The Mechanisms of HOXA9-Mediated Oncogenic Transformation

By

Yuqing Sun

A dissertation submitted in partial fulfillment
of the requirements of the degree of
Doctor of Philosophy
(Molecular and Cellular Pathology)
in the University of Michigan
2017

Doctoral Committee:

Associate Professor Yali Dou, Co-chair
Professor Jay L. Hess, Co-chair, Indiana University
Professor James D. Engel
Associate Professor Maria E. Figueroa, University of Miami
Associate Professor Ivan P. Maillard
Assistant Professor Andrew G. Muntean

Yuqing Sun

yuqings@umich.edu

ORCID iD: 0000-0003-2655-4303

Dedication

To Mom and Dad

who provide me unconditional love and support
despite the distance between us

Acknowledgements

First and foremost, I would like to express my deepest appreciation to both of my mentors, Profs. Jay Hess and Yali Dou. Jay guided me into the field of leukemia research and offered incredible support for me to grow scientifically and personally. His dedication to mentorship, openness to novel ideas and cultivation of essential research and professional skills have profoundly influenced my career path. Yali provided me the much-needed mechanistic insights and a nourishing environment for me to grow as a scientist. This work would have not been possible without her scientific vision, positive outlook and confidence in my ability. I would like to thank both for inspiring me to strive for my personal best in research, and allowing me to pursue my interests in Bioinformatics and medicine. Their understanding and encouragement have given me the opportunity to live up to my aspirations. It has truly been a privilege to have both as my mentors and committee chairs.

I would also like to thank my committee members, Profs. Ivan Maillard, Ken Figueroa, Doug Engel, Andy Muntean for all their constructive input throughout my thesis work. Their scientific insights and helpful suggestions are invaluable for me in pursuing this work. I would especially like to thank Andy for always being a great source of moral and technical support. Without him, the *in vivo* aspect of my study would not have been possible. A special thank-you goes to Ivan, who offered professional advice and set a role model for me as to beautifully balance and integrate basic research and

clinical practice in his work. Additional thanks go to Ken and Doug for the time they have set aside for answering my questions and advancing my dissertation scientifically. I am appreciative to Profs. Weiping Zou, Andy Muntean and Jeff Rual for letting me contribute to their research and valuing my input. These are great opportunities to further develop my scientific understanding and advance my scholarly efforts. I also would like to acknowledge our MCP Program directors, Profs. Nick Lucas and Zaneta Nikolovska-Coleska, and two program administrators, Laura Hessler and Laura Labut for their administrative support.

I am grateful to all the previous Hess lab members, as well as all the previous and current Dou lab members, including but not limited to Jingya Wang, Cailin Collins, Joel Bronstein, Hongzhi Miao, Peilin Ma, Bo Zhou, Shirley Lee, Hui Zhang, Aaron Dendekker, Zhenhua Zou, Jing Xu and Jie Xiong. It has been such a pleasure to work with these talented scientists and delightful colleagues, share the ups and downs with them, and experience considerable personal growth along the way. I also want to acknowledge the tremendous support and encouragement provided to me by all my friends in Ann Arbor as well as worldwide. They have lightened up my life, provided me with strength, given me courage during the heights of frustration. Without them, my last six years as a graduate student would not have been as exciting and enjoyable.

Last but not least, no words can adequately express my gratitude towards my parents, Xu Guan and Lijun Sun. They encourage me to dream big and never give up in face of obstacles. Their unwavering love and support has been with me at every moment of this journey. I would have never accomplished this work without them. I am so grateful for them being the best parents that one can ask for.

Table of Contents

List of Figures

Abstract

Hox genes encode a family of homeodomain-containing transcription factors that are critical for body plan specification and tissue morphogenesis during embryonic development. *Hoxa9*, in particular, is required for adult hematopoiesis in which it promotes stem cell renewal and expansion. Most importantly, *Hoxa9* is commonly dysregulated in various types of acute leukemia, including acute myeloid leukemia (AML), and T- and B-precursor acute lymphoblastic leukemia (B-ALL and T-ALL). Together with its co-factor MEIS1, HOXA9 plays a causal role in driving leukemic transformation. *Hoxa9* dysregulation is also linked to various types of solid tumors, and both gain and loss of function have been implicated in tumorigenesis. Despite its central role, the mechanism through which HOXA9 mediates oncogenic transformation remains poorly understood.

Previous work in our lab found that in a HOXA9/MEIS1-driven AML cell line, HOXA9 primarily binds to promoter-distal regions of the genome. Its target regions predominately carry the epigenetic signatures indicative of active enhancers. A substantial portion of HOXA9 binding sites are co-occupied by lineage-determining factors, such as C/EBPα and PU.1. However, it remains unknown 1) whether HOXA9 drives the formation of active enhancers and globally alters the enhancer landscape; 2) whether HOXA9 strictly acts downstream of other transcription factors, or it can play a pioneer role and acts upstream of all other transcription factors and chromatin regulators; 3) if its regulatory functions are conserved in other cell lineages.

To address these questions, I show that in the myeloid lineage, HOXA9/MEIS1-transformed cells are characterized by significant alterations of the enhancer landscape and exhibit prominent emergence of *de novo* enhancers. These de novo enhancers are absent of enhancer modifications in any hematopoietic cells, and are associated with activation of a leukemia-specific transcription program. HOXA9 acts as a pioneer factor at these *de novo* regions and is required for the recruitment of myeloid lineage factor C/EBPα, while it is dispensable for the formation of the normal hematopoietic enhancers. Together, these results suggest an active role of HOXA9 in altering enhancer landscapes during leukemic transformation.

To explore the mechanisms of HOXA9-mediated enhancer formation, I assessed the role of the histone H3K4 methyltransferase MLL3/MLL4 complex in this alteration of enhancer landscape. Using immunoprecipitation and ChIP-seq analysis, I found physical interaction between HOXA9 and the MLL3/MLL4 complex. In addition, I determined that the MLL3/MLL4 complex is required for formation of *de novo* enhancers, as well as for *in vivo* leukemogenesis driven by HOXA9/MEIS1. Collectively, these findings provide strong evidence for an essential role for the MLL3/MLL4 complex in HOXA9-mediated leukemic transformation.

I have also collected preliminary data pertaining to HOXA9's function in other cell lineages. I found that HOXA9 localizes to active enhancer regions in B-lineage leukemia cells and reshape the enhancer landscape; hence, confirming HOXA9's enhancer binding characteristics. Furthermore, I discovered that HOXA9 efficiently blocks the adipogenic program in pre-adipocytes by preventing the upregulation of the key adipogenesis factor,

Pparg. These data highlight a coherent role for HOXA9 in regulating gene expression and modulating cellular differentiation across different lineages.

In summary, this dissertation reveals a previously uncharacterized role of HOXA9 in leukemogenesis and cellular transformation, and provides a strong rationale for targeting the HOXA9-collaborating chromatin modulators, as well as the leukemia-specific enhancers, for the therapeutic development of acute leukemia.

Chapter 1

Introduction

1.1 *Hox* genes – an overview

1.1.1 Evolution of *Hox* genes

Hox genes encode a family of evolutionarily conserved homeodomain-containing transcription factors that control critical processes in early development, such as body plan establishment, cell identity specification and organogenesis [1, 2]. *Hox* genes were first found to cause homeotic transformation in *Drosophila*, where the mutation of a single *Hox* gene was sufficient to cause homeotic transformation - that is, changing one body segment to another [3]. Since this early discovery, homeotic transformation triggered by mutations or misexpression of *Hox* genes has been identified in several other organisms, including chicken, mice, and human [4-6]. This conserved homeotic change testifies to the importance of *Hox* genes for proper development across many species.

While invertebrates generally have one cluster of *Hox* genes, vertebrates have several *Hox* clusters, likely due to multiple duplication events of the entire cluster during evolution. In mammals, the four paralogous clusters of *Hox* genes, named A, B, C and D, are located on separate chromosomes [7]. Within each cluster, *Hox* genes are numbered based on their positions on the chromosomes from the 3’ end to the 5’ end; paralogs of

the same number have the highest sequence similarity. *Hox* genes are arranged typically according to their expression pattern along the anterior-posterior (A-P) axis, which is called “spatial collinearity”: genes towards the 3’ end of a cluster are expressed anteriorly in an embryo; those towards the 5’ end are expressed progressively more posteriorly. Their spatial arrangement also correlates with the expression temporal sequence, which is called “temporal collinearity”. The 3’ genes are activated earlier and expressed in strictly the anterior domains. The 5’ genes are expressed later, and function progressively more posteriorly in the developing embryo [8]. The coordinated spatiotemporal control of *Hox* gene expression is vital to the distinct body structure establishment and cell identity specification in early embryogenesis.

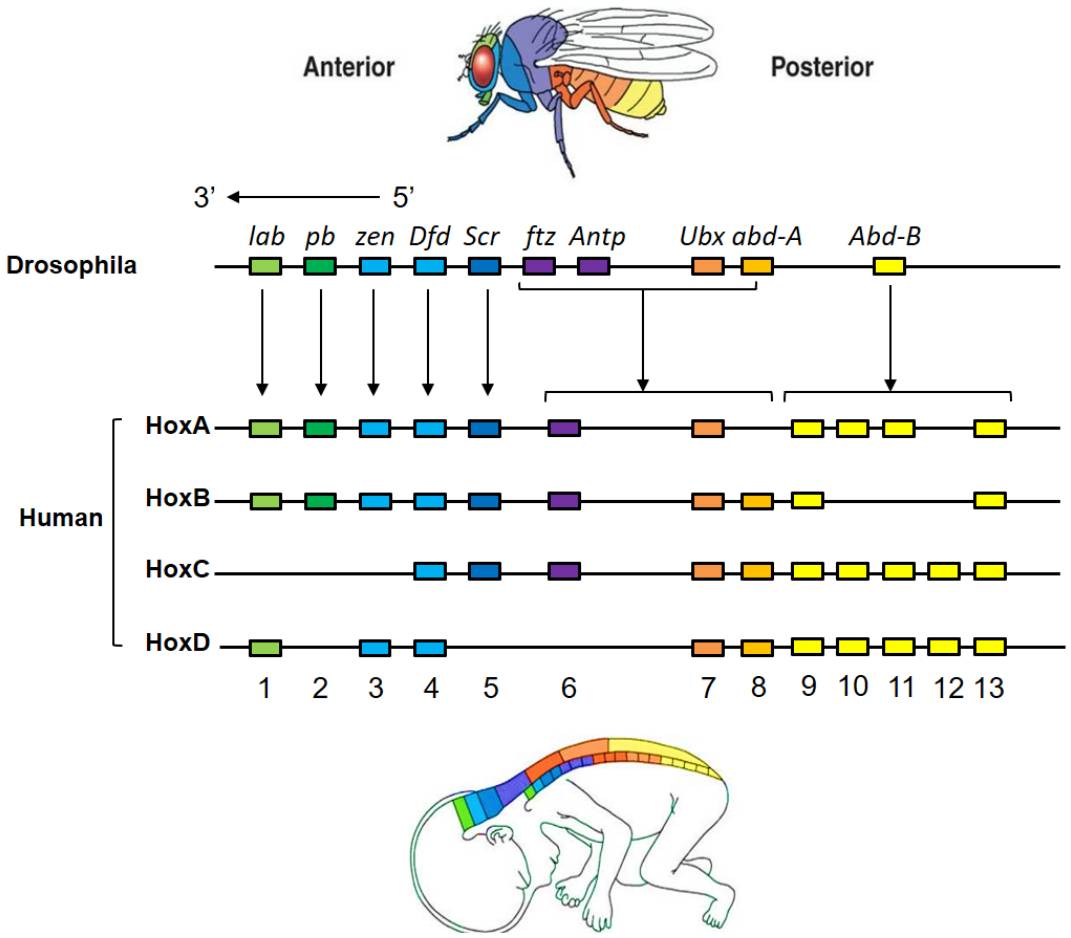

Figure 0-1 The spatial and function collinearity of *Hox* genes

A schematic showing the homology between *Drosophila* and human *Hox* genes. The four clusters of *Hox* genes are located on different chromosomes: 7p15 (A), 17q21 (B), 12q13 (C) and 2q31 (D), evolved from repeated duplications from one cluster of ancestral genes [9]. The coloration of the genes represents the correspondence between their genomic arrangements and their functional zone along the anterior-posterior axis in development. Schematic modified from Pearson Education, 2009.

1.1.2 The regulation of *Hox* genes in development

Among metazoans, each species presents divergent morphologies, but relies on the same *Hox* regulation system to establish their body plan. Thus, the expression collinearity with the genomic organization of *Hox* clusters has fascinated generations of developmental biologists. Studies have found various mechanisms accounted for the

tight, coordinated control of *Hox* gene expression. Global chromatin structure [10, 11], histone modifications [12], availability of certain transcription factors, long non-coding RNAs (lncRNAs) [13], as well as *Hox* gene themselves [14], all play a role in the highly organized expression regulation.

Polycomb (PcG) and trithorax (TrxG) group genes are the key regulators of *Hox* loci. These genes encode histone methyltransferases that implement the post-translational modifications (PTMs) on histones, and thus control the local chromatin conformation. PcG-mediated chromatin compaction is achieved through the synergized actions of Polycomb-Repressive Complex 1 and 2 (PRC1 and PRC2). PRC2 triggers histone H3 lysine 27 trimethylation (H3K27me3), an epigenetic mark closely related to gene repression, while PRC1 is recruited to the trimethylated regions and mediates ubiquitination of H2A lysine 119. By contrast, TrxG protein complexes mediate the trimethylation of H3 lysine 4 (H3K4me3) at gene promoter regions, a histone PTM commonly associated with gene activation. Therefore, PcG and TrxG complexes antagonize with each other to exert dynamic control over the target genomic regions [15]. It was first discovered in *Drosophila* that in early embryogenesis, maternally supplied factors pre-determine the chromatin configuration of *Hox* clusters [16]. In late embryogenesis, this configuration is either maintained by PcG family proteins or counteracted by TrxG family members to achieve proper gene activation or repression in a highly regulated manner [17]. In mice, sequential activation of *Hox* genes is observed at the *Hox*D cluster, as a result of the dynamic interplay between PcG and TrxG proteins. During the extension of main body A-P axis, the progressive loss of

H3K27me3 is accompanied by gain of H3K4me1, which leads to a conversion between the two epigenetic states and a shift of transcriptionally activated domain [18].

The epigenetic state switch is facilitated by the three-dimensional chromatin architecture of the *Hox* gene clusters. First discovered with Fluorescent *in situ* Hybridization (FISH) [19] then more fully characterized with Chromatin Conformation Capture (3C)-derived technologies [10, 11], *Hox* gene clusters are differentially organized in accordance with the gene transcription activity. In embryonic stem cells (ESC) where *Hox* genes are uniformly silenced, the chromatin segment hosting all *Hox* genes is compacted into a single spatial structure with bivalent epigenetic marks, as were seen with widespread interaction among various genomic loci. Later when *Hox* genes are differentially activated and required for morphogenesis, the entire chromosomal region is organized into bimodal compartments and segregated into activated or repressed domains with different histone PTMs (reviewed in [20]).

Besides, vertebrates also acquired additional regulatory mechanisms to adapt to the increasing demand of spatiotemporal gene regulation. These modalities are enhancer elements located outside of the *Hox* clusters, scattered over a large gene dessert on either side of *Hox* genes [21, 22]. One example to illustrate this additional mechanism is the regulation of posterior *Hox*D cluster genes (*Hoxd9-13*) during limb generation. These genes are required for the patterning of both proximal (arm and leg) and distal (hand and foot) segment. Initially during the development of proximal limb, *HoxD* genes form interaction loop with potential regulatory elements located on the telomeric side of the *Hox* cluster [23]. Later in the distal limb formation, a centromeric region containing groups of active regulatory elements establishes long-range

interaction with the gene cluster. These active enhancer elements first play a partially redundant role in activating posterior *HoxD* genes, but are progressively silenced as the distal limb formation completes [11].

Moreover, *Hox* gene expression is further modulated with non-coding RNAs, micro RNAs and metabolic signals, which compounds the transcriptional complexity. For instance, *HOTAIR,* the lncRNA transcribed from the *HoxC* cluster, interacts with both PRC2 and the histone demethylase LSD1 to maintain repression of the *HoxD* cluster in human [24]. Such a multilayered regulatory network may both confer a tight spatiotemporal control on *Hox* gene expression, and allow some regulatory flexibility during development and evolution [25].

1.1.3 The function of *Hox* genes in homeosis

The mutation of the *extra sex combs* (*esc*) gene, a member of the PcG family, led to the discovery of the functional hierarchy among *Hox* genes. Loss of *ecs* causes simultaneous expression of several homeotic genes in the all domains along the A-P axis. Interestingly, the most posterior genes dictate the resulting segment identity. For instance, the head and thoracic segment in larvae with *ecs* mutation experience posterior transformation and develop into the abdominal segment, A8, which is specified by the most posteriorly-acting gene, *abd-B*. In the absence of *abd-B*, the second most posterior gene, *abd-A*, conferred the segment A4 identity to all other domains. When all abdominal genes are eliminated, the larva becomes a chain of thoracic segments, which are specified by *Scr* and *Antp*. When the five most posterior genes were all removed, the larvae had cephalic structure in all segments. This stepwise loss-of-

function study suggests a functional hierarchy among the homeotic genes: the posterior genes in general had the comparatively dominant role over the anterior genes [26, 27].

Further genetic gain- and loss-of-function analyses in *Drosophila* corroborated this finding. In general, loss-of-function mutations cause anterior homeotic transformations, while gain-of-function results in posterior transformation [1, 3, 28]. This is especially true for the most posterior factors, *Antp, Ubx, abd-A,* and *abd-B* genes, although some exceptions to this pattern do exist for the anterior paralogs.

The functional hierarchy is also present in vertebrates, which is termed posterior prevalence. Systematic loss-of-function studies in mice revealed that the removal of most anterior *Hox* genes (between paralog 1-6) affects morphogenesis mainly in the cervical structures, although they are expressed in more posterior domains as well [29, 30]. In contrast, inactivation of *Hoxd13* causes a prevalent phenotypical alteration in all sites with normal *Hoxd13* expression [31]. Gain-of-function mutations primarily generate phenotypes under the same principle: when posterior *Hox* genes are ectopically expressed in anterior domains, such as *Hoxd4* in the domains of Hoxa1, those domains more likely experience posterior homeotic transformation [32], while expression of anterior *Hox* genes in posterior domains produces less drastic transformations. Deviations from this principle have also been described, likely due to the complexity of vertebrate gene regulation (systematically reviewed in [7, 33]).

Altogether, the series of cross-species genetic studies demonstrate a principal role of posterior homeotic genes. They also suggest that besides the tight spatiotemporal regulation that control their *in vivo* functions, *Hox* genes possess

intrinsic mechanistic differences to exert their instructive role on morphogenesis and body patterning in development.

1.2 The molecular functions of HOX proteins

1.2.1 Protein structure and motif recognition

Mammalian HOX proteins are relatively small, with molecular weight ranging from 25kDa to 49kDa. These proteins are encoded on two exons, with the homeodomain generally present within the second exon. The 60-amino acid homeodomain shows high sequence conservation, especially among paralogs. Another conserved motif is a hexapeptide (HX) upstream of the homeodomain that interacts with TALE (Three Amino-acid Loop Extension) family proteins, which act as cofactors for concerted DNA binding [34]. Sequences outside of the homeodomain and the HX motif diverge substantially, which include an acidic tail at the C-terminus, a linker region between the homeodomain and the hexapeptide, and a highly variable N-terminus. While their functions are largely unknown, these non-conserved regions are reported to contain PTMs and interact with other transcription regulators to orchestrate the transcriptional response [35].

The homeodomain has three alpha helices and a flexible N-terminal arm. Helix 3 makes sequence-specific contact in the DNA major groove, N-terminal arm with the minor groove, while helix 1 and 2 form the Helix-Turn-Helix (HTH) structure and lie above the DNA [34]. The homeodomains of HOX proteins in general bind to a very similar set of AT-rich sequences *in vitro*. Isoleucine (Ile) 47, glutamine (Gln) 50, asparagine (Asn) 51 and methionine (Met) 54 are conserved among all homeodomains,

and are responsible for making direct or water-mediated contacts with these sequence motifs [36, 37]. Arginine (Arg) 5 in the N-terminal arm is the most conserved residue binding in the minor groove.

The core motif for HOX recognition is 5'-TAAT[t/g][a/g]-3', with the exception of abd-B paralogs (5'-TTAT[t/g][a/g]-3') [38]. The specific binding of the first two bases pairs (5'-TTTAT-3') are imparted by the Arg 5 in the N-terminal arm. Because of the short and well-ordered linker, the N-terminal arm can also induce minor groove compression and DNA bending, a phenomenon specific to the posterior Abd-B paralogs such as HOXA9. The later three 3' base pairs (5'-TTTAT-3') are specified by helix 3 in the homeodomain. Asn 51 interacts with the adenosine at the fourth position of the motif. For Abd-B paralogs, because of DNA bending and contribution from the minor groove, Asn 51 also forms van der Waals interaction with the thymidine at the third position, which explains the slightly different motif recognition pattern [39]. Ile 47 and Glu 50 together specify the first base tolerated at base pair 3' of the core motif, and Met 54 specifies the second base in the recognition motif.

These structural studies showed that posterior Abd-B paralog HOX proteins have greater number of nonspecific interactions with DNA than the anterior proteins, which may explain the increased DNA binding affinity observed in HOXA9 as compared with HOXB1 [36, 37, 39]. This increased contact with DNA may also be an important mechanism for the functional hierarchy observed in *Drosophila* and mice. Because of this affinity difference, the stronger interaction with DNA may allow the posterior HOX proteins to compete with their anterior counterparts at shared HOX binding sites [39].

Another phenotypic study corroborates the idea that the homeodomains of different HOX paralogs may bear some intrinsic differences in gene regulation, independent from the hexapeptide-mediated cofactor interactions or N-terminus-mediated transactivation. The over expression of HOXA1 or HOXA9 drive leukemia development with significantly different aggressiveness and latency. Using these two leukemia models, Constanze *et al* found that an exchange of their homeodomain is sufficient to convert the slow progressing leukemia of HOXA1 into the aggressive one induced by HOXA9. Similar gene expression profiles were established by the homeodomain of HOXA9 regardless of the remaining protein sequences, suggesting the homeodomain controls chromatin binding and transcriptional regulation in leukemia [40].

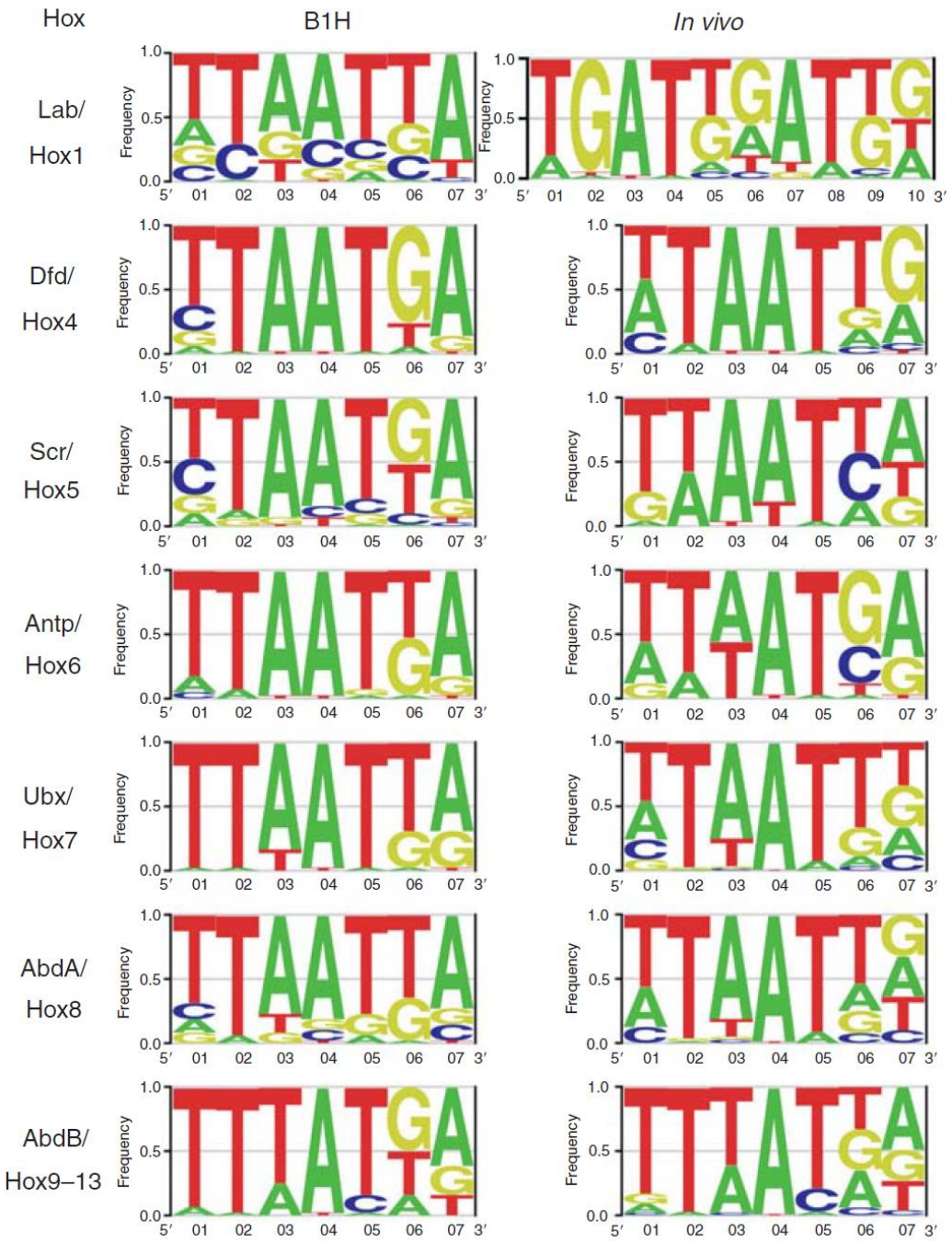

Figure 0-2 The motifs of different HOX paralogs

HOX protein DNA-binding motifs as determined by bacteria one-hybrid (B1H, left) or various *in vivo* platforms (*in vivo*, right) showing the slightly different binding preference for the most anterior (HOX1) and most posterior (HOX9-13), as compared to the rest HOX paralogs (Figure extracted from [41])

1.2.2 TALE family cofactors and HOX latent specificity

TALE family proteins were identified nearly simultaneously with HOX proteins as their cofactors and modulators for DNA binding and *in vivo* functions. These includes PBC proteins, such as CEH in *C. elegans*, Extradenticles (Ext) in *Drosophila* and PBX in mammals, and HMP proteins, such as UNC in *C. elegans*, Homothorax (Hth) in *Drosophila* as well as MEIS/PREP in mammals (Figure 1.3). TALE proteins are characterized by a highly conserved homeodomain that differs from the canonical homeodomain by an insertion of three amino-acid loop extension (TALE) motif between helix 1 and 2 in the homeodomain. The traditional view of the HOX-TALE interaction is that HOX proteins insert the conserved hexapeptide motif into hydrophobic pocket of TALE proteins, which is comprised of the TALE motif and residues of helix 1 and 3. The HOX hexapeptide motif contains a conserved Y/E-P/D-W-M sequence, in which the tryptophan (W) residue is essential for the interaction with the hexapeptide-binding pocket of TALE. However, subsequent studies show that the posterior Abd-B or HOX paralogs group 9-13 adopts a different conformation and relies on the single W to mediate the interaction. Moreover, residues in the non-conserved regions, the acidic C-terminus and the linker regions, are also shown to provide additional surface and contribute to the HOX-PBC interaction.

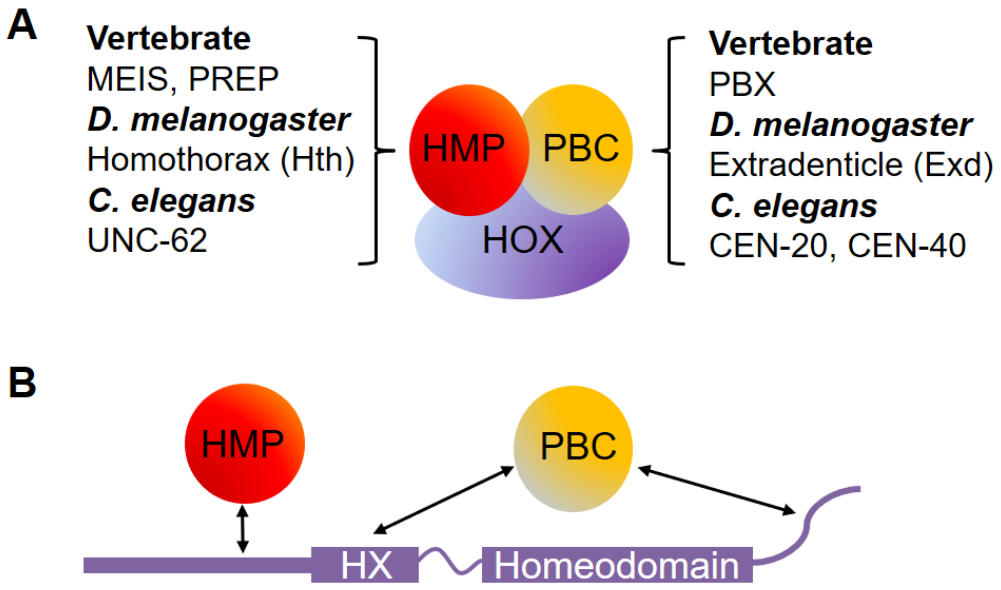

Figure 0-3 The interaction between HOX, HMP and PBC proteins
(A) Names of HOX protein co-factor orthologs in *C. elegans*, *Drosophila* and vertebrates; (B) Motifs in HOX proteins to mediate interactions with PBC and HMP proteins. (Figure modified from [35])

The interaction with TALE family factors is thought to enhance both specificity and affinity of HOX proteins' DNA binding. These HOX cofactors have been shown to form trimeric Hox/Exd/Hth (HOX/PBX/MEIS) complexes that confer functional specificity *in vivo* [42-45]. Using biochemical tools, it has been demonstrated that the addition of TALE factors can result in a shift in HOX binding preference. Thus, the term "latent specificity" was coined to describe the concept that binding with cofactors in a complex changes sequence specificity. This concept was tested comprehensively in a recent study using SELEX-seq (Systematic Evolution of Ligands by EXponential enrichment followed by sequencing). Briefly, a pool of double-stranded DNA oligomers was selected with purified *Drosophila* Hox monomers alone or Hox-Ext complexes and then sequenced to determine their preferred DNA motifs. It is revealed that although Hox monomers have very similar binding sites, Hox-Ext complexes diverge in terms of motif

preferences. The preferred core motif can be classified into three clusters along the A-P axis: the anterior (lab, pb), middle (Dfd and Scr) and posterior Hox factors (Antp, Ubx, Abd-A and Abd-B). Anterior Hox factors select sequences with a narrow minor groove, while posterior Hox factors target sequences with wider minor groove. Differentiating preferences for sequences flanking the Hox core motif have also been found, especially for anterior Hox factors. These differences are clearly dependent upon the dimerization with TALE proteins [46].

Using structural and biochemical tools, it has been demonstrated that the interaction with PBC proteins allow the variable regions of Hox proteins to contribute to DNA binding [47, 48]. One example using the Src-Exd complex demonstrated this principle. At a selective Src-Exd binding site, insertion of certain Src-specific residues into the DNA minor groove changes the electrostatic environment which favors binding of basic residues. These residues, in the N-terminal arm and immediate 3' linker region, are otherwise disordered at canonical Hox-Exd binding sites. This allosteric change suggests that correct positioning of the variable regions of Hox, in concert with Exd factors, is most critical for Hox binding at non-canonical/paralog-specific sites [49].

The study described above with Scr-Exd is also in agreement with the recent hypothesis that the specificity of Hox proteins primarily manifests at low-affinity binding sites: while canonical Hox binding sites only require the conserved regions of homeodomain, low-affinity sites depend on paralog-specific residues to mediate stable contact. Another piece of evidence comes from Croker *et al* [50]: they identified a cluster of low-affinity Ubx-Exd binding sites in enhancers of the *Drosophila* shavenbaby gene, a critical gene for early development. When the low-affinity sites are replaced with

high-affinity ones, it allows the activation of the enhancer by other Hox paralogs. These low-affinity sites may offer regulatory advantage over high-affinity sites for fine-tuning the gene expression [50-52].

Altogether, these studies have revealed several additional layers of mechanisms that HOX complexes exploit to achieve specificity beyond the simple contact with DNA sequence motifs.

