The Role of the MLL-HOXA9 Axis in Normal and Malignant

Hematopoiesis

by
Jingya Wang

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Molecular and Cellular Pathology)
in the University of Michigan
2013

Doctoral Committee:

Professor Jay L. Hess, Chair
Assistant Professor Tomasz Cierpicki
Professor Eric R. Fearon
Assistant Professor Maria E. Figueroa
Professor Tom K. Kerppola
Associate Professor Xiaochun Yu

To
My parents
and
My husband

Acknowledgement

This dissertation would not have been possible without the tremendous support and guidance from my advisor, Dr. Jay L. Hess. Jay's enthusiasm and dedication towards scientific research have always been a great source of inspiration. His scientific vision and openness to novel ideas greatly influenced my scientific career. I would also like to thank Jay for encouraging me to pursue my interest in statistics and the opportunity to apply my knowledge to biomedical research. I deeply appreciate Jay's invaluable advice and generous support for the next stage of my career.

My sincere appreciations also go to Drs. Andrew Muntean and Maria ("Ken") Figueroa. Throughout my graduate study, Andy was actively involved in many aspects of my projects and my scientific career, including providing technical assistance and scientific thinking, and most importantly, sharing his insights of science and being a better scientist. I could not have been where I am without his generous help. Ken led me into the field of high throughput sequencing data analysis, and patiently discussed experimental design to algorithm in details. Her presence also contributes greatly to the completion of this thesis.

I would like to thank my thesis committee members, Drs. Eric R. Fearon, Tom K. Kerppola, Xiaochun Yu, Maria ("Ken") E. Figueroa, and Tomasz Cierpicki for

Preface

This dissertation comprises the research I conducted in Dr. Jay Hess' lab from 2009 to 2013. The objective was to better understand the role of the MLL-HOXA9 axis in normal and malignant hematopoiesis. Specifically, I wanted to explore: How is the activity of MLL regulated during hematopoiesis? What are the mechanisms of MLL fusion mediated transformation? And how does HOXA9 regulate gene expression and mediate leukemogenesis?

The first two questions were tackled by studying the function of the PHD/Bromo region of MLL, which is invariably disrupted or deleted in oncogenic MLL fusion proteins and is incompatible to MLL fusion-mediated transformation. My studies revealed that this region facilitates MLL ubiquitination in multiple ways.

Chapter 2 describes the role of MLL bromodomain in recruitment of the ECS^{ASB2} E3 ligase complex. This work was published in *Blood*, 2012 (Wang, J., Muntean, A.G., and Hess, J.L. ECS^{ASB2} mediates MLL degradation during hematopoietic differentiation). Chapter 3 describes the intrinsic E3 ligase activity of the second PHD finger. This work was published in *The Journal of Biological Chemistry,* 2012 (Wang, J., Muntean, A.G., Wu, L., and Hess, J.L. A subset of mixed lineage leukemia proteins has PHD domain-mediated E3 ligase activity).

The third question was tackled in Chapter 4 by studying the cooperativity between Hoxa9 and its potential cofactors C/ebpα and Pu.1. This on-going project is highly collaborative: Joel Bronstein focused on protein biochemistry assays; Cailin Collins performed all the molecular and cellular assays and prepared libraries for ChIP-sequencing and RNA-sequencing; and I performed analysis of the high throughput sequencing data, which is presented in this dissertation. Chapter 4 also serves as the writing component to partially fulfill the requirements for a dual Master's degree in Statistics.

Table of Contents

List of Figures

List of Tables

xix

List of Appendices

Chapter1   Introduction

1.1   MLL in normal and malignant hematopoiesis

1.1.1   Protein structure and molecular function of wild-type MLL

The histone H3 lysine4 (H3K4) methyltransferase mixed lineage leukemia protein

MLL (MLL1) is the human homolog of the Drosophila trithorax protein (Trx). MLL

is a large protein (3,969 amino acids) with multiple domains. Full length MLL is

proteolytically cleaved by the protease Taspase1 into a 320 kDa MLL^N fragment and

a 180 kDa MLL^C fragment. MLL^N and MLL^C non-covalently associate and function

as a holocomplex (Hsieh et al., 2003a; Hsieh et al., 2003b). On MLL^N, there are three

AT hooks with DNA binding capability (Birke et al., 2002), a DNA methyltransferase

homology region (or CxxC domain) which binds to CpG DNA and also recruits the

PAF complex through its flanking regions (Cierpicki et al., 2010; Hess, 2004b;

Muntean et al., 2010), a poorly understood PHD/Bromo region that contains four

PHD fingers (PHD1-4) and a bromodomain between the third and fourth PHD finger,

and a FYRN motif which together with the C terminal FYRC motif mediates the

interaction between MLL^N and MLL^C (Hsieh et al., 2003b). On MLL^C, there is a

transactivation domain (Ernst et al., 2001), a FYRC motif, and a SET (Su(var)3–9,

enhancer of zeste, trithorax) domain which has intrinsic H3K4 methyltransferase

activity (Milne et al., 2002; Nakamura et al., 2002) (Figure 1.1).

Figure 1.1 Schematic diagram showing the domain structure of wild-type MLL,

oncogenic MLL fusion proteins and MLL PTD
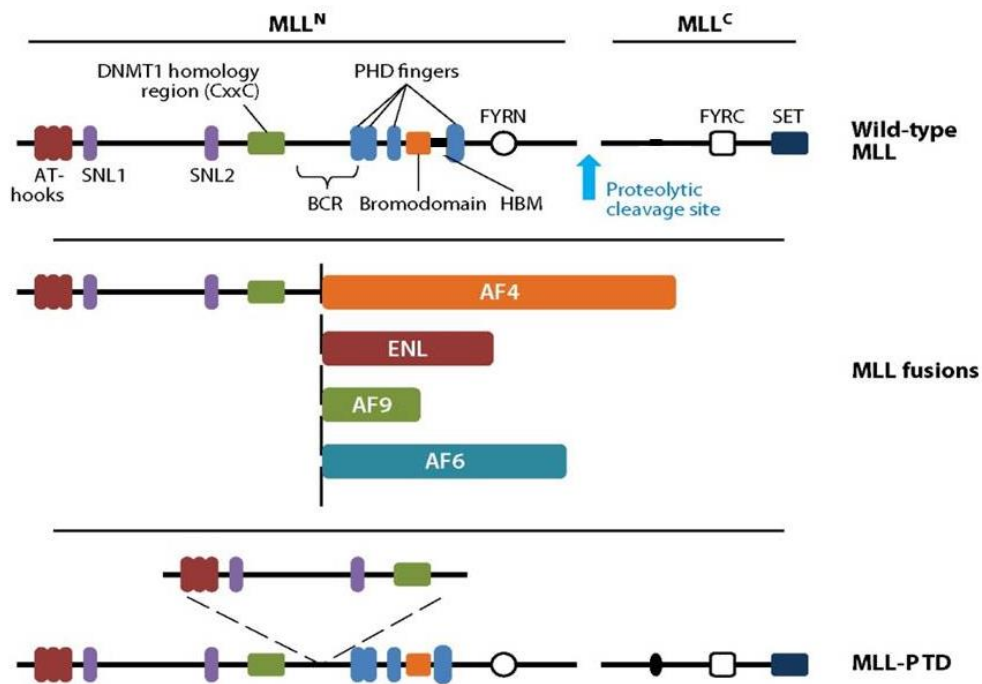
Figure 1.1. Schematic diagram showing the domain structure of wild-type MLL, oncogenic MLL fusion proteins and MLL PTD. (Top) Domain structure of wild-type MLL. Full length MLL is proteolytically cleaved by the protease Taspase1 into a 320 kDa MLL^N fragment and a 180 kDa MLL^C fragment. On MLL^N, there are three AT hooks, a DNA methyltransferase homology region (or CxxC domain), a PHD/Bromo region that contains four PHD fingers (PHD1-4) and a bromodomain between the third and fourth PHD finger, and a FYRN motif. On MLL^C, there is a transactivation domain, a FYRC motif, and a SET (Su(var)3–9, enhancer of zeste, trithorax) domain. (Middle) Chromosomal translocations involving MLL result in chimeric MLL-fusion proteins that include the N-terminal sequence of MLL up to the BCR (dotted vertical line), followed by one of several different fusion partners. Shown are examples of fusion partner proteins, including AF4, ENL, AF9, and AF6. MLL-fusion proteins invariably retain AT-hooks and the CxxC domain, while losing the downstream PHD fingers and beyond. (Bottom) MLL partial tandem duplications (MLL-PTD) leads to duplication of MLL sequences comprising the AT-hooks and the CxxC domain that are inserted at the BCR. Figure is modified from (Muntean and Hess, 2012).

MLL forms a complex with ASH2, WDR5, hDPY30 and RBBP5 and mediates methylation on histone H3 lysine4. This "core complex" is required for the maximal enzymatic activity of MLL as knockdown of these components leads to impaired MLL methyltransferase activity and deregulation of MLL target genes (Nakamura et al., 2002; Wysocka et al., 2005). Methylation of H3K4, particularly trimethylation, is closely associated with transcriptional activation, so MLL functions as a positive regulator of gene expression. MLL is not a global regulator of H3K4 methylation, as knockout of MLL in mouse embryonic fibroblast (MEF) cells only leads to moderate decrease of total H3K4me3 (Gregory et al., 2007). Further, it was shown that MLL is required for H3K4me3 of only ~5% of the promoters carrying this modification (Wang et al., 2009b). However, MLL controls H3K4me3 and maintains expression of a subset of genes that play important roles in development and hematopoiesis (Wang et al., 2009b). The best-studied MLL target genes are homeobox (*HOX*) transcription factors, especially *HOXA9* and its co-factor *MEIS1*, which play critical roles in maintaining hematopoietic stem cell identity (discussed below) (Argiropoulos and Humphries, 2007). Other important MLL target genes include the transcription factor EVI-1, another regulator of hematopoiesis (Arai et al., 2011), and the receptor tyrosine kinase EPHA7, which regulates MAPK/ERK pathway (Nakanishi et al., 2007) .

1.1.2   MLL in normal hematopoiesis

Different models have been established to study the function of MLL at different stages

of hematopoiesis. *Mll* knockout mice are embryonic lethal, while heterozygous mice

show retarded growth, decreased hematopoietic cell populations, and deregulated

expression of *Hox* genes (Yu et al., 1995). The yolk sac progenitors from the knockout

mice have defects in differentiating into multiple hematopoietic lineages (Hess et al.,

1997). Similar phenotype was also observed in another *Mll* mutant mouse model, in

which exons 12-14 was replaced with PGK-neo (Yagi et al., 1998). However, mice in

which only the SET domain is deleted from *Mll* do not show significant embryonic

lethality, although the Hox genes show reduced expression, indicating that the

methyltransferase activity does not account for the complete function of MLL

(Terranova et al., 2006). Further, inducible knockout mouse model was used to study

the role of MLL in adult hematopoiesis. Excision of *Mll* in adult mice leads to fatal

bone marrow cytopenia within three weeks, probably caused loss of self-renewal in

hematopoietic stem cells and decreased proliferation of multiple hematopoietic

progenitors (Jude et al., 2007). Together, these data show that Mll is required for both

embryonic and adult hematopoiesis.

