 2007;17(7):318–324.

62. Zhang P, Iwasaki-Arai J, Iwasaki H, et al. Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. *Immunity*. 2004;21(6):853–863.

63. Laslo P, Pongubala JMR, Lancki DW, Singh H. Gene regulatory networks directing myeloid and lymphoid cell fates within the immune system. *Semin. Immunol.* 2008;20(4):228–235.

64. Laslo P, Spooner CJ, Warmflash A, et al. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*. 2006;126(4):755–766.

65. Rosenbauer F, Tenen DG. Transcription factors in myeloid development: balancing differentiation with transformation. *Nat. Rev. Immunol.* 2007;7(2):105–117.

66. Brettingham-Moore KH, Taberlay PC, Holloway AF. Interplay between Transcription Factors and the Epigenome: Insight from the Role of RUNX1 in Leukemia. *Front. Immunol.* 2015;6:499.

67. Nutt SL, Kee BL. The transcriptional regulation of B cell lineage commitment. *Immunity*. 2007;26(6):715–725.

68. Luyten A, Zang C, Liu XS, Shivdasani RA. Active enhancers are delineated de novo during hematopoiesis, with limited lineage fidelity among specified primary blood cells. *Genes Dev.* 2014;28(16):1827–1839.

69. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science*. 2017;357(6348):.

70. Moosavi A, Motevalizadeh Ardekani A. Role of epigenetics in biology and human diseases. *Iran Biomed J*. 2016;20(5):246–258.

71. Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Res.* 2011;21(4):564–578.

72. Wouters BJ, Delwel R. Epigenetics and approaches to targeted epigenetic therapy in acute myeloid leukemia. *Blood*. 2016;127(1):42–52.

73. Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* 2015;16(2):71–84.

74. Rose NR, Klose RJ. Understanding the relationship between DNA methylation and histone lysine methylation. *Biochim. Biophys. Acta*. 2014;1839(12):1362–1372.

75. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. *Exp Mol Med*. 2017;49(4):e324.

76. Wang J, Jia ST, Jia S. New Insights into the Regulation of Heterochromatin. *Trends Genet.* 2016;32(5):284–294.

77. Rao VK, Pal A, Taneja R. A drive in SUVs: From development to disease. *Epigenetics*. 2017;12(3):177–186.

78. Kimura H. Histone modifications for human epigenome analysis. *J. Hum. Genet.* 2013;58(7):439–445.

79. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000;403(6765):41–45.

80. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Res.* 2011;21(3):381–395.

81. Berndsen CE, Denu JM. Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* 2008;18(6):682–689.

82. Vakoc CR, Sachdeva MM, Wang H, Blobel GA. Profile of histone lysine methylation across transcribed mammalian chromatin. *Mol. Cell. Biol.* 2006;26(24):9185–9195.

83. Vedadi M, Blazer L, Eram MS, et al. Targeting human SET1/MLL family of proteins. *Protein Sci.* 2017;26(4):662–676.

84. Herviou L, Cavalli G, Cartron G, Klein B, Moreaux J. EZH2 in normal hematopoiesis and hematological malignancies. *Oncotarget*. 2016;7(3):2284–

2296.

85. Suzuki S, Murakami Y, Takahata S. H3K36 methylation state and associated silencing mechanisms. *Transcription*. 2017;8(1):26–31.

86. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol.* 2012;13(2):115–126.

87. D'Oto A, Tian Q-W, Davidoff AM, Yang J. Histone demethylases and their roles in cancer epigenetics. *J. Med. Oncol. Ther.* 2016;1(2):34–40.

88. Rotili D, Mai A. Targeting Histone Demethylases: A New Avenue for the Fight against Cancer. *Genes Cancer*. 2011;2(6):663–679.

89. Tsukada Y, Ishitani T, Nakayama KI. KDM7 is a dual demethylase for histone H3 Lys 9 and Lys 27 and functions in brain development. *Genes Dev.* 2010;24(5):432–437.

90. Kang J-Y, Kim J-Y, Kim K-B, et al. KDM2B is a histone H3K79 demethylase and induces transcriptional repression via sirtuin-1-mediated chromatin silencing. *FASEB J.* 2018;fj201800242R.

91. Qiang M, Denny A, Lieu M, Carreon S, Li J. Histone H3K9 modifications are a local chromatin event involved in ethanol-induced neuroadaptation of the NR2B gene. *Epigenetics*. 2011;6(9):1095–1104.

92. Marmorstein R, Zhou M-M. Writers and readers of histone acetylation: structure, mechanism, and inhibition. *Cold Spring Harb. Perspect. Biol.* 2014;6(7):a018762.

93. Dancy BM, Cole PA. Protein lysine acetylation by p300/CBP. *Chem. Rev.* 2015;115(6):2419–2452.

94. Zhu B, Zheng Y, Pham A-D, et al. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Mol. Cell*. 2005;20(4):601–611.

95. Hahn MA, Dickson K-A, Jackson S, et al. The tumor suppressor CDC73 interacts with the ring finger proteins RNF20 and RNF40 and is required for the maintenance of histone 2B monoubiquitination. *Hum. Mol. Genet.* 2012;21(3):559–568.

96. Lang G, Bonnet J, Umlauf D, et al. The tightly controlled deubiquitination activity of the human SAGA complex differentially modifies distinct gene regulatory elements. *Mol. Cell. Biol.* 2011;31(18):3734–3744.

97. Morgan MT, Haj-Yahya M, Ringel AE, et al. Structural basis for histone H2B deubiquitination by the SAGA DUB module. *Science*. 2016;351(6274):725–728.

98. Li X, Seidel CW, Szerszen LT, et al. Enzymatic modules of the SAGA chromatin-modifying complex play distinct roles in Drosophila gene expression and development. *Genes Dev.* 2017;31(15):1588–1600.

99. Jaehning JA. The Paf1 complex: platform or player in RNA polymerase II transcription? *Biochim. Biophys. Acta*. 2010;1799(5–6):379–388.

100. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. USA*. 2010;107(50):21931–21936.

101. Calo E, Wysocka J. Modification of enhancer chromatin: what, how, and why? *Mol. Cell*. 2013;49(5):825–837.

102. Tachibana M, Ueda J, Fukuda M, et al. Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev.* 2005;19(7):815–826.

103. Mallo M, Alonso CR. The regulation of Hox gene expression during animal development. *Development*. 2013;140(19):3951–3963.

104. Argiropoulos B, Humphries RK. Hox genes in hematopoiesis and leukemogenesis. *Oncogene*. 2007;26(47):6766–6776.

105. Lewis EB. A gene complex controlling segmentation in Drosophila. *Nature*. 1978;276(5688):565–570.

106. Schuettengruber B, Bourbon H-M, Di Croce L, Cavalli G. Genome regulation by polycomb and trithorax: 70 years and counting. *Cell*. 2017;171(1):34–57.

107. Kennison JA, Tamkun JW. Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. *Proc. Natl. Acad. Sci. USA*. 1988;85(21):8136–8140.

108. Djabali M, Selleri L, Parry P, et al. A trithorax-like gene is interrupted by chromosome 11q23 translocations in acute leukaemias. *Nat. Genet.* 1992;2(2):113–118.

109. Kanduri M, Sander B, Ntoufa S, et al. A key role for EZH2 in epigenetic silencing of HOX genes in mantle cell lymphoma. *Epigenetics*. 2013;8(12):1280–1288.

110. Hanson RD, Hess JL, Yu BD, et al. Mammalian Trithorax and polycomb-group homologues are antagonistic regulators of homeotic development. *Proc. Natl. Acad. Sci. USA*. 1999;96(25):14372–14377.