1.2.3 Collaboration with epigenetic and transcriptional machineries

There is no known HOX target gene whose expression pattern is controlled by HOX proteins alone [53]. While TALE cofactors clearly contribute to the DNA motif recognition, the HOX-TALE protein complexes also employ other collaborating proteins to orchestrate the transcriptional response [54]. It has been proposed that the assembly of multi-component complexes at HOX-regulated *cis*-elements, which includes transcriptional machineries, the mediator complex and chromatin remodelers, is indispensable for modulating HOX-mediated transcriptional outcomes [41].

It was discovered early on that HOX proteins can regulate target genes by either activating or repressing transcription [55]. The distinct regulatory functions are likely executed in collaboration with different interaction partners. One well-known example is the regulation of mammalian osteocalcin expression [56-58]. Osteocalcin gene promoter contains adjacent HOX and PBX consensus motifs. In pre-osteoblast, PBX1 and HOXA10 are both bound at this promoter and together recruit histone deacetylase (HDAC), and as a result, maintain a repressive chromatin conformation. As pre-osteoblasts differentiate into osteoblasts, PBX1 is downregulated and leaves

osteocalcin promoter. HOXA10 alone recruits CBP/p300 histone acetyltransferase (HAT) instead, and activate the gene expression. This example illustrates that HOX-PBX heterodimer simultaneously recruits HAT and HDAC with the overall outcome determined by the ratio of these activities. In addition, it suggests that the availability of certain cofactors can either enhance or reverse the direction of transcription regulation mediated by HOX factors.

Multiple HOX family proteins, including HOXA9, HOXB1, HOXB7 and HOXD4 [59-61], have been shown to interact with CBP/p300 complex and mediate the histone acetylation of their target sites. Likewise, HOX cofactor MEIS1 recruits CREB1/CBP via its C-terminal regions [62]. While the activity of HOX-PBX complexes may be dependent on their relative ratio, co-binding of MEIS in the complex seems to shift the activity balance towards the active side [59, 63]. In addition, several HOX paralogs have been found to recruit repressive complexes, such as Groucho co-repressor proteins in *Drosophila* [64, 65] and G9a histone H3 lysine 9 methyltransferase in mammalian cells [66].

The direct association with the Mediator (MED) complex has also been reported for *Drosophila* Hox proteins. MED complex is a multi-subunit protein machinery: some of its subunits interact with DNA and other DNA-bound transcription factors, while others bridge RNA polymerase II (Pol II) to the transcription start site. Analysis with *Drosophila* pd and Src revealed that these Hox proteins make direct contact with Med19, a subunit of MED complex, via the homeodomains. Mutation of Med19 affects Hox target gene transcription and abolishes its developmental activity, suggesting that the MED complex may function as a HOX collaborator [67]. This study provides the first

evidence that Hox proteins can serve an interface for the targeting of general transcriptional machinery.

In addition to the role in Pol II recruitment, some members of HOX family can also modulate the activity of Pol II. In early zebrafish embryo, Hoxb1b, Pbx and Prep/Meis act in cooperation to regulate the expression of Hoxb1a. In zygote, maternally-supplied Pbx and Prep are loaded on Hoxb1a promoter, which facilitate the implementation of histone acetylation marks and the recruitment of Pol II. However, in this phase without Hoxb1b, Pol II is maintained in a poised state that is insufficient to initiate transcription. Later in embryogenesis, Hoxb1b is expressed and loaded on Hoxb1a promoter, where it promotes the phosphorylation of Pol II and the subsequent transcription elongation of Hoxb1a [68, 69]. The regulation of Pol II activity adds another level of control to Hox-dependent gene expression.

In sum, Hox family proteins utilize diverse mechanisms to modulate gene transcription involving the assembly of multi-protein complexes; these complexes are required on HOX-regulated *cis*-elements to produce the proper expression outcomes.

1.2.4 Impact on chromatin accessibility

Chromatin is by default packaged into nucleosomes that restricts protein binding and interferes with DNA-templated processes such as transcription [70]. In order for Hox factors to exert gene regulation function, they must be able to gain access to histone-free DNA. Given their transcription regulation roles, it is then a logical next-step to examine whether their functions are restricted to pre-accessible chromatin, or alternatively, they are sufficient to reposition nucleosomes to increase chromatin

accessibility and drive relaxation of chromatin. The prevailing idea is that the genome-wide targeting of HOX factors is strongly influenced by the chromatin accessibility of a specific tissue; lineage-determining transcription factors establish the chromatin landscape and thus allow a distinct subset of targets to become accessible by Hox [35, 71]. However, this view is somewhat inconsistent with the strong homeotic phenotype driven by Hox mutations or misexpression. Since cell type specification is one key role of HOX factors (see discussions above), they are presumably able to play a role in defining the lineage context for transcriptional regulation.

Direct evidence came from a chromatin accessibility test in *Drosophila* cell culture. It is revealed that Hox proteins may differ in their ability to associate with nucleosome-bound DNA [72]. In this study, a *Drosophila* embryonic cell line was transiently transfected with Ubx, Abd-A or Abd-B, and their global binding was assessed and compared with DNase I hypersensitive sites prior to transfection. While the vast majority (94%) of Ubx and Abd-A binding sites occurred within the pre-accessible regions, a significant portion (25%) of Abd-B occupied DNase I-insensitive, previously closed chromatin regions. This suggests that the ability to access closed chromatin regions varies among Hox factors.

Although the expression of Ubx alone does not induce substantial change in global chromatin accessibility, co-binding with cofactor Exd and Hth may collectively drive chromatin remodeling and have a greater impact on genome-wide DNA conformation. Indeed, when Hth is co-transfected with Ubx into the *Drosophila* cell line, the number of Ubx peaks doubled and the percentage of Ubx binding on previously inaccessible sites increased from 5% to 17%. Both Exd and Hth motifs were enriched at

the additional Ubx binding sites, suggesting a collaborative effect between Hox and its TALE cofactors [72]. Together, these findings indicate that although Hox family proteins may differ in the ability to globally remodel DNA accessibility, complexing with Exd/PBX or Hth/MEIS cofactors will increase their potential to reposition nucleosomes and facilitate their interaction with DNA sequence.

There is also evidence suggesting that HOX proteins may work in conjunction with SWI/SNF chromatin remodeling complexes. In a murine cell line, the ATPase subunit of the SWI/SNF complex (Brg1) not only co-immunoprecipitates with HOXA9, but also co-localizes with HOXA9 at hundreds of promoter-distal regulatory elements [73]. Although further evidence needs to be collected, this serves as our first clue that at specific sites, Hox proteins may act jointly with the ATP-dependent chromatin remodelers and render the nucleosome-bound DNA more accessible to other transcriptional regulators.

1.3 HOXA9 in normal development and malignant transformation

Amplification of the ancestral *Abd-B* gene likely happened prior to the duplication of the entire *Hox* cluster. This singular amplification event may be linked to the emergence of the appendicular system during evolution [74]. Together with subsequent duplication of the entire *Hox* cluster, it produced four different paralog groups (9-13) as we now know for vertebrates (Figure 1.1). These genes exert critical functions in both embryonic development and adult tissue homeostasis.

1.3.1 Normal function in hematopoiesis

HOXA9, in particular, is extensively expressed in the developing fetus [75]. Its transcripts are also detected in various adult tissues, including bone marrow, colon, kidney, prostate and skeletal muscles [76]. Like most of the posterior *Hox* genes, *Hoxa9* is highly expressed in the CD34+ population of the hematopoietic progenitors, and its down-regulation is associated with hematopoietic differentiation [77]. In embryonic stem cells (ESC), HOXA9 promotes the hematopoietic differentiation into hematopoietic stem cells (HSC), and enhances the commitment of precursors into primitive and mature blood cells [78]. While strong over expression of *Hoxa9* in the hematopoietic tissues causes embryonic lethality, lower level of enforced expression lead to an expansion of the stem cell and progenitor populations, with a concomitant block on differentiation. Over extended time (three to six months) or with additional mutations, mice with *Hoxa9* over expression in bone marrow develop acute myeloid or T-cell lymphoblastic leukemia [79, 80]. However, possibly as a consequence of functional redundancy, disruption of this gene results in minor phenotypical changes. Mice with *Hoxa9* mutations show normal health and weight, although have decreased size and cellularity in the spleen and thymus. The most drastic phenotype is displayed in the hematopoietic compartment, which has 30-40% of reduction in the number of B cells, T cells and granulocytes. *Hoxa9*-deficient cells also have repopulation deficiencies compared to wild type cells, which include reductions in common myeloid progenitors (CMP), granulocyte/monocyte precursors (GMP), common lymphoid precursors (CLP) and lymphoid precursors (pro- and pre-B cells, pro-T cells) [81-83]. Mice transplanted with *Hoxa9*$^{-/-}$ BM cells also have deficient hematopoiesis after irradiation, and display more

than 10-fold reduction of HSC in the bone marrow [80]. Together, these results indicate that the stem-cell capacity of *Hoxa9*-deficient cells is impaired, which underscores the importance of *Hoxa9* in controlling HSC expansion and self-renewal.

1.3.2 The role of HOXA9 in leukemia

Overview

HOXA9 is most intensively studied in the case of acute leukemia. More than 50% of acute myeloid leukemia cases have 2-8-fold increase of *Hoxa9* expression than healthy controls, as a result of various genetic abnormalities [54, 84]. In acute lymphoblastic leukemia (ALL) cases, such as pro-B cell ALL or pro-T cell ALL, *Hoxa9* over expression is more strictly associated with certain types of gene rearrangements [85, 86]. High *Hoxa9* expression is associated with intermediate to unfavorable prognosis [87], and one study found that *Hoxa9* is the most prognostic factor for poor prognosis [88]. Admittedly, elevated expression of *Hoxa9* is often the consequence of upstream genetic alterations, which themselves have an adverse prognosis [73]. Nonetheless, the fact that a range of leukemogenic programs all converge to drive the over expression or over activation of *Hoxa9* attests to its critical role in hematological malignancies.

Upstream regulators of Hoxa9

Acute myelogenous leukemia

A variety of genetic aberrations lead to the over expression of *Hoxa9* in AML. Among them, *MLL*-related leukemias are the most intensively studied. These leukemias are classified based on abnormalities at chromosome 11q23, a locus encoding the gene

MLL1. Wild type *MLL1* encodes a histone H3K4 methyltransferase, which is required for promoter activation of *Hoxa9* in normal hematopoiesis as well as in MLL-related leukemia [89]. Leukemia with genetic alterations at this locus constitute 70% of pediatric AML and 10% of adult AML, which include duplication, amplification and chromosome translocation [90]. Over 80 translocation partners have been discovered in MLL-related leukemia, although in most cases, the translocations involve one of the six most common partners: *AF4* [t(4;11)], AF9 [t(9;11)], *ENL* [t(11;19)(q23;p13.3)], *AF10* [t(10;11)], *ELL* [t(11;19)(q23,p13.1)], or *AF6* [t(6;11)] [91]. Among them, *AF4*, *AF9*, *ENL* and *ELL* have been implicated in transcriptional activation or elongation. In addition, DOT1L, the histone H3 lysine 79 methyltransferase, interacts with MLL-fusion proteins, which can contribute to the promoter methylation of MLL-fusion target genes [92]. Since both wild type MLL1 and MLL-fusion bind to the *Hoxa9* promoter [93], all these transcriptional activation mechanisms have been directly linked to *Hoxa9* upregulation in MLL-fusion leukemia.

HOXA9 is also involved in chromosomal translocations in leukemia. The most frequent fusion partner is NUP98, a member of the nuclear pore family. NUP98-HOXA9 induces leukemia with an extended latency (11-12 months). However, co-expression with *Meis1* significantly accelerates disease progression, suggesting a conserved mechanism for HOXA9 and HOXA9-fusion to drive leukemia.

Several other mutations have been found to correlate with elevated *Hoxa9* expression, although the mechanisms are less clear. One of the most common genetic abnormalities in AML is nucleophosmin1 mutation, a protein normally resides in the nucleus. The mutation results in cytoplasmic localization of NPM1, which, through

unknown mechanisms, contributes to *HOXA9* upregulation [73, 94]. Several additional mechanisms have been linked to *Hoxa9* upregulation, including *EZH2* mutation [95], *CDX2* over expression [96], MOZ-fusion [97] and CALM-AF10 fusion [98, 99].

Acute lymphoblastic leukemia

Dysregulation of *Hoxa9* is also reported in ALL, including both B- and T-precursor ALL (B-ALL and T-ALL), which are commonly associated with *MLL* translocations [100]. In addition, the T-ALL cases with CALM-AF10 translocation display *Hoxa9* upregulation [101]. HOXA9 can also form chimeric fusion products with T cell receptors (HOXA/TCR) [98, 102], adding to the diversity of HOXA9-related abnormalities in ALL.

Loss of function studies

MLL-related leukemia is addicted to both HOXA9 and its cofactor MEIS1 both in AML and ALL. Knocking-down of either gene exhibits largely overlapped phenotypes, including apoptosis, cell cycle arrest, cellular differentiation and reduced leukemogenic potential [103-106]. An *in vivo* study revealed that although homing capacity was intact, their ability to colonize bone marrow was undermined [107]. Therefore, maintenance of the leukemic state requires over expression of both genes, regardless of the lineage specification. Moreover, MLL-AF9 fusion protein is unable to transform HOXA9-deficient bone marrow cells. The similar phenotypes shared by knocking down both genes suggest that HOXA9 and MEIS1 collaborate and function in the same oncogenic program to drive the development of acute leukemia.

Gain of function studies

As described earlier, over expression of *Hoxa9* in murine BM results in AML, and the disease latency inversely correlates with the dose of *Hoxa9*-expressing cells [80]. Thorsteinsdottir *et al* performed a thorough analysis on the hematopoietic phenotypes of *Hoxa9*-overexpressing cells. They found that murine BM transduced with *Hoxa9* gave rise to an expanded myeloid compartment with increases in both mature cells and myeloid progenitors. In contrast, *Hoxa9* over expression suppresses B lymphopoiesis in the chimeric model as well as in two transgenic models. *Hoxa9*-overexpressing BM produces very few pre-B lymphoid progenitor cells, and those pre-B cells generate fewer and smaller pre-B colonies than their wild type controls. *Hoxa9* alone can immortalize BM progenitors *in vitro*, but they are myeloid-lineage restricted. Collectively, these findings indicate that *Hoxa9* over expression skews hematopoiesis towards the myeloid lineage in mice [80].

Although *Hoxa9* itself is only weakly oncogenic, co-expression of the Hth family cofactor *Meis1* along with *Hoxa9* induces rapid leukemia development in mice [108]. Unlike *Hoxa9*-only cells that are primarily restricted to the myeloid lineage, the *Hoxa9*/*Meis1* cells (HM cells) maintain multipotent potential and can be induced into either myeloid or lymphoid lineage [109]. It was also discovered that the tumor-initiating capacity of HM cells exists in all phenotypic compartments, including the myeloid, the lymphoid and the lineage-negative population [110]. HM cells have increased expression of HSC-specific genes, such as *Flt3* and *Cd34*, as well as the lymphoid lineage-inducing of IL-7 receptor. Consequently, these cells proliferate in response to Flt3-ligand (FL) and IL-7, which may account for their multipotent potential [109]. Over

expression of *Meis1* alone does not transform BM cells [111]. However, co-expression of *Meis1* and *Hoxa9* is frequently found in human acute leukemias, suggesting a functional cooperation between the two factors [100, 112]. The aggressive leukemia driven by *Hoxa9* and *Meis1* is also a useful tool to study the mechanisms of *Hoxa9*-dependent leukemia.

HOXA9 targets in leukemia

The study of HOXA9-mediated leukemogenic mechanisms has been strongly accelerated by the development of high-throughput technologies. These technologies enable the delineation of HOXA9-responsive regulatory elements on a genome-wide scale. In two independent studies of a HOXA9/MEIS1-transformed AML cell, the genome-wide binding sites of HOXA9 were identified using ChIP-on-chip or ChIP-seq [54, 61]. It was discovered that HOXA9 primarily binds on promoter-distal (>2kb from transcriptional start sites) regulatory sequences, which show a high degree of evolutionary conservation. These HOXA9-target sites are found to be associated with certain protooncogenes that have been implicated in hematological malignancies, such as *Erg*, *Flt3* and *Myb* [109, 113-115]. Using an inducible model of HOXA9, these studies found that nearly equal numbers of genes were up-regulated and down-regulated upon loss of HOXA9, suggesting that HOXA9 may employ distinct regulation mechanisms to modulate gene expression (see Chapter 1.2.3). Consistent with its oncogenic role, HOXA9 in general up-regulates pro-proliferative and anti-apoptotic genes, while repressing myeloid differentiation and immune response programs [73].

A number of studies have examined HOXA9 regulation of individual genes. In hematopoietic cells, HOXA9 binds to the promoter of the proto-oncogene protein kinase

Pim1, and positively regulates *Pim1* expression. It was then proposed that *Pim1* may act as a mediator and execute oncogenic and anti-apoptotic functions in leukemia [116]. In both myeloid and B-lineage leukemia, HOXA9 upregulates insulin-like growth factor, *Igf-1*, which in turn promotes survival and transformation potential of leukemia cells [115, 117]. Furthermore, HOXA9 activates the Rho family of GTPases, including the signaling protein RAC1, through upregulating *Vav2* expression [118]. Taken together, HOXA9-targets are implicated in various signaling pathways, molecular functions and metabolic processes, which implies that HOXA9 regulates a transcriptional network to promote leukemogenesis.

1.3.3 The role of HOXA9 in other types of cancer

Aside from leukemia, *Hoxa9* misexpression is also implicated in many other cancers, including ovarian, prostate, colon, breast and bladder cancer, as well as hepatocellular carcinoma and non-small cell lung cancer [119-129]. The effect of HOXA9 on carcinogenesis can be either positive or negative: in ovarian, prostate and colon cancer, elevated HOXA9 expression is found to be associated with advanced disease status or metastasis, while in bladder cancer and hepatocellular carcinoma, DNA hypermethylaton at *Hoxa9* promoter is an indicator of poor prognosis, which suggests that tumor samples had reduced *Hoxa9* expression compared with normal tissues. Furthermore, sustained expression of HOXA9 is required for modulating the breast cancer tumor suppressor BRCA1, and loss of HOXA9 promotes tumor progression, metastasis, and patient mortality [126].

It is clear from the wide variety of malignancies with dysregulated *HOXA9* expression that both loss and gain of HOXA9 function contribute to disease, and fine-

tuned HOXA9 expression is critical for the maintenance of tissue homeostasis. Therefore, understanding the regulation mechanisms of HOXA9 on downstream targets will shed light on disease mechanisms and provide novel avenues for therapeutic design. While the dysregulation of *HOX* genes can be resulted from a variety of mechanisms, defining the common and unique characteristics of HOXA9 protein function will undoubtedly give insights into the general principles of HOX biology that can be applied to various diseases with misregulated *HOX* expression.

1.4 Epigenetic regulation

1.4.1 Epigenetic regulation – an overview

In a multicellular organism, almost every single cell shares the identical genome, and yet they generate diverse cell types that have distinct and inheritable characteristics. The reason behind this paradox falls into the realm of epigenetics. The term “epigenetics” was coined in 1942 by Conrad Waddingon as “the branch of biology which studies the casual interactions between genes and their products, which brings the phenotype into being”. In line of this concept, the study of epigenetics focuses on how gene expression is regulated without alterations in DNA sequence [130]. According to Waddington, cellular differentiation can be described as the decision-making process that takes the individual cell into different trajectories. The diverse trajectories together constitute the “epigenetic landscape” (Figure 1.4) [131].

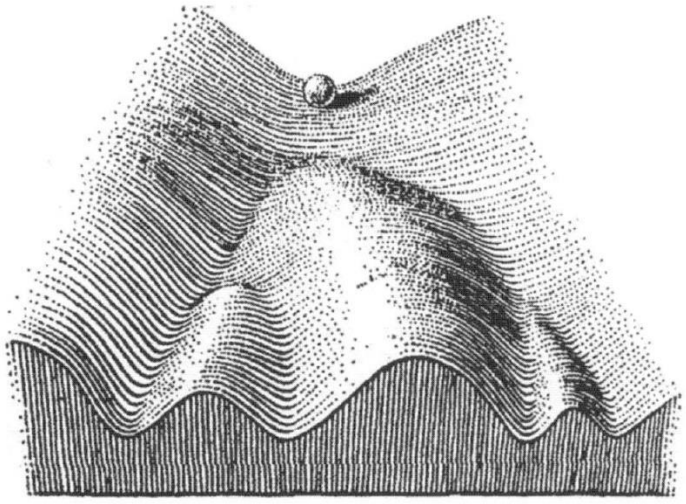

Figure 0-4 Waddington's Classical Epigenetic Landscape
This is a visual metaphor that portrays how one single cell (represented by a ball on the top) follows different permitted paths into different final cell states or fates (Figure modified from [131]).

The core molecular actors that play an indispensable role are covalent and non-covalent modifications on DNA and histones – the proteins intimately associated with DNA. These modifications are under dynamic regulations and subjected to addition and removal based on various intracellular and extracellular inputs. The key regulators that participate in these modification processes are classified into three groups: writers, reader and erasers. Epigenetic writers such as HAT and histone methyltransferases (HMTs) catalyze the addition of epigenetic marks on different residues on histone tails; readers such as the bromodomains or chromodomain containing proteins bind to these epigenetic marks. Epigenetic erasers, as the name implies, facilitate the removal of epigenetic marks [132]. These epigenetic marks control the chromatin state and represent certain "signatures" for transcription regulator proteins to interpret in order to facilitate their downstream transcriptional and biological effects.

DNA Methylation

DNA methylation is one of the best characterized chromatin modifications. It occurs on at position C5 of the cytidine ring of CpG dinucleotides. Regions of the genome where CpG dinucleotides occur at high frequency are called CpG islands, and their methylation correlates with transcriptional repression and silenced chromatin state [133]. Together with non-coding RNA and histone modifications, DNA methylation plays an important role in gene regulation and chromatin organization during development of individual organisms and maintenance of tissue homeostasis [134].

Histone modifications

Histones are subjected to modifications at over 60 different sites, including methylation, acetylation and phosphorylation, which primarily cluster on histone N-terminal tails. These modifications are involved in various chromatin-related processes, such as transcription regulation and DNA repair [135]. Especially in the recent years, there has been intense interests in determining how different chromatin modifications influence the patterns of gene expression. So far, several mechanisms have been uncovered. First, some modifications, such as acetylation and phosphorylation, alter the electrostatic properties of chromatin fibers, which ultimately lead to remodeling of the higher order structure. Second, certain histone modifications can recruit or stabilize the localization of chromatin-binding proteins, which further recruit machineries to activate or repress gene expression (reviewed in [130]). This understanding led to the proposal of the “histone code” theory: the combination of histone modifications forms a readable pattern for proteins or protein modules, to bring forth the downstream transcriptional events [136, 137]. For example, acetylation of lysine residues on histone tails are

commonly associated with transcriptional activation [138, 139], while the effect of histone tail methylation is more site-specific [140]. The *Hox* gene regulation controlled by H3K4me3 and H3K27me3 (mentioned in Chapter 1.1.2) is one of the most classic examples in this category.

Additional mechanisms, such as ATP-dependent chromatin remodeling and incorporation of histone variants, have also been uncovered. They utilize non-covalent modifications to introduce changes to chromatin conformation. For example, ATP-dependent nucleosome remodeling complexes alter the chromatin accessibility by shifting or ejecting nucleosomes, and thus give rise to a regions of relaxed, nucleosome-free chromatin for downstream transcription to occur [141]. Together, the dynamic interplay among these different mechanisms collectively guides the meaningful interpretation of genetic information encoded by the DNA sequences.

1.4.2 Epigenetic regulation in hematopoiesis

Blood is the most regenerative tissue in adults, with more than one trillion cells emerging from bone marrow every day. In this process, named hematopoiesis, HSC gives rise to all cellular components in the blood. Like all other cellular differentiation processes, hematopoiesis is governed by the dynamic interplay of different epigenetic mechanisms. Since it has a well-defined hierarchical pattern from the most primitive stem cells to the mature ones (systematically reviewed in [142]), it has served as an ideal model to study the chromatin state dynamics during cellular differentiation. In

addition, study of the chromatin regulators in hematopoiesis can also help elucidate their dysregulation and identify therapeutic targets in hematological malignancies.

Lineage-specifying transcription factors

During hematopoiesis, specific sets of transcription factors must act in a highly regulated manner to establish the proper cue for proliferation and differentiation. For example, the expression of *Gata1* is essential for the erythroid and megakaryocytic lineage, and loss of *Gata1* converts erythropoiesis to myelopoiesis [143]. For myelopoiesis specifically, the differentiation process is orchestrated by a relatively small number of transcription factors, including PU.1, CCAAT/enhancer binding proteins - C/EBPα, C/EBPβ, C/EBPε, and growth-factor independent 1 (GFI1) [144]. Among them, PU.1 and C/EBPα have been shown as the pioneer factors to build the transcriptional environment specific for myeloid differentiation [145, 146]. To execute this function, these master regulators of each lineage bind to specific DNA sequences, recruit transcription co-activators, co-repressors or chromatin-remodelers, and as a result, modulate the expression of downstream lineage-specific genes.

Histone modifications

The chromatin modification changes during hematopoiesis were initially determined in a locus- and/or modifier-specific manner. For example, it was discovered using genetic approaches and targeted ChIP-assays that GATA-1 recruits CBP and HAT, leading to the acetylation of H3 and H4 at β-globin locus [147]. Recently, with the advancement of low-cell-number ChIP-seq technologies, comprehensive histone modification maps have been drawn for the entire hematopoietic hierarchy [148, 149]. These studies showed that bivalent domains - regions with both H3K4me3 and

H3K27me3 – in ESC, partially resolve into either active or repressive domains in HSC, and those remaining bivalent domains in HSC can further resolve in mature hematopoietic cells [148]. Moreover, formation of novel enhancers also plays a role in hematopoiesis. The establishment of *de novo* enhancers, defined by the emergence of H3K4me1 mark, precedes the transcription program changes in differentiation, suggesting that the progenitor cells acquire extra regulatory potential before committing to a mature cell fate [149].

DNA-methylation

DNA methylation-mediated chromatin silencing is also crucial for hematopoiesis and blood tissue homeostasis. For example, *Hoxa9* and *Meis1*, with their important functions in HSCs, both possess Differentially Methylated Regions (DMR). These DMRs remain unmethylated until the stage of MPPs, then become hypermethylated as differentiation proceeds [150]. Loss of the DNA methyltransferase enzymes DNMT1 in HSC causes reduced self-renewal and skewed lineage commitment towards the myeloid/erythroid lineages [151, 152]. These findings imply that in hematopoiesis, DNA methylation-mediated gene silencing and maintenance of the repressive chromatin state is as critical as the gene activation process.

Chromatin accessibility

Nucleosome eviction at regulatory elements results from binding of specific regulatory factors during the establishment of chromatin landscapes [153]. Thus, accessible regions of the genome are regarded as the “footprints” of master transcription factors in the chromatin remodeling process. Based on this rationale, the

chromatin accessibility, or the “footprint” information, can provide valuable insights on the key regulators of a particular cell state [154].

This methodology has been implemented in the study of hematopoiesis. Lara-Astiaso *et al* systematically identified the critical transcription factor cohorts for various hematopoietic cell types. They found that the motifs of classical lineage-specifying factors, such C/EBPα and PU.1 in myeloid lineage, are overrepresented in their respective lineages. According to their discovery, HOXA9 motif is most highly enriched in short term-HSC (ST-HSC), multipotent progenitors (MPP) and common lymphoid progenitors (CLP), while MEIS1 motif is enriched in long term-HSC (LT-HSC) and ST-HSC. This result is consistent with their vital role in the stem cell compartment [149]. In addition, it is suggested that the chromatin landscape can precisely define a cell identity and developmental trajectory. This epigenomic information is also precious to determine the developmental context where the disease-related elements become active [155].

1.4.3 Epigenetic dysregulation in hematological malignancies

The advances in genome-wide technologies allow large scale mapping of mutations and other genomic events in malignancies. It has been increasingly recognized that many hematological malignancies, in particular, are “epigenetic diseases” – driven by mutations in chromatin modifiers as well as by genetic alterations in the non-coding regions of the genome [156]. As mentioned in 1.3.2, *MLL1* abnormalities constitute a substantial portion of acute leukemia, in both adult and pediatric patients. Mutations at tyrosine 641 in the PRC2 complex component EZH2 occur in ~30% of diffuse large B cell lymphomas [157]. Mutated EZH2 fails to implement mono- and di-methylation mark on H3K27, but collaborates with the germline EZH2 to

convert mono- and di-methylated H3K27 into H3K27me3. This mutation thus leads to a stoichiometric shift in the H3K27 methylation pattern and inappropriate silencing of EZH2 target genes [158]. Thus, abnormalities in epigenetic regulators in cancer result in global changes of epigenetic landscapes.

Mutations and local amplifications of regulatory sequences are also commonly identified in hematological neoplasia. Disease-associated Single Nucleotide Polymorphisms (SNP) are more enriched in regulatory elements, such as super enhancers, than in other regions of the genome [149, 155, 159]. For instance, the oncogene *Myc* is associated with a super enhancer 1.7 Mbp downstream of the gene promoter, which is found frequently amplified in leukemia samples [160]. Since mutated or dysregulated enhancer elements are frequently found in the proximity of disease-relevant genes and are critical for their expression, it has been proposed that these enhancers confer cell identity and disease phenotypes [159].

The unique mutational landscape renders the leukemia and lymphoma tissues more sensitive to epigenetic treatments than untransformed normal cells. There has been an exponential growth in our understanding of the role of epigenetic regulators, as well as a swell of interest in targeting them for cancer therapies [161]. For example, using CRISPR-Cas9-mediated approaches, Shi *et al.* found that acute myeloid leukemias are dependent on at least 25 epigenetic enzymes out of the 192 chromatin modulators in the screening assay [162]. Among them, BRD4, a member of the BET (bromodomain and extra terminal domain) family and an acetylated lysine reader, has received considerable attention because of the recent success in targeting the BET domain for leukemia treatment [163-165]. Another known drug target is DOT1L, a

collaborating partner of MLL-fusion proteins [166, 167]. DOT1L is recruited to MLL-fusion binding sites and mediates transcriptional activation. Its inhibitors have shown promising results in targeting MLL-rearrangement leukemias and have entered clinical trials [168-170]. These studies together offer great prospects for treating hematological malignancies by targeting epigenetic regulators.

Chapter 2

HOXA9-mediated epigenetic landscape alterations

2.1 Introduction

It has been increasingly recognized that epigenetic dysregulation is one of the key characteristics of malignancies [156, 171, 172]. Mutated or misexpressed chromatin regulators, such as transcription factors and histone modifiers, trigger the formation of an aberrant gene regulation landscape, and thus promote cancer formation [84, 173]. Enhancers, the distal regulatory elements of gene expression, are the frequent targets of malfunctional regulators, and are re-configured with aberrant histone post-translational modifications (PTMs) and/or DNA-cytosine hyper- or hypo-methylation. Notably, such alterations in enhancer landscape must involve chromatin regions that are developmentally silenced; this epigenetic reactivation process is elicited by pioneer transcription factors who engage their targets on closed regions and recruit additional epigenetic machinery that initiate the relaxation of chromatin. Studies have found a subset of transcription factors with this epigenetic remodeling ability during oncogenic transformation. It is conceivable that these transcription factors with reprogramming potential are also critical for the initial cell fate specification in embryonic development [174].

It has long been recognized that HOXA9 and its cofactor MEIS1 play a causal role in promoting leukemic transformation in both mouse models and human leukemia patients. However, the mechanisms through which HOXA9 and its cofactor MEIS1 directly regulate target genes are poorly understood. Previous studies have shown that in a AML cell line, HOXA9 binds on promoter-distal regions whose epigenetic signature is indicative of active enhancers. HOXA9 transcriptionally activates a group of oncogenes, while represses the myelopoietic pathways and inflammatory responses [61]. However, it remains to be fully elucidated how HOXA9 exerts the differential regulatory functions, and whether HOXA9 exploits the pre-existing regulatory landscape, or remodels it to adopt a leukemogenic cell fate. Here we performed coordinated analysis of chromatin states at HOXA9 binding sites in HOXA9/MEIS1-transformed leukemia cells and their normal hematopoietic counterparts. We discovered that HOXA9 reshapes the enhancer landscape and initiates the formation of a group of novel enhancers. These *de novo* enhancers, which likely represent reactivation of an embryonic development program, are critical for the oncogenic properties of HOXA9.