1.1.3   MLL fusion mediated leukemia

1.1.3.1   MLL translocations in human leukemia

The *MLL* gene is located at chromosome 11q23. Chromosomal translocations

involving *MLL* are one of the most common genetic alterations in human leukemia,

4

representing about 10% of leukemia cases overall (Marschalek, 2011). These include

up to ~80% of infant acute myeloid (AML) and lymphoid (ALL) leukemia, ~5-10% of

adult acute leukemia and ~25% of therapy related leukemia, especially in patients

treated with topoisomerase II inhibitors (Aplan, 2006; Hess, 2004a). Leukemia cases

with MLL translocations are generally associated with poor prognosis. For example,

the 5-year survival rate of infant with mixed lineage leukemia is less than 40%, while

the 5-year survival rate for non-mixed lineage leukemia is nearly 90% (Slany, 2009). In

chromosome translocations the N-terminus of MLL is fused in frame to a fusion partner

(Figure 1.1). These oncogenic fusion proteins lead to constitutive activation of *HOX*

genes, especially *HOXA7, HOXA9* and their cofactor *MEIS1*, which are required for

uncontrolled cell proliferation and occurrence of leukemia (Ayton and Cleary, 2003;

Zeisig et al., 2004). So far, more than 60 MLL fusion partners have been discovered,

with AF4 (AFF1), ENL (MLLT1), and AF9 (MLLT3) being most frequent in acute

lymphoid leukemia (ALL), accounting for 80-90% of the MLL-translocation cases, and

AF9 (MLLT3), AF10 (MLLT10), ELL and AF6 (MLLT4) being most frequent in acute

myeloid leukemia (AML), accounting for 70-80% of the cases (Marschalek, 2011;

Meyer et al., 2006). Recent studies also discovered that AF4-MLL and NUP98-MLL,

which contain the N-terminus of the fusion partner and the C terminus of MLL, can

have oncogenic properties (Bursen et al., 2010; Kaltenbach et al., 2010). However,

current observations indicate that these two fusion proteins transform through

mechanisms independent of *HOXA9* activation (Bursen et al., 2010; Kaltenbach et al.,

2010). Besides chromosomal translocations, partial tandem duplications (PTD) of

MLL, are found in AML patients with trisomy 11 (Bernard et al., 1995; Caligiuri et al.,

1996), and ~ 10% of AML patients with normal cytogenetics (Caligiuri et al., 1998; Yu

et al., 1996). In MLL PTD, *MLL* exons 5-11 or exons 5-12, which contain the AT hooks

and the CxxC domain, are duplicated and inserted into intron 4 of the full-length MLL

(Figure 1.1). However, MLL PTD only moderately up-regulates HOXA9 and requires

additional mutations, such as FLT3-internal tandem duplication (ITD), to transform,

indicating that it causes leukemia through different mechanisms from MLL fusion

proteins (Zorko et al., 2012). This overview will focus on MLL fusion proteins with

N-terminus of MLL and C-terminus of the fusion partner, as they account for the

majority of MLL associated leukemia cases.

1.1.3.2   Mechanisms for MLL-fusion mediated leukemogenesis

Oncogenic MLL fusion proteins mediate transformation through complicated

mechanisms that require both MLL and the fusion partner, and involve multiple

interactions with oncogenic cofactors. The N-terminal 43 amino acid fragment of MLL

mediates interaction with MENIN, which then interacts with the transcription cofactor

lens epithelium-derived growth factor (LEDGF) to form a tri-molecular complex

(Caslini et al., 2007; Grembecka et al., 2010; Yokoyama and Cleary, 2008). This

interaction is critical for the recruitment of MLL fusion proteins to their target gene loci.

Small molecules that specifically disrupt MLL-MENIN interaction efficiently

down-regulate HOX gene expression and reverse oncogenic activity of MLL fusion

proteins, representing an attractive therapeutic approach for MLL fusion mediated

leukemia (Grembecka et al., 2012; Shi et al., 2012). Also, the flanking regions of MLL

CxxC domain recruit the PAF complex, which synergize with MLL fusions to activate

HOX gene expression. Knockdown of PAF components or deletion of the PAF

interacting region in MLL-AF9 abolishes its transformation ability (Muntean et al.,

2010).

While the interactions preserved in MLL N terminus represent global mechanisms for

MLL fusion mediated transformation, the functions brought by the fusion partners are

more complicated due to the large number of fusion partners and diversity of their

molecular function. The MLL fusion partners can be roughly divided into two classes.

The first class is composed of nuclear proteins, mostly transcription factors, such as

AF9, AF4, ENL and ELL; while the second class is composed mostly cytoplasmic

proteins with dimerization ability, such as AF1 and GAS7.

For nuclear translocation partners, the most well studied mechanism is through

promoting transcription elongation. Multiple groups have found that MLL fusion

partners, such as AF9, AF4 and ENL, form a complex with multiple transcription

elongation factors (Table 1), which promote phosphorylation of RNA polymerase II

C-terminal domain and thus activation of MLL target genes (Lin et al., 2010; Monroe et

al., 2011; Yokoyama et al., 2010). Further, AF9 recruits the histone H3K79

methyltransferase DOT1L, which is required for MLL fusion mediated transformation (Chang et al., 2010a; Okada et al., 2005). Similar to MENIN, DOT1L inhibitors block proliferation of MLL leukemia cell lines and show therapeutic potential (Daigle et al., 2011).

Table 1-1 MLL fusion partner recruited transcription elongation complex

Complex Name	Components	Reference
Super Elongation Complex (SEC)	P-TEFb complex, ELL2, ELL3, EAF1, EAF2	(Lin et al., 2010)
AF4/ENL/P-TEFb Complex (AEP)	P-TEFb complex, AF4, ENL and	(Yokoyama et al., 2010)
Elongation Assisting Proteins (EAPs)	P-TEFb complex, AF9, ENL, AF4, LAF4, AF5Q31, ELL, AF10, DOT1L	(Monroe et al., 2010)

The oncogenic mechanisms of MLL fusion proteins with cytoplasmic fusion partners are less studied. Interestingly, a number of the cytoplasmic fusion partners contain self-association domains, such as the coiled-coil domain in GAS7 and AF1, and the Ras-interacting domain in AF6, raising the possibility that dimerization of MLL fusion proteins contributes to transformation. Indeed, mutation of the coiled-coil domain blocks MLL-GAS7 and MLL-AF1 mediated transformation (So et al., 2003). However, although artificially made MLL fusion proteins with synthetic dimerizers enhance self-renewal of hematopoietic progenitors, the cells are not completely transformed, indicating that additional mechanisms are needed (Martin et al., 2003; So et al., 2003). Also, common transformation mechanisms may be used by nuclear and cytoplasmic fusion partners. For example, although AF6 is a cytoplasmic protein, it has been shown

that the AEP complex co-localizes with MLL-AF6 on target gene loci, and MLL-AF6

mediated transformation requires the presence of DOT1L (Deshpande et al., 2013).

1.1.3.3   Mechanisms for occurrence of MLL break point

In MLL translocations, the break point occurs in a breakpoint cluster region (BCR) that

is ~8.3 kb long and most frequently in MLL intron 9 and intron 11(Meyer et al., 2009).

Understanding this region of chromosome breakage may provide insight for the

pathogenesis of MLL fusion mediated leukemia, and several mechanisms have been

proposed to explain the specific location of MLL breakpoint. Some of them focus on

the susceptibility of the DNA sequence to double strand break. For example, based on

the observation that the t(4:11) (MLL-AF4) chromosome translocation breakpoints

contain heptamer-like or nonamer-like sequences similar to the consensus sequences of

the immunoglobulin and T-cell antigen receptor (TCR) loci, Gu et al. proposed that

these translocations may be mediated by abnormal activity of VDJ recombinase.

However, these heptamer-like and nonamer-like sequences have only been reported in

two cell lines (MV4:11 and RS4:11), and direct evidence are lacking that VDJ

recombinase can mediate MLL translocation (Gu et al., 1992). Alu elements are

repetitive elements accounting for ~10% of human genome (Lander et al., 2001). When

a DNA double strand break is introduced, an Alu element can recombine with another

Alu element or non-Alu element, which leads to chromosome translocation (Elliott et

al., 2005). Strout et al. examined nine cases of AML with MLL PTD and found that six

of them result from homologous Alu recombination and two from heterozygous Alu recombination (Strout et al., 1998). However, although Alu recombination seems to be a major cause for MLL PTD, it is a rear event for MLL chromosomal translocations, as so far only two cases of MLL-AF4 translocation and one case of MLL-AF9 translocation have been reported that are possibly caused by Alu recombination (Reichel et al., 2001; Super et al., 1997). The association of MLL translocation to therapy related leukemia, especially in patients treated with DNA topoisomerase II inhibitors, raises the possibility that MLL translocation is caused by inhibition of DNA topoisomerase II activity. DNA topoisomerases bind to DNA and control its topological status by introducing double strand breaks and re-ligation of the DNA, thus playing critical roles in gene transcription and DNA replication. DNA topoisomerase II inhibitors, such as doxorubicin and etoposide, interfere with the activity of DNA topoisomerase II by promoting cleavage or preventing re-ligation of the DNA, thus leading to increased DNA double strand break and apoptosis of the highly proliferating cancer cells. These inhibitors have been widely used in the treatment of cancers such as breast cancer, lung cancer and testicular cancer (Nitiss, 2009). Not surprisingly, the increased DNA double strand breaks and non-homologous end joining (NHEJ) leads to increased possibility of chromosomal translocation. Indeed, in vitro assays have shown that etoposide and its metabolites enhance topoisomerase II mediated cleavage near the BCR of MLL gene (Ishii et al., 2002; Lovett et al., 2001), and another study shows that the breakpoint of a patient MLL-AF9 fusion protein is a preferred substrate for

10

topoisomerase II cleavage in vitro (Whitmarsh et al., 2003). Of course, in vivo

experiments are needed to prove this model, and whether it also explains the occurrence

of de novo MLL translocations remains to be determined.

Notably, besides accessibility of the DNA sequence to double strand break, another

prerequisite for chromosome translocations to occur and be detected is that the fusion

proteins gain oncogenic ability. In other words, the break point occurs so that regions

that are required for oncogenesis are preserved while regions that are incompatible with

oncogenesis are deleted. In most if not all MLL translocations, the break point happens

after the CxxC region and before or within the first PHD finger, thus the CxxC region is

invariably preserved while the PHD/Bromo region disrupted or deleted from oncogenic

MLL fusion proteins. Indeed, our lab has shown that, although insertion of the first

PHD finger into a MLL-AF9 construct does not affect its ability to transform murine

bone marrow cells, insertion of the first two PHD fingers and beyond abolishes its

transformation ability (Muntean et al., 2008). Similarly, insertion of the third PHD

finger into MLL-ENL fusion proteins blocks transformation (Chen et al., 2008). These

findings indicate that the PHD/Bromo region has important roles in the regulation of

MLL activity, which is the focus of this dissertation. At the same time, Muntean et al.

show that sequences flanking the MLL CxxC domain recruit the PAF complex, which

promotes MLL-AF9 mediated transcriptional activation and is required for the

transformation ability.

1.1.4   PHD finger

1.1.4.1   Structure of PHD fingers

Plant Homeo Domain (PHD) fingers are widely distributed zinc-binding motifs that

have been found in more than 200 human proteins, especially transcription factors and

chromatin modifier enzymes. Typical PHD fingers are 50-80 amino acids long and

contain a Cys4-His-Cys3 (C4-H-C3) consensus. The conserved cysteine and histidine

residues coordinate two Zn^{2+} ions in a cross-brace topology, and maintain the stability

of the domain structure (Bienz, 2006) (Figure 1.2). Several other classes of

zinc-binding motifs share structural and functional similarity to PHD fingers, such as

RING domains with C3-H-C4 consensus, LIM domains with C2-H-C-C/H/D/E

consensus, MYND domains with C/H-C4-C/H-H-C consensus and FYVE domains

with C8 consensus (Matthews et al., 2009). Although being structurally conserved,

the conservation of amino acid sequences among PHD fingers is generally low, which

may account for the diverse functions of PHD fingers.