111. Aplan PD. Chromosomal translocations involving the MLL gene: molecular mechanisms. *DNA Repair (Amst.)*. 2006;5(9–10):1265–1272.

112. Xu F, Liu L, Chang C-K, et al. Genomic loss of EZH2 leads to epigenetic modifications and overexpression of the HOX gene clusters in myelodysplastic syndrome. *Oncotarget*. 2016;7(7):8119–8130.

113. Calvi LM, Link DC. The hematopoietic stem cell niche in homeostasis and disease. *Blood*. 2015;126(22):2443–2451.

114. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391–2405.

115. Bonnet D, Dick JE. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* 1997;3(7):730–737.

116. Krivtsov AV, Twomey D, Feng Z, et al. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature*. 2006;442(7104):818–822.

117. Ye M, Zhang H, Yang H, et al. Hematopoietic differentiation is required for initiation of acute myeloid leukemia. *Cell Stem Cell*. 2015;17(5):611–623.

118. Guan Y, Hogge DE. Proliferative status of primitive hematopoietic progenitors from patients with acute myelogenous leukemia (AML). *Leukemia*. 2000;14(12):2135–2141.

119. Lohse I, Statz-Geary K, Brothers SP, Wahlestedt C. Precision medicine in the treatment stratification of AML patients: challenges and progress. *Oncotarget*. 2018;9(102):37790–37797.

120. Sun C-C, Li S-J, Chen Z-L, et al. Expression and Prognosis Analyses of Runt-Related Transcription Factor Family in Human Leukemia. *Mol. Ther. Oncolytics*. 2019;12:103–111.

121. Goyama S, Schibler J, Cunningham L, et al. Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. *J. Clin. Invest.* 2013;

122. Wang H, Chu T-T, Han S-Y, et al. FLT3-ITD and CEBPA Mutations Predict Prognosis in Acute Myelogenous Leukemia Irrespective of Hematopoietic Stem Cell Transplantation. *Biol. Blood Marrow Transplant.* 2018;

123. Roe J-S, Vakoc CR. C/EBPα: critical at the origin of leukemic transformation. *J. Exp. Med.* 2014;211(1):1–4.

124. Calvo KR, Knoepfler PS, Sykes DB, Pasillas MP, Kamps MP. Meis1a suppresses differentiation by G-CSF and promotes proliferation by SCF: potential mechanisms of cooperativity with Hoxa9 in myeloid leukemia. *Proc. Natl. Acad. Sci. USA*. 2001;98(23):13120–13125.

125. Fujino T. Inhibition of myeloid differentiation by Hoxa9, Hoxb8, and Meis

homeobox genes. *Exp. Hematol.* 2001;29(7):856–863.

126. Zhang W, Zhao C, Zhao J, et al. Inactivation of PBX3 and HOXA9 by down-regulating H3K79 methylation represses NPM1-mutated leukemic cell survival. *Theranostics*. 2018;8(16):4359–4371.

127. Zhou J, Wu J, Li B, et al. PU.1 is essential for MLL leukemia partially via crosstalk with the MEIS/HOX pathway. *Leukemia*. 2014;28(7):1436–1448.

128. Collins C, Wang J, Miao H, et al. C/EBPα is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis. *Proc. Natl. Acad. Sci. USA*. 2014;111(27):9899–9904.

129. Chapuis N, Tamburini J, Cornillet-Lefebvre P, et al. Autocrine IGF-1/IGF-1R signaling is responsible for constitutive PI3K/Akt activation in acute myeloid leukemia: therapeutic value of neutralizing anti-IGF-1R antibody. *Haematologica*. 2010;95(3):415–423.

130. Chen D, Yoon J-B, Gu W. Reactivating the ARF-p53 axis in AML cells by targeting ULF. *Cell Cycle*. 2010;9(15):2946–2951.

131. Brunetti L, Gundry MC, Sorcini D, et al. Mutant NPM1 Maintains the Leukemic State through HOX Expression. *Cancer Cell*. 2018;34(3):499–512.e9.

132. Abdel-Wahab O, Adli M, LaFave LM, et al. ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell*. 2012;22(2):180–193.

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

134. Feinberg AP, Koldobskiy MA, Göndör A. Epigenetic modulators, modifiers and mediators in cancer aetiology and progression. *Nat. Rev. Genet.* 2016;17(5):284–299.

135. Baylin SB, Jones PA. A decade of exploring the cancer epigenome - biological and translational implications. *Nat. Rev. Cancer*. 2011;11(10):726–734.

136. Waldmann T, Schneider R. Targeting histone modifications--epigenetics in cancer. *Curr. Opin. Cell Biol.* 2013;25(2):184–189.

137. Audia JE, Campbell RM. Histone modifications and cancer. *Cold Spring Harb. Perspect. Biol.* 2016;8(4):a019521.

138. Kurdistani SK. Histone modifications as markers of cancer prognosis: a cellular view. *Br. J. Cancer*. 2007;97(1):1–5.

139. Boila LD, Chatterjee SS, Banerjee D, Sengupta A. KDM6 and KDM4 histone lysine

demethylases emerge as molecular therapeutic targets in human acute myeloid leukemia. *Exp. Hematol.* 2018;58:44–51.e7.

140. Andricovich J, Kai Y, Tzatsos A. Lysine-specific histone demethylases in normal and malignant hematopoiesis. *Exp. Hematol.* 2016;44(9):778–782.

141. Gozdecka M, Meduri E, Mazan M, et al. UTX-mediated enhancer and chromatin remodeling suppresses myeloid leukemogenesis through noncatalytic inverse regulation of ETS and GATA programs. *Nat. Genet.* 2018;50(6):883–894.

142. Greif PA, Hartmann L, Vosberg S, et al. Evolution of cytogenetically normal acute myeloid leukemia during therapy and relapse: an exome sequencing study of 50 patients. *Clin. Cancer Res.* 2018;24(7):1716–1726.

143. Mozzetta C, Boyarchuk E, Pontis J, Ait-Si-Ali S. Sound of silence: the properties and functions of repressive Lys methyltransferases. *Nat. Rev. Mol. Cell Biol.* 2015;16(8):499–513.

144. Khan SN, Jankowska AM, Mahfouz R, et al. Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies. *Leukemia*. 2013;27(6):1301–1309.

145. Lu R, Wang GG. Pharmacologic targeting of chromatin modulators as therapeutics of acute myeloid leukemia. *Front. Oncol.* 2017;7:241.

146. Thiel AT, Feng Z, Pant DK, Chodosh LA, Hua X. The trithorax protein partner menin acts in tandem with EZH2 to suppress C/EBPα and differentiation in MLL-AF9 leukemia. *Haematologica*. 2013;98(6):918–927.

147. Allen MD, Grummitt CG, Hilcenko C, et al. Solution structure of the nonmethyl-CpG-binding CXXC domain of the leukaemia-associated MLL histone methyltransferase. *EMBO J.* 2006;25(19):4503–4512.

148. Cierpicki T, Risner LE, Grembecka J, et al. Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. *Nat. Struct. Mol. Biol.* 2010;17(1):62–68.

149. Birke M, Schreiner S, García-Cuéllar M-P, et al. The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. *Nucleic Acids Res.* 2002;30(4):958–965.

150. Chang P-Y, Hom RA, Musselman CA, et al. Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. *J. Mol. Biol.* 2010;400(2):137–144.

151. Wang Z, Song J, Milne TA, et al. Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. *Cell*. 2010;141(7):1183–1194.

152. Yu BD, Hanson RD, Hess JL, Horning SE, Korsmeyer SJ. MLL, a mammalian trithorax-group gene, functions as a transcriptional maintenance factor in morphogenesis. *Proc. Natl. Acad. Sci. USA*. 1998;95(18):10632–10636.