2.2 Materials and Methods

Animals

All animal experiments were performed as approved by the University of Michigan Committee on the Use and Care of Animals and Unit for Laboratory Animal Medicine. For *in vivo* leukemogenesis assays, 8-10 week-old female C57BL/6 (WT)

mice (JAX no. 000664; The Jackson Laboratory) were purchased and used as transplantation recipients.

Antibodies

For Western blot analysis, anti-HOXA9 (07-178, Millipore), anti-PTIP and anti-MLL4 #3 antibodies generated in rabbits were used [175]. For ChIP, anti-HA (ab9110; Abcam), anti-C/EBPα (sc-61X; Santa Cruz Biotechnology), anti-MLL4 #3, anti-H3K4me1 (ab8895; Abcam), anti-H3K27ac (ab4729; Abcam), anti-H3K27me3 (07-449; Millipore), and IgG (sc-2027; Santa Cruz Biotechnology) were used.

For flow cytometry, allopycocyanin (APC) anti-c-Kit (105812; Biolegend), APC anti-Gr1 (108412; Biolegend), APC/Cy7 anti-B220 (103224, Biolegend), phycoerythrin(PE) anti-CD43 (143205, Biolegend), PE anti-CD11b (101208, Biolegend), PE anti-CD16/32 (101307, Biolegend), Peridinin-chlorophyll-Cy5.5 (PerCP-Cy5.5) anti-CD19 (45-0193-82, eBioscience), PerCP-Cy5.5 anti-Sca1 (45-5981-80), eFluor® 450 anti-CD34 (48-0341-80, eBioscience) and DAPI (Sigma) were used.

Cell Lines

Bone marrow from 6- to 10-week-old C57BL mice was harvested 5 d after treatment with 5-flurouracil (150 mg/kg). Lineage-negative (Lin^-) bone marrow cells were first flushed from femora and tibiae with 25G needles, and then isolated using the EasySep Mouse Hematopoietic Progenitor Cell Enrichment Kit (19856, Stem Cell Technologies). Lin^- cells were maintained in Iscove's modified Dulbecco's medium (IMDM) supplemented with 15% Fetal Bovine Serum (FBS, Sigma F4135), 10 ng/mL Interleukin (IL) -3, and 100 ng/mL stem cell factor (SCF).

To package retroviruses, Plat-E cells (RV-101, Cell Biolabs) were transfected with MIGR1-HA-*Hoxa9* (*Hoxa9*) or MIGR1-HA-*Hoxa9*-estrogen receptor tag (*Hoxa9-ER*) and with MIGR1-Flag-*Meis1* (*Meis1*) retroviral vectors (plasmids previously described in [61]) using FuGENE 6 (E2691, Promega). Cell-free supernatant was collected 48 hours after transfection.

To generate the HMM cell line, *Lin*$^-$ BM were spinoculated with *Hoxa9* or *Hoxa9-ER* and *Meis1* retrovirus together at 3200rpm for 90mins at room temperature on two consecutive days. *Hoxa9-ER*/*Meis1*-transduced cells were then cultured in continuous 100nM 4-hydroxytamoxifen (OHT) to maintain transformation. After transduction, *Lin*$^-$ or *Hoxa9(-ER)/Meis1*-transduced progenitors were cultured in IMDM with continuous 10ng/ml IL-3, while SCF was gradually withdrawn in 7 days.

shRNA knock-down

To knock-down *Aldh1a3*, bone marrow from *rtTA* knock-in mice (no. 006965; The Jackson Laboratory) was retrovirally transduced with *Hoxa9* and *Meis1* to generate HM tet-on cells. *Aldh1a3* siRNA sequences were cloned into dsRed-expressing TRMPVneo vector [176] and transduced in to HM tet-on cells. Infected cells were then selected with 1mg/ml G418 (10131, Gibco) for 1 week. Resistant cells were treated with 1µg/ml doxycycline to activate the shRNA. RNA samples were taken at 1, 2, 4 and 6 days. qRT-PCR was run to check the knock-down efficiency.

Cellular assays

For HOXA9-inactivation studies, HMM cells were washed three times with culture media and then maintained in either 100 nM OHT or equal volume of ethanol for 3 days before being used for downstream analysis.

Competitive proliferation assay was carried out by mixing the parental HM tet-on cells with the shRNA-expressing (dsRed+) HM tet-on cells in a 1:3 ratio. The percentage of dsRed+ cells was monitored by flow cytometry analysis over the course of 10 days.

For CFU assay, HM tet-on expressing Aldh1a3 shRNA or non-targeting (Renilla) shRNA were seeded at 1000 cells/ml in semi-solid methylcellulose-based media (Methocult, M3234 STEM CELL Technologies) with 10ng/ml IL-3 and were allowed to grow for 6 days. Colonies were stained with iodonitrotetrazolium chloride (I10406, Sigma) for 30mins at 37°C, followed by scanning and imaging.

Cellular morphology was assessed using cytospin followed by HEMA 3 staining (22-122-911, Fisher Scientific). Whole-cell lysates were collected by directly lysing washed cells in SDS loading buffer plus β-mercaptoethanol. Protein levels were visualized using SDS/PAGE and Western blotting on PVDF membranes. RNA was collected and purified using a Qiagen RNeasy Kit with on-column DNase treatment. cDNA was generated using SuperScript III Reverse Transcriptase (18080093, Invitrogen), and target gene expression was determined relative to Gapdh using Power SYBR Green (4368708, Thermo Fisher)

Flow Cytometry

For surface marker expression, cells were washed and resuspended in standard media (2% FBS and 0.1% NaN3 in DPBS), and incubated for 30 min on ice with 0.2µg of the appropriate antibodies. After incubation, cells were washed twice before analysis on a Becton Dickinson LSR II. Data collected from at least 20,000 events from biological replicate experiments were analyzed using FlowJo Version 10 (TreeStar).

ChIP

ChIP assays were conducted as described in [54]. Briefly, cells were fixed for 15 min at room temperature with 1% paraformaldehyde, lysed in lysis buffer [1% SDS, 10 mM EDTA, 50mM Tris·HCl (pH 8)] for 10min on ice, and sonicated using Bio-raptor. Diluted chromatin was incubated with 2.5µg of appropriate antibody overnight at 4 °C with rotation. Immunoprecipitation was then performed using protein G Dynabeads (10004D, Thermo Fisher). Immunoprecipitates were washed for 5 min in low-salt (150 mM), high-salt (500 mM), and lithium chloride (0.25 M) buffers, and twice with Tris/EDTA buffer. Captured chromatin was eluted by incubating beads in 250µL of elution buffer (1% SDS, 100 mM $NaHCO_3$) for 30 min at 42 °C. Cross-linking was reversed by the addition of 50µM NaCl and overnight incubation at 65 °C. Chromatin was then RNase A-treated and purified using a Qiagen PCR purification kit. Binding was quantified relative to input by quantitative PCR (7500 PCR System; Applied Biosystems) using SYBR green fluorescent labeling and primers designed using the Integrated DNA Technologies PrimerQuest program.

CRISPR-mediated enhancer deletion

The constitutive Cas9 expression construct MSCV-hCas9-PGK-Puro and the single guide RNA (sgRNA) expression construct U6-sgRNA-EFS-mCherry were gifts kindly provided by Dr. Christopher Vakoc, Cold Spring Harbor Laboratory. All sgRNA sequences in this study were designed using https://benchling.com/crispr. The sequences with minimal off-target score were chosen. A 5' guanine (G) nucleotide was added to all sgRNA sequences that did not start with a 5' G, as suggested in [162].

To derive constitutive hCas9-expressing HMM cells, HMM cells were spinoculated with retrovirus packaged from MSCV-hCas9-PGK-Puro at 3200rpm for 90mins at room temperature. Transduced cells were selected with 1µg/ml Puromycin, and isolated by limiting dilution. Cas9 expression in the monoclonal cell lines were tested by Western blotting using CRISPR/Cas9 Monoclonal Antibody [7A9] (A-9000-050, Epigentek) and two clones with high hCas9 expression were chosen.

sgRNAs targeting the four *Aldh1a3*-proximal enhancer regions were designed to remove the ATAAA binding motif of HOXA9. To improve the efficiency of enhancer deletion, pairs of sgRNAs flanking the same HOXA9 motif were cloned into one U6-sgRNA-EFS-mCherry vector, by sequentially inserting the two U6-sgRNA cassettes via the BsmBI site and the EcoRI/XhoI site. The hCas9-expressing HMM cells were then transduced with the sgRNAs-expressing viruses. At Day 3 post-transduction, mCherry-expressing cells were individually sorted into 96-well-plates by flow cytometry and were expanded into monoclonal cell lines.

Genomic DNA was isolated from the clonal CRISPR-targeted HMM cell lines using Radiant™ Extract & Amplify Tissue PCR kit (C462, Alkali Scientific Inc.). To screen clones with successful enhancer deletion, enhancer-spanning fragments were

amplified by PCR, with the supposed length of 600-1000bp in wildtype cells. Gel electrophoresis was applied to identify the clones with shorter PCR products. The sequences of the target regions were confirmed by Sanger DNA sequencing. *Aldh1a3* mRNA expression of the deletion clones was compared with control clones with CRISPR-targeted *Rosa26* promoter [177].

4C-seq

4C-seq was performed following the protocol published by [178]. Briefly, $1x10^7$ MP, HMM or HOXA9-inact cells were crosslinked in 1% paraformaldehyde at room temperature and lysed in lysis buffer [10mM Tris-CL, pH7.5; 10mM NaCl; 0.2% NP-40; 1x protease inhibitor (Roche complete mini)] on ice for 10 min. Nuclei were pelleted by spinning down at 400g for 5 min, resuspended in 0.5ml 1.2x restriction buffer plus 0.25% SDS and shaken at 37C for 1h at 900rpm. The nuclei were digested with 400U of HindIII (NEB) at 37°C overnight. The restriction enzyme was inactivated by adding 80µl 10% SDS and incubating at 65°C for 25 min while shaking at 900rpm. Ligation was then performed by adding 6.125ml 1.15x ligation buffer, 375µl 20% Triton X-100 and 5µl T4 DNA ligase and incubating at 16C overnight. Crosslinking was reversed by incubating overnight at 65°C with Proteinase K. Chromatin was then RNase A-treated and purified using phenol chloroform. Secondary digestion was performed with DpnII, followed by heat inactivation of the restriction enzyme and ligation with T4 DNA ligation at 16°C overnight. DNA was then purified first by ethanol precipitation and then with Qiagen PCR Purification Kit. 100ng chromatin was amplified with Expand Long Template Polymerase (ELONG-RO Roche). The primers used to amplify 4C-seq libraries anneal to the bait sequence, Aldh1a3 TSS regions, and have overhangs that contain the

sequences of barcode, Illumina adaptor and sequencing primer. Barcoded libraries were pooled at equal molar ratio and subjected to massively parallel sequencing using a HiSeq 2000 instrument (Illumina) using single-end 50-bp sequencing.

ChIP-Seq, peak calling and peak annotation

Multiplexed ChIP-seq libraries were prepared at the University of Michigan DNA Sequencing Core. 50-cycle single-end sequencing runs were performed on Illumina HiSeq 2500 at a sequencing depth of 10-50 million aligned reads per sample. Sequenced reads were preprocessed to trim adaptor sequences (Trimmomatic) and then aligned to mouse reference genome (mm9) using BWA software (version 0.6.2). Only uniquely mapped reads were used in downstream analyses. Model-based Analysis for ChIP-seq (MACS) was used the identification of ChIP-seq peaks with p value=10^{-4} for HOXA9. Peaks of C/EBPα and histone modifications were identified with the default parameters. Peaks were annotated to their nearby genes using the default parameters of GREAT[179]. Genes that associated with both HOXA9+ DE and HOXA9+ PE were considered only as DE-regulated genes. Peak overlap was calculated with the criterion that there is at least 1bp overlap between the test peaks. Pathway analysis was performed using the DAVID Functional Annotation web tool [180]. The HOXA9, H3K27me3 and one set of H3K4me1 ChIP-seq data in HMM cells were obtained from our previous study [54], the rest were obtained in this study.

RNA-seq analysis

Poly-A enriched RNA-seq libraries were prepared at the University of Michigan DNA Sequencing Core. 50-cycle single-end sequencing runs were performed on

Illumina HiSeq 2500 at a sequencing depth of 40-50 million reads per sample. Sequencing reads were aligned to mm9 genome using Tophat (version 2.0.11). Transcript counts were generated with HTSeq (version 2.20.1). Differential expression analysis was performed with edgeR.

Global enhancer profile analysis

The analysis method was slightly modified from [149]. Peaks from 20 H3K4me1 ChIP-seq samples (MP, HMM, HMB from this study, Pro-B from [138], 16 normal hematopoietic cell types from [149]) were combined into one unified catalog. Overlapping peaks between replicates (peak center distance < 500bp) were merged, and peaks identified in only one replicate were discarded. Any peaks overlapping with annotated mouse promoter regions (-2/+1kb of TSS) were removed. Redundant peaks that occurred in more than one cell types were reduced to one representative peak with the largest fold enrichment within any 2000bp window. This resulted in a catalog with 116,182 putative peaks.

We counted H3K4me1 reads within 2kb around peak centers using the annotatePeaks.pl function in HOMER suite (http://homer.salk.edu/homer/), normalizing to 10^7 reads per library. Cell types were clustered by hierarchical clustering using the log2-transformed counts. Putative peaks in the enhancer catalog were per-peak normalized before clustered by K-mean (K=16). Hierarchical clustering and K-means clustering were both done using Cluster 3 [181].

Differential H3K4me1 analysis to identify *de novo* enhancers

The differential H3K4me1 analysis was performed using DiffBind [182], following its reference manual. To consider all normal peaks in the myeloid lineage, MACS H3K4me1 peaksets derived from untransformed myeloid CMP, GMP, Monocytes (Mono), Macrophages (Mφ), GN and myeloid progenitor (MP) cells were compared with peaksets from three replicates of HMM cells ([54] and this study). The dba.count function was used with minOverlap=1 and summits=1000. Differential analysis was done with DESeq2. Regions with FDR<0.05 were considered as differentially methylated. We consider a HOXA9 peak overlapping with a gained or lost H3K4me1 region if its center is within 2kb from the H3K4me1 summit.

2.3 Results

2.3.1 HOXA9 binds to active distal regulatory elements in myeloid leukemia cells.

We previously discovered that the genome-wide binding pattern of HOXA9 in HOXA9/MEIS1-transformed myeloid leukemia cells (HMM cells) is consistent with active distal regulatory elements [61]. Using the same cell line, we performed chromatin immunoprecipitation sequencing (ChIP-seq) to map the global binding of HOXA9, as well as histone modifications associated with enhancer status ([54] and this study), and confirmed the previous result. Enhancer status is commonly defined with histone post-translational modifications (PTMs), including histone H3 lysine 4 monomethylation (H3K4me1), H3 lysine 27 acetylation (H3K27ac), and H3 lysine 27 trimethylation (H3K27me3) signals, so that their status can be classified into primed (H3K4me1 only), active (H3K4me1/H3K27ac) or poised (H3K4me1/H3K27me3) [183]. Consistent with our previous findings, HOXA9 binding sites were highly enriched of H3K4me1 and H3K27ac,

but were depleted of H3K27me3, indicating active regulatory potential (Figure 2.1 A). When these binding sites were mapped to genomic features, only 6% (n = 405) occurred within promoter regions (-2/+1 kb relative to transcription start sites (TSS)). For the 6173 promoter-distal peaks, the majority (63.5%) overlapped with putative active enhancers and 18.1% with the primed regions, while limited binding (1.0%) was found on poised enhancers (Figure 2.1 B).

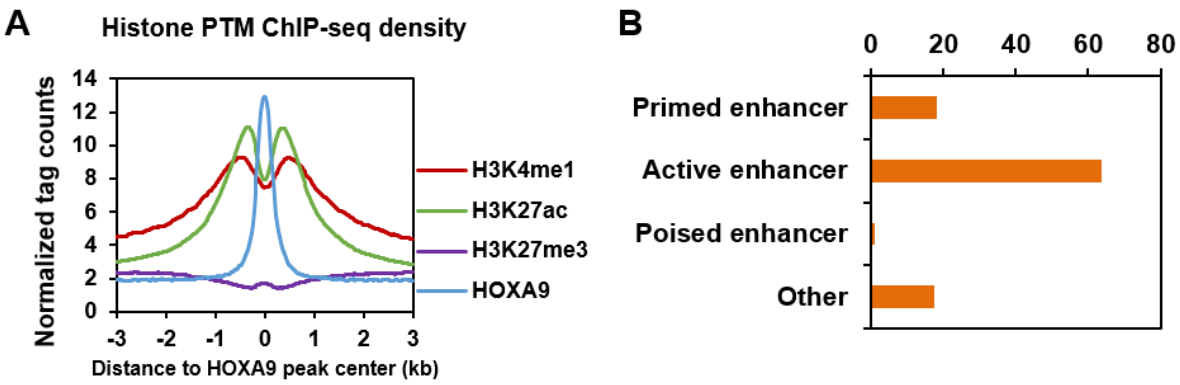

Figure 0-1 HOXA9 binding sites are enriched of active enhancer signature
(A) Composite plot showing average per base pair (bp) density of *Hoxa9*, H3K4me1, H3K27ac and H3K27me3 at *Hoxa9* binding sites in HMM cells. Library size normalized to 1E7 reads. *Hoxa9* peaks are enriched with H3K4me1 and H3K27ac globally, but depleted of H3K27me3. (B) Percentage distribution among different enhancer states at *Hoxa9*'s promoter-distal binding sites. These sites are preponderantly associated with active enhancer signatures. Promoter: -2 ~ +1 kb of transcription start site (TSS). Active enhancer: H3K4me1+;H3K27ac+. Primed enhancer: H3K4me1+;H3K27 unmarked. Poised enhancer: H3K4me1+;H3K27me3+. Other: H3K4me1 unmarked; H3K27 unmarked.

2.3.2 HOXA9-mediated transformation reshapes the enhancer landscape in myeloid leukemia cells

Previous studies suggested that distal enhancers are key for maintaining cell type-specific transcriptome and as a result, cell identity [159, 184]. To test whether HOXA9-induced leukemic transformation is accompanied by alterations of the cell-specific

enhancer landscape, we generated a myeloid culture cells (MP cells) under the same *in vitro* culture condition as HMM cells, and used them as the control cell line for HMM cells. Both flow cytometric and morphological analysis showed that these two cell types shared considerable resemblance, although their transformation status differed (Figure 2.2).

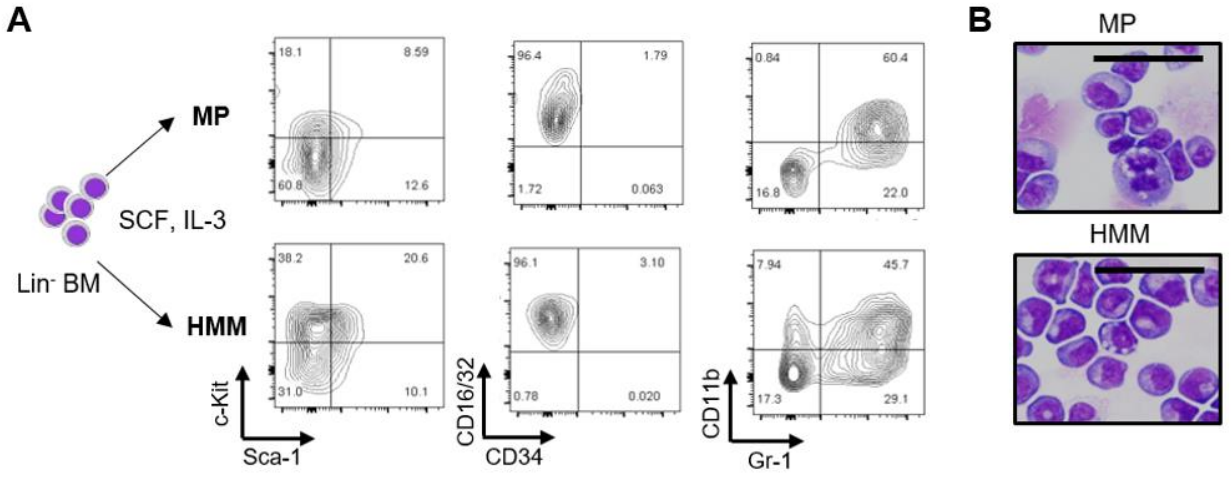

Figure 0-2 The Immunophenotype and histology of MP and HMM cells
(A) Flow plot showing the surface expression of c-Kit, Sca-1, CD16/32, CD34, CD11b and Gr-1 of HMM and MP cells, indicating that they are both mixed populations in the myeloid lineage. (B) Cytospin result showing the morphology of HMM cells 4 weeks after transformation, and MP cells 1 week in culture. Scale bar: 50μm

To relate MP and HMM cells to the hematopoietic hierarchy, we conducted hierarchical clustering using the H3K4me1 ChIP-seq data from HMM and MP cells, as well as from 16 hematopoietic cell types characterized in Lara-Astiaso, D. *et al* [149]. This analysis included LT-HSC and ST-HSC, oligopotent progenitors (CLP, CMP, GMP, MEP), as well as terminally differentiated mature cells, such as monocytes (Mono), macrophages (Mφ) and B lymphocytes (B). Consistent with the immunophenotypic and histological characterization (Figure 2.2), phylogenetic tree based on the global H3K4me1 dynamic profile revealed that HMM cells were closely related to the cells in myeloid lineage, suggesting that they largely maintained the enhancer landscapes of their cell-of-

origin (Figure 2.3 A). Notably, HMM cells were not only related to MP cells, but instead, were classified with the entire myeloid lineage. Therefore, in order to systematically examine the gained and lost enhancers specifically for leukemic transformation, we performed differential analysis on the H3K4me1 ChIP-seq data from three preparations of HMM, as compared with those of all untransformed cells in myeloid lineage (CMP, GMP, Mono, Mφ, GN and MP cells). We identified 11,814 regions that were differentially H3K4 monomethylated between HMM and the untransformed cells, among which 3760 regions consistently gained H3K4me1 density, while 8054 lost intensity (Figure 2.3B, 2.4A left panel). Representative loci with gained or lost enhancer(s) are shown in Figure 2.4 C (bottom four tracks).

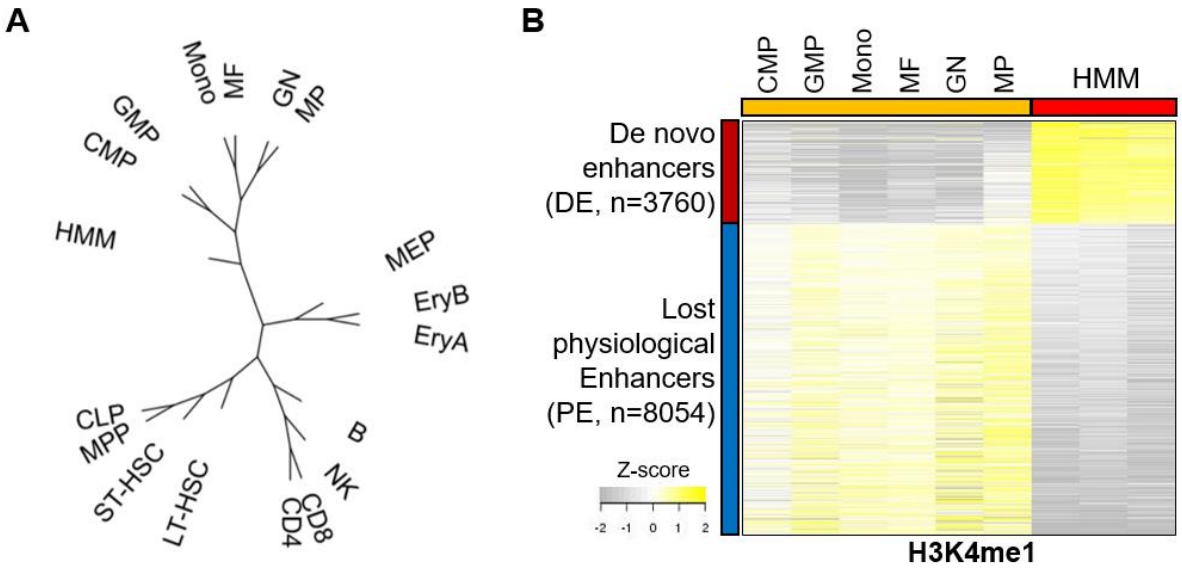

Figure 0-3 The enhancer landscape changes in HMM cells
(A) Clustering dendrogram of cell types based on H3K4me1 profiles showing the association of HMM cells with the untransformed myeloid cells. (B) Heatmap showing the 11,816 differentially enriched H3K4me1 regions between HMM cells and normal myeloid cells. Cut-off: FDR < 0.05.

Interestingly, HOXA9 showed significantly higher occupancy on gained enhancers than on both lost enhancers and non-differentially H3K4 monomethylated (unchanged)

loci (Figure 2.4B). In fact, 20% (n=764) of the gained enhancer were bound by HOXA9; in contrast, 1.8% (n=146) of the lost enhancer and 5.8% (n=4558) of the unchanged loci overlapped with HOXA9 peaks (Figure 2.4A, right panel), suggesting that the gained enhancers are possibly favorable targets of HOXA9. We refer to the gained enhancer regions in HMM cells as *de novo* enhancers (DE), and those that had prior H3K4me1 modification in the untransformed cells (either lost signal density or remained unchanged) as physiological enhancers (PE). Figure 2.4C shows two representative loci with a DE and two lost PE, showing that HOXA9 is more likely to occupy DE than PE.

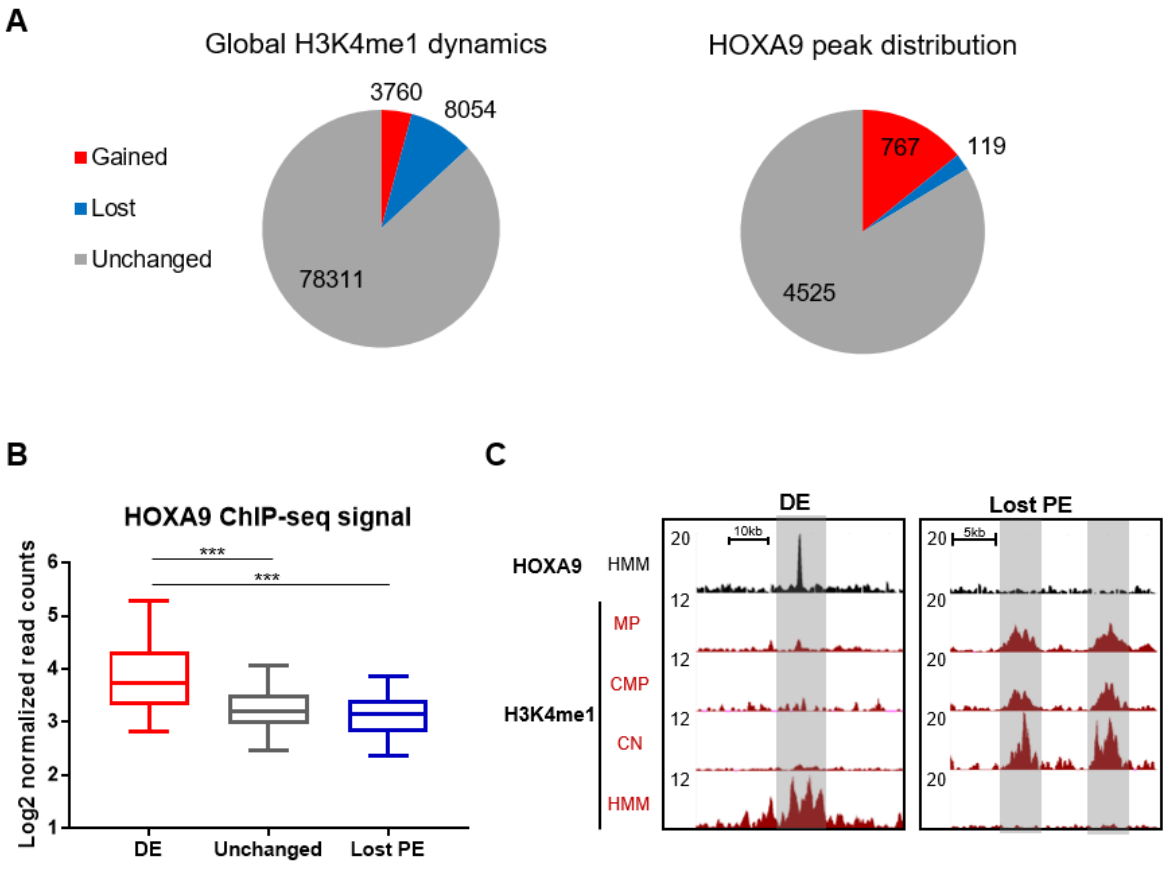

Figure 0-4 HOXA9 binding is enriched at *de novo* enhancers
(A) Pie charts illustrating the percent of enhancers are altered significantly or unchanged between HMM and normal myeloid cells (left) and the percent of HOXA9 peak distribution among the three types of enhancers (B) Average HOXA9 ChIP-seq tag

density over DE, lost PE and unchanged regions, illustrating that HOXA9 signal is significantly enriched at DE. Statistics were obtained from Mann-Whitney U-test. (C) UCSC browser views of H3K4me1 profiles from three representative myeloid lineage cells -Common Myeloid Progenitor (CMP), Granulocyte (GN), MP cells - and HMM cells, illustrating a *de novo* enhancer (left) and two lost physiological enhancers (right). HOXA9 binding at these regions shown on top. Library size normalized to 1E7 reads.

2.3.3 HOXA9+ DE and HOXA9+ PE have different characteristics

Given that HOXA9 binds both on the *de novo* enhancers and on the physiological enhancers, we hypothesized that the two types of HOXA9 binding sites would display distinct behaviors in leukemic transformation. HOXA9 target sites were then classified into two groups based on whether they overlap with *de novo* enhancers (HOXA9+ DE) or physiological enhancers (HOXA9+ PE). As shown in Figure 2.5A, the 4644 HOXA9+ PE have the same or reduced H3K4me1 in HMM as compared to MP cells, while the 764 HOXA9+ DE have H3K4me1 only in HMM cells. Consistent with the fact that MP cells were a mixture of cells at different stages of myeloid differentiation, HOXA9+ PEs were also H3K4 monomethylated at various level in normal myeloid lineage cells, suggesting that they have potential functions in hematopoiesis (Figure 2.5B, lower panel). In contrast, HOXA9+ DE were exclusively enriched in transformed HMM cells, but were not identified in any hematopoietic cell types. Notably, the modification at these regions was unlikely remnants of *cis*-regulatory elements in early hematopoiesis, as they were not present in the undifferentiated progenitors (LT-HSC, ST-HSC and MPP) (Figure 2.5B, upper panel). The majority of HOXA9+ PE in both MP and HMM cells, as well as HOXA9+ DE in HMM cells, were active enhancers with high H3K27ac. Furthermore, no H3K27ac was found at DE in MP cells, suggesting that H3K27ac was established downstream of H3K4me1 (Figure 2.5C).