Figure 1.2 Structure of PHD finger and RING finger
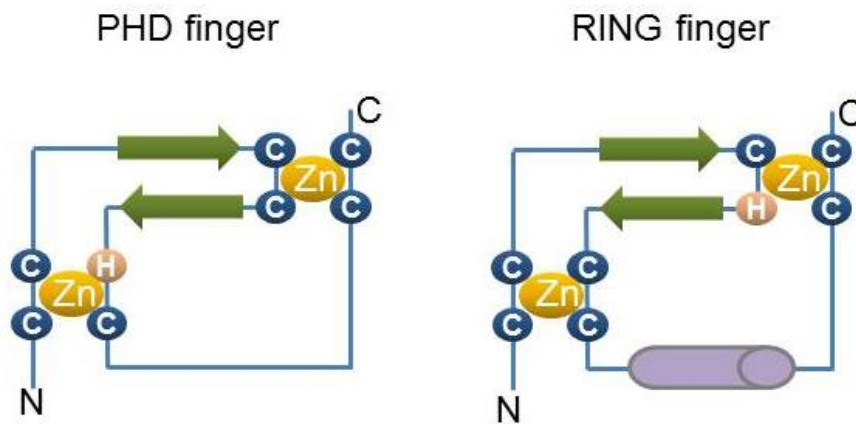
Figure 1.2 Schematic showing the structure of PHD finger and RING finger. The conserved Cysteine and Histidine residues coordinate Zn^{2+} ions in a cross-brace topology, and maintain the stability of the domain. Zn: Zn ion. C and H: conserved cysteine and histidine residues. Green arrow: β-strand. Purple: α-helix.

1.1.4.2 Function of PHD fingers

1.1.4.2.1 Binding to histone marks

Post-translational modifications of histones, such as methylation, acetylation, phosphorylation and ubiquitination, mark the chromosome status and play important roles in recruiting transcription factors and chromosome modifiers (Jenuwein and Allis, 2001). The best-studied function of PHD fingers is to bind to histone marks, mostly on histone H3, and function as "epigenetic readers". For example, the PHD fingers of AIRE and UHRF1 bind to unmodified histone H3 (Chakravarty et al., 2009; Chignola et al., 2009; Lallous et al., 2011; Org et al., 2008; Rajakumara et al., 2011; Wang et al., 2011a); the PHD fingers of ING proteins (ING1-5), BPTF, YNG1 and RAG2 bind to H3K4me3 (Champagne et al., 2008; Li et al., 2006; Liu et al., 2007b;

Martin et al., 2006; Matthews et al., 2007; Pena et al., 2006; Pena et al., 2009; Ramon-Maiques et al., 2007; Taverna et al., 2006; Wysocka et al., 2006); the PHD finger of CHD4 binds to H3K9me3 (Mansfield et al., 2011); the PHD finger of NTO1 binds to H3K36me3 (Shi et al., 2007); and the PHD finger of DPF3b and MOZ bind to H3K14ace (Qiu et al., 2012; Zeng et al., 2010). Besides individual binding, multiple PHD fingers can bind cooperatively to one histone modification. For example, the two tandem PHD fingers in DPF3 function as one unit and bind to H3K14ac (Zeng et al., 2010).

In most cases, binding of PHD finger to histone marks recruits protein or protein complex to the chromosome, while the biological effect of this recruitment, such as chromatin modification, transcription activation or repression, is largely decided by other effector domains of the protein or other subunits of the protein complex. One example is the inhibitor of growth (ING) family proteins. All the ING family members (ING1-5) contain a C terminal PHD finger that show ~80% sequence homology and bind to H3K4me3 in highly similar conformations through an aromatic cage. This binding recruits the ING containing complexes to the target gene loci. ING1 and ING2 associate with Sin3A/HDAC1/2 repressor complex, which lead to transcriptional repression. On the contrary, ING3, ING4 and ING5 associate with NuA4 complex, HBO1 histone acetyltransferases (HAT) and MOZ/MORF HAT

complex, respectively, which lead to transcriptional activation (Champagne and

Kutateladze, 2009).

PHD fingers recognize histone modifications and are responsible for the proper

recruitment of transcription factors and chromatin modifiers, so the specificity of

PHD finger-histone mark interaction must be accurately controlled. This is achieved

on multiple levels. First, structures have been solved for a number of PHD fingers

interacting with different histone peptides, which reveal that the binding specificity

depends on sequence and conformation of PHD fingers. For example, the H3K4me3

binding PHD fingers, such as those of BPTF and RAG2, contain an aromatic cage

formed by two to four aromatic and hydrophobic residues, including an invariable Trp

residual, which accommodates the methylated H3K4 side chain (Li et al., 2006;

Ramon-Maiques et al., 2007). Mutation of the Trp residual in the PHD finger of

BPTF reduces its binding affinity for more than 20 fold. Similar effects are also

observed when the other three amino acids lining the aromatic cage are mutated (Li et

al., 2006). However, in H3K4me0 binding PHD fingers, such as those of AIRE and

DNMT3L, the aromatic cage is replaced by acidic and hydrophobic residues that form

multiple hydrogen bonds to unmethylated H3K4 (Chakravarty et al., 2009; Chignola

et al., 2009). The H3K4me0-binding PHD finger of UHRF1 represents an atypical

PHD finger. Although it does not have the conserved C terminal acidic residue, an

extra zinc finger preceding the PHD finger helps to accommodate the H3K4 side

chain (Lallous et al., 2011). Secondly, adjacent residues on histone H3 may have positive or negative effects on the binding affinity of PHD fingers. For example, Arg2 on H3 affects the binding of PHD fingers, especially those bind to unmethylated or methylated H3K4, in different ways. For the PHD finger of AIRE, Arg2 occupies a groove adjacent to the groove occupied by H3K4me3 and form direct contact with the PHD finger through ionic and hydrogen-bonding interactions. This is required for the PHD-H3K4me3 interaction, as methylation of Arg2 abolishes the interaction (Chignola et al., 2009). On the contrary, in the PHD finger of RAG2, the carboxylate (an Asp or Glu) that interacts with Arg2 was replaced by a Tyr, which forms interaction with methylated Arg2, so its binding to H3K4me3 is enhanced rather than inhibited by dimethylation of Arg2. Notably, mutation of this Tyr reverses this binding preference as the mutant binds to the singly modified (K4me3) over doubly modified (K4me3/R2me2) H3 peptide by 3- to 4-fold (Ramon-Maiques et al., 2007). Finally, PHD fingers generally exist in proteins in a tandem manner with other PHD fingers, or adjacent to other chromosome binding motifs, such as bromodomains, chromodomains and tandem tudor domains. These adjacent domains may cooperate with PHD finger to define its binding specificity. One example is UHRF1, in which the PHD finger binds to unmodified R2 of histone H3 tail, while the adjacent tandem tudor domain binds to H3K9me3. Together, they recognize specific H3 tails and are critical for the proper recruitment of UHRF1. Similarly, in TIF1r, the PHD finger recognizes the N-terminal region of the H3 tail with unmethylated K4 and R2, while

the adjacent bromodomain binds to the C terminus of the same H3 tails with

acetylated lysines (Agricola et al., 2011).

1.1.4.2.2 Binding to other proteins

Besides histone marks, PHD fingers can also bind to other proteins. Examples include

the PHD finger of Pygo2, which binds to the B9L/BCL9 cofactor, and the third PHD

finger of MLL, which binds to Cyp33. Interestingly, for both of them, the PHD finger

binds to histone tail and another protein simultaneously through different surfaces and

binding of the protein affects the binding affinity to histone tail. The PHD finger of

Pygo2 binds to the HD1 domain of B9L/BCL9 cofactor on one surface, and

methylated histone H3 on another surface. Structural studies show that binding of

HD1 leads to an allosteric conformational change of the PHD finger, especially the

conserved isoleucine residue (I344) between the two grooves that accommodate

H3A1 and H3K4me, which promotes its binding to methylated H3K4. Consistent

with this, the binding affinity is 2~3 fold higher in the presence of HD1 compared to

that without HD1; and mutation of I344 and W377, the residues that mediate

interaction with HD1, leads to decreased binding affinity to methylated H3K4 (Fiedler

et al., 2008; Miller et al., 2010). Similarly, the third PHD finger of MLL binds to

H3K4me3 and Cyp33 simultaneously through different surfaces. This is achieved by a

PPIase-mediated cis-to-trans conformational transition of the Pro1629 of MLL, which

exposes the surface on PHD3 for recognition by the RRM domain of Cyp33. Binding

17

of Cyp33 leads to recruitment of HDAC and switches MLL from a transcription activator to a transcription repressor. Unlike Pygo2, binding of Cyp33 leads to ~ 2 fold decrease in the binding affinity of MLL to histones (Park et al., 2010; Wang et al., 2010). Other examples of protein binding PHD fingers include the PHD finger of Hepatitis B virus X-associated protein (HBXAP) which interact with Hepatitis B virus pX, and the PHD finger of Polycomb-Like Protein 1 (PHF1) which binds to EZH2 (O'Connell et al., 2001; Shamay et al., 2002). Further, all the PHD fingers with E3 ligase activity (discussed below) interact with E2 conjugating enzymes.

1.1.4.2.3 Binding to lipids and DNA

Besides binding to proteins, the PHD finger of ING2 has been shown to bind to Phosphoinositides (PtdInsPs), which is required for ING2 mediated p53 acetylation and activation of p53 related pathways. This binding ability is probably due to the structural similarity of the PHD finger to the FYVE domain, a well-known Phosphoinositides binding module, and is mediated by two basic patches in the middle and C terminus of the PHD finger. Similar binding ability was also observed for PHD fingers of ACF, MLZF and WSTF (Gozani et al., 2003). Notably, the same PHD finger of ING2 can also bind to H3K4me3 (Shi et al., 2006). Although the key amino acid residues identified for PtdInsPs binding are different from those identified for H3K4me3 binding, it remains to be determined if the two interactions occur simultaneously and if there is any crosstalk (Gozani et al., 2003; Pena et al., 2006).

Although PHD fingers have long been believed to lack direct DNA binding ability, recently Liu L et al., reported that the second PHD finger of BRPF2 can bind DNA non-specifically, while having no affinity for unmethylated or methylated histones (Liu et al., 2012). In this PHD finger, the binding to DNA is mediated through a positively charged surface; the characteristic aromatic cage responsible for H3K4me3 binding is replaced by short-side-chain residues like Thr or Ser, and the segment that forms hydrogen bonds to the N termimus of H3 is also replaced by a β-sheet. This may represent a new subclass of PHD fingers.

1.1.4.2.4 PHD finger as ubiquitin E3 ligase

Ubiquitination is a post-translational modification that plays an important role in protein degradation and signal transduction. During ubiquitination, one or more ubiquitin molecules are attached to lysine residues of a substrate through sequential activity of three enzymes, an E1 activating enzyme, an E2 conjugating enzyme and an E3 ligase. The substrate is then targeted to proteasome-mediated degradation or other signal transduction pathways (Pickart, 2001). E3 ligases directly bind to target proteins, thus defining the specificity of the reaction. The structural similarities between PHD fingers and RING fingers (Figure 1.2), which are known to function as E3 ligases, raise the possibility that PHD fingers may also have similar activity (Capili et al., 2001). The E3 ligase activity of PHD finger was first discovered in Kaposi's sarcoma-associated herpesvirus (KSVH). Two PHD finger containing proteins, MIR1

and MIR2, mediate down-regulation of cell surface molecules, such as MHC-I, B7.2

and ICAM-1, thus impairing the recognition of infected cells by the immune system.

This down-regulation depends on the lysine residues of the surface molecules, and is

through ubiquitination and degradation by the endocytosis system. Further, in vitro

ubiquitination assays showed that the PHD fingers of MIR1 and MIR2 contain E3

ligase activity in the presence of ATP, E1 activating enzyme and E2 conjugating

enzyme, while disruption of the PHD finger abolishes the activity and

down-regulation of cell surface molecules (Coscoy et al., 2001). Later, the PHD

fingers of several eukaryotic proteins were also discovered to contain E3 ligase activity,

including the yeast protein Msc1, and the mammalian proteins MEKK1, AIRE. In the

current study, we found that the second PHD finger of MLL also has intrinsic E3

ligase activity (Dul and Walworth, 2007; Lu et al., 2002; Uchida et al., 2004; Wang et

al., 2012). All the proteins with E3 ligase PHD fingers tested so far leads to

ubiquitination and degradation of the protein itself or other targets. However, these

proteins are quite diverse in their localization and molecular functions. For example,

MEKK1 localizes in the cytoplasm and regulate the MAP kinase pathway by

degrading ERK1/2; while AIRE localizes to the nucleus and functions as transcription

factor. Besides ubiquitin E3 ligase activity, the PHD finger of KAP1 has been shown

to have sumo (Small Ubiquitin-like Modifier) E3 ligase activity (Ivanov et al., 2007).