153. McMahon KA, Hiew SY-L, Hadjur S, et al. Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. *Cell Stem Cell*. 2007;1(3):338–345.

154. Ernst P, Mabon M, Davidson AJ, Zon LI, Korsmeyer SJ. An Mll-dependent Hox program drives hematopoietic progenitor expansion. *Curr. Biol.* 2004;14(22):2063–2069.

155. Mishra BP, Zaffuto KM, Artinger EL, et al. The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis. *Cell Rep.* 2014;7(4):1239–1247.

156. Winters AC, Bernt KM. MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches. *Front. Pediatr.* 2017;5:4.

157. Muntean AG. Mechanisms of mixed-lineage leukemia. *Int J Hematol Oncol*. 2013;2(3):207–217.

158. Steinhilber D, Marschalek R. How to effectively treat acute leukemia patients bearing MLL-rearrangements ? *Biochem. Pharmacol.* 2018;147:183–190.

159. Ziemin-van der Poel S, McCabe NR, Gill HJ, et al. Identification of a gene, MLL, that spans the breakpoint in 11q23 translocations associated with human leukemias. *Proc. Natl. Acad. Sci. USA*. 1991;88(23):10735–10739.

160. Meyer C, Hofmann J, Burmeister T, et al. The MLL recombinome of acute leukemias in 2013. *Leukemia*. 2013;27(11):2165–2176.

161. Whitman SP, Liu S, Vukosavljevic T, et al. The MLL partial tandem duplication: evidence for recessive gain-of-function in acute myeloid leukemia identifies a novel patient subgroup for molecular-targeted therapy. *Blood*. 2005;106(1):345–352.

162. Kumar AR, Li Q, Hudson WA, et al. A role for MEIS1 in MLL-fusion gene leukemia. *Blood*. 2009;113(8):1756–1758.

163. Yokoyama A, Somervaille TCP, Smith KS, et al. The menin tumor suppressor protein is an essential oncogenic cofactor for MLL-associated leukemogenesis. *Cell*. 2005;123(2):207–218.

164. Yokoyama A, Cleary ML. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell*. 2008;14(1):36–46.

165. Mueller D, García-Cuéllar M-P, Bach C, et al. Misguided transcriptional elongation causes mixed lineage leukemia. *PLoS Biol.* 2009;7(11):e1000249.

166. Lin C, Smith ER, Takahashi H, et al. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell*. 2010;37(3):429–437.

167. Yokoyama A, Lin M, Naresh A, Kitabayashi I, Cleary ML. A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription. *Cancer Cell*. 2010;17(2):198–212.

168. Filippakopoulos P, Qi J, Picaud S, et al. Selective inhibition of BET bromodomains. *Nature*. 2010;468(7327):1067–1073.

169. Dawson MA, Prinjha RK, Dittmann A, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. *Nature*. 2011;478(7370):529–533.

170. Yu M, Yang W, Ni T, et al. RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science*. 2015;350(6266):1383–1386.

171. Chen FX, Xie P, Collings CK, et al. PAF1 regulation of promoter-proximal pause release via enhancer activation. *Science*. 2017;357(6357):1294–1298.

172. Chen FX, Woodfin AR, Gardini A, et al. PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. *Cell*. 2015;162(5):1003–1015.

173. Shi X, Chang M, Wolf AJ, et al. Cdc73p and Paf1p are found in a novel RNA polymerase II-containing complex distinct from the Srbp-containing holoenzyme. *Mol. Cell. Biol.* 1997;17(3):1160–1169.

174. Yart A, Gstaiger M, Wirbelauer C, et al. The HRPT2 tumor suppressor gene product parafibromin associates with human PAF1 and RNA polymerase II. *Mol. Cell. Biol.* 2005;25(12):5052–5060.

175. Rozenblatt-Rosen O, Hughes CM, Nannepaga SJ, et al. The parafibromin tumor suppressor protein is part of a human Paf1 complex. *Mol. Cell. Biol.* 2005;25(2):612–620.

176. Wade PA, Werel W, Fentzke RC, et al. A novel collection of accessory factors associated with yeast RNA polymerase II. *Protein Expr. Purif.* 1996;8(1):85–90.

177. Mueller CL, Jaehning JA. Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol. Cell. Biol.* 2002;22(7):1971–1980.

178. Chu X, Qin X, Xu H, et al. Structural insights into Paf1 complex assembly and histone binding. *Nucleic Acids Res.* 2013;41(22):10619–10629.

179. Zhu B, Mandal SS, Pham A-D, et al. The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes Dev.*

2005;19(14):1668–1673.

180. Amrich CG, Davis CP, Rogal WP, et al. Cdc73 subunit of Paf1 complex contains C-terminal Ras-like domain that promotes association of Paf1 complex with chromatin. *J. Biol. Chem.* 2012;287(14):10863–10875.

181. Xu Y, Bernecky C, Lee C-T, et al. Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex. *Nat. Commun.* 2017;8:15741.

182. Qiu H, Hu C, Wong C-M, Hinnebusch AG. The Spt4p subunit of yeast DSIF stimulates association of the Paf1 complex with elongating RNA polymerase II. *Mol. Cell. Biol.* 2006;26(8):3135–3148.

183. Hartzog GA, Fu J. The Spt4-Spt5 complex: a multi-faceted regulator of transcription elongation. *Biochim. Biophys. Acta*. 2013;1829(1):105–115.

184. He N, Chan CK, Sobhian B, et al. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *Proc. Natl. Acad. Sci. USA*. 2011;108(36):E636-45.

185. Pavri R, Zhu B, Li G, et al. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell*. 2006;125(4):703–717.

186. Bedard LG, Dronamraju R, Kerschner JL, et al. Quantitative Analysis of Dynamic Protein Interactions during Transcription Reveals a Role for Casein Kinase II in Polymerase-associated Factor (PAF) Complex Phosphorylation and Regulation of Histone H2B Monoubiquitylation. *J. Biol. Chem.* 2016;291(26):13410–13420.

187. Ng HH, Dole S, Struhl K. The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *J. Biol. Chem.* 2003;278(36):33625–33628.

188. Krogan NJ, Dover J, Wood A, et al. The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol. Cell*. 2003;11(3):721–729.

189. Van Oss SB, Shirra MK, Bataille AR, et al. The Histone Modification Domain of Paf1 Complex Subunit Rtf1 Directly Stimulates H2B Ubiquitylation through an Interaction with Rad6. *Mol. Cell*. 2016;64(4):815–825.

190. Piro AS, Mayekar MK, Warner MH, Davis CP, Arndt KM. Small region of Rtf1 protein can substitute for complete Paf1 complex in facilitating global histone H2B ubiquitylation in yeast. *Proc. Natl. Acad. Sci. USA*. 2012;109(27):10837–10842.

191. Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J. Biol. Chem.*

2003;278(37):34739–34742.

192. Kim J, Roeder RG. Direct Bre1-Paf1 complex interactions and RING finger-independent Bre1-Rad6 interactions mediate histone H2B ubiquitylation in yeast. *J. Biol. Chem.* 2009;284(31):20582–20592.

193. Gaertner B, Zeitlinger J. RNA polymerase II pausing during development. *Development*. 2014;141(6):1179–1183.

194. Penheiter KL, Washburn TM, Porter SE, Hoffman MG, Jaehning JA. A posttranscriptional role for the yeast Paf1-RNA polymerase II complex is revealed by identification of primary targets. *Mol. Cell*. 2005;20(2):213–223.