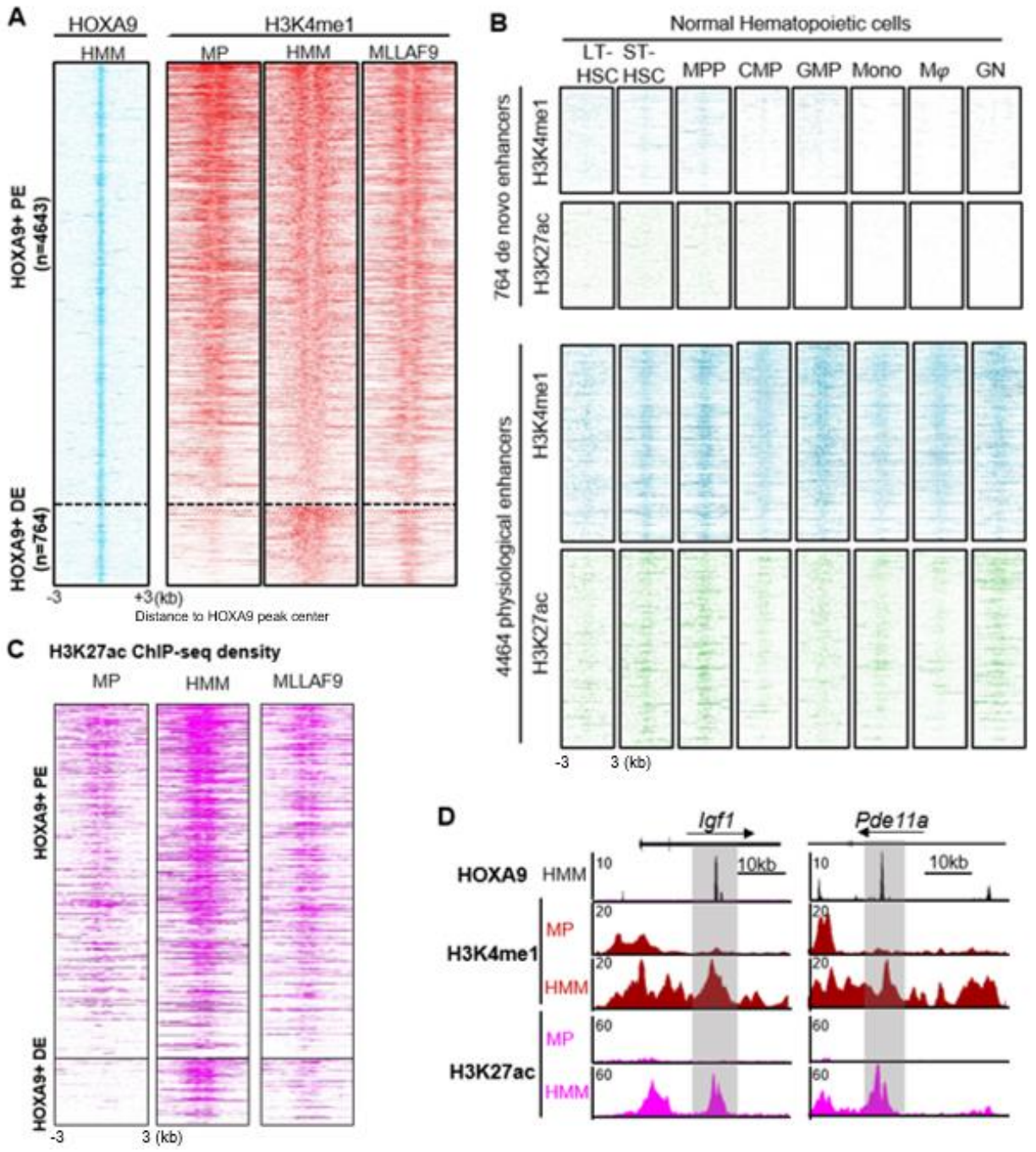

Figure 0-5 HOXA9+ DE are specifically active in HOXA9-dependent AML cells
(A) Heatmap depicting the corresponding signal intensity of HOXA9 in HMM cells and H3K4me1 in MP, HMM and MLLAF9 cells at 5,407 HOXA9-bound distal regulatory elements (HOXA9+ PE and DE). The rows show 3kb upstream and downstream of HOXA9 peak center. Peaks sorted based the total normalized H3K4me1 tag counts within each category. (B) Heatmap depicting signal intensity of H3K4me1 and H3K27ac at 788 HOXA9+ DE and 4643 HOXA9+ PE in normal hematopoietic cells of different differentiation stages into the myeloid lineage. (C) Heatmap depicting the corresponding H3K27ac signal intensity in MP, HMM and MLLAF9 cells at HOXA9+ PE and DE. (D)

UCSC browser view of HOXA9, H3K4me1 and H3K27ac profile before and after HOXA9-mediated transformation at HOXA9+ DE in *Igf1* (left) and *Pde11a* (right) intronic regions.

To further confirm the functional relevance of HOXA9-mediated DE in leukemogenesis, we examined H3K4me1 at HOXA9+ PE and DE in MLL-AF9 leukemia cells, which are driven by MLL-AF9 fusion proteins and aberrantly over express HOXA9 [112]. As shown in Figure 2.5A and C, both H3K4me1 and H3K27ac distribution patterns at HOXA9-bound enhancers were strikingly similar to that of HMM cells. The lower H3K4me1 at DE in MLL-AF9 cells was probably due to relatively low HOXA9 level in these cells as compared to HMM cells (data not shown). These results suggest that establishment of DE is a fundamental mechanism of transformation by HOXA9.

2.3.4 DE and PE have different dependency on HOXA9

To establish the causal relationship between HOXA9 and the establishment of DE in HOXA9/MEIS1-driven leukemia, we employed the inducible form of HMM cells. In this cell line, HOXA9 is constitutively overexpressed but activated only in the presence of tamoxifen (OHT). As a result, OHT withdrawal causes HOXA9 inactivation and cellular differentiation [61]. Importantly, we observed significant reduction of H3K4me1 at HOXA9+ DE 72 hours after HOXA9 inactivation (Figure 2.6A, right panel, p=2.40E-14). On the contrary, H3K4me1 slightly increased at HOXA9+ PE after HOXA9 inactivation (Figure 2.6A, left panel, p=8.02E-5). Distinct and reversible requirement of HOXA9 for H3K4me1 at DE, as compared to PE, prompted us to examine the binding of other transcription factors at these genomic loci. Previous studies show that C/EBPα mediates the creation of myeloid specific enhancers [145, 146, 185] and functionally collaborates with HOXA9 for the initiation of AML [54, 186] . To examine C/EBPα binding at HOXA9+

DE and PE, we performed ChIP-seq analyses for C/EBPα in HMM cells with or without HOXA9 inactivation. ChIP-seq in the untransformed MP cells were used as the control. As shown in Figure 2.6B, C/EBPα bound at relatively high level at HOXA9+ PE in both MP and HMM cells regardless of the transformation status. In accordance, its binding at PE was not significantly affected by HOXA9 inactivation. This is foreseeable as C/EBPα occupies a large subset of enhancers in lineage-committed myeloid cells [185]. Interestingly, C/EBPα binding at HOXA9+ DE was completely dependent on HOXA9: C/EBPα bound at an extremely low level at DE in MP cells and significantly increased upon HOXA9 over expression and leukemic transformation. Furthermore, C/EBPα binding at DE was drastically reduced upon HOXA9 inactivation (Figure 2.6B, right panel). Figure 2.6C shows two representative genes, *Ikzf2* and *Crhbp*, regulated by HOXA9+ DE with both increased H3K4me1 and C/EBPα binding upon transformation. Together, these results suggest that HOXA9 has distinct functions at two classes of enhancers: 1) at DE, it acts as a pioneer transcription factor to recruit other transcription factor(s) and mediate the incorporation of enhancers marks in leukemia cells; and 2) at PE, it exhibits opportunistic binding and is dispensable for the formation and maintenance of these physiological regulatory elements.

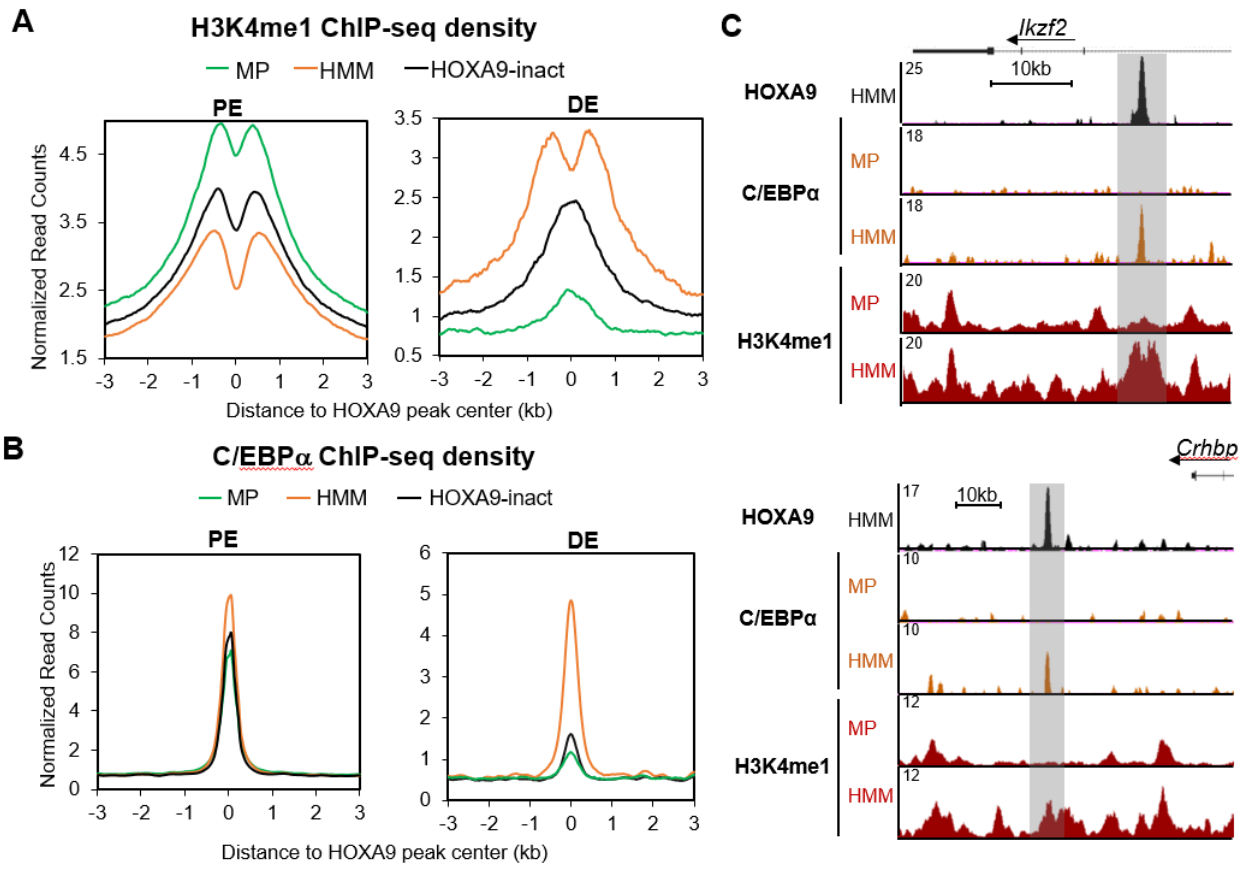

Figure 0-6 HOXA9 is essential for H3K4me1 maintenance and C/EBPα binding at DE
(A) Composite plots showing the change of average H3K4me1 signal at HOXA9+ PE and DE in HMM and in HOXA9-inactivated HMM cells (HOXA9-inact). Statistics obtained by *K-S* test (B) Composite plots showing the change of average C/EBPα signal at HOXA9+ PE and DE in MP, HMM and HOXA9-inact cells. (C) UCSC browser view of HOXA9, H3K4me1 and C/EBPα profile before and after HOXA9-mediated transformation at HOXA9+ DE at Igf1 (left) and Pde11a (right) intronic regions.

2.3.5 HOXA9 ectopically activates developmental programs with *de novo* enhancers

To evaluate whether the two classes of HOXA9-bound enhancers exert differential functions, we performed Gene Ontology (GO) analyses on the associated genes. This showed that while PE and DE were equally enriched for several common pathways such as regulation of metabolic processes and cell signaling, they diverged significantly for others. The most conspicuous difference was strong enrichment of embryonic or early

developmental pathways for HOXA9-bound DE (Figure 2.7A). This result was consistent with reactivation of the embryonic gene program in MLL rearranged leukemia [187] as well as the critical roles of *Hox* genes in early development [188]. RNA-seq analyses for MP and HMM cells showed that genes associated with DE were significantly upregulated in HMM cells as compared to MP cells (Figure 2.7B), and this upregulation was mitigated 72 hours after HOXA9 inactivation (HOXA9-inact). In contrast, expression of genes associated with PE had no significant differences among MP, HMM and HOXA9-inact cells. The strong association of DE, but not PE, with embryonic gene programs, as well as prominent up regulation of DE regulated genes in HMM cells, suggests that the HOXA+ DE-associated genes are the main effectors of HOXA9-regulated program in leukemogenesis.

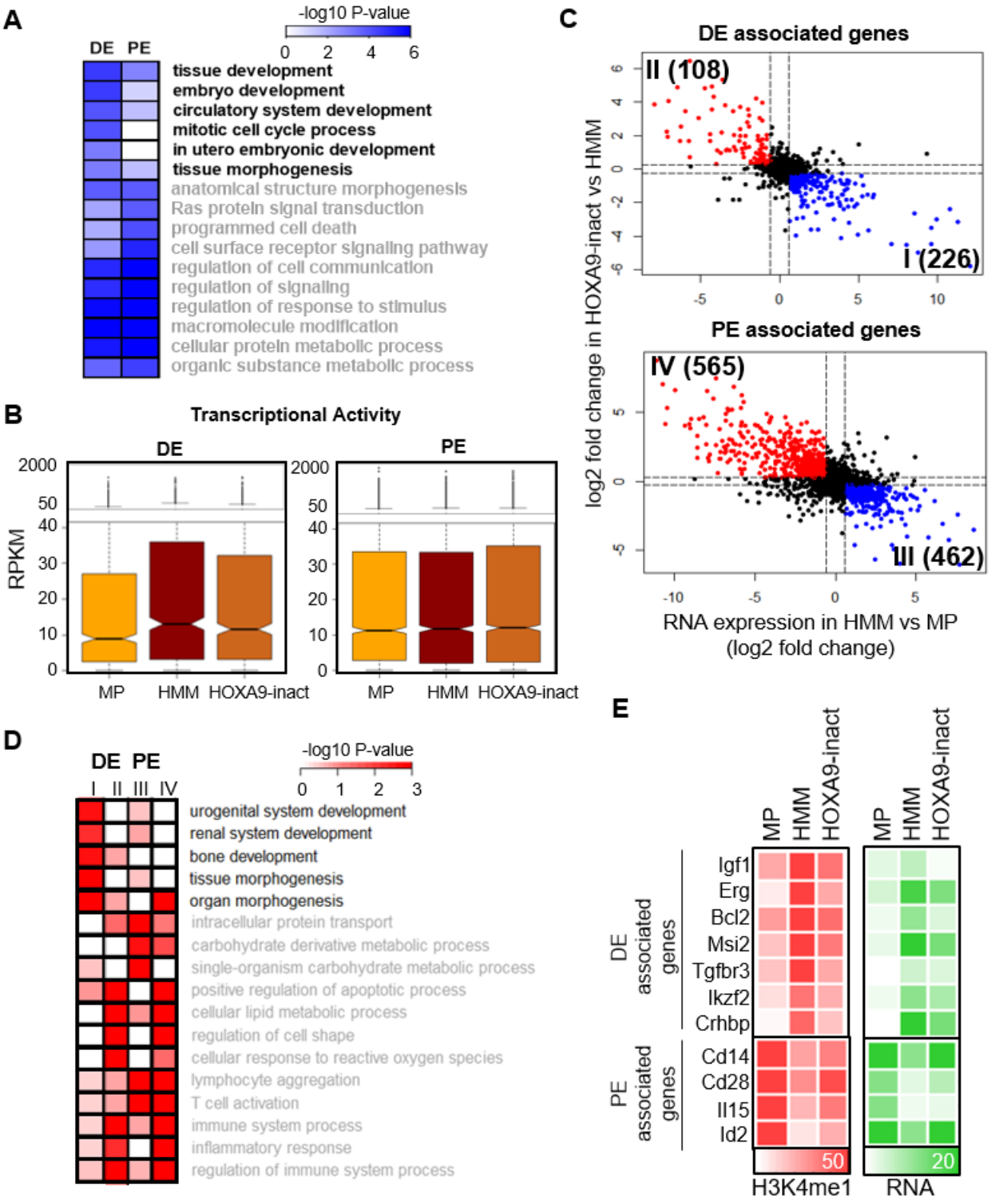

Figure 0-7 HOXA9 activates embryonic development and organogenesis pathways with *de novo* enhancers.

(A) Gene Ontology Biological Processes for HOXA9+ PE and DE-associated genes with the corresponding Benjamini p-values. Pathways specifically enriched with DE are highlighted. (B) Transcriptional activity (RPKM) of PE and DE-associated genes in MP, HMM and HOXA9-inact cells, respectively. P-values are obtained from Mann-Whitney U-test. (C) Transcriptional changes in genes associated with DE and PE are shown in the scatter plots. The log2 fold change in HMM vs. MP is plotted on the x axis, and log2 fold change in HOXA9-inact vs HMM on the y-axis. Upper panel: DE-associated genes. Lower panel: PE-associated genes. Blue: $x > \log2(1.5) = 0.585$ and $y < -\log2(1.2) = -0.263$. Red: $x < -\log2(1.5) = -0.585$ and $y < -\log2(1.2) = -0.263$. I: DE-associated HOXA9 activated genes. II: DE-associated HOXA9 repressed genes. III: PE-associated HOXA9 activated genes. IV: PE-associated HOXA9 repressed genes. (D) Gene Ontology Biological Process terms associated with the Class I, II, III and IV genes. Pathways specifically enriched with I are highlighted. (E) Heatmap showing the normalized read counts of H3K4me1 at enhancers and RPKM of representative genes in Class I and Class IV. If multiple HOXA9+ enhancers are found, the one nearest to gene promoter was selected.

As shown in Figure 2.7C, for both HOXA9+ DE and PE, most genes upregulated by HOXA9/MEIS1-mediated transformation were down-regulated by HOXA9 inactivation (HOXA9-activated targets). Genes down-regulated by the transformation were more likely to be up-regulated by HOXA9 inactivation (HOXA9-repressed targets). This reversion in the transcriptome is consistent with the finding that HOXA9 inactivation leads to an exit of the leukemogenic status in HMM cells, as well as partial myeloid differentiation [61]. Based on this result, we classified HOXA9+ DE- and PE-associated genes in to four subgroups based on their responses to HOXA9. For DE, more genes were activated by HOXA9 than repressed (Group I: 226 vs. Group II: 108); for PE, fewer genes were activated by HOXA9 than repressed (Group III: 462 vs. Group IV: 565). This distinction in transcription changes between DE and PE is consistent with the change in H3K4me1. Furthermore, when GO analysis was performed on the four groups of genes, we found that multiple organogenesis pathways were specifically enriched with Group I genes, while immune and apoptotic responses were enriched with Group II and IV (Figure 2.7D). Representative genes from group I as well as group II/IV were shown in Figure 2.7E. These results suggest that *Hoxa9* over expression mediates leukemic transformation by

establishing a new epigenomic landscape in support of a primitive/embryonic transcriptome. Notably, several genes previously implicated in hematological malignancies (e.g *Igf-1* [114, 115], *Bcl-2* [189, 190], and *Erg* [73, 113]) were identified as direct HOXA9 downstream targets, further supporting the central role of HOXA9 in leukemic transformation.

2.3.6 HOXA9 regulates *Aldh1a3* expression with *de novo* enhancers

The gene that showed strongest activation with HOXA9+ DE is *Aldh1a3,* a key component of the retinoic acid metabolism pathway [191] and a therapeutic target in solid tumors [192-195]. We used *Aldh1a3* as a model to examine long distance gene regulation by HOXA9-dependent *de novo* enhancers. As shown in Figure 2.8A, *Aldh1a3* expression was significantly upregulated in HMM cells and downregulated upon HOXA9 inactivation. To examine the functional requirement of *Aldh1a3* in HMM cells, we employed an inducible knock-down system to specifically target its expression. Decreased expression of *Aldh1a3* led to reduced cell proliferation (Figure 2.8B) and colony formation on methylcellulose (Figure 2.8C), suggesting a critical role of this gene for the survival and stemness of HMM cells.

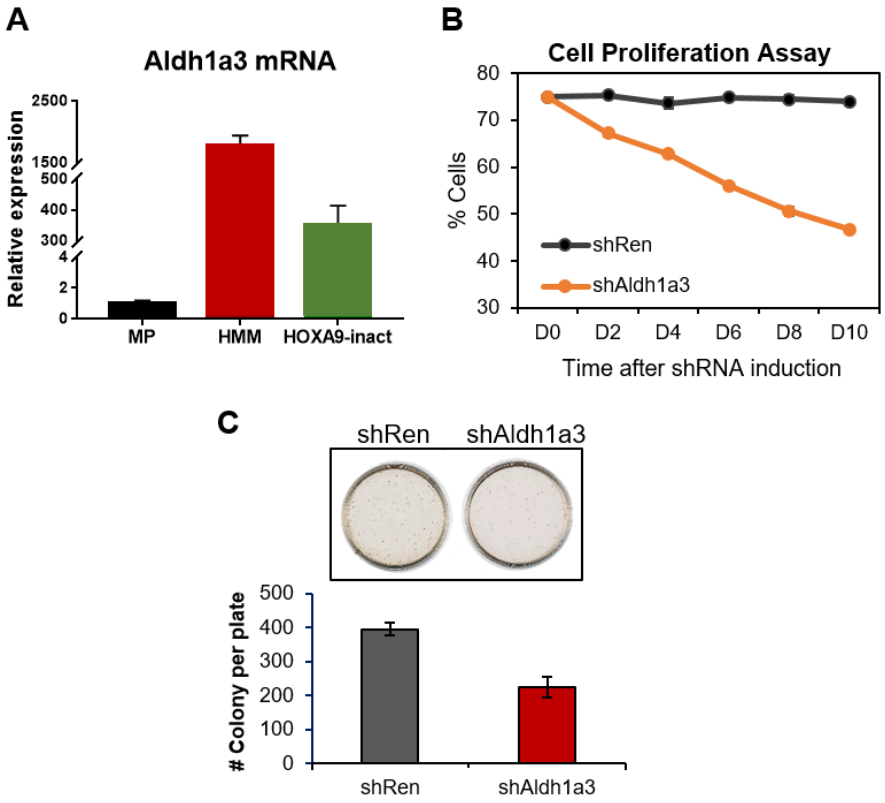

Figure 0-8 Loss of Aldh1a3 impairs cell proliferation and colony formation
(A) Bar plot showing the normalized expression level of *Aldh1a3* in MP, HMM and HOXA9-inact. Transcription level of *Aldh1a3* in MP cells was set as 1. (B) Percentage of shRNA positive cells in the 10-day course after the induction of Aldh1a3 shRNA or non-targeting Renilla shRNA. shRNA-expressing cells are 3:1 mixed with parental cells at Day 0, and their percentage constitution was monitored with flow cytometry. (C) CFU-assay in Methylcellulose showing the reduction of colony-forming units after *Aldh1a3* knockdown.

We identified three HOXA9+ DE upstream of *Aldh1a3* gene promoter at -53kb, -77kb and -118kb respectively. All three DEs were highly enriched for the active enhancer signatures, H3K4me1 and H3K27ac, in HMM cells (Figure 2.9A). Notably, several chromatin regions near the *Aldh1a3* promoter also had HOXA9-dependent H3K4me1 up regulation during leukemic transformation but had little or no direct HOXA9 binding (Figure 2.9A, regions 1). Interestingly, region 1, which was indirectly regulated by HOXA9, had low H3K27ac as compared to regions 2-4. To establish that HOXA9-bound

enhancers play a causal role in regulation of *Aldh1a3* expression, we first performed Circular Chromatin Conformation Capture with High-throughput Sequencing (4C-seq) in MP, HMM and HMM cells with HOXA9 inactivation. With *Aldh1a3* TSS as the view point, we found that regions 1, 2, and 4 had significant long-distance interactions with the TSS (Figure 2.9B, top panel). These interactions were HOXA9-dependent since HOXA9 inactivation reduced their interaction frequency (Figure 2.9B, bottom panel), indicating that HOXA9 is required for the effective looping of region 1, 2 and 4 to *Aldh1a3* promoter. In contrast, region 3 showed no interaction with *Aldh1a3* TSS in any cell line, suggesting that it is not a distal regulatory enhancer for this gene. We next used CRISPR/Cas9-mediated genome editing to delete each of regions 1 to 4 in HMM cells. We also targeted the *Rosa26* promoter with the same CRISPR strategy and used it as a negative control. After confirmation of specific deletion of the genomic regions (Supplemental Figure 6B), we examined *Aldh1a3* expression by qRT-PCR. As shown in Figure 5F, deletion of regions 2 and 4 completely abolished *Aldh1a3* expression, suggesting that they play non-redundant functions in *Aldh1a3* regulation. Interestingly, deletion of region 1, which was not bound by HOXA9, only modestly affected *Aldh1a3* expression. Together, these results strongly support that HOXA9 plays a direct and causal role in promoting long-distance interactions between distal regulatory enhancers and gene promoters, which is necessary for activation of the leukemic transcriptional program.

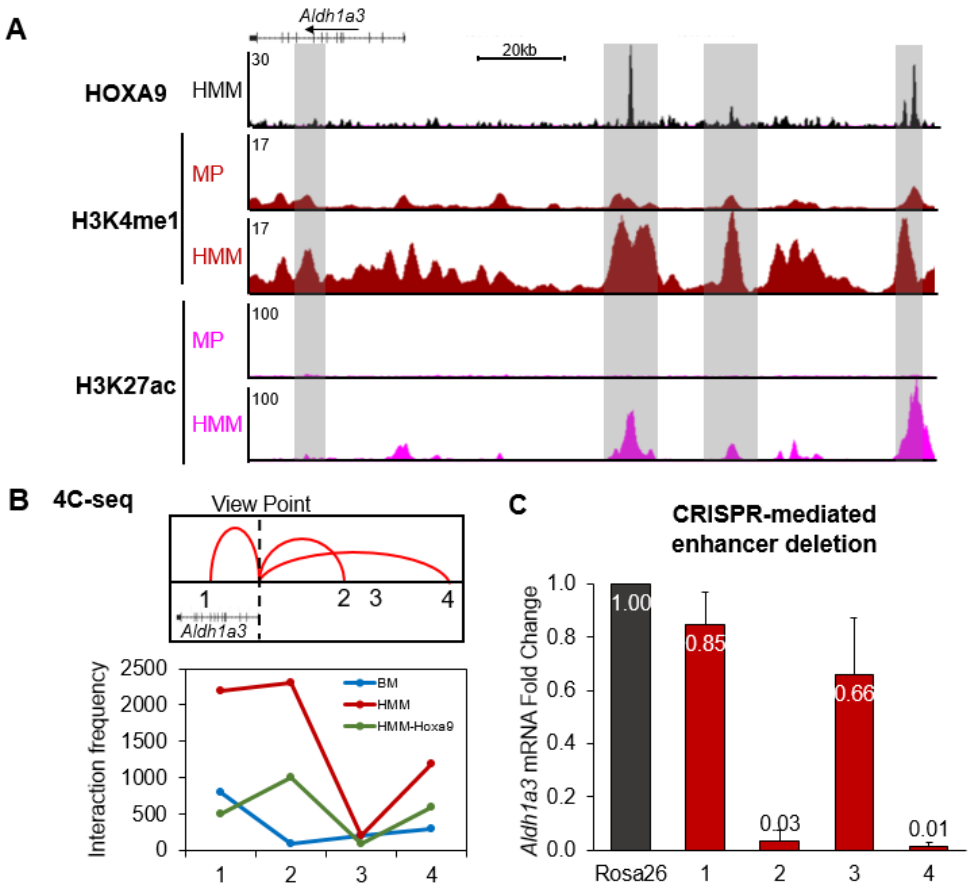

Figure 0-9 HOXA9 regulates *Aldh1a3* expression via two HOXA9+ DE
(A) ChIP-seq binding profile of HOXA9, H3K4me1 and H3K27ac around *Aldh1a3* TSS and 150Mbp upstream region. Library size normalized to 1E7 reads. Three HOXA9+ DE annotated with *Aldh1a3* TSS (2, 3 and 4) as well as a no-HOXA9 enhancer region (1) are shaded in grey. (B) 4C-seq analysis showing the interaction frequencies between *Aldh1a3* TSS and the four enhancer regions of interest. Library size normalized to 1E7 reads. (C) qRT-PCR analysis of relative *Aldh1a3* expression after CRISPR-mediated deletion of enhancer 1, 2, 3 and 4. Transcription level of *Aldh1a3* in the *Rosa26*-targeting cells was set as 1.

2.4 Discussion

In this Chapter, we systematically characterized the genome-wide enhancer state alterations in HOXA9/MEIS1-transformed AML cells, and discovered that the leukemic transformation induces a global shift in enhancer state. By recruiting other transcription

factors, HOXA9 mediates the formation of *de novo* enhancers, which are in turn required for the activation of the specific transcription program in leukemia.

2.4.1 The pioneer role of HOXA9 in the creation of *de novo* enhancers

It has been previously reported that HOXA9 binding sites are enriched for active enhancer modifications in leukemia cells [61]. However, it was unclear whether HOXA9 passively binds on pre-accessible enhancer regions or targets inactive chromatin to induces the establishment of novel enhancers during the transformation process. Since HOX proteins all recognize a very similar set of ‘‘AT’’-rich sequence motifs, which seem to be insufficient for its precise instructive role *in vivo* [41], their binding is considered highly dependent on DNA accessibility and collaborator proteins. Several co-binding partners of HOXA9 have been identified as pioneer factors in different cellular contexts, such as PBX1 in breast cancer [196] and C/EBPa in myelopoiesis [145]. Based on our current findings, we argue that HOXA9 also acts as a pioneer factor at specific loci, when it is pathologically overexpressed. While a large fraction of HOXA9 binds on pre-established enhancers and is dispensable for the H3K4me1 implementation at those sites, a subset of HOXA9 travels to novel regions with no prior H3K4me1 and directly mediates the creation of active enhancers. Induction of these sites is achieved despite the lack of enhancer formation in MP cells, its putative cell-of-origin, or any other progenitor and mature cells in hematopoiesis. At these loci, it recruits, instead of being recruited by, C/EBPa and mediates H3K4 monomethylation. HOXA9’s pioneer binding is demonstrated as vital for the stable assembly of transcriptional machinery such as C/EBPa at the *de novo* sites.

HOXA9 on its own is unlikely sufficient to remodel chromatin and recruit methyltransferase and acetyltransferase; it likely plays this role in concert with its transcription cofactors, such as MEIS1 and PBX1, since co-over expression with MEIS1 is required for its leukemogenesis. Multimeric complex assembly is likely essential for the complete transition from closed to open chromatin configuration. To fully characterize the pioneer role of HOXA9, chromatin accessibility changes upon HOXA9 activation and the percent of HOXA9 targeting inaccessible regions need to be evaluated.

2.4.2 The functions of *de novo* enhancers

The genes activated by HOXA9+ *de novo* enhancers are specifically enriched for tissue morphogenesis and multiple organ developmental pathways. Similar activation of embryonic program has also been found in HOXA9-dependent MLL-AF9 leukemia. This is consistent with *Hox* genes' role in specifying cell fate and controlling organogenesis in early development [7], and that misexpression of *Hox* genes causes homeotic transformation [4, 5]. In neoplasia, abnormal expression of *Hox* genes leads to differentiation failure, altered characteristics and adoption of an alternative cell fate [197]. In this model, over expression of *Hoxa9* ectopically activates the transcriptional program underlying a different cell fate, such as the development of vasculature system, while suppresses hematopoiesis fate in the bone marrow progenitors. Thus, the *de novo* enhancers in HOXA9-mediated leukemias are likely to bear regulatory functions in other developmental lineages. Indeed, 80% of the HOXA9+ *de novo* enhancers are identified with H3K4 monomethylation in at least one normal tissue type curated in the ENCODE database (Data not shown), suggesting that 1) these leukemia cells

abnormally acquire characteristics of other lineages and 2) *de novo* regions possess regulatory potential in other lineage contexts or developmental stages. With the exception of *Aldh1a3* enhancers tested here, the exact function of each individual HOXA9+ *de novo* enhancer in both development and leukemogenic transformation remains largely unclear, and warrants further investigation.

Chapter 3

The MLL3/MLL4 complex collaborates with HOXA9 to promote the development of leukemia

3.1 Introduction

Histone modifications, such as methylation and acetylation, constitute one of the most critical players in the regulation of gene expression. The monomethylated histone H3 lysine 4 (H3K4me1) broadly marks multiple classes of enhancer regions, which are further demarcated by the modifications on H3 lysine 27 into active, primed and poised subtypes. Genome-wide aberrations of H3K4me1 have been noted in several types of malignancies [198-201], which are linked to the specific transcriptional programs in oncogenic transformation.

Histone-lysine *N*-methyltransferase (KMT2) family proteins catalyze the addition of methyl group(s) on lysine 4 of the histone H3 tail. They are also named MLL family proteins, because their founding member KMT2A (MLL1) was first identified undergoing genetic rearrangement in mixed-lineage leukemia. For the sake of consistency, here I use the old nomenclature, mouse MLL1-4 or human MLL1,4,3,2, to refer to KMT2 A-D. The SET domain of MLL family members have histone methyltransferase activity. Although their substrate specificity is still being investigated, a multitude of studies have shown that MLL3 and MLL4 are predominantly monomethyltransferases, while MLL1

and MLL2 are capable of mediating H3K4 mono-, di- and tri-methylation [202]. MLL proteins act in large complexes composed of both common and specific subunits. The four subunits common to all MLL complexes are WD repeat protein 5 (WDR5), Set1/Ash2 histone methyltransferase complex subunit ASH2 (ASH2L), retinoblastoma-binding protein 5 (RBBP5) and DPY30 [203]. Specific subunits, such as menin in the MLL1/MLL2 complex and the PTIP in the MLL3/MLL4 complex, are employed to target these complexes to certain genomic loci [204].