Sumoylation is a post-translational modification that is very similar to ubiquitination.

A sumo molecule is attached to the lysine residual of target proteins through action of

sumo E1 activating enzyme, E2 conjugating enzyme (UBC9) and E3 ligase, which

then regulates transcriptional regulation, protein stability and signal transduction (Hay,

2005). KAP1 functions to recruit the nucleosome remodeling and deacetylation

(NuRD) complex and SETDB1 histone H3K9 methyltransferase to target loci, and

mediates transcription repression (Schultz et al., 2002; Schultz et al., 2001). It has

been shown that the PHD finger of KAP1 interacts with UBC9 sumo E2 conjugating

enzyme, and mediates sumoylation of the adjacent bromodomain in vitro and in vivo.

Both SETDB1 and CHD3, the subunits of NuRD complex, contain a

SUMO-interacting motif and interact with KAP1 through sumo binding. So this

activity is critical for the recruitment of NuRD (Capili et al., 2001; Ivanov et al., 2007;

Zeng et al., 2008).

1.1.4.3   PHD fingers in human disease

As PHD fingers play important roles in recruitment, stability and activity of proteins,

mutations of these domains are associated with various human diseases.

One example is RAG2. RAG2 is a core component of the RAG complex, which

mediates DNA cleavage during V(D)J recombination. The PHD finger of RAG2 has

been shown to bind to H3K4me3, which recruits the complex to the chromosome and

is required for efficient DNA recombination (Oettinger et al., 1990). Deletion of

RAG2 in mouse model fails to initiate V(D)J recombination and have severe defect in

T cell and B cell maturation, leading to a Severe Combined Immunodeficiency (SCID)

phenotype (Shinkai et al., 1992). Mutations of RAG2 have been found in patients with SCID and Omenn immunodeficiency syndrome with several of them occuring in the PHD domain (Gomez et al., 2000). One of the mutations found in Omenn syndrome occurs on W453, which is one of the highly conserved aromatic residues on the binding surface to H3K4me3. It has been shown that this mutation abrogates H3K4me3-binding by RAG2 (Matthews et al., 2007). Similarly, mutations of C478Y and H481P, two of the conserved Cysteine and Histidine residues are found in SCID patients (Noordzij et al., 2002; Schwarz et al., 1996).

On the contrary, gain of function of PHD fingers is also associated with human diseases. Nucleoporin 98 (NUP98) is a component of the nuclear pore complex which mediates mRNA export from the nucleus. Chromosome translocations of NUP98, which fuses NUP98 to another fusion partner, are commonly found in hematopoietic malignancies, including acute myeloid and lymphoid leukemia and myelodysplastic syndrome. So far more than 20 fusion partners of NUP98 have been discovered and one class of them is PHD fingers, including the PHD finger of JARID1A or PHF23 (Moore et al., 2007; Wang et al., 2009a). These PHD fingers have been shown to bind to H3K4me3. Indeed, the NUP98-PHD fusion proteins are recruited to the genomic loci with high H3K4me3 signals, with a number of them being genes involved in hematopoietic stem cell and progenitor self-renewal, such as *HOX* genes and *MEIS1*. Binding of NUP98-PHD prevents removal of the active H3K4me3 marks and

maintains the constitutive activation of these genes, which mediate leukemogenesis.

The oncogenic activity of NUP98-PHD fusions is dependent on the intact function of

the PHD fingers, as mutation of the PHD fingers abolishes their transformation ability

(Wang et al., 2009a). Another fusion partner of NUP98, NSD3, also contains a PHD

finger, although whether the oncogenic ability of this fusion protein depends on the

activity of the PHD finger remains to be determined.

So far, no disease-associated mutations have been discovered that specifically target

the E3 ligase activity. However, this possibility cannot be excluded as disruption of

these PHD fingers thus their E3 ligase activity definitely impairs the normal function

of the proteins (Lu et al., 2002; Uchida et al., 2004). Notably, the first PHD finger of

AIRE has been shown to have both E3 ligase activity and histone binding ability.

Missense mutations of this PHD finger (C311Y and P326Q) have been discovered as

the causal mutations for autoimmune-polyendocrinopathy-candidiasis ectodermal

dystrophy (APECED). However, it remains to be determined if the disease is caused

by loss of the E3 ligase activity, or loss of the histone binding activity, or both

(Uchida et al., 2004). Another example may be MLL, as the PHD/Bromo region,

including the second PHD finger with E3 ligase activity, is invariably deleted from

oncogenic MLL fusion proteins.

1.1.5 Bromodomain

Bromodomain is another class of motifs that commonly bind to histone marks and function as epigenetic reader. Bromodomains are generally ~100 amino acid long and contain a bundle of four α-helices and inter-helical ZA and BC loops (Mujtaba et al., 2007). So far, 121 human proteins have been discovered to contain one or more bromodomains, with most of these proteins being transcription factors and chromatin modifiers. Like PHD fingers, bromodomains recognize and bind to acetylated histones, and recruit proteins or protein complexes to the genome. Examples include the bromodomain of the histone acetyltransferase GCN5, the transcription co-activator PCAF (p300/CREB-binding protein-associated factor) and TAF1, the SWI/SNF chromosome remodeling complex component BRG1, and the BET family protein BRDs (Dhalluin et al., 1999; Hudson et al., 2000; Jacobson et al., 2000; Nakamura et al., 2007; Shen et al., 2007). Structural studies show that bromodomain binds to acetylated lysine residual through its hydrophobic groove, which is highly conserved among acetyl-binding bromodomains (Owen et al., 2000). Most of the bromodomains studied have broad recognition specificity and bind to multiple acetylated lysine residues on histone H2B, H3 and H4 in vitro, while in vivo, the binding specificity may be achieved by cooperation of adjacent domains (Zeng et al., 2008).

Besides acetylated histones, bromodomains also bind to other acetylated proteins. For example, the bromodomain of the co-activator CBP binds to the C-terminal acetylated

K382 of p53, which is responsible for p53 recruitment of CBP upon DNA damage, and p53-induced transcriptional activation (Mujtaba et al., 2004). Also, the bromodomain of PCAF binds to acetylated lysine of the human immunodeficiency virus type 1 (HIV-1) trans-activator protein Tat in a very similar mode as binding to acetylated histones, which stimulates the transcription of the integrated virus genome (Mujtaba et al., 2002).

Although the hydrophobic groove is highly conserved among acetyl-lysine binding bromodomains, the overall amino acid sequences, especially the ZA and BC loops can be very diverse (Jeanmougin et al., 1997). A sequence analysis of 56 human bromodomains was reported, which classifies these bromodomains into 9 classes. Notably, the MLL bromodomain was defined as an outlier, due to significant sequence variation (Sanchez et al., 2000; Sanchez and Zhou, 2009). Consistent with this, although this bromodomain exhibits a typical bromodomain fold, it does not recognize acetylated histones, perhaps due to loss of a conserved asparagine found in acetyl-lysine binding bromodomains (Wang et al., 2010). In this study, we discovered that the bromodomain of MLL has a novel function, which is to recruit the E3 ligase ASB2 and mediates ubiquitination and degradation of MLL.

Adjacent PHD finger and bromodomain are commonly found in proteins and have intensive functional interactions. One example is the PHD –bromodomain module in TIF1γ. The PHD finger recognizes the N-terminal region of the H3 tail with

unmethylated K4 and R2, while the adjacent bromodomain binds to the C terminus of the same H3 tails with acetylated lysines. Together, the two modules define a specific binding pattern for the recruitment of TIF1γ (Agricola et al., 2011). Another example is KAP1, in which the PHD finger has sumo ligase activity that mediates sumoylation of the adjacent bromodomain, which is required for the recruitment and proper function of KAP1 (Ivanov et al., 2007).

Bromodomains are closely associated with pathogenesis of human diseases. One example is BRD4, which is a key regulator of cell cycle by interacting with cyclin T1 and CDK9 of the P-TEFb complex (Jang et al., 2005). The t(15;19)(q15;p13) translocation fuses the nuclear protein in testis (NUT) to BRD4, which is believed to be the driving mutation of NUT midline carcinoma. Although this is a rare tumor, it is highly aggressive and is almost uniformly lethal (French et al., 2003). Studies discovered that the bromodomain from BRD4 binds to acetylated histones on the chromosome, while the NUT fragment sequester p300, which leads to histone hyper-acetylation mediated chromosome compaction and transcriptional inactivation of key target genes (Reynoird et al., 2010). Besides being a direct driving mutation for tumorigenesis, an RNAi screening identified BRD4 as the top gene that is required for MLL fusion mediated leukemogenesis. This is possibly due to its important role in cell cycle regulation, thus BRD4 represents an example of non-oncogene addiction. Consistent with this, two BRD4 inhibitors, JQ1 and I-BET151, block MLL fusion

mediated leukemogenesis in vitro and in vivo, thus representing an attractive

therapeutic strategy (Dawson et al., 2011; Zuber et al., 2011). Further, it was shown

that BRD4 inhibitor can target the MYC pathway, and is effective for a broad range of

AML leukemia cases not limited to MLL translocation (Zuber et al., 2011).

1.2   HOXA9 is a key target gene of MLL

1.2.1   HOXA9 in MLL mediated leukemia

HOXA9 and its cofactor MEIS1 are the best-studied MLL target genes and are required

for MLL fusion mediated leukemogenesis (Ayton and Cleary, 2003; Wong et al., 2007).

HOXA9 is a member of the *HOX* transcription factor family. It specifies segment

identity and cell fate during development and plays an essential role during

hematopoiesis (Argiropoulos and Humphries, 2007). HOXA9 maintains the

self-renewal ability of hematopoietic stem cells and progenitors. Hoxa9-deficient mice

show severe impairment of the repopulation ability of HSCs and have defects in

development of multiple hematopoietic lineages. During hematopoietic differentiation,

HOXA9 expression decreases, this is required for the normal progression of

hematopoiesis (Pineault et al., 2002).

MLL positively regulates the expression of *HOXA9* and *MEIS1* in hematopoietic stem

cells and progenitors (Jude et al., 2007; McMahon et al., 2007; Wang et al., 2009b).

Mll-null mice are defective in maintaining *Hox* gene expression resulting in embryonic

lethality by E10.5, while re-expression of *Hox* genes in *Mll*-deficient progenitors

rescues hematopoietic colony formation (Ernst et al., 2004; Yu et al., 1995). On the

other hand, MLL fusion proteins constitutively activate *HOXA9* and *MEIS1*, which

leads to uncontrolled cell proliferation and is required for MLL fusion mediated

leukemia. Consistent with this, MLL fusion proteins are unable to transform HOXA9

or MEIS1 knockout murine bone marrows in vitro or induce leukemia in vivo (Ayton

and Cleary, 2003; Wong et al., 2007). Further, overexpression of HOXA9 and MEIS1

could replace the transformation ability of MLL-ENL fusion protein (Zeisig et al.,

2004).

Besides MLL fusion mediated leukemia, up-regulation of HOXA9 is also associated

with a variety of human leukemias, such as NPM1 mutation-associated AML and

CDX2/CDX4 mediated leukemia (Bansal et al., 2006; Mullighan et al., 2007; Rawat

et al., 2008). In fact, overexpression of HOXA9 is present in approximately 50% of all

AML cases (Lawrence et al., 1999). HOXA9 was identified as the most highly

correlated gene for poor prognosis in AML (Golub et al., 1999; Lawrence et al., 1999),

while low expression of HOXA9 associates with favorable AMLs (Debernardi et al.,

2003). Also, HOXA9-NUP98 fusion resulting from chromosome translocation leads

to aberrant gene expression profiling and is associated with AML, trilineage MDS and

chronic myelomonocytic leukemia (Borrow et al., 1996; Hatano et al., 1999; Inaba et

al., 1996; Kroon et al., 2001; Nakamura et al., 1996; Wong et al., 1999).