195. Mueller CL, Porter SE, Hoffman MG, Jaehning JA. The Paf1 complex has functions independent of actively transcribing RNA polymerase II. *Mol. Cell*. 2004;14(4):447–456.

196. Shi Y, Di Giammartino DC, Taylor D, et al. Molecular architecture of the human pre-mRNA 3' processing complex. *Mol. Cell*. 2009;33(3):365–376.

197. Rozenblatt-Rosen O, Nagaike T, Francis JM, et al. The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3' mRNA processing factors. *Proc. Natl. Acad. Sci. USA*. 2009;106(3):755–760.

198. Porter SE, Washburn TM, Chang M, Jaehning JA. The yeast pafl-rNA polymerase II complex is required for full expression of a subset of cell cycle-regulated genes. *Eukaryotic Cell*. 2002;1(5):830–842.

199. Tan M-H, Morrison C, Wang P, et al. Loss of parafibromin immunoreactivity is a distinguishing feature of parathyroid carcinoma. *Clin. Cancer Res.* 2004;10(19):6629–6637.

200. Simonds WF, Robbins CM, Agarwal SK, et al. Familial isolated hyperparathyroidism is rarely caused by germline mutation in HRPT2, the gene for the hyperparathyroidism-jaw tumor syndrome. *J. Clin. Endocrinol. Metab.* 2004;89(1):96–102.

201. Yang Y-J, Han J-W, Youn H-D, Cho E-J. The tumor suppressor, parafibromin, mediates histone H3 K9 methylation for cyclin D1 repression. *Nucleic Acids Res.* 2010;38(2):382–390.

202. Rather MI, Swamy S, Gopinath KS, Kumar A. Transcriptional repression of tumor suppressor CDC73, encoding an RNA polymerase II interactor, by Wilms tumor 1 protein (WT1) promotes cell proliferation: implication for cancer therapeutics. *J. Biol. Chem.* 2014;289(2):968–976.

203. Lin L, Zhang J-H, Panicker LM, Simonds WF. The parafibromin tumor suppressor protein inhibits cell proliferation by repression of the c-myc proto-oncogene. *Proc.*

Natl. Acad. Sci. USA. 2008;105(45):17420–17425.

204. Zhao J, Yart A, Frigerio S, et al. Sporadic human renal tumors display frequent allelic imbalances and novel mutations of the HRPT2 gene. *Oncogene*. 2007;26(23):3440–3449.

205. Rather MI, Nagashri MN, Swamy SS, Gopinath KS, Kumar A. Oncogenic microRNA-155 down-regulates tumor suppressor CDC73 and promotes oral squamous cell carcinoma cell proliferation: implications for cancer therapeutics. *J. Biol. Chem.* 2013;288(1):608–618.

206. Shen D-F, Liu X, Yang X-F, et al. The roles of parafibromin expression in ovarian epithelial carcinomas: a marker for differentiation and prognosis and a target for gene therapy. *Tumour Biol.* 2016;37(3):2909–2924.

207. Selvarajan S, Sii LH, Lee A, et al. Parafibromin expression in breast cancer: a novel marker for prognostication? *J. Clin. Pathol.* 2008;61(1):64–67.

208. Takahashi A, Tsutsumi R, Kikuchi I, et al. SHP2 tyrosine phosphatase converts parafibromin/Cdc73 from a tumor suppressor to an oncogenic driver. *Mol. Cell*. 2011;43(1):45–56.

209. Kikuchi I, Takahashi-Kanemitsu A, Sakiyama N, et al. Dephosphorylated parafibromin is a transcriptional coactivator of the Wnt/Hedgehog/Notch pathways. *Nat. Commun.* 2016;7:12887.

210. Mosimann C, Hausmann G, Basler K. Parafibromin/Hyrax activates Wnt/Wg target gene transcription by direct association with beta-catenin/Armadillo. *Cell*. 2006;125(2):327–341.

211. Dey P, Rachagani S, Vaz AP, Ponnusamy MP, Batra SK. PD2/Paf1 depletion in pancreatic acinar cells promotes acinar-to-ductal metaplasia. *Oncotarget*. 2014;5(12):4480–4491.

212. Moniaux N, Nemos C, Schmied BM, et al. The human homologue of the RNA polymerase II-associated factor 1 (hPaf1), localized on the 19q13 amplicon, is associated with tumorigenesis. *Oncogene*. 2006;25(23):3247–3257.

213. Vaz AP, Deb S, Rachagani S, et al. Overexpression of PD2 leads to increased tumorigenicity and metastasis in pancreatic ductal adenocarcinoma. *Oncotarget*. 2016;7(3):3317–3331.

214. Zhi X, Giroux-Leprieur E, Wislez M, et al. Human RNA polymerase II associated factor 1 complex promotes tumorigenesis by activating c-MYC transcription in non-small cell lung cancer. *Biochem. Biophys. Res. Commun.* 2015;465(4):685–690.

215. Karmakar S, Seshacharyulu P, Lakshmanan I, et al. hPaf1/PD2 interacts with OCT3/4 to promote self-renewal of ovarian cancer stem cells. *Oncotarget*.

2017;8(9):14806–14820.

216. Vaz AP, Ponnusamy MP, Rachagani S, et al. Novel role of pancreatic differentiation 2 in facilitating self-renewal and drug resistance of pancreatic cancer stem cells. *Br. J. Cancer*. 2014;111(3):486–496.

217. Zeng H, Lu L, Chan NT, et al. Systematic identification of Ctr9 regulome in ERα-positive breast cancer. *BMC Genomics*. 2016;17(1):902.

218. Zeng H, Xu W. Ctr9, a key subunit of PAFc, affects global estrogen signaling and drives ERα-positive breast tumorigenesis. *Genes Dev.* 2015;29(20):2153–2167.

219. Hanks S, Perdeaux ER, Seal S, et al. Germline mutations in the PAF1 complex gene CTR9 predispose to Wilms tumour. *Nat. Commun.* 2014;5:4398.

220. Tarkkanen M, Larramendy ML, Böhling T, et al. Malignant fibrous histiocytoma of bone: analysis of genomic imbalances by comparative genomic hybridisation and C-MYC expression by immunohistochemistry. *Eur. J. Cancer*. 2006;42(8):1172–1180.

221. Camps J, Armengol G, del Rey J, et al. Genome-wide differences between microsatellite stable and unstable colorectal tumors. *Carcinogenesis*. 2006;27(3):419–428.

222. Da Costa GG, Gomig THB, Kaviski R, et al. Comparative proteomics of tumor and paired normal breast tissue highlights potential biomarkers in breast cancer. *Cancer Genomics Proteomics*. 2015;12(5):251–261.

223. Tenney K, Gerber M, Ilvarsonn A, et al. Drosophila Rtf1 functions in histone methylation, gene expression, and Notch signaling. *Proc. Natl. Acad. Sci. USA*. 2006;103(32):11970–11974.

224. Serio J, Ropa J, Chen W, et al. The PAF complex regulation of Prmt5 facilitates the progression and maintenance of MLL fusion leukemia. *Oncogene*. 2018;37(4):450–460.

225. Hetzner K, Garcia-Cuellar M-P, Büttner C, Slany RK. The interaction of ENL with PAF1 mitigates polycomb silencing and facilitates murine leukemogenesis. *Blood*. 2018;131(6):662–673.

226. Chong PSY, Zhou J, Cheong L-L, et al. LEO1 is regulated by PRL-3 and mediates its oncogenic properties in acute myelogenous leukemia. *Cancer Res.* 2014;74(11):3043–3053.