Not only are MLL family proteins frequently mutated in human solid and blood cancer, their wild type form has also been implicated in the epigenetic dysregulations underlying several types of hematological neoplasia. The role of MLL proteins in tumorigenesis seems to be complex and context-dependent, since genetic alterations resulting in loss of the protein function are able to both accelerate and repress oncogenic transformation. This dual role can be well illustrated with MLL4 (KMT2D). In diffuse large B cell lymphoma, its genetic ablation or loss of function mutations promote lymphoma development, and it thus serves as a tumor suppressor, in both mice and in human patients [201, 205]. However, in the MLL-AF9-driven acute myeloid leukemia model, MLL4 is required for the maintenance of cancer stem cell properties and the rapid onset of leukemia in mice [206]. Moreover, specific targeting of the SET domain of MLL4 impairs the proliferation of MLL-AF9-transformed cells, indicating that the methyltransferase activity of MLL4 is indispensable for this type of leukemia [162]. By contrast, SET-domain deletion in MLL1 fails to inhibit MLL-AF9-induced leukemia initiation [207], suggesting that members of MLL family may be differentially utilized in the leukemic transformation process.

In the previous chapter, I systematically assessed the leukemic enhancer landscape changes induced by HOXA9 over expression. I found that the leukemia-specific, *de novo* enhancers drive a leukemia-specific transcription program and are most responsive to HOXA9 over expression and inactivation. Subsequently, I set out to explore the mechanisms by which the *de novo* enhancers are established, and identify a collaborating chromatin regulating complex. As HOXA9 itself is difficult to target therapeutically, determining the collaborating epigenetic modulators may provide further mechanistic insights for development of epigenetic therapy for acute myeloid leukemia.

3.2 Materials and Methods

Mice

For *Ptip* or *MLL4* SET domain deletion *in vitro* and *in vivo* assays, C57BL/6 mice (WT), *Ptip* $^{f/f}$ [208] or *Mll4*-SET$^{f/f}$ (Kai Ge lab, unpublished) mice were crossed with B6.129-*Gt(ROSA)26Sor*$^{tm1(cre/ERT2)Tyj}$/J mice (JAX no. 008463; The Jackson Laboratory) to obtain *Ptip* $^{f/f}$;*CRE-ER* $^{+/-}$, *Mll4*-SET$^{f/f}$; *CRE-ER*$^{+/-}$ or WT;*CRE-ER* $^{+/-}$ mice.

Antibodies

For Western blot analysis, anti-HOXA9 (07-178, Millipore), rabbit anti-PTIP and anti-MLL4 #3 antibody were used [175]. For ChIP, anti-HA (ab9110; Abcam), anti-MLL4 #3, anti-H3K4me1 (ab8895; Abcam), and IgG (sc-2027; Santa Cruz Biotechnology) were used. For flow cytometry, allopycocyanin (APC) anti-Gr1 (108412; Biolegend), PE anti-CD11b (101208, Biolegend), PE anti-CD135 (135305,Biolegend) and DAPI (Sigma) were used.

Cell Lines

Bone marrow from 6- to 10-week-old *Ptip* $^{f/f}$;*CRE*$^{-}$, *Ptip* $^{f/f}$; *Cre-ER*$^{+/-}$, WT;*Cre-ER*$^{+/-}$, mice was harvested 5 d after treatment with 5-flurouracil (150 mg/kg). Lineage-negative bone marrow cells were first flushed from femora and tibiae with 25G needles, and then isolated using the EasySep Mouse Hematopoietic Progenitor Cell Enrichment Kit (19856, Stem Cell Technologies). *Lin*$^{-}$ cells were maintained in Iscove's modified Dulbecco's medium (IMDM) supplemented with 15% Fetal Bovine Serum (FBS, Sigma F4135), 10 ng/mL Interleukin (IL) -3, and 100 ng/mL stem cell factor (SCF).

To package retroviruses, Plat-E cells (RV-101, Cell Biolabs) were transfected with MIGR1-HA-*Hoxa9* or with MIGR1-Flag-*Meis1* retroviral vectors [plasmids previously described by [61]] using FuGENE 6 (E2691, Promega). Cell-free supernatant was collected 48 hours after transfection. To overexpress *Hoxa9* and *Meis1*, *Lin*$^{-}$ BM were spinoculated with *Hoxa9* and *Meis1* retrovirus together at 3200rpm for 90mins at room temperature on two consecutive days.

Immunoprecipitation

HMM or MLLAF9 cells were washed with DPBS and lysed in M-PER regent (78501, Thermo Fisher) supplemented with 10µl/ml Halt Protease Inhibitor Cocktail (87786, ThermoFisher) and 10µl/ml Halt Phosphatase Inhibitor Cocktail (78428, ThermoFisher) at 4°C for 30mins. Nuclear extracts were pelleted and treated with 500U/ml Benzonase Nuclease (07664, Millipore) at 4°C for 30mins. Nuclease activity was terminated with 5mM EDTA. After the debris was pelleted at 4°C for 30mins at 14,000g, the supernatant was pre-cleared with rat IgG-AC (sc-2344, Santa Cruz) at 4°C for 2 hours, then incubated with anti-HA agarose beads (11815016001, Roche)

overnight with rotation. The beads were washed twice each with M-PER reagent containing zero, 150mM and 300mM NaCl. The beads were eluted by heat denaturation at 95°C for 5mins. The protein contents were examined by western blotting.

Flow Cytometry

For surface marker expression, cells were washed and resuspended in standard media (2% FBS and 0.1% NaN3 in DPBS), and incubated for 30 min on ice with 0.2µg of the appropriate antibodies. After incubation, cells were washed twice before analyzed on a Becton Dickinson LSR II. Data collected from at least 20,000 events from biological replicate experiments were analyzed using FlowJo Version 10 (TreeStar).

***in vivo* leukemogenesis assay**

To induce *Ptip* deletion prior to transplantation, *Hoxa9*/*Meis1*-transformed *Ptip*$^{f/f}$;*Cre-ER*$^{+/-}$ cells were treated continuously with 5 nM tamoxifen (OHT, H7904; Sigma) to generate *Ptip*$^{-/-}$ cells. Then *Ptip*$^{-/-}$ or *Ptip*$^{f/f}$ cells (no CRE) were injected via tail vein in cohorts of lethally irradiated (900rad) 8-week-old female mice (1.25×10^5 cells per mouse). To induce *Ptip* deletion after engraftment, WT;*CreER*$^{+/-}$ cells or *Ptip*$^{f/f}$;*Cre-ER*$^{+/-}$ were directly injected into irradiated C57BL/6 female recipient mice. Mice in all groups were maintained on antibiotics for 2 weeks post-irradiation. After 14 days, mice were treated with intraperitoneal injections of OHT (200 mg/kg) twice a week until becoming moribund. Moribund mice were euthanized. Liver, spleen, heart and bone were harvested from control and leukemic mice at the time of death for paraffin embedding and HEMA 3 staining. Bone marrow was flushed for collecting RNA, whole cell lysate,

and cytospin samples. Survival curves were plotted in Prism (GraphPad), and statistical significance was evaluated by log rank test.

ChIP

ChIP assays were conducted as described in [54]. Briefly, cells were fixed for 15 min at room temperature with 1% paraformaldehyde, lysed in lysis buffer [1% SDS, 10 mM EDTA, 50mM Tris·HCl (pH 8)] for 10min on ice, and sonicated for 15 min twice using Bio-raptor. Diluted chromatin was incubated with 2.5µg of appropriate antibody overnight at 4 °C with rotation. Immunoprecipitation was then performed using protein G Dynabeads (10004D, Thermo Fisher). Immunoprecipitates were washed for 5 min in low-salt (150 mM), high-salt (500 mM), and lithium chloride (0.25 M) buffers, and twice with Tris/EDTA buffer. Captured chromatin was eluted by incubating beads in 250µL of elution buffer (1% SDS, 100 mM $NaHCO_3$) for 30 min at 42 °C. Cross-linking was reversed by the addition of 50µM NaCl and overnight incubation at 65 °C. Chromatin was then RNase A-treated and purified using a Qiagen PCR purification kit. Binding was quantified relative to input by quantitative PCR (7500 PCR System; Applied Biosystems) using SYBR green fluorescent labeling and primers designed using the Integrated DNA Technologies PrimerQuest program.

ChIP-Seq, peak calling and density calculation

Multiplexed ChIP-seq libraries were prepared at the University of Michigan DNA Sequencing Core. 50-cycle single-end sequencing runs were performed on Illumina HiSeq 2500 at a sequencing depth of 10-50 million aligned reads per sample. Sequenced reads were preprocessed to trim adaptor sequences (Trimmomatic) and

then aligned to mouse reference genome (mm9) using BWA software (version 0.6.2). Only uniquely mapped reads were used in downstream analyses. Model-based Analysis for ChIP-seq (MACS) was used the identification of ChIP-seq peaks with p value=10^{-4} for MLL3/MLL4. Peaks of histone modifications were identified with the default parameters. For composite plots, we counted the per bp tag density within 6kb around peak centers using the annotatePeaks.pl function in HOMER suite (http://homer.salk.edu/homer/), normalizing to 1E7 reads per library.

3.3 Results

3.3.1 HOXA9 recruits H3K4 methyltransferase MLL3/MLL4 to the *de novo* enhancer sites.

The ability of HOXA9 to establish active enhancer signatures at *de novo* binding sites raised the question of which histone modifying enzymes are responsible for HOXA9-dependent chromatin changes. Although previous studies have shown that HOXA9 is able to recruit histone acetyltransferase p300 via its cofactor MEIS1 [62], it remains unclear which H3K4 methyltransferase is required for HOXA9-mediated enhancer modification. Among the six MLL family H3K4 methyltransferases in mammals, MLL1, MLL3 and MLL4 are able to deposit H3K4me1 at distal enhancers [202]. Furthermore, our previous studies showed that inhibiting MLL1 methyltransferase activity had minimal effects on HOXA9/MEIS1 leukemia [89]. Based on these results, we posited that HOXA9 functions through the MLL3/MLL4 complex to establish H3K4me1 enhancer signature in HMM cells. Indeed, when we immunoprecipitated HA-HOXA9 from HMM cells, both MLL3 and MLL4 as well as their cofactor PTIP were detected with immunoblot (Figure 3.1A).

This interaction was specific, since neither MLL3/MLL4 nor PTIP was detected from HA-IP using control MLL-AF9 cells (Figure 3.1A).

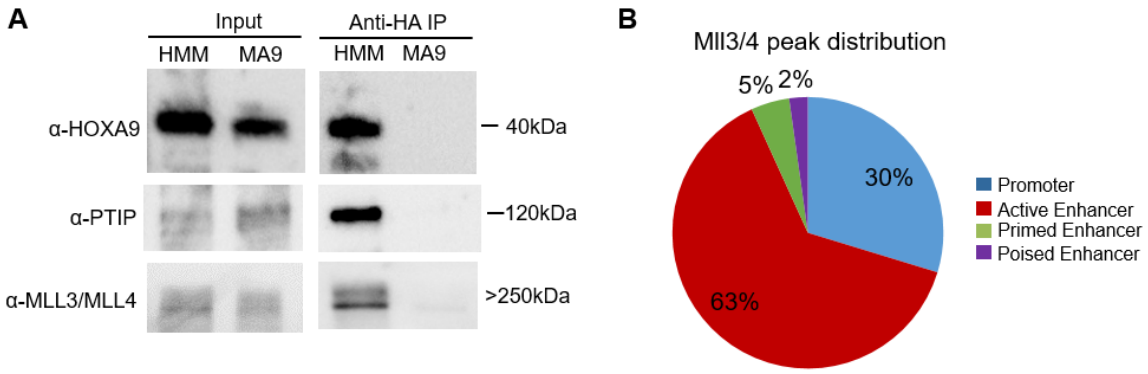

Figure 0-1 MLL3/MLL4 binds active enhancers and co-purifies with HOXA9.
(A) Co-immunoprecipitations performed with anti-HA antibody in HMM cells. Immunoblots show that PTIP and MLL3/MLL4 both immunoprecipitate with HOXA9. HOXA9 is HA-tagged in HMM cells, but is untagged in MA9 cells. (B) The genomic distribution of MLL3/MLL4 peaks

To confirm that the MLL3/MLL4 complex functions in conjunction of HOXA9/MEIS1 in HMM cells, we performed ChIP-seq for MLL3/MLL4 using a previously characterized antibody [209], which recognizes both MLL3 and MLL4 proteins (Kai Ge, unpublished data). Consistent with previous studies [210, 211], the majority of MLL3/MLL4 chromatin binding was at active enhancers in HMM cells (Figure 3.1B). Strikingly, a significant portion (39%) of MLL3/MLL4 peaks overlapped with those of HOXA9 (Figure 3.2A). Representative ChIP-seq signals for HOXA9 and MLL3/MLL4 are shown in Figure 3.2B. Co-localization of HOXA9 and MLL3/MLL4 was further validated by quantitative ChIP-qPCR (Figure 3.2C and D, black bars). To confirm that MLL3/MLL4 functions downstream of HOXA9 in enhancer regulation, we performed MLL3/MLL4 ChIP-seq in HMM cells after HOXA9 inactivation. Interestingly, we found that while MLL3/MLL4 binding were not changed at HOXA9-bound PE, they were drastically reduced at HOXA9-bound DE upon

HOXA9 inactivation. This distinction is shown at selected HOXA9+ DE and PE loci (Figure 3.2 D). Regions 1-4 were examined for *Aldh1a3* regulation activity in Chapter 2.3.6. Among them, regions 2 and 4 were confirmed to be HOXA9-dependent *de novo* enhancers that are required for *Aldh1a3* upregulation in HMM cells. Interestingly, these loci had reduced MLL3/MLL4 binding after HOXA9 inactivation (Figure 3.2 D, red arrows). By contrast, the same treatment had no impact on selected HOXA9+ PE loci (Figure 3.2 D, black box), although HOXA9 binding significantly decreased (Figure 3.2 C, black box). Given that MLL3/MLL4 are considered responsible for implementing the H3K4me1 mark at enhancers, this observation is consistent with the finding in Chapter 2.3.4 that HOXA9 inactivation specifically reduces H3K4me1 at the *de novo* enhancers. Collectively, these results suggest that HOXA9 physically interacts with MLL3/MLL4, and is essential for MLL3/MLL4 recruitment and/or maintenance at HOXA9+ DE.

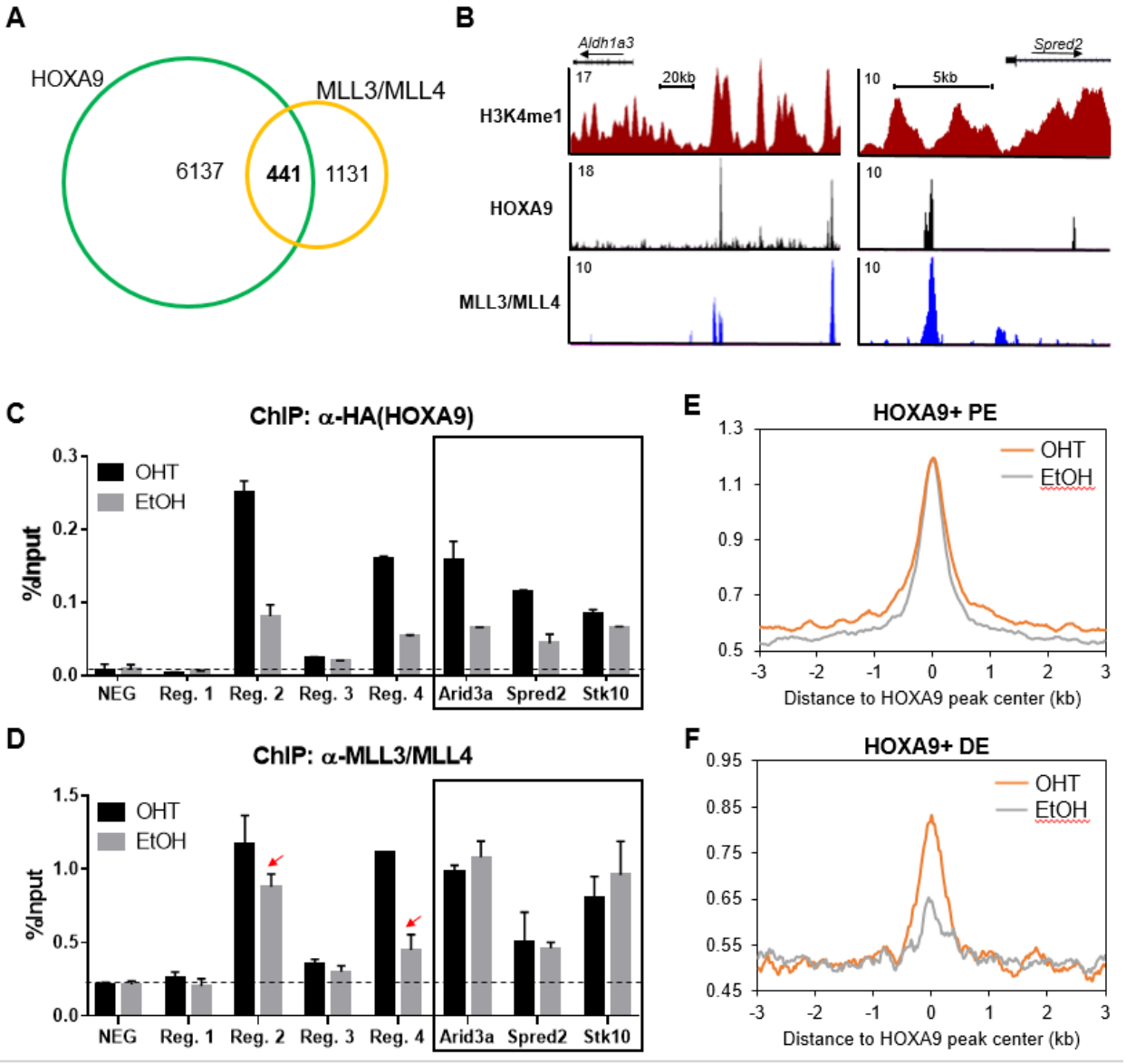

Figure 0-2 HOXA9 colocalizes with the MLL3/MLL4 histone methyltransferase complex, and is required for its binding on *de novo* enhancers

(A) Venn diagram showing the ChIP-seq peak physical overlap between HOXA9 and MLL3/MLL4. Overlap criteria set to be 1bp. (B) Representative plots showing the co-localization of HOXA9 and MLL3/MLL4 at HOXA9+ DE across the *Aldh1a3* and *Spred2* loci. (C-D) Quantitative ChIP assay of HOXA9 and MLL3/MLL4 at *Adh1a3*-proximal regulatory regions 1-4, as well as HOXA9+ PE loci genes (black box), validating MLL3/MLL4's dependency on HOXA9 at DE. Red arrows indicate significant reduction of MLL3/4 binding signal. (E-F) Composite plots depicting the per bp average MLL3/MLL4 signal on HOXA9+ PE and DE after HOXA9 inactivation.

3.3.2 Disruption of the MLL3/MLL4 complex impairs HOXA9/MEIS1-mediated leukemogenesis

Given that HOXA9 physically interacts with MLL3/MLL4, we next examined if MLL3/MLL4 was required for HOXA9/MEIS1-mediated leukemogenesis. To this end, we used the conditional *Ptip*$^{f/f}$; *CRE-ER* mouse model, in which *Ptip* can be conditionally deleted by OHT treatment. PTIP is a core component of the MLL3/MLL4 complex and is essential for MLL3/MLL4 chromatin binding and activity [175, 209, 212]. We first transduced *Ptip* $^{f/f}$ or *Ptip* $^{-/-}$ Lin^{-} BM cells with *Hoxa9*/*Meis1* viruses, and then transplanted them into lethally irradiated recipient mice. All mice (n=8) transplanted with *Ptip* $^{f/f}$ cells developed severe symptoms of leukemia within 38 days, including weight loss, shortness of breath, splenomegaly (Figure 3.3B), heart and liver infiltration (Figure 3.3C), and had high blast percentage in circulation (Figure 3.3C). In contrast, mice transplanted with *Ptip* $^{-/-}$ cells had a significant delay in disease onset and did not succumb to leukemia until after 60 days (Figure 3.3A). Extension of the leukemia latency was also observed even when *Ptip* was excised two weeks after cells were transplanted into syngeneic recipient mice (Figure 3.3D), suggesting that the delayed disease onset was not due to impaired homing or engraftment. Importantly, genotyping results showed that leukemia cells isolated from these mice still had intact *Ptip* $^{f/f}$ alleles; indicating that they were escapees from *CRE-ER*-induced *Ptip* deletion (data not shown). Re-expression of *Ptip* in the leukemia cells was also confirmed by RT-qPCR and immunoblot (Figure 3.4A), which suggests that the few clones that escaped the deletion treatment had higher survival advantage than the bulk of the leukemia cells that had undergone *Ptip* deletion.

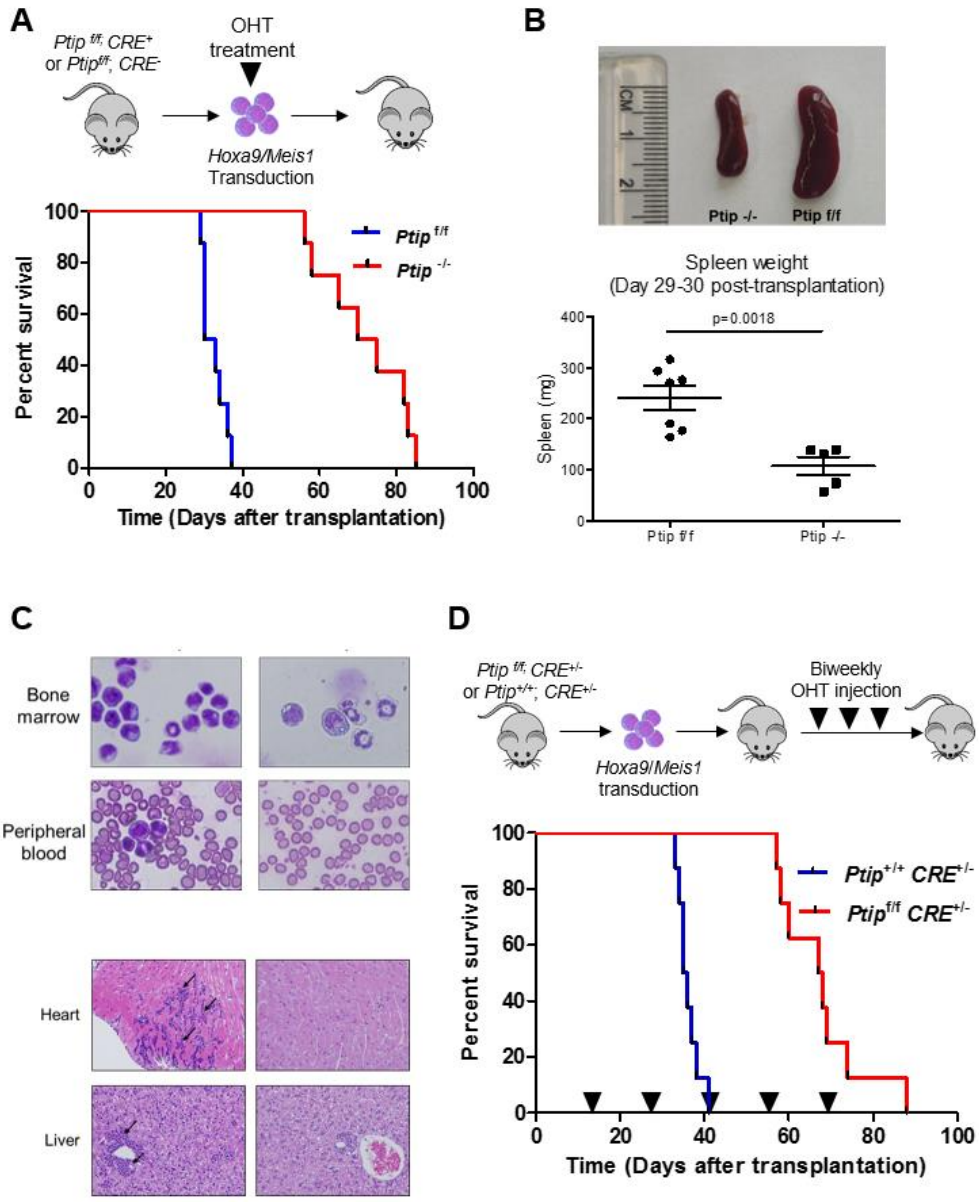

Figure 0-3 Loss of Ptip impairs the leukemogenic ability of HOXA9/MEIS1 cells
(A) Transplantation schematic and survival curve for HOXA9/MEIS1-mediated leukemogenesis in *Ptip* $^{f/f}$ and *Ptip* $^{-/-}$ bone marrow cells. p<0.0001 (log-rank test)
(B) Spleens taken from mice 30-days post transplantation. Mice transplanted with *Ptip*$^{f/f}$ HOXA9/MEIS1 cells had significantly larger spleens than those transplanted with *Ptip*$^{-/-}$ HOXA9/MEIS1 cells.
(C) Tissue histology of liver and heart, and cytospins of bone marrow and peripheral blood taken from mice 30-day post transplantation. (Scale bars: 50μm)
(D) Transplantation schematic and survival curve for mice received *Ptip*$^{+/+}$; *CRE-ER*$^{+/-}$ or *Ptip* $^{f/f}$;*CRE-ER*$^{+/-}$ HOXA9/MEIS1 leukemia cells. Starting from Day 14, OHT was injected into both groups of mice every 14 days until death. Black arrows indicate the time of OHT injection. p<0.0001 (log-rank test)

Flow cytometric analysis of the bone marrows from both groups reveal a delayed bone marrow repopulation of *Ptip* $^{-/-}$ cells: at Day 30, 29% of the bone marrow in the *Ptip* $^{-/-}$ recipient mice were GFP positive, whereas nearly 100% of the bone marrow cells in *Ptip* $^{f/f}$ cell recipient mice expressed GFP. Moreover, *Ptip* $^{-/-}$ cells presented higher levels of CD11b and Gr-1 (Figure 3.4A), suggesting that they were phenotypically more differentiated than the *Ptip*-expressing cells (Figure 3.4B, upper panel). Interestingly, *Ptip* $^{-/-}$ cells had lower staining of Flt3 (CD135) (Figure 3.4B, lower panel), a known target of HOXA9, suggesting that *Ptip* deletion impaired the transcription activation induced by HOXA9 at selected targets. *Ptip* $^{-/-}$ leukemic cells also exhibited decreased proliferation and increased apoptosis when cultured *in vitro*. Altogether, these findings demonstrate that loss of *Ptip* reduces the leukemogenic potential of HOXA9/MEIS1-transformed cells. Together, these data highlighted the importance of PTIP in HOXA9/MEIS1-mediated leukemia development.

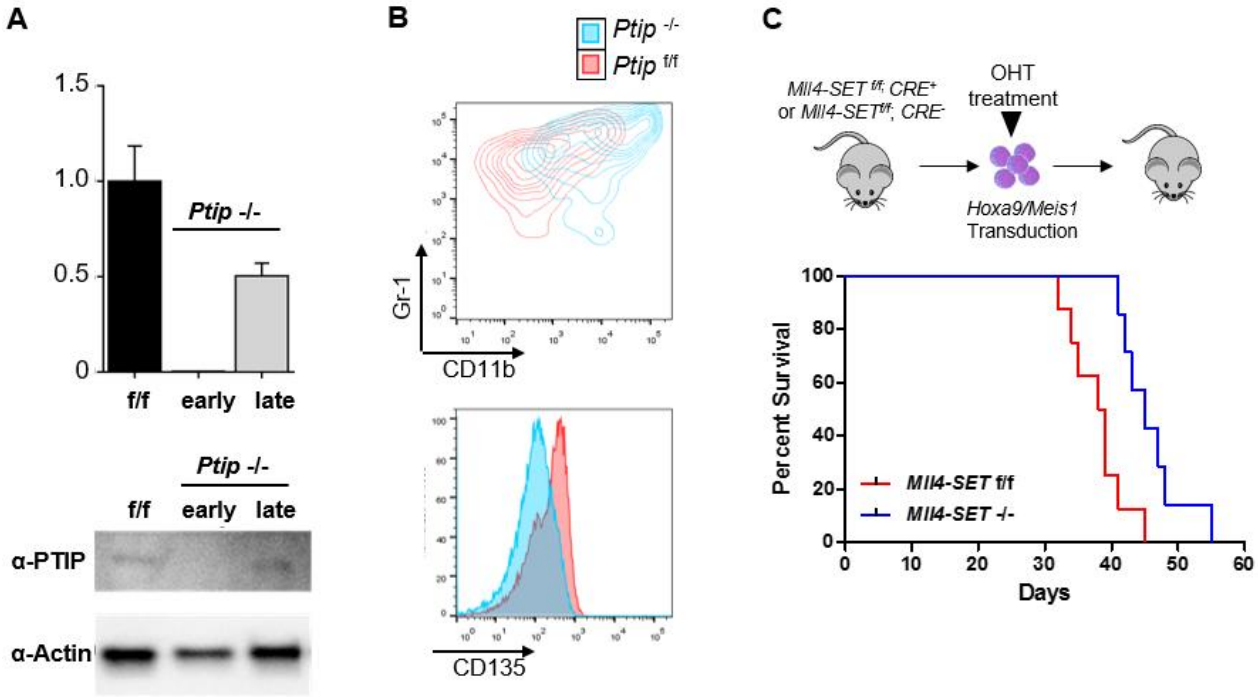

Figure 0-4 Both PTIP and MLL4 SET domain are required for development of acute leukemia in mice
(A) qRT-PCR and Western-blot showing the level of *Ptip* transcripts and protein expression in Ptip f/f and Ptip -/- HOXA9/MEIS1 cells at an early and a late time point in the leukemogenesis process.
(B) Flow plot comparing the surface presentation of CD11b, Gr-1 and CD135 (FLT3) on *Ptip* $^{f/f}$ and *Ptip* $^{-/-}$ HOXA9/MEIS1 cells harvested at Day 30 after transplantation.
(C) Transplantation schematic and survival curve for HOXA9/MEIS1-mediated leukemogenesis in *Mll4-SET* $^{f/f}$ and *Mll4-SET* $^{-/-}$ bone marrow cells. $p<0.0001$ (log-rank test)

PTIP is also reported to participate in nuclear events other than MLL3/MLL4-mediated histone methylation [213-215]. To verify that the prolonged survival is due to the disruption of MLL3/MLL4 complex and is dependent on the histone methyltransferase activity of MLL3/MLL4, we conducted *in vivo* leukemogenesis assay with targeted deletion of the SET-domain of MLL4, which specifically abolishes its catalytic activity. Similarly, *Mll4-SET* $^{f/f}$ or *Mll4-SET* $^{-/-}$ Lin$^-$ BM cells were transduced with *Hoxa9*/*Meis1* viruses, and then transplanted into lethally irradiated mice. The mice received *Mll4-SET* $^{-/-}$ *Hoxa9*/*Meis1* cells (n=8) also had significantly delayed disease progression than those

received *Mll4-SET* $^{f/f}$ cells (Figure 3.4C, p=0.0047), although the latency difference was less drastic than with *Ptip* deletion. This increase in survival indicates that the histone methyltransferase activity of MLL4 is required for HOXA9-induced leukemogenesis. The smaller survival improvement suggests that MLL3 and MLL4 possibly play a partially redundant role in modulating the enhancer activity in leukemic development.

3.3.3 Loss of *Ptip* compromises the incorporation of H3K4me1 at HOXA9+ DE

To alleviate the possible effects of methyltransferase redundancy, we again turned to the *Ptip* deletion model to target both MLL3 and MLL4 in the formation of *de novo* enhancers. To this end, we isolated both *Ptip* $^{f/f}$ cells and *Ptip* $^{-/-}$ cells with HOXA9/MEIS1-mediated transformation from the mouse bone marrow at Day 30 when 50% of the *Ptip* $^{f/f}$ mice became morbid. We again performed H3K4me1 ChIP-seq on these cells to determine how *Ptip* deletion and consequential disruption of MLL3/MLL4 complex affected the enhancer landscape in HMM leukemia cells. Analysis at HOXA9+ PE and DE revealed that although the averaged read density remained unaffected by *Ptip* deletion at HOXA9+ PE, it decreased specifically at DE in the *Ptip*-deficient cells (Figure 3.5). This distinction between the two groups of enhancers of *Ptip* deletion phenocopied that of HOXA9 inactivation, thus, it suggests that HOXA9 and *Ptip*-dependent H3K4 methyltransferase collaborate to form *de novo* enhancers but not physiological enhancers in the leukemogenesis process.