1.2.2   Molecular function of HOXA9

Although it has been established for a long time as the key regulator of hematopoiesis and important mediator of leukemia, how HOXA9 regulates gene expression remains poorly understood. For example, all the HOX proteins bind to similar AT rich sequences through their highly homologous homeodomain, so how does HOXA9 specifically recognize its target genes (Laughon, 1991)? Also, as HOXA9 can both activate and repress gene expression, how is the specific activity determined (Huang et al., 2012)? To solve these questions, the concept of "Hoxasome" was brought up, meaning that transcription factors and chromatin modifiers (HOXA9 cofactors) form a complex with HOXA9 and bind to DNA in a cooperative pattern, thus increasing the binding specificity, affinity or transactivation ability of HOXA9 (Mann et al., 2009). One example of HOXA9 cofactor is the non-Hox homeodomain protein MEIS1. MEIS1 can dimerize with HOXA9, or form a triple complex with HOXA9 and PBX proteins (PBX1-3), another class of HOXA9 cofactors that contain a homeodomain with ~40% identity to that of MEIS1. These interactions stabilize the HOXA9 DNA binding complex. In in vitro binding assays, addition of MEIS1 leads to a 10-fold increase of PBX2-HOXA9 interaction; while addition of PBX2 leads to a 5-fold increase of MEIS1-HOXA9 interaction. Further, addition of MEIS1 binding consensus (TGACAG) and PBX binding consensus (TGAT) increases the specificity of HOXA9 recognition on target loci (Shen et al., 1997; Shen et al., 1999).

To better understand the mechanisms of HOXA9 regulated gene expression and discover other HOXA9 cofactors, the Hess lab performed ChIP-seq experiments to

define the genome-wide binding pattern of Hoxa9 using a murine cell line

transformed with Hoxa9 and Meis1. De novo motif analysis reveals that Hoxa9 peaks

are enriched with binding motifs of several transcription factors and chromatin

modifiers, such as PU.1, C/EBP, RUNX1 and STAT5 (Huang et al., 2012). Further,

mass spectrometry analysis identified that C/EBP and STAT5 interact with HOXA9,

which was also confirmed by in vitro binding assays and/or co-immunoprecipitation.

These findings support that these proteins are potential HOXA9 cofactors. The

functional interaction between HOXA9, C/EBPA and PU.1 will be further explored in

the current study.

1.3   Function of MLL family proteins and their role in human diseases

The MLL protein family contains five members, including MLL1 (MLL), MLL2 (chr.

12q13.12, previously referred to as MLL4), MLL3, MLL4 (chr. 19q13.1, previously

referred to as MLL2), and MLL5. MLL1 through MLL4 are homologous to the yeast

methyltransferase Set1, and share structural and functional similarities, while MLL5 is

considered to be the homolog of the yeast protein Set3/Set4 and is more different

compared to other MLL proteins. MLL1-4 all contain multiple PHD fingers on their N

terminus and a SET domain on their C terminus. Also, they all form a "core

methyltransferase complex" with WDR5, RBBP5, hDPY30 and ASH2, and mediate

methylation on histone H3K4. However, their functions are not redundant, with MLL1

and MLL4 being structurally and functionally more similar, and MLL2 and MLL3

being more similar (Shilatifard, 2012). MLL1 and MLL4 both interact with MENIN

and LEDGF on their N terminus, which is crucial for the proper recruitment of the

complexes to target gene promoters, including *HOXA* genes (Dhar et al., 2012;

Yokoyama and Cleary, 2008). On the contrary, MLL2 and MLL3 form complexes with

UTX, PTIP, PA1 and NCOA6. Unlike MLL1 and MLL4, loss of MLL2 or MLL3 has

little effect on the expression of HOXA genes (Wang et al., 2009b). Instead, MLL2 and

MLL3 function as co-activator of estrogen receptor alpha, and are also involved in the

regulation of adipogenesis and immunoglobulin class switch recombination (Lee et al.,

2009; Mohan et al., 2011; Shilatifard, 2012; Smith et al., 2011; Takahashi et al., 2011).

Further, recent study shows that Trr, the Drosophila homolog of MLL2 and MLL3, is

responsible for H3K4me1, suggesting the role of MLL2 and MLL3 in enhancer

establishment.

Figure 1.3 Structure of MLL protein complexes

Figure 1.3. Structure of MLL protein complexes. All MLL proteins interact with WDR5, RBBP5, hDPY30 and ASH2 on their C terminus and form a "core methyltransferase complex". However, MLL1 and MLL4 interact with MENIN and LEDGF on their N terminus (left), while MLL2 and MLL3 interact with UTC, PTIP, PA1 and ASC-2 on their N terminus (right). Blue indicates MLL proteins; green indicates common components; purple indicates spefic components.

Besides MLL1, mutations and chromosomal rearrangements of other MLL family

members are also associated with malignancies and other disorders. MLL2 mutations

have been discovered in Kabuki Syndrome, leukemia and lymphoma; mutations of

MLL3 are associated with leukemia, colorectal cancer, pancreatic cancer; and

mutations of MLL4 have been found in multiple solid tumors (Huntsman et al., 1999;

Ng et al., 2010; Ruault et al., 2002; Saigo et al., 2008). Notably, like *MLL1*, *MLL4* also

undergoes cancer associated gene rearrangements. In one study, intronic insertions of

hepatitis B virus (HBV) genomic sequences or chromosomal translocations on the

MLL4 locus were found in 26 out of 42 human hepatocellular carcinoma (HCC) cases

suggesting MLL4 fusions plays a key role in the pathogenesis of HCC (Saigo et al.,

2008).

1.4   ECS^{ASB} (Elongin B/C-Cullin–SOCS box protein) E3 ligase complex

Ubiquitin E3 ligases can be divided into two classes according to the mechanism of

their function, HECT type E3 ligases and RING type E3 ligases. HECT type E3

ligases have catalytic activity that accept the ubiquitin molecule from the E2

conjugating enzyme and transfer it to the substrate. On the contrary, RING type E3

ligases do not have catalytic activity, but function as a bridge that link E2 conjugating

enzyme to the substrate. The ubiquitin molecule is directly transferred from the E2

conjugating enzyme to the substrate. ECS^{ASB} E3 ligase is a class of RING type E3

ligase complexes. Cullin5 functions as a scaffold. It interacts with the E2 conjugating

enzyme RBX2 on its C terminus; while on its N terminus, it interacts with the adaptor

protein ElonginC, which dimerizes with ElonginB. ElonginC/B then interacts with an

ASB protein, which recruits the substrate. So in this complex, the ASB proteins define

the substrate specificity (Kohroki et al., 2005).

Figure 1.4 Structure of ECS^{ASB} E3 ligase complex
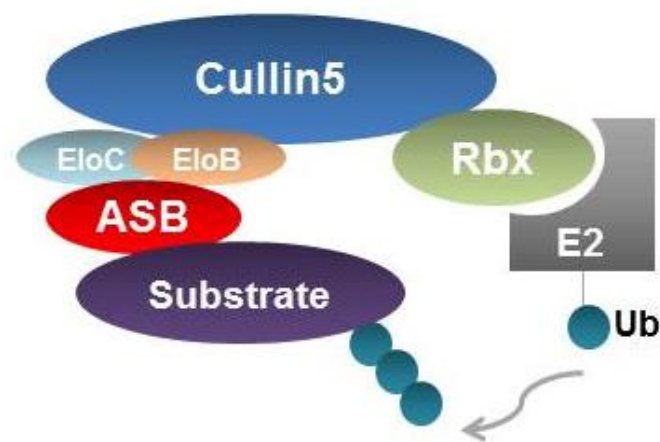
Figure 1.3. Schematic showing the structure of ECS^{ASB} E3 ligase complex. Cullin5 functions
as a scaffold. It interacts with the E2 conjugating enzyme RBX2 on its C terminus, and with
the adaptor protein ElonginC on its N terminus. ElonginC dimerizes with ElonginB.
ElonginC/B then interacts with an ASB protein, which recruits the substrate. ELOB: EloginB.
ELOC: EloginC. Ub: Ubiquitin. E2: Ubiquitin E2 conjugating enzyme.

The ASB protein family contains 18 members (ASB1-18). The founding member

ASB2 was discovered as a gene induced by *all-trans* retinoid acid (ATRA) via a

retinoid receptor (RAR) binding element present in the *ASB2* promoter. *ASB2*

expression is rapidly induced during certain ATRA induced differentiation of leukemia

cell lines, including APL cells containing the PML-RARα fusion protein, as well as

HL60 and NB4 cell lines, while ectopic expression of ASB2 promotes growth

inhibition and cell differentiation (Guibal et al., 2002; Kohroki et al., 2001). Filamins A

and B are two known targets of ASB2 (Heuze et al., 2008). ASB2 expression is also

activated by Notch signaling, which leads to degradation of Jak2 directly and

degradation of Jak3 and E2A indirectly by bridging the formation of a Cul1-Cul5

dimeric E3 ligase complex (Nie et al., 2011; Wu and Sun, 2011). Besides a role in

hematopoiesis, ASB2 can also regulate muscle differentiation (Bello et al., 2009).

Other ASB proteins also function in various biological processes. For example, ASB3

mediates ubiquitination and degradation of tumor necrosis factor receptor II, ASB1

regulates spermatogenesis, and ASB15 regulates myoblasts differentiation (Chung et

al., 2005; Heuze et al., 2005; Kile et al., 2001; McDaneld and Spurlock, 2008). In this

study, we provide evidence that ASB2 regulates hematopoiesis by mediating MLL

ubiquitination and degradation.

Chapter2   ECS^{ASB2} Mediates MLL Degradation during Hematopoietic Differentiation[1]

2.1   Abstract

MLL is a key epigenetic regulator of normal hematopoietic development and chromosomal translocations involving *MLL* are one of the most common genetic alterations in human leukemia. Here we show that ASB2, a component of the ECS^{ASB} E3 ubiquitin ligase complex, mediates MLL degradation through interaction with the PHD/Bromodomain region of MLL. Forced expression of ASB2 degrades MLL and reduces MLL transactivation activity. In contrast, the MLL-AF9 fusion protein does not interact with ASB2 and is resistant to ASB2 mediated degradation. Increased expression of ASB2 during hematopoietic differentiation is associated with decreased levels of MLL protein and down-regulation of MLL target genes. Knockdown of ASB2 leads to increased expression of *HOXA9* and delayed cell differentiation. Our data support a model whereby ASB2 contributes to hematopoietic differentiation, in part, through MLL degradation and *HOX* gene down-regulation. Moreover, deletion of the

---

[1] This work has been published as Wang, J., Muntean, A.G., and Hess, J.L. (2012). ECSASB2 mediates MLL degradation during hematopoietic differentiation. Blood 119, 1151-1161.

PHD/Bromo region renders MLL fusion proteins resistant to ASB2 degradation and

likely contributes to leukemogenesis.

2.2   Introduction

The histone H3 lysine 4 (H3K4) methyltransferase Mixed Lineage Leukemia (MLL) is

necessary for the maintenance of *HOX* patterning and essential for normal

hematopoiesis. Full length MLL is a multi-domain protein that is proteolytically

cleaved into a N-terminal fragment (MLL^N) and a C-terminal fragment (MLL^C) that

non-covalently associate to form a stable complex (Hsieh et al., 2003b). MLL^N contains

several DNA binding domains including three AT-hooks and a CxxC domain, as well

as a poorly understood PHD/Bromodomain (PHD/Bromo) region that contains four

plant homeodomain fingers (PHD1-4) and a bromodomain between PHD3 and PHD4.