227. Chong PSY, Zhou J, Chooi J-Y, et al. Non-canonical activation of β-catenin by PRL-3 phosphatase in acute myeloid leukemia. *Oncogene*. 2018;

228. Chen H, Shi N, Gao Y, et al. Crystallographic analysis of the conserved C-terminal

domain of transcription factor Cdc73 from Saccharomyces cerevisiae reveals a GTPase-like fold. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* 2012;68(Pt 8):953–959.

229. Bilodeau S, Kagey MH, Frampton GM, Rahl PB, Young RA. SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. *Genes Dev.* 2009;23(21):2484–2489.

230. Trojer P, Reinberg D. Facultative heterochromatin: is there a distinctive molecular signature? *Mol. Cell*. 2007;28(1):1–13.

231. Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129(4):823–837.

232. Peters AH, O'Carroll D, Scherthan H, et al. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell*. 2001;107(3):323–337.

233. Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 2005;6(11):838–849.

234. Fritsch L, Robin P, Mathieu JRR, et al. A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex. *Mol. Cell*. 2010;37(1):46–56.

235. Saksouk N, Simboeck E, Déjardin J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin*. 2015;8:3.

236. Timms RT, Tchasovnikarova IA, Antrobus R, Dougan G, Lehner PJ. ATF7IP-Mediated Stabilization of the Histone Methyltransferase SETDB1 Is Essential for Heterochromatin Formation by the HUSH Complex. *Cell Rep.* 2016;17(3):653–659.

237. Wang T, Xu C, Liu Y, et al. Crystal structure of the human SUV39H1 chromodomain and its recognition of histone H3K9me2/3. *PLoS One*. 2012;7(12):e52977.

238. O'Carroll D, Scherthan H, Peters AH, et al. Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. *Mol. Cell. Biol.* 2000;20(24):9423–9433.

239. Aagaard L, Laible G, Selenko P, et al. Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3-9 encode centromere-associated proteins which complex with the heterochromatin component M31. *EMBO J.* 1999;18(7):1923–1938.

240. Ogawa H, Ishiguro K-I, Gaubatz S, Livingston DM, Nakatani Y. A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells.

Science. 2002;296(5570):1132–1136.

241. Tachibana M, Sugimoto K, Fukushima T, Shinkai Y. Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J. Biol. Chem.* 2001;276(27):25309–25317.

242. Tachibana M, Sugimoto K, Nozaki M, et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev.* 2002;16(14):1779–1791.

243. Tachibana M, Matsumura Y, Fukuda M, Kimura H, Shinkai Y. G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO J.* 2008;27(20):2681–2690.

244. Yang L, Xia L, Wu DY, et al. Molecular cloning of ESET, a novel histone H3-specific methyltransferase that interacts with ERG transcription factor. *Oncogene*. 2002;21(1):148–152.

245. Falandry C, Fourel G, Galy V, et al. CLLD8/KMT1F is a lysine methyltransferase that is important for chromosome segregation. *J. Biol. Chem.* 2010;285(26):20234–20241.

246. Kim K-C, Geng L, Huang S. Inactivation of a histone methyltransferase by mutations in human cancers. *Cancer Res.* 2003;63(22):7619–7623.

247. Pinheiro I, Margueron R, Shukeir N, et al. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell*. 2012;150(5):948–960.

248. Eom GH, Kim K, Kim S-M, et al. Histone methyltransferase PRDM8 regulates mouse testis steroidogenesis. *Biochem. Biophys. Res. Commun.* 2009;388(1):131–136.

249. Wong C-M, Wei L, Law C-T, et al. Up-regulation of histone methyltransferase SETDB1 by multiple mechanisms in hepatocellular carcinoma promotes cancer metastasis. *Hepatology*. 2016;63(2):474–487.

250. Chen B, Wang J, Wang J, et al. A regulatory circuitry comprising TP53, miR-29 family, and SETDB1 in non-small cell lung cancer. *Biosci. Rep.* 2018;38(5):.

251. Ceol CJ, Houvras Y, Jane-Valbuena J, et al. The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. *Nature*. 2011;471(7339):513–517.

252. Regina C, Compagnone M, Peschiaroli A, et al. Setdb1, a novel interactor of ΔNp63, is involved in breast tumorigenesis. *Oncotarget*. 2016;7(20):28836–28848.

253. Du D, Katsuno Y, Meyer D, et al. Smad3-mediated recruitment of the methyltransferase SETDB1/ESET controls Snail1 expression and epithelial-mesenchymal transition. *EMBO Rep.* 2018;19(1):135–155.

254. Hua K-T, Wang M-Y, Chen M-W, et al. The H3K9 methyltransferase G9a is a marker of aggressive ovarian cancer that promotes peritoneal metastasis. *Mol. Cancer*. 2014;13:189.

255. Dong C, Wu Y, Yao J, et al. G9a interacts with Snail and is critical for Snail-mediated E-cadherin repression in human breast cancer. *J. Clin. Invest.* 2012;122(4):1469–1486.

256. Chen M-W, Hua K-T, Kao H-J, et al. H3K9 histone methyltransferase G9a promotes lung cancer invasion and metastasis by silencing the cell adhesion molecule Ep-CAM. *Cancer Res.* 2010;70(20):7830–7840.

257. Rada M, Vasileva E, Lezina L, et al. Human EHMT2/G9a activates p53 through methylation-independent mechanism. *Oncogene*. 2017;36(7):922–932.

258. Kim G, Kim J-Y, Lim S-C, et al. SUV39H1/DNMT3A-dependent methylation of the RB1 promoter stimulates PIN1 expression and melanoma development. *FASEB J.* 2018;fj201700645RRRRR.

259. Wang Q, Ma S, Song N, et al. Stabilization of histone demethylase PHF8 by USP7 promotes breast carcinogenesis. *J. Clin. Invest.* 2016;126(6):2205–2220.

260. Zhang L, Hui T-L, Wei Y-X, et al. The expression and biological function of the PHF2 gene in breast cancer. *RSC Adv.* 2018;8(69):39520–39528.

261. Berry WL, Janknecht R. KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. *Cancer Res.* 2013;73(10):2936–2942.

262. Mallette FA, Richard S. JMJD2A promotes cellular transformation by blocking cellular senescence through transcriptional repression of the tumor suppressor CHD5. *Cell Rep.* 2012;2(5):1233–1243.

263. Ueda J, Ho JC, Lee KL, et al. The hypoxia-inducible epigenetic regulators Jmjd1a and G9a provide a mechanistic link between angiogenesis and tumor growth. *Mol. Cell. Biol.* 2014;34(19):3702–3720.

264. Koide S, Oshima M, Takubo K, et al. Setdb1 maintains hematopoietic stem and progenitor cells by restricting the ectopic activation of nonhematopoietic genes. *Blood*. 2016;128(5):638–649.

265. Cho S, Park JS, Kang Y-K. Dual functions of histone-lysine N-methyltransferase Setdb1 protein at promyelocytic leukemia-nuclear body (PML-NB): maintaining PML-NB structure and regulating the expression of its associated genes. *J. Biol. Chem.* 2011;286(47):41115–41124.

266. Pasquarella A, Ebert A, Pereira de Almeida G, et al. Retrotransposon derepression leads to activation of the unfolded protein response and apoptosis in pro-B cells. *Development*. 2016;143(10):1788–1799.

267. Cuellar TL, Herzner A-M, Zhang X, et al. Silencing of retrotransposons by SETDB1 inhibits the interferon response in acute myeloid leukemia. *J. Cell Biol.* 2017;216(11):3535–3549.

268. Mabuchi H, Fujii H, Calin G, et al. Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. *Cancer Res.* 2001;61(7):2870–2877.

269. Lehnertz B, Pabst C, Su L, et al. The methyltransferase G9a regulates HoxA9-dependent transcription in AML. *Genes Dev.* 2014;28(4):317–327.