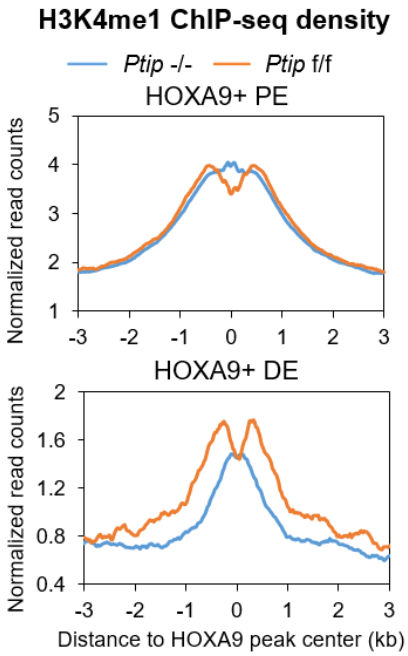

Figure 0-5 PTIP is required for the formation of HOXA9+ DE
Composite plots showing average H3K4me1 signal at HOXA9+ PE and DE in HOXA9/MEIS1 leukemic cells with and without *Ptip* deletion. Bone marrow samples were taken from mice 30-days post-transplantation, fixed right after red blood cell lysis.

3.4 Discussion

This chapter focuses on the candidate histone methyltransferases that collaborate with HOXA9 to implement the enhancer modification. We discover that HOXA9 directly mediates the establishment of HOXA9-regulated *de novo* enhancers by recruiting MLL3/MLL4 methyltransferases that are dispensable for the formation and maintenance of physiological enhancers. The molecular and functional collaboration between HOXA9 and MLL3/MLL4 complex is essential for the rapid onset of acute leukemia.

3.4.1 The different recruitment mechanisms of MLL3/MLL4

Our data suggest that HOXA9 is differentially required for targeting of the MLL3/MLL4 complex, and is dispensable for its binding to the physiological enhancers. The physiological enhancers are present in the cell-of-origin without HOXA9-mediated transformation, and likely exert regulatory function in hematopoietic development. Previous studies have shown that lineage-specific transcription factors, such as C/EBPα and PU.1 in macrophage [145], collectively recruit epigenetic modifiers and define the enhancer landscape in each lineage. The collaboration between C/EBPα and MLL3/MLL4 has been reported in adipogenesis [210]. Given that C/EBPα exercises strong binding at physiological enhancers with or without activated HOXA9, it is highly probable that MLL3/MLL4 is brought to those sites by C/EBPα. In fact, it is also possible that the recruitment of MLL3/MLL4 to HOXA9-regulated *de novo* enhancers is mediated by C/EBPα, as C/EBPα is also shown to directly interact with *Hoxa9* and is required for transformation by *Hoxa9* [61].

3.4.2 The histone methyltransferases at *de novo* enhancers

We find intensive enrichment of both H3K4me1 and H3K27ac at HOXA9 binding sites, as well as preferential binding of HOXA9 on active enhancers. Previous studies have shown that several HOX proteins as well as MEIS1 recruit the CBP-p300 complex at HOXA9 target sites and facilitate the incorporation of H3K27ac [61, 62], likely explaining the strong correlation between HOXA9 binding and H3K27ac modification. However, it was unknown which histone methyltransferase implements H3K4me1 and primes the HOXA9+ DE for further activation. We primarily considered MLL3/MLL4 due

to the following reasons: 1) *in vitro* assays demonstrated that MLL1, MLL3/MLL4 are the major H3K4 mono-methyltransferase, while MLL1 shows slight preference for trimethylation [202]. 2) MLL3/MLL4 preferentially bind on enhancer regions *in vivo* [210, 211, 216-219]. 3) Inhibition of MLL1 has little or no effect on cell growth in HOXA9/MEIS1 leukemia cells [89]. To support this hypothesis, we performed immunoprecipitation-Western blot, ChIP-seq assay, as well as functional analysis to confirm the collaboration between MLL3/MLL4 and HOXA9 at both molecular and functional level. Despite all these findings, we do not preclude the involvement of other H3K4 methyltransferases in the establishment of *de novo* enhancers. In fact, the observation that the formation of HOXA9+ DE was not completely abrogated upon *Ptip* deletion suggests a possible role for other lysine methyltransferase. However, our discovery of the collaboration between histone methyltransferases and HOXA9 provides therapeutic insights on targeting the epigenetic regulators and enhancer alterations for the treatment of acute leukemia.

Chapter 4

HOXA9-mediated transformation in other lineages

4.1 *Hoxa9* over expression in B lineage

4.1.1 Background

Although most extensively studied in AML, *Hoxa9* over expression occurs in several subtypes of acute leukemia, as well as in a fraction of myelodysplastic syndromes (MDS). A large-scale microarray-based assay generated gene expression profiles of 2,096 leukemia and MDS patients from three continents showed that *Hoxa9* is substantially elevated both in a large subset of pro-B acute lymphoblastic leukemias and in a minor population of T-cell acute lymphoblastic leukemia (Figure 4.1A) [220]. This cross-lineage involvement suggests that HOXA9 permits and participates in the oncogenic transformation in a lineage-independent manner. Indeed, *in vivo* forced expression of *Hoxa9* in pro-B cells blocks differentiation along the B-lineage and inhibits B lymphopoiesis, a phenotype similar to what is observed in the myeloid lineage [80]. Therefore, understanding the biological functions and target regulation of HOXA9 in preventing differentiation and promoting stemness will shed lights on the mechanisms of acute leukemias across different subclassifications.

When B-ALL cases are subdivided based on their driving mutations, it is revealed that *Hoxa9* over expression is strictly associated with MLL-rearranged (*MLL-r*) B-ALL

and is typically silenced in other subtypes (Figure 4.1B, by Dr. Figueroa) [221]. Although B-ALLs with rearranged *MLL* or germline *MLL* are similar in most morphological characteristics, they differ significantly for their gene expression profiles. Over expression of *Hoxa9*, as well as several other *Hox* genes, is one of the unique characteristics that separate *MLL-r* B-ALL from other B-ALLs. In fact, *MLL-r* B-ALL bears more resemblance to *MLL-r* AML, since both are shown to be arrested at earlier hematopoietic stages than conventional AML and ALL. In addition, they also exhibit multilineage gene expression [112]. In light of this discovery, it is intriguing to investigate if the specific lineage environment affects the regulatory function of HOXA9. In this study, I show that the over expression of *Hoxa9* and *Meis1*, as well as *Hoxa9* alone, transforms progenitors into the B-lineage, although they experience lineage conversion when transplanted into recipient mice. In the B-precursor cell line, HOXA9 predominantly binds to active distal elements, which are also enriched for B-lineage specific transcription factor motifs. Moreover, *Hoxa9* over expression in the B lineage also alters the epigenetic state on a genome-wide scale.

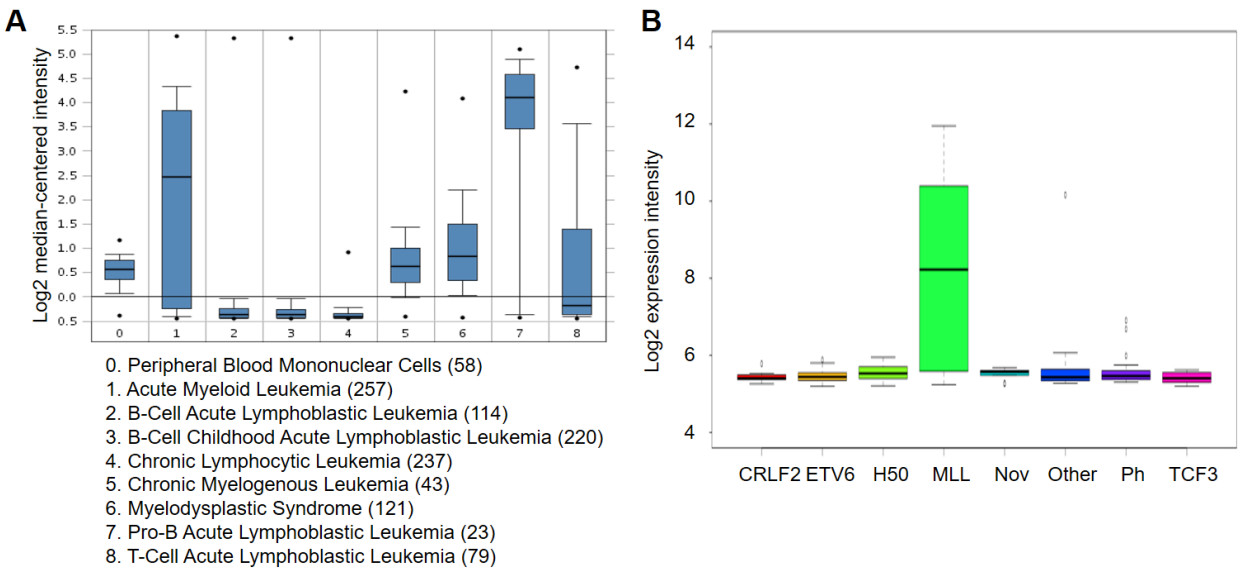

Figure 0-1 The expression of *Hoxa9* in acute leukemia
(A) Box plot showing *Hoxa9* expression in different subtypes of acute leukemia. Data set from [220]; plot generated by https://www.oncomine.org/.
(B) Box plot showing *Hoxa9* expression in B-ALLs with different genetic aberrations. TCF3: t(1;19)(q23;p13.3) translocation. Ph: t(9;22)(q34;q11) translocation. CRLF2: cytokine receptor–like factor 2 rearrangement, MLLr: *MLL* rearrangement at 11q23. ETV6: t(12;21)(p13;q22) translocation. Bar plot generated by Dr. Figueroa.

4.1.2 Materials and Methods

B cell differentiation

Lin$^-$ BM was harvested according to the protocol in Chapter 2.2: Cell lines. OP9 cells were seeded 24 hours before the experiment. Lin- BM culture was then transferred on to OP9 cells with cytokine supplements: Flt3L 5ng/mL, IL-7 5ng/mL. The culture was transferred onto fresh OP9 cells every 3-4 days. CD19+, B220+ cells emerged after around two weeks.

in vivo leukemogenesis assay

This assay protocol was modified from in Chapter 3.2: *in vivo* leukemogenesis assay. Briefly, $1.25x10^5$ HMB cells were injected into lethally irradiated (900rads) female

C57BL/6J mice together with $2.5x10^4$ Ficoll (F4375, Sigma) purified fresh bone marrow. Mice were maintained on antibiotic water for two weeks. The mice were sacrificed when they became moribund. Visceral organs (liver, lung heart and spleen) and bone (tibia and sternum) were collected for histological and flow cytometric analysis.

ChIP, ChIP-seq and RNA-seq analysis

These analyses were performed as described in Chapter 2.2: ChIP, ChIP-seq analysis, Peak annotation and RNA-seq analysis.

4.1.3 Results and Discussions

Generation of Hoxa9-transformation B-precursor cell lines

First, we tested the previous claim that progenitors immortalized by *Hoxa9* alone are myeloid-restricted, while co-expression with *Meis1* confers lymphoid differentiation potential on the immortalized cells [109]. To this end, I used an *in vitro* B cell induction protocol to force differentiation along the B lineage. This was achieved by transducing *Hoxa9* or *Hoxa9* plus *Meis1* into *Lin*- BM, and culturing them in B cell specific conditions (with IL-7, FLT3L and OP9 stromal cells) (Figure 4.1A). The differentiation progress was compared to that of the empty vector (EV)-transduced progenitors. As shown in Figure 4.1B, at Day 7, a greater portion of *Hoxa9*/*Meis1* co-transduced cells remained lineage-negative, while large percentage of EV and *Hoxa9* alone were positive for myeloid markers. At Day 11, subpopulations of CD19-positive B cell started to emerge in all three groups, whereas the myeloid population first disappeared in the *Hoxa9*/*Meis1*-transduced group. At Day 15, all three groups became predominately positive for CD19, and their distributions between the myeloid, B-lymphoid and lineage-negative

compartment remained unchanged until EV cells underwent apoptosis and died out (Figure 4.1B). Based on this result, it is clear that both *Hoxa9* alone and *Hoxa9*/*Meis1* co-transduced cells retain the potential of lymphopoiesis. *Hoxa9* cells developed into the B-lineage at a similar rate as the control cells, suggesting that HOXA9 does not actively prevent early B-lymphopoiesis. The observed difference between *Hoxa9* only and *Hoxa9*/*Meis1* cells in the previous study was possibly reflected in their differentiation kinetics: by upregulating IL-7r and FLT3 [109], *Meis1* over expression provides more surface receptors for B cell-inducing cytokines or other environmental cues, and thus accelerates differentiation along the B-lineage.

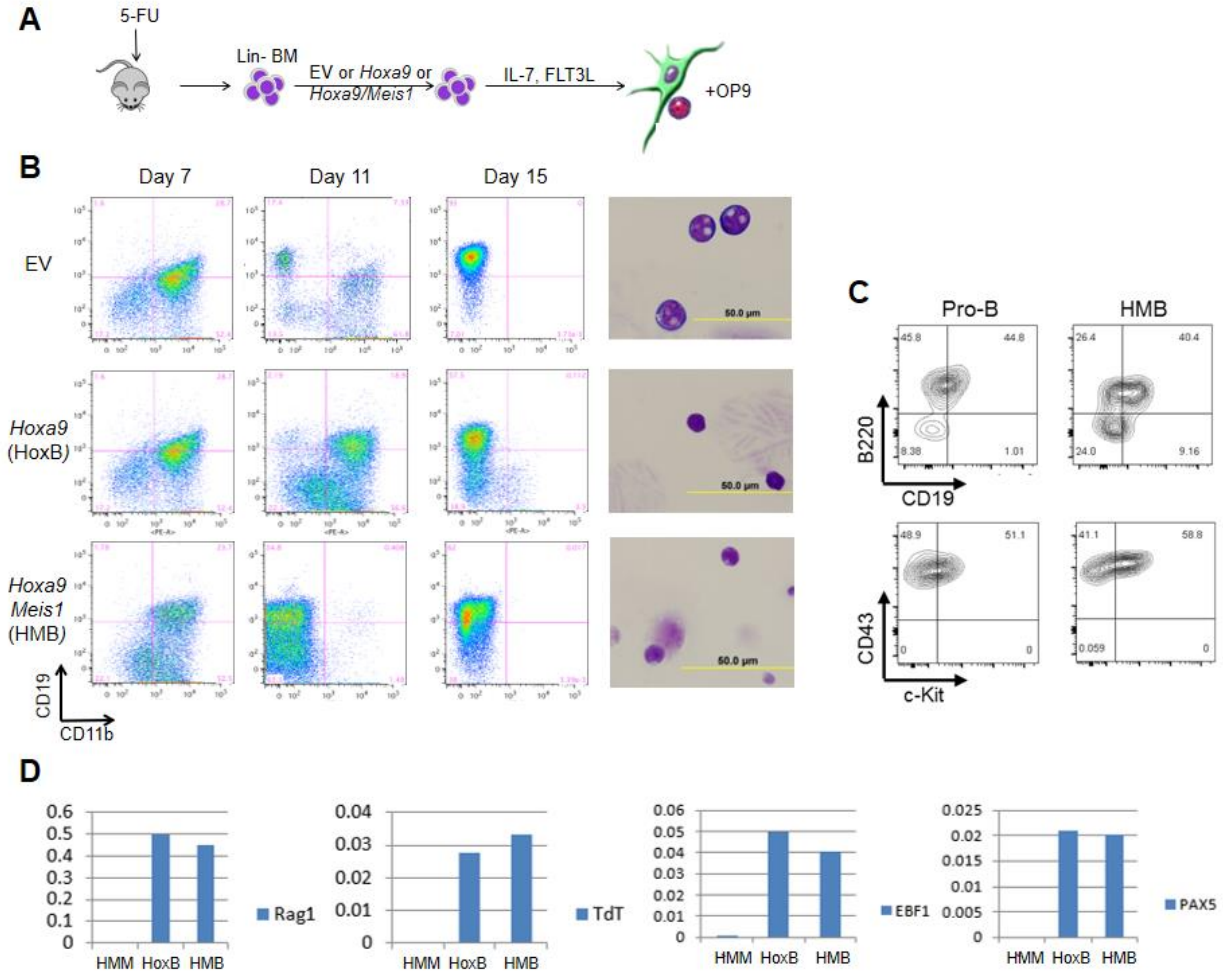

Figure 0-2 The generation and characterization of HOXA9-transformed B lineage cells (A) The differentiation schematic. (B) Flow cytometry analysis on the differentiation progress of the three types of *in vitro* induced B lineage cells. The population constitution indicated with CD11b (myeloid) and CD19 (B cells) (C) The surface marker characterization of pro-B cells and HMB cells with CD19, B220, c-Kit and CD43. (D) Relative RNA expression of *Rag1*, *TdT*, *Ebf1* and *Pax5* in HMM, HoxB and HMB cells.

Hoxa9 and *Hoxa9*/*Meis1* cells were then characterized using a series of cell surface markers and transcription factors to indicate their differentiation stages. Both cells were largely B220+ and CD19+ (Figure 4.1C and data not shown). They also expressed similar level of *TdT*, *Rag1*, *Ebf1* and *Pax5* (Figure 4.1D), while these genes were completely silenced in HMM cells. The surface presentation of CD19, as a result

of *Pax5* expression, suggests that the two types of cells have entered the lineage-committed stage of B cell differentiation. Based on the c-Kit positivity, we classified both cells as at pro-B stage. *Hoxa9* alone and *Hoxa9*/*Meis1*-transformed B precursors are thus referred to as HoxB and HMB cells respectively. Although having similar immunophenotype, HoxB cells constantly experienced apoptosis and were completely dependent on OP9 stromal cells for proliferation. By contrast, HMB cells proliferated well with or without OP9, and could propagate *in vitro* for at least three months without undergoing apoptosis, suggesting that these cells were fully immortalized. HMB cells thus were utilized in the downstream analysis.

Lineage conversion of HMB cells

The leukemogenic ability of HMB cells were then analyzed using *in vivo* assays. These cells were transplanted into lethally irradiated syngeneic mice together with 1/10 count of fresh bone marrow to overcome the irradiation-induced lethality. HMB recipient mice developed acute leukemia within the same timeframe as compared with HMM recipient mice (Figure 4.3A). Surprisingly, the leukemic bone marrow harvested from the HMB mice showed a different phenotype than the HMB transplant: they lost the B cell marker, CD19, and acquired mature myeloid markers, CD11b and Gr-1 (Figure 4.3 B and C). This lineage conversion was inconsistent with the pro-B differentiation stage, since in normal hematopoiesis, *Pax5*-expressing pro-B cells would have lost multilineage developmental potential. To determine whether it was *bona fide* transdifferentiation, or simple clonal expansion of the few lineage-uncommitted/myeloid leukemia stem cells, we tested for the genetic rearrangement at the immunoglobulin heavy chain locus (*IgH*). Presumably, if this acts as true transdifferentiation, the post-

transplantation leukemic cells would have rearranged genetic configuration at the *IgH* locus, since they have entered the B cell fate previously; in contrast, the straight myeloid differentiation would not alter the genetic configuration at this locus. Surprisingly, HMB cells indeed experienced the transdifferentiation phenomenon which was confirmed with this clonality test. Both pre- and post-transplantation HMB cells had the D_H-J_H rearrangement, while HMM had the non-rearranged configuration as the germline cells (Figure 4.3D). This result indicates that HMB cells are featured with lineage infidelity, and their immunophenotype can be switched by the environmental inducers. It also suggests that *Hoxa9*/*Meis1*-mediated transformation provides the cells multi-lineage potential, even when they are phenotypically more differentiated. Similar lineage conversion has been reported with FACS-isolated lymphoid lineage *Hoxa9*/*Meis1* cells from leukemic mice, and it was suggested that the tumor-initiating activity is independent of their stable immunophenotypes [110]. Altogether, it is highly plausible that *Hoxa9*/*Meis1*-mediated leukemias have a conserved transformation mechanism that is uncoupled to the lineage specification. The exact mechanisms by which HMB cells entered the myeloid fate *in vivo* is currently unknown.

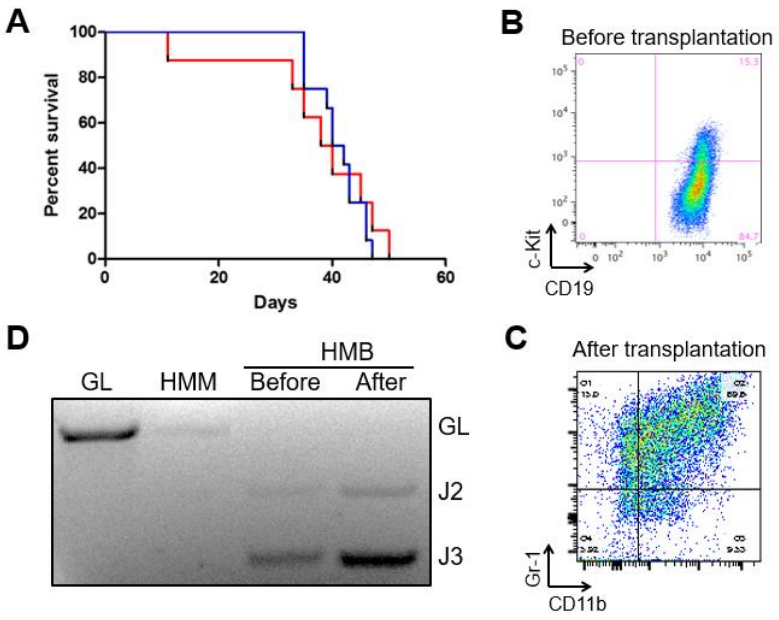

Figure 0-3 The *in vivo* lineage conversion of HMB cells
(A) The survival curves of mice transplanted with HMM and HMB cells respectively. (B) The flow plot showing CD19 positivity, indicative of B lineage commitment, of HMB cells before transplantation. (C) The cell surface markers of the post-transplantation HMB leukemic cells, showing the positivity of CD11b and Gr-1. (D) The genetic configuration of IgH locus in germline (GL), HMM, pre- and post-transplantation HMB cells.

The genome-wide binding of HOXA9

Previously, HOXA9 has been found in HMM cells to primarily bind promoter-distal regulatory elements [54, 61]. To determine the genome-wide localization of HOXA9 in HMB cells, we performed chromatin immunoprecipitation sequencing (ChIP-seq) using an anti-HA antibody to target the HA-tag fused to HOXA9. With the 18,601 high-confidence peaks, we observed similar preferential binding to promoter-distal regions: while 26% (n = 4849) occurred within promoter regions, 70% (n=13033) occupied intergenic and intronic regions (Figure 4.4A). To relate these binding sites to chromatin states, we performed ChIP-seq for key histone modifications, including H3K4me1, H3K27ac, and H3K27me3. We found that HOXA9 binding sites were

enriched with H3K4me1 and H3K27ac, but depleted of H3K27me3 (Figure 4.4B), similar to the signature in HMM cells. Indeed, the majority (80%) of the promoter-distal peaks of HOXA9 overlapped with putative active enhancers, while limited binding (10% and 2%) was found on primed enhancers and poised enhancers (Figure 4.4C). Comparison of HOXA9 binding in the two cell types identified only 1,839 common peaks, which represent 28% and 9.9% of total HOXA9 target sites in HMM and HMB cells, respectively (Figure 4.4D). Lineage-specific HOXA9 binding was further confirmed with motif enrichment analysis. As shown in Figure 4.4E, sequence motifs of B-cell specific transcription factors EBF1 and TCF3 were overrepresented at HOXA9 binding sites in HMB cells, while those of myeloid specific transcription factors, such as C/EBP and ETS family factors, were enriched in HMM cells, suggesting that the lineage-specific transcriptional milieu plays a role in HOXA9's localization.

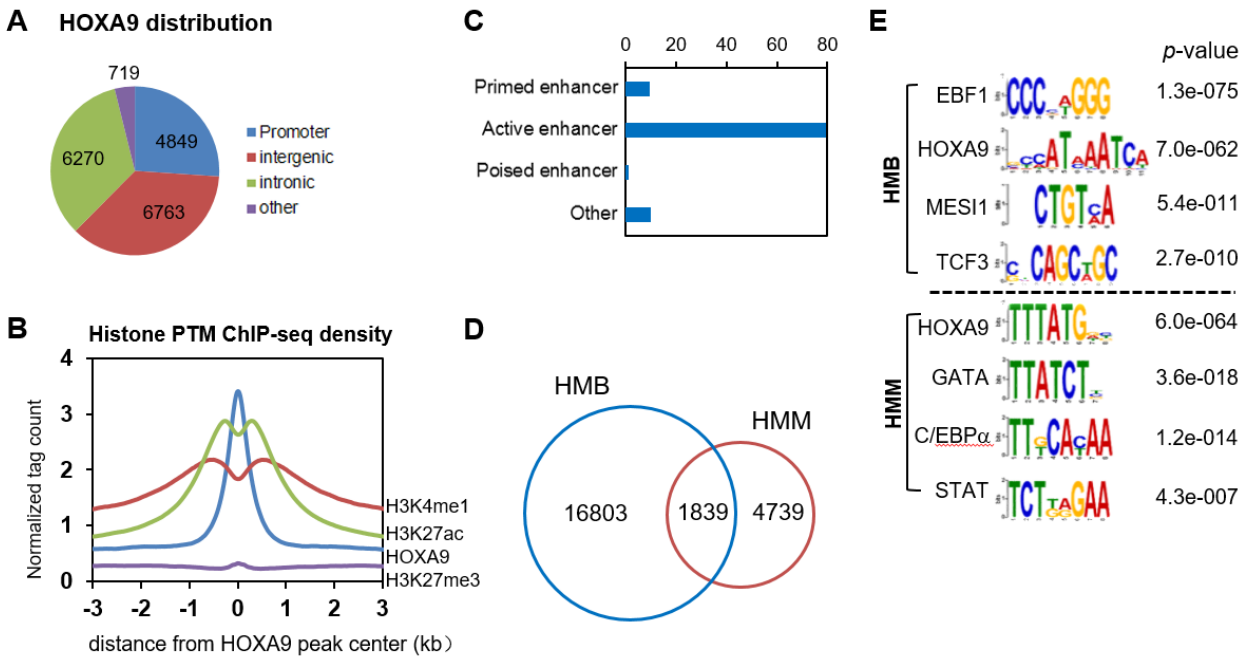

Figure 0-4 The genome-wide binding of HOXA9 in HMB cells
(A) The genomic distribution of HOXA9 in HMB cells. (B) Composite plot showing average per bp density of HOXA9, H3K4me1, H3K27ac and H3K27me3 at HOXA9 binding sites in HMB cells. Library size normalized to 1E7 reads. (C) Percentage

distribution among different enhancer states at HOXA9's promoter-distal binding sites. These sites are preponderantly associated with active enhancer signatures. (D) Physical overlap of HOXA9 peaks in HMM and HMB cells. (E) Motif analysis of 6,578 and 18,601 HOXA9 peaks in HMM and HMB cells.

Global epigenome changes in HMB cells

Enhancers constitute the regulatory code that drives cell type-specific gene expression. Since HMB cells are capable of undergoing lineage conversion, we wondered whether the lineage infidelity was reflected in their enhancer landscape. To this end, we compared the global H3K4me1 profiles of HMB and HMM cells with their respective culture controls, as well as the cell types in normal hematopoietic hierarchy. Surprisingly, hierarchical clustering based on the global H3K4me1 dynamic profile showed that HMB cells share the highest similarity with Pro-B cells, but not HMM cells, suggesting that HMB cells still largely maintained the enhancer landscapes of the cell-of-origins (Figure 4.5A). When closely inspecting the 116,182 H3K4me1 regions in all 20 cell types, we discovered that two non-overlapping clusters of regions exhibited high signal enrichment within HMB (Cluster 1) and HMM cells (Cluster 2), respectively, but were unmarked in any other cell types, which are consistent with the *de novo* enhancers identified in Chapter 2 (Figure 4.5B). These data imply that although HOXA9/MEIS1-dependent AML and B-ALL may be able to transdifferentiate and induce aggressive leukemia in mice, their epigenomes are still largely determined by their lineage specification. Most importantly, these two types of cells form distinct repertoires of *de novo* enhancers after transformation, suggesting that the changes in enhancer landscapes are likewise cell type-specific.

HOXA9 facilitates the formation of cancer-specific enhancers in HMM cells. In accord with this finding, we found that among the 12,378 HOXA9-bound enhancers in HMB cells, 1,473 had no H3K4me1 prior to transformation in pro-B cells, the normal lineage counterparts for HMB (Figure 4.5C). Interestingly, these novel enhancers also lacked H3K4me1 mark in other normal hematopoietic cells (Figure 4.5D), suggesting that like the HOXA9+ DE in HMM cells (Chapter 2.3.3), these enhancers are created *de novo* specifically during the leukemic transformation process. Most strikingly, GO analysis showed that comparing with the HOXA9-regulated physiological enhancers, these *de novo* enhancers were also enriched for multiple embryonic development pathways (Figure 4.5E). This indicates that the ectopic activation of early developmental pathways is a conserved feature of HOXA9-mediated transformation. It is fascinating that HOXA9 resorts to the same pathways to drive transformation, but employs different players of the same pathways in different lineages. The exact mechanisms by which the two transformation programs converge are still largely unknown.

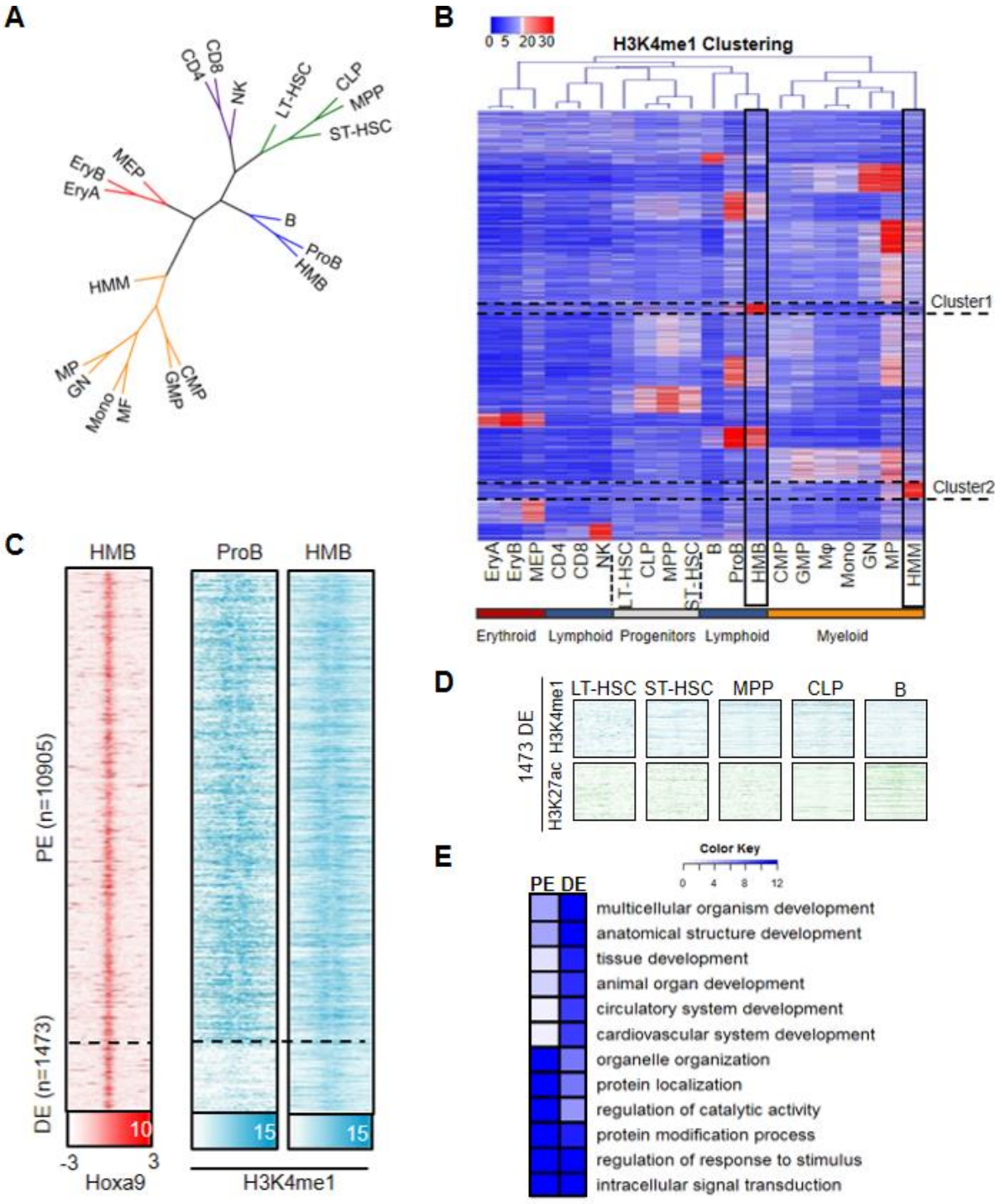

Figure 0-5 HOXA9 targets a subset of *de novo* enhancer in HMB cells
(A) Clustering dendrogram of cell types based on H3K4me1 profiles showing the association of HMM and HMB cells with the untransformed counterparts in their respective lineages. Color code: green for multipotent progenitors, orange for myeloid lineage (including oligopotent progenitor CMP and GMP), blue for lymphoid lineage, and

red for erythroid lineage. (B) Heatmap showing 101,413 hematopoiesis and leukemogenesis enhancers clustered with K-means (K=16) using the normalized read count at each enhancer region. Cluster 1 in HMB and Cluster 2 in HMM indicate the unique H3K4me1 signature in the transformed HM cells. (C) Heatmap depicting the signal intensity of HOXA9 in HMB and H3K4me1 in ProB and HMB at 12,389 HOXA9-bound distal regulatory elements. The rows show ±3kb regions around HOXA9 peak center. The 12,389 regions are separated into two categories, 1473 HOXA9+ DE and 10905 HOXA9+ PE based on the differential H3K4me1 modification status. (D) Heatmap depicting signal intensity of H3K4me1 at 10905 HOXA9+ PE (upper panel) and 1473 HOXA9+ DE (lower panel) in normal hematopoietic cells of different differentiation stages into the B lymphocyte maturation. (E) Gene Ontology terms for PE and DE-associated genes with the corresponding Benjamini p-values. Pathways specifically enriched with DE are shown on top.