MLL^C contains a transactivation domain and a SET domain with intrinsic H3K4

methyltransferase activity (Figure 2.1A). MLL mediates methylation of H3K4 and

positively regulates target gene expression, including homeobox (*HOX*) genes and their

co-factors such as *MEIS1* (Wang et al., 2009b). MLL is responsible for maintaining

expression of *HOX* and *MEIS1* in hematopoietic stem cells and progenitors, which is

required for stem cell self-renewal and progenitor expansion (Jude et al., 2007;

McMahon et al., 2007; Wang et al., 2009b). *HOX* expression decreases concurrent with

hematopoietic differentiation. This is crucial for normal hematopoiesis, as constitutive

activation of *HOX* genes is associated with leukemia and other malignancies

(Argiropoulos and Humphries, 2007).

Chromosomal translocations involving *MLL* are one of the most common genetic

alterations in human leukemia, accounting for up to ~80% of infant leukemia and about

5-10% of adult leukemia overall (Aplan, 2006; Hess, 2004a). Most of the leukemogenic

MLL fusion proteins contain the N-terminus of MLL fused in frame to the C-terminus

of a translocation partner, thus forming a chimeric protein with abnormal

transactivation ability (Hess, 2004a). Both *in vitro* and *in vivo* studies have

demonstrated that these MLL fusion proteins induce leukemogenesis mainly through

constitutive activation of *HOXA9* and *MEIS1* (Armstrong et al., 2002; Ayton and

Cleary, 2003; Wong et al., 2007). Notably, translocations of *MLL* invariably occur

within the breakpoint cluster region (BCR) which leads to the deletion or disruption of

the PHD/Bromo region (Zhang and Rowley, 2006). Further, insertion of PHD/Bromo

into MLL-AF9 and MLL-ENL fusion proteins abolishes their transformation ability

suggesting this region may be important for the regulation of MLL (Chen et al., 2008;

Muntean et al., 2008).

The ankyrin repeat and suppressor of cytokine signaling (SOCS) box-containing (ASB)

protein family contains 18 members (ASB1-18), which function as the substrate

recognition module in the ECS^{ASB} (Elongin B/C-Cullin–SOCS box protein) E3

ubiquitin ligase complex (Kohroki et al., 2005). Through interaction with ElonginC

(EloC), the ASB proteins associate with ElonginB (EloB), Cullin5 (Cul5) and Rbx2, and target the substrate for ubiquitination (Kohroki et al., 2005). A number of studies have shown that ASB proteins function in a wide range of biological processes (Chung et al., 2005; Heuze et al., 2005; Kile et al., 2001; McDaneld and Spurlock, 2008). ASB2 was discovered as a gene induced by *all-trans* retinoid acid (ATRA) via a retinoid receptor (RAR) binding element present in the *ASB2* promoter. *ASB2* expression is rapidly induced during ATRA induced differentiation of leukemia cell lines, including APL cells containing the PML-RARα fusion protein, as well as HL60 and NB4 cell lines, while ectopic expression of ASB2 promotes growth inhibition and cell differentiation (Guibal et al., 2002; Kohroki et al., 2001). Filamins A and B are two known targets of ASB2 (Heuze et al., 2008). ASB2 expression is also activated by Notch signaling, which leads to degradation of Jak2 directly and degradation of Jak3 and E2A indirectly by bridging the formation of a Cul1-Cul5 dimeric E3 ligase complex (Nie et al., 2011; Wu and Sun, 2011). Besides a role in hematopoiesis, ASB2 can also regulate muscle differentiation (Bello et al., 2009).

In this study, we explored the function of the MLL PHD/Bromo region and identified ASB2 as a novel regulator of MLL stability during hematopoietic differentiation.

2.3   Experimental Procedures

2.3.1 Cell culture

Human embryonic kidney 293 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Hoxa9-ER cells were cultured in Iscove's modified Dulbecco's medium (IMDM) supplemented with 15% FBS (Stem Cell Technologies) and 100nM 4-OHT. NB4 and K562 cells were cultured in RPMI-1640 medium supplemented with 10% FBS. MLL-AF9, MLL-ENL and E2A-HLF cells were cultured in IMDM supplemented with 15% FBS (Stem Cell Technologies) and 10ng/ml IL3.

2.3.2 Vector construction

pCXN2-Flag-MLL, pCXN2-Flag-MLL-AF9 have been described previously (Milne et al., 2002). Flag-MLL-AF9 was cloned into MSCV vector through digestion with restriction enzyme and ligation. CxxC, CxxC-PHD/Bromo and the deletion constructs were cloned from full length MLL, Flag or Myc tag and the nuclear localization signal were added by PCR reaction. The fragments were then ligated to MigR1 or pCMV vector (Clontech). pCMV-ASB2-3XFlag, pCMV-ASB6-3XFlag, pCMV-ASB7-3XFlag expression vectors were kindly provided by Dr. Junya Kohroki (Tokyo University of Science). The Flag tag was replaced with HA tag by PCR-based mutagenesis. ASB2 deletion constructs were cloned from pCMV-ASB2-HA and ligated into the pCMV vector. pcDNA5/FRT-Flag-ElonginB,

pcDNA5/FRT-Flag-Cullin5, and pCI-neo-ElonginC expression vectors were kindly

provided by Dr. Joan Conaway (Stowers Institute for Medical Research).

2.3.3   Immunoprecipitation and Western blotting

Immunoprecipitation and western blotting were performed as described previously

(Chen et al., 2008). 293 cells were transfected with FuGene 6 (Roche) according to the

manufacturer's instructions and cultured for 48hrs. Cells were then lysed with BC-300

lysis buffer (20mMTris-HCl (pH 7.4), 10% glycerol, 300mM KCl, 0.1% NP-40) and

incubated with agarose affinity beads overnight at 4C. Beads were washed 3 times with

BC-300 buffer. Proteins were eluted by boiling in SDS-loading buffer, resolved by

SDS-PAGE, and detected by western blotting. Primary antibodies against MLL, ASB2,

ElonginB, SIII p15 (ElonginC) andβ-ACTIN were obtained from Bethyl and Upstate,

Abcam and Imgenex, Santa Cruz biotechnology, BD Transduction Laboratories, and

Sigma respectively. Antibodies against Myc and HA tag were obtained from Abcam.

Anti-Flag antibody and M2 anti-Flag agarose affinity beads were purchased from

Sigma. Anti-Myc agarose affinity beads were purchased from Clontech.

2.3.4   Protein identification by LC-Tandem Mass Spectroscopy

293 cells were transfected with empty vector or Flag tagged CxxC or

CxxC-PHD/Bromo. Cells were lysed and immunoprecipitations were performed as

described above. Bound material was eluted with Flag peptide and concentrated with a

Micron YM-30 centrifugal filter column (Millipore). Proteins were then separated by

SDS-PAGE. Mass spectroscopy analysis was performed using a LTQ-Orbitrap XL

mass spectrometer (ThermoFisher). Proteins were identified by searching the data

against Human IPI database (v 3.41, 72,254 entries) appended with decoy (reverse)

sequences using X!Tandem/Trans-Proteomic Pipeline (TPP) software suite

(Nesvizhskii et al., 2003). All proteins with a ProteinProphet probability score of >0.9

(error rate <2%) were considered positive identifications.

2.3.5 Dual luciferase assay

293 cells were transfected with MLL or MLL-AF9, ASB2, Renilla luciferase reporter

(internal control), and *Hoxa9*-LUC reporter with FuGene 6 according to the

manufacturer's instructions. Cells were serum-starved in 0.5% FBS in OPTI-MEM

media for 48 hr. Luciferase assays were performed using the Dual Luciferase Assay Kit

(Promega) according to manufacturer's instructions. Emission was detected using a

Monolight 3010 Luminometer (BD Biosciences).

2.3.6 Chromatin immunoprecipitation (ChIP)

ChIP assay was performed as described previously.(Milne et al., 2005) Antibodies

against Flag tag and H3K4me3 were obtained from Sigma and Abcam respectively.

Primary rabbit antibody against MLL[C] was kindly provided by Dr. Yali Dou

(University of Michigan). qPCR was performed on the precipitated DNAs with

TaqMan primers and probes from Applied Biosystems. Binding was quantitated as

follows: $\Delta CT = CT(input) - CT(Chromatin\ IP)$, % total = $2^{\Delta CT}$. Primer and probe

sequences are listed as below.

Primer1:	Taqman Probe:	CCTGCGGTGGCAACCTCAGATCC
	Forward Primer:	GCCATCAAGGCCTAATCGTG
	Reverse Primer:	AAGACCCGAAGCTCCTCCTG
Primer2:	Taqman Probe:	CCCACATCGAGGGCAGGAAACACT
	Forward Primer:	CACCCGCGGCGTCTT
	Reverse Primer:	CGAACCAATGGATCTGGCA
Primer3:	Taqman Probe:	TGAATTTTCCCCCTTTTGGGCCAC
	Forward Primer:	TAGACTCACAAGGACAATATCTCCTTTT
	Reverse Primer:	AGGTACTGAGTATTAAGCAGCTGTTTACA
Primer4:	Taqman Probe:	CCACCGCCCCTCCCATTAACAACA
	Forward Primer:	CTGTTGCTTTGTGTTCCAGATTG
	Reverse Primer:	AAGTGAGAAGGCCACAGCCA
Primer5:	Taqman Probe:	CCTCTTGATGTTGACTGGCGATTTTCCC
	Forward Primer:	TGACCCCTCAGCAAGACAAAC
	Reverse Primer:	TCCCGCTCCCCAGACTG
Primer6:	Taqman Probe:	AAGCGCCTGGCTGGCTTTCCA
	Forward Primer:	AGGGTGATCTGGCCGATGT
	Reverse Primer:	AAAATGGGCTACCGACCCTAGT
Primer7:	Taqman Probe:	TGTTGGTCGCTCCTGACTTTCCACC
	Forward Primer:	CACAGCGAGGCAAACGAAT
	Reverse Primer:	TTATTGTTTCGGAAGCCACACA
Primer8:	Taqman Probe:	ATTATGACTGCAAAACACCGGGCCATT
	Forward Primer:	CGCGATCCCTTTGCATAAAA
	Reverse Primer:	CGTAAATCACTCCGCACGCT
Primer9:	Taqman Probe:	CTTCAGTCCTTGCAGCTTCCAGTCCAA
	Forward Primer:	CAGCTCTGGCCGAACACC
	Reverse Primer:	TTCCACGAGGCACCAAACA
Primer10:	Taqman Probe:	TACCCTCCAGCCGGCCTTATGGC
	Forward Primer:	GGTGCGCTCTCCTTCGC
	Reverse Primer:	GCATAGTCAGTCAGGGACAAAGTGT
Primer11:	Taqman Probe:	CCAACTGGCTACATGCTCGCTCCA
	Forward Primer:	TCTCTCTCCCTCCGCAGATAAC
	Reverse Primer:	GGGCATCGCTTCTTCCG

2.3.7   Real-time quantitative reverse transcription PCR (RT-qPCR)

RNA was extracted from cells using TRIzol reagent (Invitrogen) and cDNA was synthesized using Superscript III Reverse Transcriptase (Invitrogen) according to manufacturer's instructions. RT-qPCR was performed with Taqman gene expression assays (Applied Biosystems) and ABI 7500 PCR Detection System. Data was analyzed using comparative ΔΔCt method (described in ABI Prism 7700 Sequence Detection System User Bulletin No. 2). Taqman primer probe sets for mouse *Asb2*, *Mll*, *Hoxa9*, *Meis1* and human *ASB2*, *MLL*, *HOXA9* and MEIS1 were purchased from Applied Biosystems.