270. Loh SW, Ng WL, Yeo KS, Lim Y-Y, Ea C-K. Inhibition of euchromatic histone methyltransferase 1 and 2 sensitizes chronic myeloid leukemia cells to interferon treatment. *PLoS One*. 2014;9(7):e103915.

271. Savickiene J, Treigyte G, Stirblyte I, Valiuliene G, Navakauskiene R. Euchromatic histone methyltransferase 2 inhibitor, BIX-01294, sensitizes human promyelocytic leukemia HL-60 and NB4 cells to growth inhibition and differentiation. *Leuk. Res.* 2014;38(7):822–829.

272. Pappano WN, Guo J, He Y, et al. The Histone Methyltransferase Inhibitor A-366 Uncovers a Role for G9a/GLP in the Epigenetics of Leukemia. *PLoS One*. 2015;10(7):e0131716.

273. Salzberg AC, Harris-Becker A, Popova EY, et al. Genome-wide mapping of histone H3K9me2 in acute myeloid leukemia reveals large chromosomal domains associated with massive gene silencing and sites of genome instability. *PLoS One*. 2017;12(3):e0173723.

274. Alves-Silva JC, de Carvalho JL, Rabello DA, et al. GLP overexpression is associated with poor prognosis in Chronic Lymphocytic Leukemia and its inhibition induces leukemic cell death. *Invest. New Drugs*. 2018;36(5):955–960.

275. Son H-J, Kim J-Y, Hahn Y, Seo S-B. Negative regulation of JAK2 by H3K9 methyltransferase G9a in leukemia. *Mol. Cell. Biol.* 2012;32(18):3681–3694.

276. Djeghloul D, Kuranda K, Kuzniak I, et al. Age-Associated Decrease of the Histone Methyltransferase SUV39H1 in HSC Perturbs Heterochromatin and B Lymphoid Differentiation. *Stem Cell Rep.* 2016;6(6):970–984.

277. Carvalho Alves-Silva J, do Amaral Rabello D, Oliveira Bravo M, et al. Aberrant levels of SUV39H1 and SUV39H2 methyltransferase are associated with genomic instability in chronic lymphocytic leukemia. *Environ Mol Mutagen*. 2017;58(9):654–

661.

278. Vajen B, Modlich U, Schienke A, et al. Histone methyltransferase Suv39h1 deficiency prevents Myc-induced chromosomal instability in murine myeloid leukemias. *Genes. Chromosomes Cancer*. 2013;52(4):423–430.

279. Carbone R, Botrugno OA, Ronzoni S, et al. Recruitment of the histone methyltransferase SUV39H1 and its role in the oncogenic properties of the leukemia-associated PML-retinoic acid receptor fusion protein. *Mol. Cell. Biol.* 2006;26(4):1288–1296.

280. Chaib H, Nebbioso A, Prebet T, et al. Anti-leukemia activity of chaetocin via death receptor-dependent apoptosis and dual modulation of the histone methyl-transferase SUV39H1. *Leukemia*. 2012;26(4):662–674.

281. Lai YS, Chen JY, Tsai HJ, Chen TY, Hung WC. The SUV39H1 inhibitor chaetocin induces differentiation and shows synergistic cytotoxicity with other epigenetic drugs in acute myeloid leukemia cells. *Blood Cancer J.* 2015;5:e313.

282. Tran HTT, Kim HN, Lee I-K, et al. Improved therapeutic effect against leukemia by a combination of the histone methyltransferase inhibitor chaetocin and the histone deacetylase inhibitor trichostatin A. *J. Korean Med. Sci.* 2013;28(2):237–246.

283. Lakshmikuttyamma A, Scott SA, DeCoteau JF, Geyer CR. Reexpression of epigenetically silenced AML tumor suppressor genes by SUV39H1 inhibition. *Oncogene*. 2010;29(4):576–588.

284. Han X, Han Y, Zheng Y, et al. Chaetocin induces apoptosis in human melanoma cells through the generation of reactive oxygen species and the intrinsic mitochondrial pathway, and exerts its anti-tumor activity in vivo. *PLoS One*. 2017;12(4):e0175950.

285. Isham CR, Tibodeau JD, Jin W, et al. Chaetocin: a promising new antimyeloma agent with in vitro and in vivo activity mediated via imposition of oxidative stress. *Blood*. 2007;109(6):2579–2588.

286. Cherblanc FL, Chapman KL, Brown R, Fuchter MJ. Chaetocin is a nonspecific inhibitor of histone lysine methyltransferases. *Nat. Chem. Biol.* 2013;9(3):136–137.

287. Mutonga M, Tamura K, Malnassy G, et al. Targeting Suppressor of Variegation 3-9 Homologue 2 (SUV39H2) in Acute Lymphoblastic Leukemia (ALL). *Transl Oncol*. 2015;8(5):368–375.

288. Piao L, Suzuki T, Dohmae N, Nakamura Y, Hamamoto R. SUV39H2 methylates and stabilizes LSD1 by inhibiting polyubiquitination in human cancer cells. *Oncotarget*. 2015;6(19):16939–16950.

289. Lakshmikuttyamma A, Takahashi N, Pastural E, et al. RIZ1 is potential CML tumor

suppressor that is down-regulated during disease progression. *J Hematol Oncol*. 2009;2:28.

290. Pastural E, Takahashi N, Dong WF, et al. RIZ1 repression is associated with insulin-like growth factor-1 signaling activation in chronic myeloid leukemia cell lines. *Oncogene*. 2007;26(11):1586–1594.

291. Mori N, Yoshinaga K, Tomita K, et al. Aberrant methylation of the RIZ1 gene in myelodysplastic syndrome and acute myeloid leukemia. *Leuk. Res.* 2011;35(4):516–521.

292. Shimura H, Mori N, Wang Y-H, Okada M, Motoji T. Aberrant methylation and decreased expression of the RIZ1 gene are frequent in adult acute lymphoblastic leukemia of T-cell phenotype. *Leuk. Lymphoma*. 2012;53(8):1599–1609.

293. Xie W, Li X, Chen X, Huang S, Huang S. Decreased expression of*PRDM2* (*RIZ1*) and its correlation with risk stratification in patients with myelodysplastic syndrome. *Br. J. Haematol.* 2010;

294. Zhao F, Chen Y, Zeng L, et al. Effects of triptolide on RIZ1 expression, proliferation, and apoptosis in multiple myeloma U266 cells. *Acta Pharmacol Sin*. 2010;31(6):733–740.

295. Lin C-H, Wong SH-K, Kurzer JH, et al. SETDB2 Links E2A-PBX1 to Cell-Cycle Dysregulation in Acute Leukemia through CDKN2C Repression. *Cell Rep.* 2018;23(4):1166–1177.

296. Shackelford D, Kenific C, Blusztajn A, Waxman S, Ren R. Targeted degradation of the AML1/MDS1/EVI1 oncoprotein by arsenic trioxide. *Cancer Res.* 2006;66(23):11360–11369.

297. Métais J-Y, Dunbar CE. The MDS1-EVI1 gene complex as a retrovirus integration site: impact on behavior of hematopoietic cells and implications for gene therapy. *Mol. Ther.* 2008;16(3):439–449.

298. Ripperger T, Hofmann W, Koch JC, et al. MDS1 and EVI1 complex locus (MECOM): a novel candidate gene for hereditary hematological malignancies. *Haematologica*. 2018;103(2):e55–e58.

299. Spensberger D, Delwel R. A novel interaction between the proto-oncogene Evi1 and histone methyltransferases, SUV39H1 and G9a. *FEBS Lett.* 2008;582(18):2761–2767.