4.2 *Hoxa9* over expression in adipogenesis

4.2.1 Background

A previous study in our lab discovered that in HOXA9/MEIS1-driven AML cells (HMM cells), adipogenesis and PPARγ signaling are two of the most enriched pathways upregulated upon HOXA9 inactivation (Figure 4.6 A and B). It suggests that in the myeloid lineage, HOXA9 represses the adipogenic fate, and its inactivation results in reactivation of the entire pathway.

It is an interesting observation, since adipogenesis involves sequential activation of a series of transcription factors, including C/EBPα, C/EBPβ, C/EBPγ and STAT5 [222], which are also lineage-specifying transcription factors in myelopoietic differentiation. In fact, although macrophages and adipocytes are morphologically different, studies have shown that they may share some similar features. First, numerous inflammatory factors, such as tumor necrosis factor alpha (TNFα), are secreted by both pre-adipocytes and macrophages [223, 224]. Adipocyte progenitors

also develop phagocytic activity upon infection, and this response disappears when they differentiate into adipose tissue [225]. Moreover, not only do preadipocytes and macrophages share common immunophenotypic properties, but they can also transdifferentiate into the other lineage under regular physiological conditions [226, 227]. Most importantly, PU.1 and C/EBPα/β efficiently convert pre-adipocytes into macrophage-like cells [228], exemplifying the significance of availability and dosage of lineage-determining factors in cell fate specification.

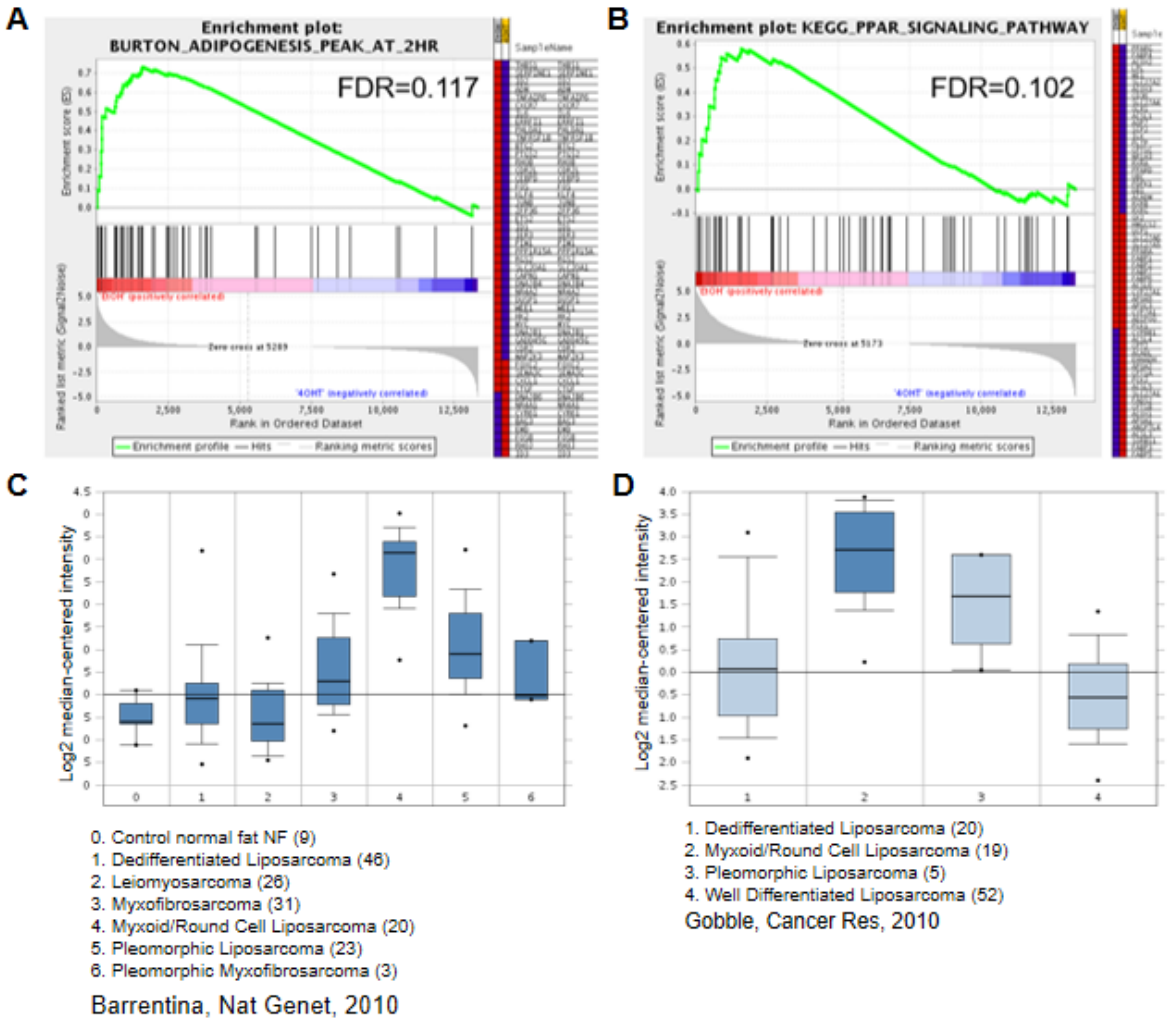

Figure 0-6 HOXA9 inhibits adipogenesis pathways in HMM cells, and is upregulated in a particular type of liposarcoma.
(A-B) Gene Set Enrichment Analysis of the expression profiles of HOXA9 inactivated HMM cells versus that of normal HMM cells, showing the enrichment of adipogenesis

pathway (A) and PPAR signaling pathway (B). Figure generated by Cailin Collins and Jingya Wang, unpublished results. (C-D) Over expression of HOXA9 is consistently associated with myxoid/round cell liposarcoma in two studies. Box plot generated by Oncomine.

Hox genes are also involved in the regulation of adipogenesis. *Hoxa9*, in particular, is highly expressed in the brown adipose tissue; in white adipose tissue, it is found significantly upregulated after fat loss [229, 230]. Apart from its role in adipocyte development, *Hoxa9* over expression is specifically associated with one subtype of liposarcoma, the myxoid/round cell liposarcoma (Figure 4.6 C and D). Thus, understanding the role of HOXA9 in the adipogenesis process has both physiological and pathological significance. In light of these studies, we used the adipogenesis system to study the functions of HOXA9, and sought to understand whether HOXA9 employs cell type specific or common mechanisms to regulation gene expression and cellular differentiation.

4.2.2 Materials and Methods

Induction of 3T3-L1 into adipocyte-like cells

3T3-L1 cells (CL-173, ATCC) cells were seeded in a six-well plate at a density of $3.5x10^4$ cells/well, and grown until reaching 70% confluency. To initiate differentiation, cells were remove from DMEM media and added MDI induction medium [0.5 mM IBMX, 1 µM dexamethasone, 10 µg/mL insulin] 2 mL/well. After 3 days, MDI medium was removed and replaced with insulin medium [10 µg/mL insulin] 2mL/well. After 7 to 10 days, cells were differentiated into adipocyte like cells.

HOXA9-ER expression and activation

HOXA9-ER plasmid was transfected in 3T3-L1 fibroblast cells using FuGene 6. Transfection medium was replaced after 24hrs. To activate HOXA9, 100nM OHT was added together with MDI medium or at different time points during differentiation to monitor the effect on adipogenic differentiation.

Oil-Red-O staining

The Oil-Red-O stock solution was prepared by dissolving 300 mg of Oil-Red-O powder (0-0625, Sigma) in 100 ml of 99% isopropanol. Working solution was then prepared by adding 20mL of water to 30mL of Oil-Red-O stock solution and filtered with filter paper. Cells were fixed with 10% formalin at R.T. for 30min, and washed with distilled water twice. Then isopropanol was added to the cells for 5min, followed by the Oil-Red-O working solution for 5min. The monolayer was then rinsed with water. Oil-Red-O can be solubilized with isopropanol and quantified with spectrometer.

RNA extraction and RT-qPCR

RNA was purified using the Qiagen RNeasy kit with on-column DNase treatment. cDNA was generated using Superscript III RT and target gene expression was determined relative to β-actin using Invitrogen Taqman probe sets .

4.2.3 Results and Discussion

Since HOXA9 inhibits the adipogenic program as well as *Pparγ* expression in myeloid lineage cells, we first examined whether HOXA9 plays the same role in pre-adipocytes. This was achieved by overexpressing *Hoxa9* in mouse embryonic fibroblasts 3T3-L1 cells, which can be induced into adipocyte-like cells with chemical treatment. Strikingly, HOXA9 activation prior to the differentiation induction completely

inhibited adipogenic development. As shown in Figure 4.7A. After 12 days, without HOXA9 activation, 3T3-L1 fully adopted the adipogenic fate and accumulated fatty acid droplets in the cytoplasm. In contrast, HOXA9 overexpressing cells retained the fibroblast morphology even after chemical induction. Using oil red o staining, the inhibition effects were quantified. It was shown in Figure 4.7 B and C that the lipid accumulation was significantly impaired in the group with HOXA9 activation, whereas the HOXA9 inactivated group as well as the empty vector control cells had similar level of cytosolic lipid content. This evidence indicates that HOXA9 indeed inhibits the differentiation progression of pre-adipocytes, which is consistent with the pathway inhibition observed in myeloid lineage cells.

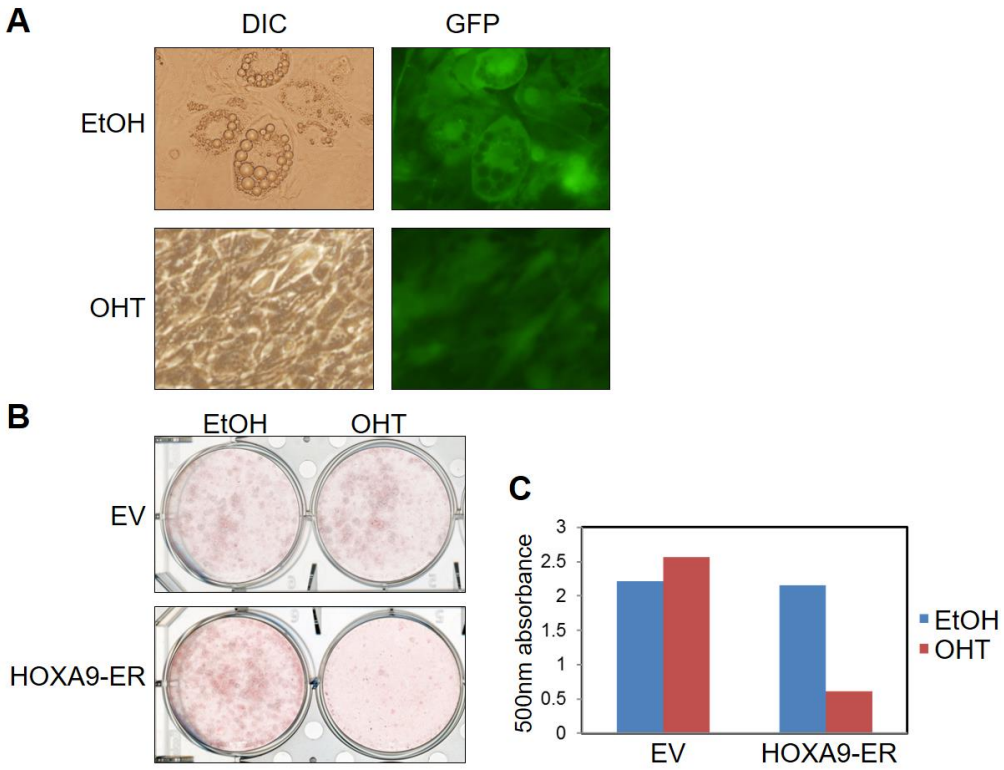

Figure 0-7 HOXA9 inhibits adipogenesis in pre-adipocytes.
(A) The cell morphology of the HOXA9 inactivated (EtOH) and HOXA9 activated (OHT) cells after adipogenic differentiation. Pictures were taken under white field (DIC) and

fluorescent channel (GFP). (B) Oil red o staining of the EtOH and OHT treated cells after adipogenic differentiation. (C) Quantification of the staining result in (B).

In normal adipogenesis, a series of transcriptional events are initiated right after differentiation induction. The expression of *Myc*, *Jun* and *Fos* surges within 1 hour of MDI medium addition, while *Cebpb* and *Cebpd* are also mildly upregulated. Subsequently, C/EBPβ and C/EBPδ together induce the expression of *Cebpa* and *Pparg*, which in their turn form a positive feedback loop and induce the expression of adipocyte specific genes [231]. Given that HOXA9 inhibits the adipogenesis program, we next examined whether HOXA9 over expression affected the transcription cascade in the differentiation process. RNA expression profiling at different time points showed that the EtOH treatment group followed the standard pattern of transcriptional regulation. However, with HOXA9 activation, *Cebpb* was significantly upregulated, while the upregulation of *Pparg* was hampered (Figure 4.8). Since PPARγ is one of the most critical determining factor of adipogenesis, and is indispensable for adipocyte-specific gene expression and fat droplet formation, its reduced activation may explain the lack of differentiation potential of the *Hoxa9*-overexpressing fibroblasts. Although the mechanism by which HOXA9 inhibits the expression of *Pparg* is not clear, we postulate that the over abundant HOXA9 may impede the correct assembly of transcriptional machineries at *Pparg* regulatory elements which is essential for the transcription initiation. *Pparg* gene has multiple functional enhancers, whose transformation from close to open state requires a series of chromatin events [222]. The overactivated HOXA9 perhaps interferes with the critical steps of the chromatin state modulation at *Pparg* enhancers, so that *Pparg* upregulation was not achieved properly. The exact mechanism of HOXA9-mediated adipogenic dysregulation needs further investigation.

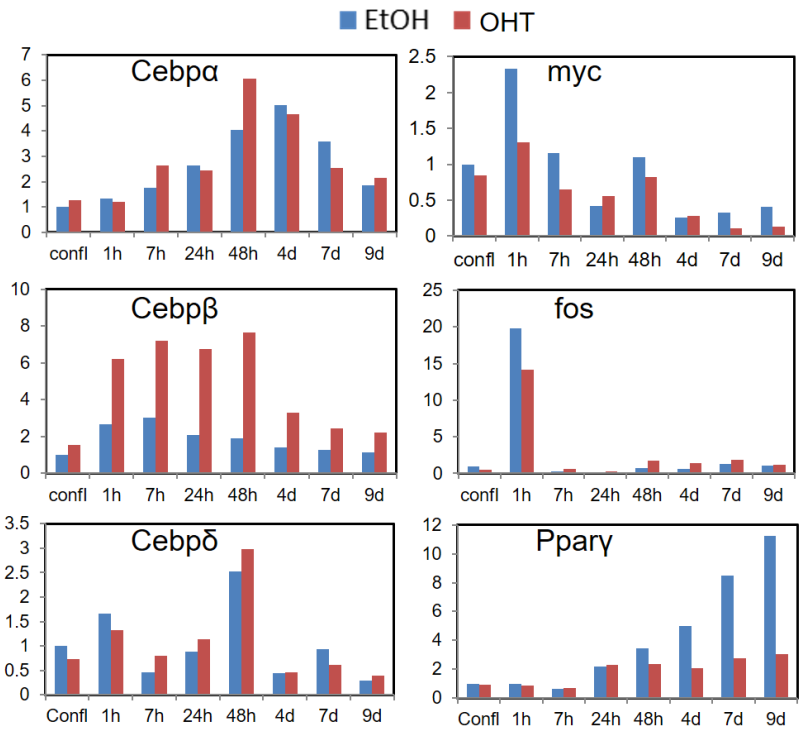

Figure 0-8 The expression of key adipogenic transcription factors with or without HOXA9 activation

Over the 9-day course, the expression of indicated genes were measured. Only *Pparg* had the change in expression consistent with the lack of adipogenic differentiation.

Chapter 5

Concluding remarks and future directions

Despite the increasingly refined knowledge in the targetable mutations and mechanisms of acute leukemia, there have not been major advances in targeted therapies for leukemias with HOXA9 over expression. Thus, there is an urgency to understand how HOXA9 achieves transformation in multiple subclassifications of leukemia to guide future therapeutic development. Recently, several studies employed high-throughput sequencing technologies to delineate the genome-wide binding sites of HOX proteins and gene regulatory networks, which have undoubtedly shed light on HOX biology. However, these works have been more descriptive than mechanistic, and fundamental understanding of HOXA9-mediated transformation mechanisms is still mostly lacking.

Our lab has worked to better define the mechanisms of HOXA9-mediated transcriptional regulation and transformation. These studies showed that HOXA9 predominately targets active enhancer regions, and a large subset of its binding sites are co-bound by lineage-specific transcription factors. To investigate these co-occurrences and correlations, I sought to establish a causative role of HOXA9 in altering the enhancer landscapes. To this end, I found that in both myeloid and lymphoid lineage, HOXA9/MEIS1-transformed cells largely inherit hematopoietic enhancers from

their cell-of-origin, but also acquire novel enhancers at specific regions that have no enhancer modifications or lineage-factor binding in any normal hematopoietic cells. I found that these novel enhancers regulate the activation of a leukemia-specific transcription program. Further, HOXA9 executes a pioneer role and its binding is essential for the establishment of these novel regulatory elements: it mediates the assembly of an enhancer-binding complex by recruiting both the myeloid lineage factor, C/EBPa, and the histone methyltransferase complex, MLL3/MLL4 complex. The complex assembly at these *de novo* regions is critical for the implementation of the enhancer characteristics and the activation of the nearby genes. In contrast, HOXA9 is dispensable for both transcription factor binding and incorporation of the H3K4me1 mark at normal physiological enhancers; it likely exhibits opportunistic binding and is nonessential for the associated gene activation. This interesting distinction can be summarized into a diagram in Figure 5-1. Together, these findings highlight HOXA9's critical role in establishing an enhancer landscape in support of the leukemia identity during oncogenic transformation.

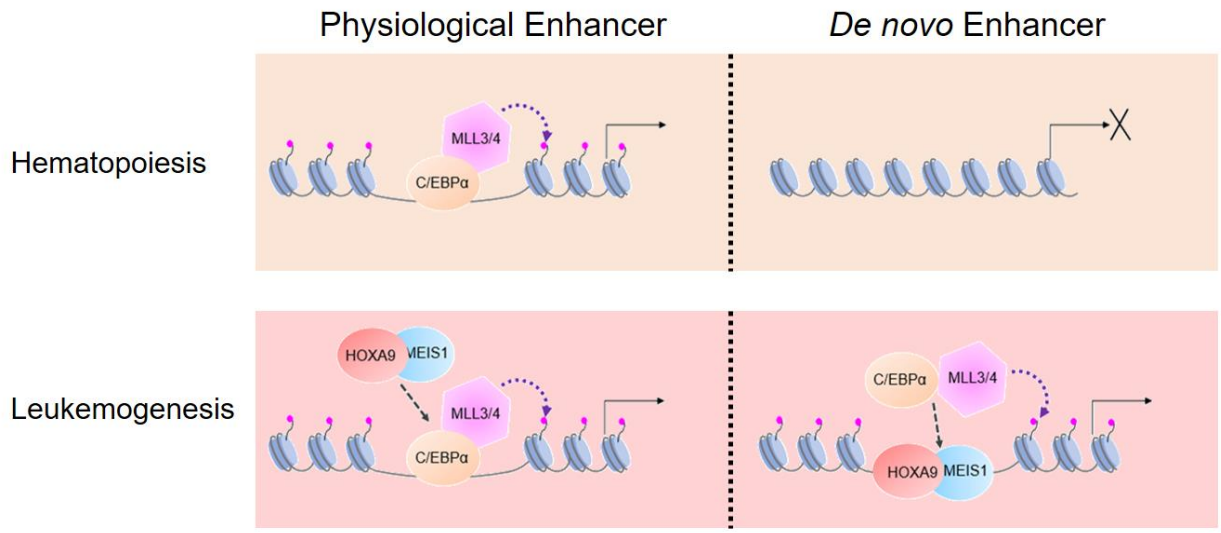

Figure 0-1 Model for HOX-regulated enhancer formation and gene regulation in leukemia development.
The work presented in this thesis brought forward the hypothesis that HOXA9's regulatory role in leukemia is associated with the establishment of *de novo* enhancers, where it recruits other factors and chromatin regulators in the activation of a leukemogenic transcription program.

There remain open questions pertaining to HOXA9's pioneer role in the establishment of active regulatory elements. First, to confidently claim HOXA9 as a pioneer factor in leukemia, the gold standard – the DNase accessibility analysis – has to be applied. Based on the results of Abd-B in *Drosophila* (Chapter 1.2.4), it is highly likely that HOXA9 acts similarly as its ortholog and targets both pre-accessible DNA and nucleosome-bound DNA. Given that HOXA9 forms multi-subunit complexes with MEIS1 and PBX1, and collaborates with them for DNA binding specificity and target gene regulation, it is thus interesting to examine how the two co-factors contribute to the chromatin remodeling function of HOXA9. Given that HOXA9 alone has limited transformation potential and only gives rise to leukemia after an extended period of time, I speculate that the synergistic interaction with MEIS1, if not PBX1 as well, is required for HOXA9 to fully exert its pioneer function in reshaping the epigenome. As a

heterodimer, the HOXA9-MEIS1 complex may have more diversified functions, form stronger interaction with transcriptional and epigenetic machineries, and regulate the downstream targets in concert.

Second, it is now relatively clear that HOXA9 mediates the establishment of *de novo* regulatory elements in leukemogenesis, and is required for their stability. However, HOXA9's role at the normal hematopoietic enhancers (i.e. HOXA9+ PE in Chapter 2) remains to be fully understood. The cumulative effect at HOXA9-bound physiological enhancers is insignificant, which means that HOXA9 over expression or inactivation does not result in a prominent up- or down-regulation of those genes as a whole. However, the overall effect can be cancelled out when similar numbers of genes are up- and down-regulated. Considering that a significant portion of HOXA9+ PE-associated genes are differentially regulated, it is plausible that HOXA9 employs divergent mechanisms to drive either activation or repression of those target genes. Along this line, it is unclear what gene repression mechanism HOXA9 may use, if at all, in mediating the inhibition of myeloid differentiation program or immune responses in AML cells. Alternatively, it is also possible that HOXA9 exhibits pure opportunistic binding at the hematopoietic enhancers, and is fully dispensable for gene regulation at those sites (Figure 5-1).

Besides offering mechanistic insights, this thesis study has profound implications in the therapeutic development for leukemias with HOXA9 over expression. It suggests that targeting the multicomponent complex assembly at the *de novo* enhancers constitutes a novel avenue in the treatment of these acute leukemias. As shown in Chapter 3.3.3, disruption of the MLL3/MLL4 complex specifically impairs the formation

of *de novo* enhancers, and significantly delays leukemia progression in mice. Therefore, inhibitors can be designed to target the catalytic activity of MLL3/MLL4, as well as its recruitment by HOXA9, and consequently block the activation of leukemogenic transcription program. Notably, MLL3/4 has a critical role in gene regulation as well as in tissue development and homeostasis. For this reason, therapeutic inhibition of its enzymatic activity would have to be highly specific and tightly controlled. A more targeted approach is to block the interaction surface between HOXA9 and the MLL3/4 complex, which requires further studies to fully characterize. Moreover, as mentioned in Chapter 3, a wide variety of mechanisms can lead to the incorporation of H3K4me1 mark at enhancer sites, and redundant mechanisms may function in concert to establish the leukemia-specific enhancer landscape. Thus, the clinical efficacy of targeting the MLL3/MLL4 complex for leukemia treatment needs careful investigation. Despite these caveats, this thesis study offers immense prospects for targeting the MLL3/MLL4 histone methyltransferase in leukemia, and provides mechanistic support for developing epigenetic therapies for hematopoietic malignancies.

References

1. McGinnis, W. and R. Krumlauf, *Homeobox genes and axial patterning.* Cell, 1992. **68**(2): p. 283-302.
2. Cerda-Esteban, N. and F.M. Spagnoli, *Glimpse into Hox and tale regulation of cell differentiation and reprogramming.* Dev Dyn, 2014. **243**(1): p. 76-87.
3. Lewis, E.B., *A gene complex controlling segmentation in Drosophila.* Nature, 1978. **276**(5688): p. 565-70.
4. Morgan, B.A., et al., *Targeted misexpression of Hox-4.6 in the avian limb bud causes apparent homeotic transformations.* Nature, 1992. **358**(6383): p. 236-9.
5. Small, K.M. and S.S. Potter, *Homeotic transformations and limb defects in Hox A11 mutant mice.* Genes Dev, 1993. **7**(12A): p. 2318-28.
6. Goodman, F.R., *Limb malformations and the human HOX genes.* Am J Med Genet, 2002. **112**(3): p. 256-65.
7. Krumlauf, R., *Hox genes in vertebrate development.* Cell, 1994. **78**(2): p. 191-201.
8. Gaunt, S.J., *The significance of Hox gene collinearity.* Int J Dev Biol, 2015. **59**(4-6): p. 159-70.
9. Lappin, T.R., et al., *HOX genes: seductive science, mysterious mechanisms.* Ulster Med J, 2006. **75**(1): p. 23-31.
10. Noordermeer, D., et al., *The dynamic architecture of Hox gene clusters.* Science, 2011. **334**(6053): p. 222-5.
11. Montavon, T., et al., *A regulatory archipelago controls Hox genes transcription in digits.* Cell, 2011. **147**(5): p. 1132-45.
12. Min, H., J.Y. Lee, and M.H. Kim, *Hoxc gene collinear expression and epigenetic modifications established during embryogenesis are maintained until after birth.* Int J Biol Sci, 2013. **9**(9): p. 960-5.
13. Wang, K.C., et al., *A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression.* Nature, 2011. **472**(7341): p. 120-4.
14. Sheth, R., et al., *"Self-regulation," a new facet of Hox genes' function.* Dev Dyn, 2014. **243**(1): p. 182-91.
15. Schuettengruber, B., et al., *Genome regulation by polycomb and trithorax proteins.* Cell, 2007. **128**(4): p. 735-45.
16. Maeda, R.K. and F. Karch, *The bithorax complex of Drosophila an exceptional Hox cluster.* Curr Top Dev Biol, 2009. **88**: p. 1-33.
17. Schwartz, Y.B. and V. Pirrotta, *Polycomb silencing mechanisms and the management of genomic programmes.* Nat Rev Genet, 2007. **8**(1): p. 9-22.
18. Soshnikova, N. and D. Duboule, *Epigenetic temporal control of mouse Hox genes in vivo.* Science, 2009. **324**(5932): p. 1320-3.
19. Chambeyron, S. and W.A. Bickmore, *Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription.* Genes Dev, 2004. **18**(10): p. 1119-30.
20. Noordermeer, D., et al., *Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci.* Elife, 2014. **3**: p. e02557.

21. Di-Poi, N., J. Zakany, and D. Duboule, *Distinct roles and regulations for HoxD genes in metanephric kidney development.* PLoS Genet, 2007. **3**(12): p. e232.
22. Godwin, A.R. and M.R. Capecchi, *Hoxc13 mutant mice lack external hair.* Genes Dev, 1998. **12**(1): p. 11-20.
23. Tschopp, P. and D. Duboule, *A regulatory 'landscape effect' over the HoxD cluster.* Dev Biol, 2011. **351**(2): p. 288-96.
24. Kogo, R., et al., *Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers.* Cancer Res, 2011. **71**(20): p. 6320-6.
25. Montavon, T. and D. Duboule, *Chromatin organization and global regulation of Hox gene clusters.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1620): p. 20120367.
26. Morata, G., et al., *Homeotic transformations of the abdominal segments of Drosophila caused by breaking or deleting a central portion of the bithorax complex.* J Embryol Exp Morphol, 1983. **78**: p. 319-41.
27. Struhl, G., *Role of the esc+ gene product in ensuring the selective expression of segment-specific homeotic genes in Drosophila.* J Embryol Exp Morphol, 1983. **76**: p. 297-331.
28. Akam, M., *The molecular basis for metameric pattern in the Drosophila embryo.* Development, 1987. **101**(1): p. 1-22.
29. Lufkin, T., et al., *Disruption of the Hox-1.6 homeobox gene results in defects in a region corresponding to its rostral domain of expression.* Cell, 1991. **66**(6): p. 1105-19.
30. Chisaka, O., T.S. Musci, and M.R. Capecchi, *Developmental defects of the ear, cranial nerves and hindbrain resulting from targeted disruption of the mouse homeobox gene Hox-1.6.* Nature, 1992. **355**(6360): p. 516-20.
31. Dolle, P., et al., *Disruption of the Hoxd-13 gene induces localized heterochrony leading to mice with neotenic limbs.* Cell, 1993. **75**(3): p. 431-41.
32. Lufkin, T., et al., *Homeotic transformation of the occipital bones of the skull by ectopic expression of a homeobox gene.* Nature, 1992. **359**(6398): p. 835-41.
33. Duboule, D. and G. Morata, *Colinearity and functional hierarchy among genes of the homeotic complexes.* Trends Genet, 1994. **10**(10): p. 358-64.
34. Lufkin, T., *Murine homeobox gene control of embryonic patterning and organogenesis*. 1st ed. Advances in developmental biology and biochemistry. 2003, Amsterdam ; Boston: Elsevier. xiii, 250 p., [13] p. of plates.
35. Zandvakili, A. and B. Gebelein, *Mechanisms of Specificity for Hox Factor Activity.* J Dev Biol, 2016. **4**(2).
36. Piper, D.E., et al., *Structure of a HoxB1-Pbx1 heterodimer bound to DNA: role of the hexapeptide and a fourth homeodomain helix in complex formation.* Cell, 1999. **96**(4): p. 587-97.
37. Passner, J.M., et al., *Structure of a DNA-bound Ultrabithorax-Extradenticle homeodomain complex.* Nature, 1999. **397**(6721): p. 714-9.
38. Affolter, M., M. Slattery, and R.S. Mann, *A lexicon for homeodomain-DNA recognition.* Cell, 2008. **133**(7): p. 1133-5.