2.3.8   shRNA knockdown and Characterization of cell differentiation

pSM2 shRNA against human ASB2 (clone ID V2HS_97797) and mouse Asb2 (clone ID V2MM_54107) and pTRIPZ shRNA against human ASB2 (clone ID V2THS_97800) were purchased from Open Biosystems. Lentiviral vector pTRIPZ shRNAs were packaged by the University of Michigan Vector Core. Retroviral vector pSM2 shRNAs were packaged in Plat-E cells as described previously(Muntean et al., 2010). Cells were transduced by spinoculation and selected with puromycin at 0.5 ug/ml (for Hoxa9-ER Hoxa9-ERcells) and 0.4 ug/ml (for NB4 cells).

To evaluate cell differentiation, pTRIPZ shRNA-transduced NB4 cells were treated with 0.5ug/ml doxycycline for 24hrs to induce shRNA expression. Then ATRA or DMSO was added to the medium at 1 μM to induce differentiation. For flow cytometry

analysis, cells were collected and washed with PBS, and stained with APC-mouse anti-human CD11b antibody or isotype control (BD Pharmingen) for 30min. Cells were then washed with Standard Buffer (1X PBS, 0.1% sodium azide, 1% heat inactivated FBS) twice and resuspended in Standard Buffer. FACS data was collected on an LSRII (BD Pharmingen) and analyzed with FlowJo flow cytometry analysis software. For nitroblue tetrazoleum (NBT) assay, 2×10^5 cells were suspended in 200μL medium with 1 mg/mL NBT and 30 ng/mL 12-O-tetradecanoylphorbol-13-acetate (Sigma), and incubated at 37 ℃ for 30 minutes. Cytospins were then prepared and the percentage of NBT positive cells was determined by microscopically counting at least 500 cells per experimental condition.

2.3.9   Protein purification

His-MBP-PHD3 (aa1551-1629) and His-MBP-PHD3-4 (aa1551-2025) were generated by cloning the fragments into the pMCSG9 vector (Donnelly et al., 2006). His-GST-ASB2 was generated by cloning ASB2-HA into pMCSG10. Expression plasmids were transformed into BL21 bacteria and screened at the University of Michigan High-Throughput Protein Lab. Bacteria were grown at 37C in TB medium at 250 rpms with 50ug/ml spectinomycin and 100ug/ml ampicillin to an OD600 of approximately 0.7, and protein expression was induced with 200μM IPTG overnight at 20C. Bacteria were lysed in CelLytic B buffer (Sigma) supplemented with 0.15M NaCl, 1 mM PMSF, 1 mM DTT, 0.4 mg/ml lysozyme, 2 mM MgCl2, 100 units/ml Benzonase,

and Complete protease inhibitor cocktail tablet (Roche). Proteins were purified over a

Ni column (GE Pharmacia) using an AKTA Purifier liquid chromatography system

(GE Pharmacia) and eluted with 20 CV of elution buffer (20mM Tris-HCl, pH 7.5, 10%

glycerol, 1mM DTT, 0.15M NaCl, 20-500mM imidazole). For His-MBP-PHD3-4 and

His-MBP-PHD3, secondary purification was performed using a amylose column (GE

Pharmacia) and identical elution buffers except 0.15M-1M NaCl gradient. Purified

proteins were visualized by Coomassie staining. For the purification of ECS^{ASB2}

complex, pCMV-ASB2-3XFlag was transfected into 293 cells using FuGene6

according to the manufacturer's instructions and cultured for 48hrs. Cells were then

lysed with BC-300 lysis buffer. ECS^{ASB2} complex was purified using Anti-FLAG   M2

affinity gel (Sigma) according to the manufacturer's instructions. Proteins are eluted

with elution buffer (20mMTris-HCl (pH 7.4), 10% glycerol, 150mM KCl, 0.1% NP-40)

containing 150 ng/ul Flag peptides and visualized by Coomassie staining.

2.3.10 In vitro binding assay

Equal amounts of purified His-MBP/ His-MBP–PHD3/ His-MBP-PHD3-4 and

His-GST/His-GST-ASB2 (1ug) were incubated in 500ul Binding Buffer (50mM

Tris-HCl pH7.9, 100mM KCl, 0.05% NP40, 0.1% BSA) for 2 hours at 4C. 20ul

Amylose resin (New England Biolabs) was then added to the binding buffer and

incubated for 1hr at 4C. Bound material was washed with Wash Buffer (50mM

Tris-HCl pH 7.9, 500 mM KCl, 0.3% NP40) 3 times and eluted from beads by boiling

in SDS loading buffer. Eluted material was visualized by western blot.

2.3.11 In vitro ubiqutination assay

Equal amounts of purified ECS^{ASB2} and His-MBP/ His-MBP–PHD3/

His-MBP-PHD3-4 (1ug) were incubated with 180ng UBE1, 560ng UbcH5C, 1ug

HA-ubiquitin and 2mM ATP in 30ul reaction butter (50mM Tris-HCl pH7.5, 5mM

MgCl2, 2mM NaF, 0.5mM DTT) at room temperature for 3hrs. Reactions were

terminated by adding 30ul 2X SDS sample buffer and boil for 5-10min. Western blot

was then performed to detect ubiquitination.

2.4 Results

2.4.1 ASB2 interacts with the PHD/Bromo region of MLL

The PHD/Bromo region of MLL is located downstream of the BCR and invariably

deleted from MLL fusion proteins (Figure 2.1A). To elucidate its function, we

performed mass spectroscopy to identify proteins that associate with the PHD/Bromo

region of MLL. A Flag-tagged MLL construct that encompasses the

CxxC-PHD/Bromo region containing a nuclear localization signal (NLS) was

transiently expressed in human embryonic kidney 293 cells and immunoprecipitated

with M2 anti-Flag agarose beads (Figure 2.1B). Coeluted proteins were resolved by

SDS-PAGE and subjected to mass spectroscopy analysis. An empty vector containing

the Flag tag and NLS, and Flag tagged CxxC with NLS were immunoprecipitated in

parallel as controls to exclude nonspecific binding and proteins that associate with the

CxxC domain. Multiple peptides corresponding to components of the ECS^{ASB} E3

ubiquitin ligase complex were identified that specifically interact with

CxxC-PHD/Bromo, including Cullin 5 (CUL5), ElonginC (ELOC) and several ASB

proteins, such as ASB7, ASB10, ASB14 and ASB18 (Figure 2.1C).

Figure 2.1 Identification of CxxC-PHD/Bromo binding proteins by mass
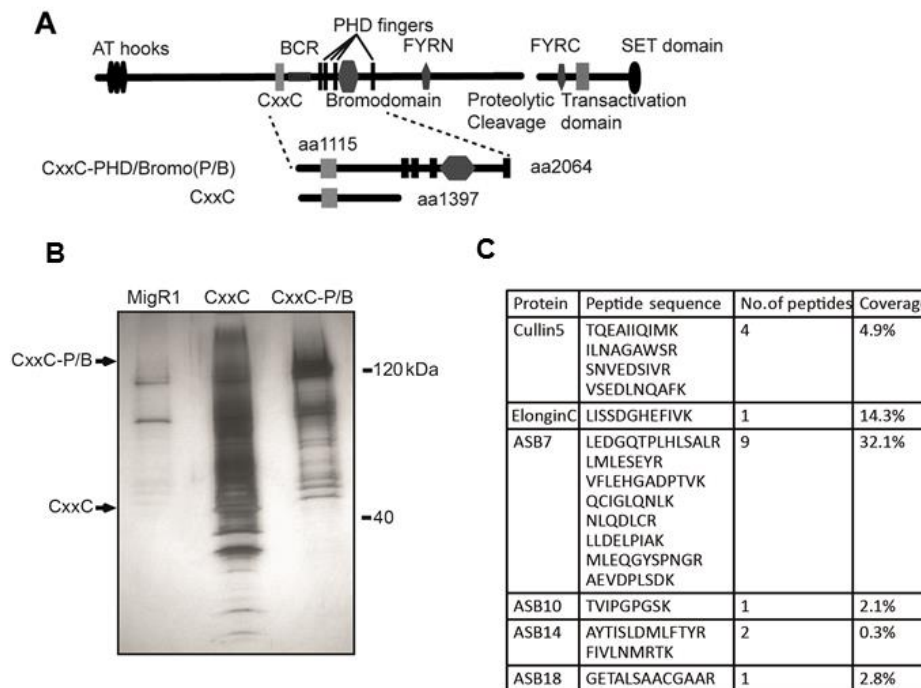
spectroscopy

Figure2.1 Identification of CxxC-PHD/Bromo binding proteins by mass spectroscopy. (A)
Schematic diagram of the structure of wild-type MLL. The CxxC and CxxC-PHD/Bromo
fragment used in immunoprecipitation are shown with the first and last MLL amino acid
retained in the constructs indicated. (B) Flag tagged CxxC or CxxC-PHD/Bromo was
transfected into 293 cells followed by anti-Flag immunoprecipitation. Silver staining shows

CxxC, CxxC-PHD and co-eluted proteins. (C) Summary of components of the ECS^{ASB} complex identified by mass spectroscopy that specifically interact with CxxC-PHD/Bromo but not CxxC.

The ASB protein family contains 18 members. As multiple ASB proteins were identified by mass spectroscopy, we focused our attention on ASB2 because of its specific activity against MLL protein (see below) and known roles in hematopoiesis. ASB2 is expressed in hematopoietic cells and has been shown to promote both differentiation and cell cycle arrest in certain myeloid leukemia cell lines, consistent with a function in hematopoiesis (Guibal et al., 2002; Kohroki et al., 2001). To test the interaction between the PHD/Bromo region and the ECS^{ASB2} complex, Myc-tagged CxxC or CxxC-PHD/Bromo and HA tagged ASB2 were coexpressed in 293 cells in the presence of the proteasome inhibitor MG132. ASB2 co-precipitated specifically with CxxC-PHD/Bromo but not CxxC confirming an interaction dependent on the PHD/Bromo region. Moreover, endogenous ELOB and ELOC also co-precipitated with CxxC-PHD/Bromo (Figure 2.2A). To determine if MLL and ASB2 interact under physiological conditions, NB4 cells were treated with ATRA to induce ASB2 expression and with MG132 to inhibit MLL degradation. Immunoprecipitation of endogenous MLL showed that ASB2 co-precipitated with MLL, further supporting a physiological interaction in hematopoietic cells (Figure 2.2B). As MLL fusion proteins generally lack the PHD/Bromo region, we tested if the interaction is disrupted in the context of MLL-AF9 by co-expressing MLL or MLL-AF9 with ASB2 in 293 cells. ASB2 co-precipitated with full length MLL but not the MLL-AF9 fusion protein again

suggesting the PHD/Bromo region is necessary for ASB2 interaction with MLL (Figure

2.2C). Based on these observations, we predicted that ubiquitination of MLL would be

at least partially dependent on the presence of the PHD/Bromo region. To this end,

CxxC or CxxC-PHD/Bromo was coexpressed with ubiquitin in 293 cells.

Immunoprecipitation revealed an ubiquitin smear associated with the

CxxC-PHD/Bromo but not the CxxC domain alone (Figure 2.2D). These data indicate

that the PHD/Bromo region interacts with the ECS^{ASB2} E3 ubiquitin ligase and results

in the ubiquitination of MLL.