300. Goyama S, Nitta E, Yoshino T, et al. EVI-1 interacts with histone methyltransferases SUV39H1 and G9a for transcriptional repression and bone marrow immortalization. *Leukemia*. 2010;24(1):81–88.

301. Zhou B, Wang J, Lee SY, et al. PRDM16 suppresses mll1r leukemia via intrinsic

histone methyltransferase activity. *Mol. Cell*. 2016;62(2):222–236.

302. Corrigan DJ, Luchsinger LL, Justino de Almeida M, et al. PRDM16 isoforms differentially regulate normal and leukemic hematopoiesis and inflammatory gene signature. *J. Clin. Invest.* 2018;128(8):3250–3264.

303. Cattaneo F, Nucifora G. EVI1 recruits the histone methyltransferase SUV39H1 for transcription repression. *J. Cell Biochem.* 2008;105(2):344–352.

304. Lee M-C, Kuo Y-Y, Chou W-C, et al. Gfi-1 is the transcriptional repressor of SOCS1 in acute myeloid leukemia cells. *J. Leukoc. Biol.* 2014;95(1):105–115.

305. Li S, Ali S, Duan X, et al. JMJD1B demethylates h4r3me2s and h3k9me2 to facilitate gene expression for development of hematopoietic stem and progenitor cells. *Cell Rep.* 2018;23(2):389–403.

306. Ikeda S, Kitadate A, Abe F, Takahashi N, Tagawa H. Hypoxia-inducible KDM3A addiction in multiple myeloma. *Blood Advances*. 2018;2(4):323–334.

307. Ohguchi H, Hideshima T, Bhasin MK, et al. The KDM3A-KLF2-IRF4 axis maintains myeloma cell survival. *Nat. Commun.* 2016;7:10258.

308. Kim J-Y, Kim K-B, Eom GH, et al. KDM3B is the H3K9 demethylase involved in transcriptional activation of lmo2 in leukemia. *Mol. Cell. Biol.* 2012;32(14):2917–2933.

309. Xu X, Nagel S, Quentmeier H, et al. KDM3B shows tumor-suppressive activity and transcriptionally regulates HOXA1 through retinoic acid response elements in acute myeloid leukemia. *Leuk. Lymphoma*. 2018;59(1):204–213.

310. Sroczynska P, Cruickshank VA, Bukowski J-P, et al. shRNA screening identifies JMJD1C as being required for leukemia maintenance. *Blood*. 2014;123(12):1870–1882.

311. Zhu N, Chen M, Eng R, et al. MLL-AF9- and HOXA9-mediated acute myeloid leukemia stem cell self-renewal requires JMJD1C. *J. Clin. Invest.* 2016;126(3):997–1011.

312. Lynch JR, Salik B, Connerty P, et al. JMJD1C-mediated metabolic dysregulation contributes to HOXA9-dependent leukemogenesis. *Leukemia*. 2019;

313. Agger K, Miyagi S, Pedersen MT, et al. Jmjd2/Kdm4 demethylases are required for expression of Il3ra and survival of acute myeloid leukemia cells. *Genes Dev.* 2016;30(11):1278–1288.

314. Cheung N, Fung TK, Zeisig BB, et al. Targeting aberrant epigenetic networks mediated by PRMT1 and KDM4C in acute myeloid leukemia. *Cancer Cell*. 2016;29(1):32–48.

315. Rui L, Emre NCT, Kruhlak MJ, et al. Cooperative epigenetic modulation by cancer amplicon genes. *Cancer Cell*. 2010;18(6):590–605.

316. Scott LM, Gandhi MK. Deregulated JAK/STAT signalling in lymphomagenesis, and its implications for the development of new targeted therapies. *Blood Rev*. 2015;29(6):405–415.

317. Chang Y, Zhang X, Horton JR, et al. Structural basis for G9a-like protein lysine methyltransferase inhibition by BIX-01294. *Nat. Struct. Mol. Biol.* 2009;16(3):312–317.

318. Vedadi M, Barsyte-Lovejoy D, Liu F, et al. A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. *Nat. Chem. Biol.* 2011;7(8):566–574.

319. Wang X, Wang S, Yao G, et al. Identification of the histone lysine demethylase KDM4A/JMJD2A as a novel epigenetic target in M1 macrophage polarization induced by oxidized LDL. *Oncotarget*. 2017;8(70):114442–114456.

320. Avgustinova A, Symeonidi A, Castellanos A, et al. Loss of G9a preserves mutation patterns but increases chromatin accessibility, genomic instability and aggressiveness in skin tumours. *Nat. Cell Biol.* 2018;20(12):1400–1409.

321. Wang P, Bowl MR, Bender S, et al. Parafibromin, a component of the human PAF complex, regulates growth factors and is required for embryonic development and survival in adult mice. *Mol. Cell. Biol.* 2008;28(9):2930–2940.

322. Barrett T, Suzek TO, Troup DB, et al. NCBI GEO: mining millions of expression profiles--database and tools. *Nucleic Acids Res.* 2005;33(Database issue):D562-6.

323. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739–1740.

324. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data.

325. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* 2011;17(1):10.

326. Chen S, Huang T, Zhou Y, et al. AfterQC: automatic filtering, trimming, error removing and quality control for fastq data. *BMC Bioinformatics*. 2017;18(Suppl 3):80.

327. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.

328. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. 2012;9(4):357–359.

329. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–2079.

330. Picard Tools - By Broad Institute.

331. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008;9(9):R137.

332. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 2014;42(Web Server issue):W187-91.

333. Anders S, Pyl PT, Huber W. HTSeq — a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–169.

334. Nicol JW, Helt GA, Blanchard SG, Raja A, Loraine AE. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics*. 2009;25(20):2730–2731.

335. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

336. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–140.

337. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 2012;40(10):4288–4297.

338. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47.

339. Zhu LJ, Gazin C, Lawson ND, et al. ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics*. 2010;11:237.

340. Zhu LJ. Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods Mol. Biol.* 2013;1067:105–124.

341. Stark R, Brown G. DiffBind: Differential binding analysis of ChIP- Seq peak data. *Bioconductor Vignettes*. 2018;

342. Ross-Innes CS, Stark R, Teschendorff AE, et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*. 2012;481(7381):389–393.

343. Wu D, Smyth GK. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res.* 2012;40(17):e133.

344. Wu D, Lim E, Vaillant F, et al. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics*. 2010;26(17):2176–2182.

345. Kim J, Guermah M, Roeder RG. The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell*. 2010;140(4):491–503.

346. Costa PJ, Arndt KM. Synthetic lethal interactions suggest a role for the Saccharomyces cerevisiae Rtf1 protein in transcription elongation. *Genetics*. 2000;156(2):535–547.

347. Woodard GE, Lin L, Zhang J-H, et al. Parafibromin, product of the hyperparathyroidism-jaw tumor syndrome gene HRPT2, regulates cyclin D1/PRAD1 expression. *Oncogene*. 2005;24(7):1272–1276.

348. Carpten JD, Robbins CM, Villablanca A, et al. HRPT2, encoding parafibromin, is mutated in hyperparathyroidism-jaw tumor syndrome. *Nat. Genet.* 2002;32(4):676–680.

349. Newey PJ, Bowl MR, Cranston T, Thakker RV. Cell division cycle protein 73 homolog (CDC73) mutations in the hyperparathyroidism-jaw tumor syndrome (HPT-JT) and parathyroid tumors. *Hum. Mutat.* 2010;31(3):295–307.