39. LaRonde-LeBlanc, N.A. and C. Wolberger, *Structure of HoxA9 and Pbx1 bound to DNA: Hox hexapeptide and DNA recognition anterior to posterior.* Genes Dev, 2003. **17**(16): p. 2060-72.
40. Breitinger, C., et al., *The homeodomain region controls the phenotype of HOX-induced murine leukemia.* Blood, 2012. **120**(19): p. 4018-27.
41. Mann, R.S., K.M. Lelli, and R. Joshi, *Hox specificity unique roles for cofactors and collaborators.* Curr Top Dev Biol, 2009. **88**: p. 63-101.
42. Chan, S.K., et al., *The DNA binding specificity of Ultrabithorax is modulated by cooperative interactions with extradenticle, another homeoprotein.* Cell, 1994. **78**(4): p. 603-15.
43. Chang, C.P., et al., *Pbx modulation of Hox homeodomain amino-terminal arms establishes different DNA-binding specificities across the Hox locus.* Mol Cell Biol, 1996. **16**(4): p. 1734-45.
44. Shen, W.F., et al., *Hox homeodomain proteins exhibit selective complex stabilities with Pbx and DNA.* Nucleic Acids Res, 1996. **24**(5): p. 898-906.
45. Sun, B., et al., *Ultrabithorax protein is necessary but not sufficient for full activation of decapentaplegic expression in the visceral mesoderm.* EMBO J, 1995. **14**(3): p. 520-35.
46. Slattery, M., et al., *Cofactor binding evokes latent differences in DNA binding specificity between Hox proteins.* Cell, 2011. **147**(6): p. 1270-82.
47. Kim, S., et al., *Probing allostery through DNA.* Science, 2013. **339**(6121): p. 816-9.
48. Narasimhan, K., et al., *DNA-mediated cooperativity facilitates the co-selection of cryptic enhancer sequences by SOX2 and PAX6 transcription factors.* Nucleic Acids Res, 2015. **43**(3): p. 1513-28.
49. Joshi, R., et al., *Functional specificity of a Hox protein mediated by the recognition of minor groove structure.* Cell, 2007. **131**(3): p. 530-43.
50. Crocker, J., et al., *Low affinity binding site clusters confer hox specificity and regulatory robustness.* Cell, 2015. **160**(1-2): p. 191-203.
51. Crocker, J., E.P. Noon, and D.L. Stern, *The Soft Touch: Low-Affinity Transcription Factor Binding Sites in Development and Evolution.* Curr Top Dev Biol, 2016. **117**: p. 455-69.
52. Ramos, A.I. and S. Barolo, *Low-affinity transcription factor binding sites shape morphogen responses and enhancer evolution.* Philos Trans R Soc Lond B Biol Sci, 2013. **368**(1632): p. 20130018.
53. Hombria, J.C. and B. Lovegrove, *Beyond homeosis--HOX function in morphogenesis and organogenesis.* Differentiation, 2003. **71**(8): p. 461-76.
54. Collins, C., et al., *C/EBPalpha is an essential collaborator in Hoxa9/Meis1-mediated leukemogenesis.* Proc Natl Acad Sci U S A, 2014. **111**(27): p. 9899-904.
55. Ladam, F. and C.G. Sagerstrom, *Hox regulation of transcription: more complex(es).* Dev Dyn, 2014. **243**(1): p. 4-15.
56. Gordon, J.A., et al., *Pbx1 represses osteoblastogenesis by blocking Hoxa10-mediated recruitment of chromatin remodeling factors.* Mol Cell Biol, 2010. **30**(14): p. 3531-41.

57. Hassan, M.Q., et al., *HOXA10 controls osteoblastogenesis by directly activating bone regulatory and phenotypic genes.* Mol Cell Biol, 2007. **27**(9): p. 3337-52.
58. Gordon, J.A., et al., *Epigenetic regulation of early osteogenesis and mineralized tissue formation by a HOXA10-PBX1-associated complex.* Cells Tissues Organs, 2011. **194**(2-4): p. 146-50.
59. Saleh, M., et al., *Cell signaling switches HOX-PBX complexes from repressors to activators of transcription mediated by histone deacetylases and histone acetyltransferases.* Mol Cell Biol, 2000. **20**(22): p. 8623-33.
60. Chariot, A., et al., *CBP and histone deacetylase inhibition enhance the transactivation potential of the HOXB7 homeodomain-containing protein.* Oncogene, 1999. **18**(27): p. 4007-14.
61. Huang, Y., et al., *Identification and characterization of Hoxa9 binding sites in hematopoietic cells.* Blood, 2012. **119**(2): p. 388-98.
62. Wang, Z., et al., *GSK-3 promotes conditional association of CREB and its coactivators with MEIS1 to facilitate HOX-mediated transcription and oncogenesis.* Cancer Cell, 2010. **17**(6): p. 597-608.
63. Choe, S.K., et al., *Meis cofactors control HDAC and CBP accessibility at Hox-regulated promoters during zebrafish embryogenesis.* Dev Cell, 2009. **17**(4): p. 561-7.
64. Andrioli, L.P., et al., *Groucho-dependent repression by sloppy-paired 1 differentially positions anterior pair-rule stripes in the Drosophila embryo.* Dev Biol, 2004. **276**(2): p. 541-51.
65. Gebelein, B., D.J. McKay, and R.S. Mann, *Direct integration of Hox and segmentation gene inputs during Drosophila development.* Nature, 2004. **431**(7009): p. 653-9.
66. Lehnertz, B., et al., *The methyltransferase G9a regulates HoxA9-dependent transcription in AML.* Genes Dev, 2014. **28**(4): p. 317-27.
67. Boube, M., et al., *Drosophila melanogaster Hox transcription factors access the RNA polymerase II machinery through direct homeodomain binding to a conserved motif of mediator subunit Med19.* PLoS Genet, 2014. **10**(5): p. e1004303.
68. Choe, S.K., F. Ladam, and C.G. Sagerstrom, *TALE factors poise promoters for activation by Hox proteins.* Dev Cell, 2014. **28**(2): p. 203-11.
69. Choe, S.K. and C.G. Sagerstrom, *Variable Meis-dependence among paralog group-1 Hox proteins.* Biochem Biophys Res Commun, 2005. **331**(4): p. 1384-91.
70. Bell, O., et al., *Determinants and dynamics of genome accessibility.* Nat Rev Genet, 2011. **12**(8): p. 554-64.
71. Choo, S.W., R. White, and S. Russell, *Genome-wide analysis of the binding of the Hox protein Ultrabithorax and the Hox cofactor Homothorax in Drosophila.* PLoS One, 2011. **6**(4): p. e14778.
72. Beh, C.Y., et al., *Roles of cofactors and chromatin accessibility in Hox protein target specificity.* Epigenetics Chromatin, 2016. **9**: p. 1.
73. Collins, C.T. and J.L. Hess, *Role of HOXA9 in leukemia: dysregulation, cofactors and essential targets.* Oncogene, 2016. **35**(9): p. 1090-8.

74. Zeltser, L., C. Desplan, and N. Heintz, *Hoxb-13: a new Hox gene in a distant region of the HOXB cluster maintains colinearity.* Development, 1996. **122**(8): p. 2475-84.
75. Kim, M.H., et al., *Genomic structure and sequence analysis of human HOXA-9.* DNA Cell Biol, 1998. **17**(5): p. 407-14.
76. Fagerberg, L., et al., *Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics.* Mol Cell Proteomics, 2014. **13**(2): p. 397-406.
77. Pineault, N., et al., *Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny.* Exp Hematol, 2002. **30**(1): p. 49-57.
78. Ramos-Mejia, V., et al., *HOXA9 promotes hematopoietic commitment of human embryonic stem cells.* Blood, 2014. **124**(20): p. 3065-75.
79. Beachy, S.H., et al., *Isolated Hoxa9 over expression predisposes to the development of lymphoid but not myeloid leukemia.* Exp Hematol, 2013. **41**(6): p. 518-529 e5.
80. Thorsteinsdottir, U., et al., *Over expression of the myeloid leukemia-associated Hoxa9 gene in bone marrow cells induces stem cell expansion.* Blood, 2002. **99**(1): p. 121-9.
81. Izon, D.J., et al., *Loss of function of the homeobox gene Hoxa-9 perturbs early T-cell development and induces apoptosis in primitive thymocytes.* Blood, 1998. **92**(2): p. 383-93.
82. Lawrence, H.J., et al., *Mice bearing a targeted interruption of the homeobox gene HOXA9 have defects in myeloid, erythroid, and lymphoid hematopoiesis.* Blood, 1997. **89**(6): p. 1922-30.
83. Magnusson, M., et al., *Hoxa9/hoxb3/hoxb4 compound null mice display severe hematopoietic defects.* Exp Hematol, 2007. **35**(9): p. 1421-8.
84. Figueroa, M.E., et al., *DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia.* Cancer Cell, 2010. **17**(1): p. 13-27.
85. Imamura, T., et al., *Frequent co-expression of HoxA9 and Meis1 genes in infant acute lymphoblastic leukaemia with MLL rearrangement.* Br J Haematol, 2002. **119**(1): p. 119-21.
86. Grier, D.G., et al., *The pathophysiology of HOX genes and their role in cancer.* J Pathol, 2005. **205**(2): p. 154-71.
87. Andreeff, M., et al., *HOX expression patterns identify a common signature for favorable AML.* Leukemia, 2008. **22**(11): p. 2041-7.
88. Golub, T.R., et al., *Molecular classification of cancer: class discovery and class prediction by gene expression monitoring.* Science, 1999. **286**(5439): p. 531-7.
89. Cao, F., et al., *Targeting MLL1 H3K4 methyltransferase activity in mixed-lineage leukemia.* Mol Cell, 2014. **53**(2): p. 247-61.
90. Hess, J.L., *MLL: a histone methyltransferase disrupted in leukemia.* Trends Mol Med, 2004. **10**(10): p. 500-7.
91. Winters, A.C. and K.M. Bernt, *MLL-Rearranged Leukemias-An Update on Science and Clinical Approaches.* Front Pediatr, 2017. **5**: p. 4.
92. Dou, Y. and J.L. Hess, *Mechanisms of transcriptional regulation by MLL and its disruption in acute leukemia.* Int J Hematol, 2008. **87**(1): p. 10-8.

93. Xu, J., et al., *MLL1 and MLL1 fusion proteins have distinct functions in regulating leukemic transcription program.* Cell Discov, 2016. **2**: p. 16008.
94. Mullighan, C.G., et al., *Pediatric acute myeloid leukemia with NPM1 mutations is characterized by a gene expression profile with dysregulated HOX gene expression distinct from MLL-rearranged leukemias.* Leukemia, 2007. **21**(9): p. 2000-9.
95. Khan, S.N., et al., *Multiple mechanisms deregulate EZH2 and histone H3 lysine 27 epigenetic changes in myeloid malignancies.* Leukemia, 2013. **27**(6): p. 1301-9.
96. Bansal, D., et al., *Cdx4 dysregulates Hox gene expression and generates acute myeloid leukemia alone and in cooperation with Meis1a in a murine model.* Proc Natl Acad Sci U S A, 2006. **103**(45): p. 16924-9.
97. Shima, H., et al., *Bromodomain-PHD finger protein 1 is critical for leukemogenesis associated with MOZ-TIF2 fusion.* Int J Hematol, 2014. **99**(1): p. 21-31.
98. Speleman, F., et al., *A new recurrent inversion, inv(7)(p15q34), leads to transcriptional activation of HOXA10 and HOXA11 in a subset of T-cell acute lymphoblastic leukemias.* Leukemia, 2005. **19**(3): p. 358-66.
99. Novak, R.L., et al., *Gene expression profiling and candidate gene resequencing identifies pathways and mutations important for malignant transformation caused by leukemogenic fusion genes.* Exp Hematol, 2012. **40**(12): p. 1016-27.
100. Ferrando, A.A., et al., *Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation.* Blood, 2003. **102**(1): p. 262-8.
101. Dik, W.A., et al., *CALM-AF10+ T-ALL expression profiles are characterized by over expression of HOXA and BMI1 oncogenes.* Leukemia, 2005. **19**(11): p. 1948-57.
102. Soulier, J., et al., *HOXA genes are included in genetic and biologic networks defining human acute T-cell leukemia (T-ALL).* Blood, 2005. **106**(1): p. 274-86.
103. Roth, J.J., R.C. Crist, and A.M. Buchberg, *Might as well face it: MLL's addicted to HOX.* Blood, 2009. **113**(11): p. 2372-3.
104. Faber, J., et al., *HOXA9 is required for survival in human MLL-rearranged acute leukemias.* Blood, 2009. **113**(11): p. 2375-85.
105. Kumar, A.R., et al., *A role for MEIS1 in MLL-fusion gene leukemia.* Blood, 2009. **113**(8): p. 1756-8.
106. Wong, P., et al., *Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential.* Genes Dev, 2007. **21**(21): p. 2762-74.
107. Orlovsky, K., et al., *Down-regulation of homeobox genes MEIS1 and HOXA in MLL-rearranged acute leukemia impairs engraftment and reduces proliferation.* Proc Natl Acad Sci U S A, 2011. **108**(19): p. 7956-61.
108. Kroon, E., et al., *Hoxa9 transforms primary bone marrow cells through specific collaboration with Meis1a but not Pbx1b.* EMBO J, 1998. **17**(13): p. 3714-25.
109. Wang, G.G., M.P. Pasillas, and M.P. Kamps, *Meis1 programs transcription of FLT3 and cancer stem cell character, using a mechanism that requires interaction with Pbx and a novel function of the Meis1 C-terminus.* Blood, 2005. **106**(1): p. 254-64.

110. Gibbs, K.D., Jr., et al., *Decoupling of tumor-initiating activity from stable immunophenotype in HoxA9-Meis1-driven AML.* Cell Stem Cell, 2012. **10**(2): p. 210-7.
111. Argiropoulos, B., E. Yung, and R.K. Humphries, *Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis.* Genes Dev, 2007. **21**(22): p. 2845-9.
112. Armstrong, S.A., et al., *MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia.* Nat Genet, 2002. **30**(1): p. 41-7.
113. Pigazzi, M., et al., *Presence of high-ERG expression is an independent unfavorable prognostic marker in MLL-rearranged childhood myeloid leukemia.* Blood, 2012. **119**(4): p. 1086-7; author reply 1087-8.
114. Chapuis, N., 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): p. 415-23.
115. Steger, J., et al., *Insulin-like growth factor 1 is a direct HOXA9 target important for hematopoietic transformation.* Leukemia, 2015. **29**(4): p. 901-8.
116. Hu, Y.L., et al., *Evidence that the Pim1 kinase gene is a direct target of HOXA9.* Blood, 2007. **109**(11): p. 4732-8.
117. Whelan, J.T., D.L. Ludwig, and F.E. Bertrand, *HoxA9 induces insulin-like growth factor-1 receptor expression in B-lineage acute lymphoblastic leukemia.* Leukemia, 2008. **22**(6): p. 1161-9.
118. Breitinger, C., et al., *HOX genes regulate Rac1 activity in hematopoietic cells through control of Vav2 expression.* Leukemia, 2013. **27**(1): p. 236-8.
119. Reinert, T., et al., *Diagnosis of bladder cancer recurrence based on urinary levels of EOMES, HOXA9, POU4F2, TWIST1, VIM, and ZNF154 hypermethylation.* PLoS One, 2012. **7**(10): p. e46297.
120. Wu, Q., et al., *DNA methylation profiling of ovarian carcinomas and their in vitro models identifies HOXA9, HOXB5, SCGB3A1, and CRABP1 as novel targets.* Mol Cancer, 2007. **6**: p. 45.
121. Kuo, C.C., et al., *Frequent methylation of HOXA9 gene in tumor tissues and plasma samples from human hepatocellular carcinomas.* Clin Chem Lab Med, 2014. **52**(8): p. 1235-45.
122. Sun, M., et al., *HMGA2/TET1/HOXA9 signaling pathway regulates breast cancer growth and metastasis.* Proc Natl Acad Sci U S A, 2013. **110**(24): p. 9920-5.
123. Rauch, T., et al., *Homeobox gene methylation in lung cancer studied by genome-wide analysis with a microarray-based methylated CpG island recovery assay.* Proc Natl Acad Sci U S A, 2007. **104**(13): p. 5527-32.
124. Hwang, J.A., et al., *HOXA9 inhibits migration of lung cancer cells and its hypermethylation is associated with recurrence in non-small cell lung cancer.* Mol Carcinog, 2015. **54 Suppl 1**: p. E72-80.
125. Ko, S.Y., et al., *HOXA9 promotes ovarian cancer growth by stimulating cancer-associated fibroblasts.* J Clin Invest, 2012. **122**(10): p. 3603-17.
126. Gilbert, P.M., et al., *HOXA9 regulates BRCA1 expression to modulate human breast tumor phenotype.* J Clin Invest, 2010. **120**(5): p. 1535-50.

127. Kim, Y.J., et al., *HOXA9, ISL1 and ALDH1A3 methylation patterns as prognostic markers for nonmuscle invasive bladder cancer: array-based DNA methylation and expression profiling.* Int J Cancer, 2013. **133**(5): p. 1135-42.
128. Bhatlekar, S., et al., *Over expression of HOXA4 and HOXA9 genes promotes self-renewal and contributes to colon cancer stem cell overpopulation.* J Cell Physiol, 2017.
129. Malek, R., et al., *TWIST1-WDR5-Hottip Regulates Hoxa9 Chromatin to Facilitate Prostate Cancer Metastasis.* Cancer Res, 2017. **77**(12): p. 3181-3193.
130. Goldberg, A.D., C.D. Allis, and E. Bernstein, *Epigenetics: a landscape takes shape.* Cell, 2007. **128**(4): p. 635-8.
131. Waddington, C.H., *The strategy of the genes; a discussion of some aspects of theoretical biology*. 1957, London,: Allen & Unwin. ix, 262 p.
132. Falkenberg, K.J. and R.W. Johnstone, *Histone deacetylases and their inhibitors in cancer, neurological diseases and immune disorders.* Nat Rev Drug Discov, 2014. **13**(9): p. 673-91.
133. Goll, M.G. and T.H. Bestor, *Eukaryotic cytosine methyltransferases.* Annu Rev Biochem, 2005. **74**: p. 481-514.
134. Surani, M.A., K. Hayashi, and P. Hajkova, *Genetic and epigenetic regulators of pluripotency.* Cell, 2007. **128**(4): p. 747-62.
135. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.
136. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.
137. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.
138. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21931-6.
139. Struhl, K., *Histone acetylation and transcriptional regulatory mechanisms.* Genes Dev, 1998. **12**(5): p. 599-606.
140. Rando, O.J., *Combinatorial complexity in chromatin structure and function: revisiting the histone code.* Curr Opin Genet Dev, 2012. **22**(2): p. 148-55.
141. Smith, C.L. and C.L. Peterson, *ATP-dependent chromatin remodeling.* Curr Top Dev Biol, 2005. **65**: p. 115-48.
142. Orkin, S.H. and L.I. Zon, *Hematopoiesis: an evolving paradigm for stem cell biology.* Cell, 2008. **132**(4): p. 631-44.
143. Galloway, J.L., et al., *Loss of gata1 but not gata2 converts erythropoiesis to myelopoiesis in zebrafish embryos.* Dev Cell, 2005. **8**(1): p. 109-16.
144. Rosenbauer, F. and D.G. Tenen, *Transcription factors in myeloid development: balancing differentiation with transformation.* Nat Rev Immunol, 2007. **7**(2): p. 105-17.
145. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.

146. Jin, F., et al., *PU.1 and C/EBP(alpha) synergistically program distinct response to NF-kappaB activation through establishing monocyte specific enhancers.* Proc Natl Acad Sci U S A, 2011. **108**(13): p. 5290-5.
147. Letting, D.L., et al., *Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1.* Mol Cell Biol, 2003. **23**(4): p. 1334-40.
148. Adli, M., J. Zhu, and B.E. Bernstein, *Genome-wide chromatin maps derived from limited numbers of hematopoietic progenitors.* Nat Methods, 2010. **7**(8): p. 615-8.
149. Lara-Astiaso, D., et al., *Immunogenetics. Chromatin state dynamics during blood formation.* Science, 2014. **345**(6199): p. 943-9.
150. Ji, H., et al., *Comprehensive methylome map of lineage commitment from haematopoietic progenitors.* Nature, 2010. **467**(7313): p. 338-42.
151. Broske, A.M., et al., *DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction.* Nat Genet, 2009. **41**(11): p. 1207-15.
152. Trowbridge, J.J., et al., *DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells.* Cell Stem Cell, 2009. **5**(4): p. 442-9.
153. Henikoff, S., *Nucleosome destabilization in the epigenetic regulation of gene expression.* Nat Rev Genet, 2008. **9**(1): p. 15-26.
154. Boyle, A.P., et al., *High-resolution genome-wide in vivo footprinting of diverse transcription factors in human cells.* Genome Res, 2011. **21**(3): p. 456-64.
155. Corces, M.R., et al., *Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution.* Nat Genet, 2016. **48**(10): p. 1193-203.
156. Ntziachristos, P., O. Abdel-Wahab, and I. Aifantis, *Emerging concepts of epigenetic dysregulation in hematological malignancies.* Nat Immunol, 2016. **17**(9): p. 1016-24.
157. Morin, R.D., et al., *Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin.* Nat Genet, 2010. **42**(2): p. 181-5.
158. Jiang, Y. and A. Melnick, *The epigenetic basis of diffuse large B-cell lymphoma.* Semin Hematol, 2015. **52**(2): p. 86-96.
159. Hnisz, D., et al., *Super-enhancers in the control of cell identity and disease.* Cell, 2013. **155**(4): p. 934-47.
160. Shi, J., et al., *Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation.* Genes Dev, 2013. **27**(24): p. 2648-62.
161. Gallipoli, P., G. Giotopoulos, and B.J. Huntly, *Epigenetic regulators as promising therapeutic targets in acute myeloid leukemia.* Ther Adv Hematol, 2015. **6**(3): p. 103-19.
162. Shi, J., et al., *Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains.* Nat Biotechnol, 2015. **33**(6): p. 661-7.
163. Loven, J., et al., *Selective inhibition of tumor oncogenes by disruption of super-enhancers.* Cell, 2013. **153**(2): p. 320-34.

164. Roe, J.S., et al., *BET Bromodomain Inhibition Suppresses the Function of Hematopoietic Transcription Factors in Acute Myeloid Leukemia.* Mol Cell, 2015. **58**(6): p. 1028-39.
165. Liu, Z., et al., *Drug Discovery Targeting Bromodomain-Containing Protein 4.* J Med Chem, 2017. **60**(11): p. 4533-4558.
166. Jo, S.Y., et al., *Requirement for Dot1l in murine postnatal hematopoiesis and leukemogenesis by MLL translocation.* Blood, 2011. **117**(18): p. 4759-68.
167. Daigle, S.R., et al., *Potent inhibition of DOT1L as treatment of MLL-fusion leukemia.* Blood, 2013. **122**(6): p. 1017-25.
168. Okuda, H., et al., *Cooperative gene activation by AF4 and DOT1L drives MLL-rearranged leukemia.* J Clin Invest, 2017. **127**(5): p. 1918-1931.
169. Dafflon, C., et al., *Complementary activities of DOT1L and Menin inhibitors in MLL-rearranged leukemia.* Leukemia, 2017. **31**(6): p. 1269-1277.
170. Morera, L., M. Lubbert, and M. Jung, *Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy.* Clin Epigenetics, 2016. **8**: p. 57.
171. Herz, H.M., D. Hu, and A. Shilatifard, *Enhancer malfunction in cancer.* Mol Cell, 2014. **53**(6): p. 859-66.
172. Sur, I. and J. Taipale, *The role of enhancers in cancer.* Nat Rev Cancer, 2016. **16**(8): p. 483-93.
173. Riggi, N., et al., *EWS-FLI1 utilizes divergent chromatin remodeling mechanisms to directly activate or repress enhancer elements in Ewing sarcoma.* Cancer Cell, 2014. **26**(5): p. 668-81.
174. Iwafuchi-Doi, M. and K.S. Zaret, *Pioneer transcription factors in cell reprogramming.* Genes Dev, 2014. **28**(24): p. 2679-92.
175. Cho, Y.W., et al., *Histone methylation regulator PTIP is required for PPARgamma and C/EBPalpha expression and adipogenesis.* Cell Metab, 2009. **10**(1): p. 27-39.
176. Zuber, J., et al., *Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi.* Nat Biotechnol, 2011. **29**(1): p. 79-83.
177. Li, Y., et al., *CRISPR reveals a distal super-enhancer required for Sox2 expression in mouse embryonic stem cells.* PLoS One, 2014. **9**(12): p. e114485.
178. Stadhouders, R., et al., *Multiplexed chromosome conformation capture sequencing for rapid genome-scale high-resolution detection of long-range chromatin interactions.* Nat Protoc, 2013. **8**(3): p. 509-24.
179. McLean, C.Y., et al., *GREAT improves functional interpretation of cis-regulatory regions.* Nat Biotechnol, 2010. **28**(5): p. 495-501.
180. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources.* Nat Protoc, 2009. **4**(1): p. 44-57.
181. de Hoon, M.J., et al., *Open source clustering software.* Bioinformatics, 2004. **20**(9): p. 1453-4.
182. Ross-Innes, C.S., et al., *Differential oestrogen receptor binding is associated with clinical outcome in breast cancer.* Nature, 2012. **481**(7381): p. 389-93.
183. Calo, E. and J. Wysocka, *Modification of enhancer chromatin: what, how, and why?* Mol Cell, 2013. **49**(5): p. 825-37.

184. Natoli, G., *Maintaining cell identity through global control of genomic organization.* Immunity, 2010. **33**(1): p. 12-24.
185. van Oevelen, C., et al., *C/EBPalpha Activates Pre-existing and De Novo Macrophage Enhancers during Induced Pre-B Cell Transdifferentiation and Myelopoiesis.* Stem Cell Reports, 2015. **5**(2): p. 232-47.
186. Ye, M., et al., *Hematopoietic Differentiation Is Required for Initiation of Acute Myeloid Leukemia.* Cell Stem Cell, 2015. **17**(5): p. 611-23.
187. Somervaille, T.C., et al., *Hierarchical maintenance of MLL myeloid leukemia stem cells employs a transcriptional program shared with embryonic rather than adult stem cells.* Cell Stem Cell, 2009. **4**(2): p. 129-40.
188. Mallo, M., D.M. Wellik, and J. Deschamps, *Hox genes and regional patterning of the vertebrate body plan.* Dev Biol, 2010. **344**(1): p. 7-15.
189. Campos, L., et al., *High expression of bcl-2 protein in acute myeloid leukemia cells is associated with poor response to chemotherapy.* Blood, 1993. **81**(11): p. 3091-6.
190. Kontro, M., et al., *HOX gene expression predicts response to BCL-2 inhibition in acute myeloid leukemia.* Leukemia, 2017. **31**(2): p. 301-309.
191. Rhinn, M. and P. Dolle, *Retinoic acid signalling during development.* Development, 2012. **139**(5): p. 843-58.
192. Duan, J.J., et al., *ALDH1A3, a metabolic target for cancer diagnosis and therapy.* Int J Cancer, 2016. **139**(5): p. 965-75.
193. Marcato, P., et al., *Aldehyde dehydrogenase activity of breast cancer stem cells is primarily due to isoform ALDH1A3 and its expression is predictive of metastasis.* Stem Cells, 2011. **29**(1): p. 32-45.
194. Perez-Alea, M., et al., *ALDH1A3 is epigenetically regulated during melanocyte transformation and is a target for melanoma treatment.* Oncogene, 2017.
195. Sullivan, K.E., et al., *The stem cell/cancer stem cell marker ALDH1A3 regulates the expression of the survival factor tissue transglutaminase, in mesenchymal glioma stem cells.* Oncotarget, 2017. **8**(14): p. 22325-22343.
196. Magnani, L., et al., *PBX1 genomic pioneer function drives ERalpha signaling underlying progression in breast cancer.* PLoS Genet, 2011. **7**(11): p. e1002368.
197. Lewis, M.T., *Homeobox genes in mammary gland development and neoplasia.* Breast Cancer Res, 2000. **2**(3): p. 158-69.
198. Akhtar-Zaidi, B., et al., *Epigenomic enhancer profiling defines a signature of colon cancer.* Science, 2012. **336**(6082): p. 736-9.
199. Suzuki, A., et al., *Aberrant transcriptional regulations in cancers: genome, transcriptome and epigenome analysis of lung adenocarcinoma cell lines.* Nucleic Acids Res, 2014. **42**(22): p. 13557-72.
200. Rogenhofer, S., et al., *Histone methylation defines an epigenetic entity in penile squamous cell carcinoma.* J Urol, 2013. **189**(3): p. 1117-22.
201. Zhang, J., et al., *Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis.* Nat Med, 2015. **21**(10): p. 1190-8.
202. Rao, R.C. and Y. Dou, *Hijacked in cancer: the KMT2 (MLL) family of methyltransferases.* Nat Rev Cancer, 2015. **15**(6): p. 334-46.

203. Mohan, M., et al., *Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis.* Nat Rev Cancer, 2010. **10**(10): p. 721-8.
204. Wang, P., et al., *Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II.* Mol Cell Biol, 2009. **29**(22): p. 6074-85.
205. Ortega-Molina, A., et al., *The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development.* Nat Med, 2015. **21**(10): p. 1199-208.
206. Santos, M.A., et al., *DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier.* Nature, 2014. **514**(7520): p. 107-11.
207. Mishra, B.P., et al., *The histone methyltransferase activity of MLL1 is dispensable for hematopoiesis and leukemogenesis.* Cell Rep, 2014. **7**(4): p. 1239-47.
208. Ranghini, E.J. and G.R. Dressler, *Evidence for intermediate mesoderm and kidney progenitor cell specification by Pax2 and PTIP dependent mechanisms.* Dev Biol, 2015. **399**(2): p. 296-305.
209. Cho, Y.W., et al., *PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex.* J Biol Chem, 2007. **282**(28): p. 20395-406.
210. Lee, J.E., et al., *H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation.* Elife, 2013. **2**: p. e01503.
211. Wang, C., et al., *Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition.* Proc Natl Acad Sci U S A, 2016. **113**(42): p. 11871-11876.
212. Daniel, J.A., et al., *PTIP promotes chromatin changes critical for immunoglobulin class switch recombination.* Science, 2010. **329**(5994): p. 917-23.
213. Starnes, L.M., et al., *A PTIP-PA1 subcomplex promotes transcription for IgH class switching independently from the associated MLL3/MLL4 methyltransferase complex.* Genes Dev, 2016. **30**(2): p. 149-63.
214. Wang, X., K. Takenaka, and S. Takeda, *PTIP promotes DNA double-strand break repair through homologous recombination.* Genes Cells, 2010. **15**(3): p. 243-54.
215. Wu, J., et al., *PTIP regulates 53BP1 and SMC1 at the DNA damage sites.* J Biol Chem, 2009. **284**(27): p. 18078-84.
216. Kaikkonen, M.U., et al., *Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription.* Mol Cell, 2013. **51**(3): p. 310-25.
217. Dorighi, K.M., et al., *Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation.* Mol Cell, 2017. **66**(4): p. 568-576 e4.
218. Lai, B., et al., *MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis.* Nucleic Acids Res, 2017.
219. Sze, C.C. and A. Shilatifard, *MLL3/MLL4/COMPASS Family on Epigenetic Regulation of Enhancer Function and Cancer.* Cold Spring Harb Perspect Med, 2016. **6**(11).
220. Haferlach, T., et al., *Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International*

Microarray Innovations in Leukemia Study Group. J Clin Oncol, 2010. **28**(15): p. 2529-37.
221. Figueroa, M.E., et al., *Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia.* J Clin Invest, 2013. **123**(7): p. 3099-111.
222. Cristancho, A.G. and M.A. Lazar, *Forming functional fat: a growing understanding of adipocyte differentiation.* Nat Rev Mol Cell Biol, 2011. **12**(11): p. 722-34.
223. White, R.T., et al., *Human adipsin is identical to complement factor D and is expressed at high levels in adipose tissue.* J Biol Chem, 1992. **267**(13): p. 9210-3.
224. Peake, P.W., et al., *Detection and quantification of the control proteins of the alternative pathway of complement in 3T3-L1 adipocytes.* Eur J Clin Invest, 1997. **27**(11): p. 922-7.
225. Cousin, B., et al., *A role for preadipocytes as macrophage-like cells.* FASEB J, 1999. **13**(2): p. 305-12.
226. Cinti, S., *Transdifferentiation properties of adipocytes in the adipose organ.* Am J Physiol Endocrinol Metab, 2009. **297**(5): p. E977-86.
227. Charriere, G., et al., *Preadipocyte conversion to macrophage. Evidence of plasticity.* J Biol Chem, 2003. **278**(11): p. 9850-5.
228. Feng, R., et al., *PU.1 and C/EBPalpha/beta convert fibroblasts into macrophage-like cells.* Proc Natl Acad Sci U S A, 2008. **105**(16): p. 6057-62.
229. Cantile, M., et al., *HOX gene network is involved in the transcriptional regulation of in vivo human adipogenesis.* J Cell Physiol, 2003. **194**(2): p. 225-36.
230. Dankel, S.N., et al., *Switch from stress response to homeobox transcription factors in adipose tissue after profound fat loss.* PLoS One, 2010. **5**(6): p. e11033.
231. Ntambi, J.M. and K. Young-Cheul, *Adipocyte differentiation and gene expression.* J Nutr, 2000. **130**(12): p. 3122S-3126S.