350. Liu T, Chen X, Li T, et al. Histone methyltransferase SETDB1 maintains survival of mouse spermatogonial stem/progenitor cells via PTEN/AKT/FOXO1 pathway. *Biochim. Biophys. Acta*. 2017;1860(10):1094–1102.

351. Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol. Cell*. 2005;19(3):381–391.

352. Choi H, Larsen B, Lin Z-Y, et al. SAINT: probabilistic scoring of affinity purification-mass spectrometry data. *Nat. Methods*. 2011;8(1):70–73.

353. Mellacheruvu D, Wright Z, Couzens AL, et al. The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods*. 2013;10(8):730–736.

354. Nesvizhskii AI. Computational and informatics strategies for identification of specific protein interaction partners in affinity purification mass spectrometry experiments. *Proteomics*. 2012;12(10):1639–1655.

355. Warde-Farley D, Donaldson SL, Comes O, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 2010;38(Web Server issue):W214-20.

356. Chatr-Aryamontri A, Breitkreutz B-J, Oughtred R, et al. The BioGRID interaction database: 2015 update. *Nucleic Acids Res.* 2015;43(Database issue):D470-8.

357. Lin X, Tirichine L, Bowler C. Protocol: Chromatin immunoprecipitation (ChIP) methodology to investigate histone modifications in two model diatom species. *Plant Methods*. 2012;8(1):48.

358. Wang Y, Krivtsov AV, Sinha AU, et al. The Wnt/beta-catenin pathway is required for the development of leukemia stem cells in AML. *Science*. 2010;327(5973):1650–1653.

359. Yeung J, Esposito MT, Gandillet A, et al. β-Catenin mediates the establishment and drug resistance of MLL leukemic stem cells. *Cancer Cell*. 2010;18(6):606–618.

360. Siriboonpiputtana T, Zeisig BB, Zarowiecki M, et al. Transcriptional memory of cells of origin overrides β-catenin requirement of MLL cancer stem cells. *EMBO J.* 2017;36(21):3139–3155.

361. Simon JM, Parker JS, Liu F, et al. A role for widely interspaced zinc finger (WIZ) in retention of the g9a methyltransferase on chromatin. *J. Biol. Chem.* 2015;290(43):26088–26102.

362. Cho S, Park JS, Kang Y-K. Regulated nuclear entry of over-expressed Setdb1. *Genes Cells*. 2013;18(8):694–703.

363. O'Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 2016;44(D1):D733-45.

364. Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* 2002;16(8):919–932.

365. Ugarte F, Sousae R, Cinquin B, et al. Progressive chromatin condensation and H3K9 methylation regulate the differentiation of embryonic and hematopoietic stem cells. *Stem Cell Rep.* 2015;5(5):728–740.

366. Chen X, Skutt-Kakaria K, Davison J, et al. G9a/GLP-dependent histone H3K9me2 patterning during human hematopoietic stem cell lineage commitment. *Genes Dev.* 2012;26(22):2499–2511.

367. Wen B, Wu H, Shinkai Y, Irizarry RA, Feinberg AP. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* 2009;41(2):246–250.

368. Ropa J, Saha N, Chen Z, et al. PAF1 complex interactions with SETDB1 mediate promoter H3K9 methylation and transcriptional repression of Hoxa9 and Meis1 in acute myeloid leukemia. *Oncotarget*. 2018;9(31):22123–22136.

369. Müller-Tidow C, Klein H-U, Hascher A, et al. Profiling of histone H3 lysine 9 trimethylation levels predicts transcription factor activity and survival in acute myeloid leukemia. *Blood*. 2010;116(18):3564–3571.

370. Bagger FO, Sasivarevic D, Sohi SH, et al. BloodSpot: a database of gene expression profiles and transcriptional programs for healthy and malignant haematopoiesis. *Nucleic Acids Res.* 2016;44(D1):D917-24.

371. Rapin N, Bagger FO, Jendholm J, et al. Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. *Blood*. 2014;123(6):894–904.

372. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 2013;6(269):pl1.

373. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012;2(5):401–404.

374. Jiang Y, Loh Y-HE, Rajarajan P, et al. The methyltransferase SETDB1 regulates a large neuron-specific topological chromatin domain. *Nat. Genet.* 2017;49(8):1239–1250.

375. Wang L, Xu S, Lee J-E, et al. Histone H3K9 methyltransferase G9a represses PPARγ expression and adipogenesis. *EMBO J.* 2013;32(1):45–59.

376. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods*. 2013;10(12):1213–1218.

377. Kim Y, Lee H-M, Xiong Y, et al. Targeting the histone methyltransferase G9a activates imprinted genes and improves survival of a mouse model of Prader-Willi syndrome. *Nat. Med.* 2017;23(2):213–222.

378. Somervaille TCP, Cleary ML. Identification and characterization of leukemia stem cells in murine MLL-AF9 acute myeloid leukemia. *Cancer Cell*. 2006;10(4):257–268.

379. Krivtsov AV, Figueroa ME, Sinha AU, et al. Cell of origin determines clinically relevant subtypes of MLL-rearranged AML. *Leukemia*. 2013;27(4):852–860.

380. Hess JL, Bittner CB, Zeisig DT, et al. c-Myb is an essential downstream target for homeobox-mediated transformation of hematopoietic cells. *Blood*. 2006;108(1):297–304.

381. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.

Proc. Natl. Acad. Sci. USA. 2005;102(43):15545–15550.

382. Chávez-Galán L, Olleros ML, Vesin D, Garcia I. Much More than M1 and M2 Macrophages, There are also CD169(+) and TCR(+) Macrophages. *Front. Immunol.* 2015;6:263.

383. Celik S, Tangi F, Oktenli C. Increased frequency of Mediterranean fever gene variants in multiple myeloma. *Oncol. Lett.* 2014;8(4):1735–1738.

384. Vales A, Kondo R, Aichberger KJ, et al. Myeloid leukemias express a broad spectrum of VEGF receptors including neuropilin-1 (NRP-1) and NRP-2. *Leuk. Lymphoma*. 2007;48(10):1997–2007.

385. Lee S-H, Chiu Y-C, Li Y-H, et al. High expression of dedicator of cytokinesis 1 (DOCK1) confers poor prognosis in acute myeloid leukemia. *Oncotarget*. 2017;8(42):72250–72259.

386. Zhang Y, Li Y, Zhang L, et al. Prognostic Importance of DOCK1 Transcript Levels, and Biologic Insights from DOCK1 -Associated Gene and Microrna Expression Signatures in De Novo acute Myeloid Leukemia. *Blood*. 2017;

387. Monroe SC, Jo SY, Sanders DS, et al. MLL-AF9 and MLL-ENL alter the dynamic association of transcriptional regulators with genes critical for leukemia. *Exp. Hematol.* 2011;39(1):77–86.e1.

388. Volpe G, Walton DS, Grainger DE, et al. Prognostic significance of high GFI1 expression in AML of normal karyotype and its association with a FLT3-ITD signature. *Sci. Rep.* 2017;7(1):11148.

389. Sadeghi L, Prasad P, Ekwall K, Cohen A, Svensson JP. The Paf1 complex factors Leo1 and Paf1 promote local histone turnover to modulate chromatin states in fission yeast. *EMBO Rep.* 2015;16(12):1673–1687.

390. Verrier L, Taglini F, Barrales RR, et al. Global regulation of heterochromatin spreading by Leo1. *Open Biol*. 2015;5(5):150045–150045.

391. Anwar T, Arellano-Garcia C, Ropa J, et al. p38-mediated phosphorylation at T367 induces EZH2 cytoplasmic localization to promote breast cancer metastasis. *Nat. Commun.* 2018;9(1):2801.

392. Image from Taconic Biosciences, Inc. <https://www.taconic.com>