Figure 2.2 PHD/Bromo region interacts with ASB2 and mediates MLL
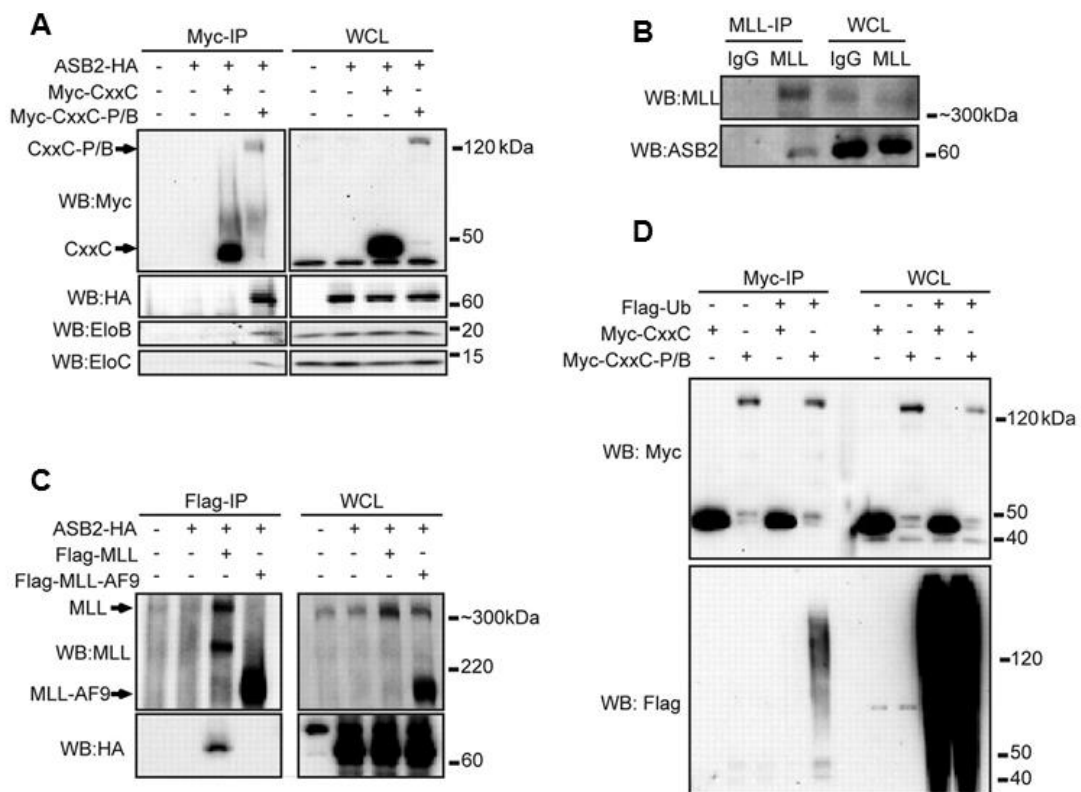
ubiquitination

Figure 2.2. PHD/Bromo region interacts with ASB2 and mediates MLL ubiquitination. (A) Myc tagged CxxC or CxxC-PHD/Bromo (CxxC-P/B) was coexpressed in 293 cells with HA tagged ASB2. Cells were treated with MG132 for 6hrs. Immunoprecipitation of CxxC or CxxC-PHD/Bromo followed by western blotting shows that ASB2 and endogenous ELOB and ELOC specifically interact with CxxC-PHD/Bromo. (B) NB4 cells were treated with ATRA for 48hrs and with MG132 for 16hrs. MLL was immunoprecipitated with an anti-MLL antibody, and western blotting using an anti-ASB2 antibody shows that endogenous ASB2 co-precipitates with MLL. (C) Flag tagged MLL or MLL-AF9 was coexpressed in 293 cells with HA tagged ASB2. After MG132 treatment, anti-Flag immunoprecipitation was performed followed by western blotting, showing that ASB2 interacts with MLL but not with MLL-AF9. (D) Myc tagged CxxC or CxxC-PHD/Bromo and Flag tagged ubiquitin were coexpressed in 293 cells. Immunoprecipitation of CxxC or CxxC-PHD/Bromo was performed after MG132 treatment. Western blotting shows that CxxC-PHD/Bromo is conjugated with ubiquitin.

2.4.2 ASB2 degrades MLL and reduces MLL transactivation ability

To test the effect of ASB2 on MLL protein turnover, full length MLL was coexpressed

in 293 cells with ASB2 or two other ASB proteins, ASB6 and ASB7. Western blotting

with whole cell lysate revealed that overexpression of ASB2 led to significant

degradation of MLL, while overexpression of ASB6 and ASB7 did not affect MLL

protein levels (Figure 2.3A). The effect of ASB2 was also confirmed on endogenous

MLL as ASB2 expression in 293 cells led to reduced endogenous MLL protein level

detected by western blotting (Figure 2.3B). Because ASB2 interacts with wild-type

MLL but not MLL-AF9 fusion protein (Figure 2.2C), we predicted that overexpression

of ASB2 would degrade MLL while largely not affecting MLL-AF9 protein levels. To

test this, equimolar amounts of plasmids expressing MLL or MLL-AF9 were

cotransfected into 293 cells with an increasing dosage of ASB2. Western blotting

revealed that ASB2 led to degradation of MLL in a dose-dependent manner, while the

MLL-AF9 level was largely unaffected (Figure 2.3C).

50

Figure 2.3 ASB2 expression reduces MLL protein level
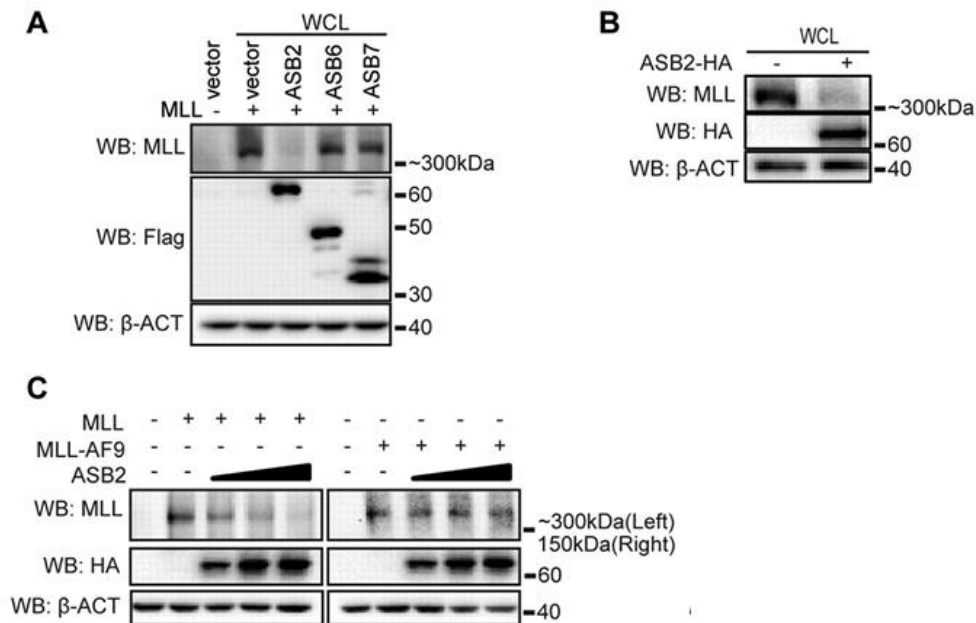
Figure 2.3. ASB2 expression reduces MLL protein level. (A) MLL was coexpressed with Flag tagged ASB2, ASB6 and ASB7 in 293 cells. Western blotting with whole cell lysate shows that ASB2 specifically degrades MLL. β-ACTIN blot shows equal loading of the samples. (B) Expression of ASB2 in 293 cells followed by western blotting with whole cell lysate shows that ASB2 leads to degradation of endogenous MLL. (C) Equimolar amounts of MLL or MLL-AF9 expression plasmids were cotransfected with ASB2 at the ratio of 6:1, 2.5, 4 in 293 cells. Western blotting with whole cell lysate shows that MLL degradation is ASB2 dose-dependent, while the level of MLL-AF9 is not affected. β-ACTIN blot indicates equal loadings.

As ASB2 functions in the ECS^{ASB} E3 ligase complex that contains multiple subunits, we tested the effect of the other components on the stability of MLL and MLL-AF9. ELOB, ELOC or CUL5 was coexpressed with MLL or MLL-AF9 in 293 cells. Western blotting with whole cell lysate revealed that CUL5 did not affect MLL stability and ELOB and ELOC had a limited effect on the degradation of MLL (Figure 2.4A, left panel), demonstrating the importance of substrate recognition by ASB2. In contrast, the stability of MLL-AF9 was not significantly changed (Figure 2.4A, right panel).

Figure 2.4 The effect of other ECSASB2 E3 ligase complex components on the stability of MLL and MLL-AF9
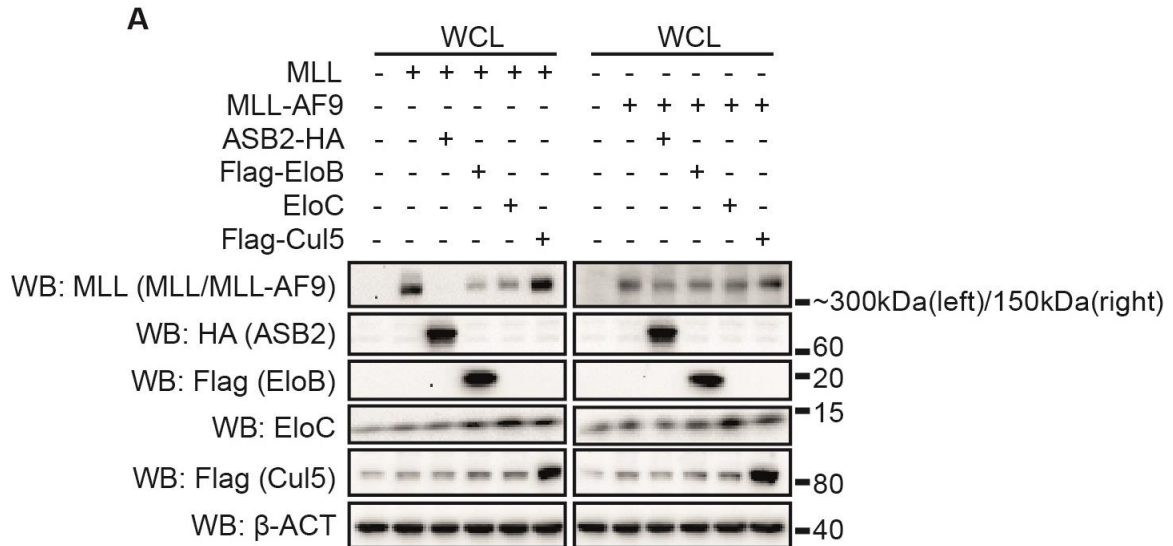
Figure 2.4. The effect of other ECS^{ASB} E3 ligase complex components on the stability of MLL and MLL-AF9. (A) HA tagged ASB2, Flag tagged ELOB, Flag tagged CUL5 or ELO C was cotransfected into 293 cells with MLL or MLL-AF9. Western blotting with whole cell lysate shows that ASB2 led to degradation of MLL. ELOB and ELOC led to degradation of MLL to a lesser extent. CUL5 did not affect the stability of MLL. None of the components have a significant effect on the stability of MLL-AF9.

Next, we examined the effect of ASB2 on MLL mediated transcriptional transactivation. Dual luciferase assays were performed with a firefly luciferase reporter driven by the murine *Hoxa9* promoter (*Hoxa9*-LUC reporter) (Figure 2.5A). Expression of MLL led to a greater than 7-fold activation of reporter expression (Figure 2.5B, lane3). While ASB2 alone had no effect on reporter expression, coexpressed ASB2 reduced MLL mediated transactivation on the *Hoxa9*-LUC reporter in a dose-dependent manner (Figure 2.5B, lane 2 and 4-6). Only a slight decrease in the reporter expression was observed when MLL-AF9 was coexpressed with ASB2 (Figure

2.5B, lane 7-10). We also tested the recruitment of MLL to the promoter region of

Hoxa9-LUC upon expression of ASB2. 293 cells were transfected with MLL, ASB2

and the *Hoxa9*-LUC reporter, and a chromatin immunoprecipitation (ChIP) assay was

performed with probes that specifically recognize the murine *Hoxa9* promoter (Figure

2.5A). Consistent with the luciferase assay, MLL showed robust binding to the *Hoxa9*

promoter, which decreased to background levels upon expression of ASB2 (Figure

2.5C).

Figure 2.5 ASB2 expression leads to reduced MLL transactivation ability
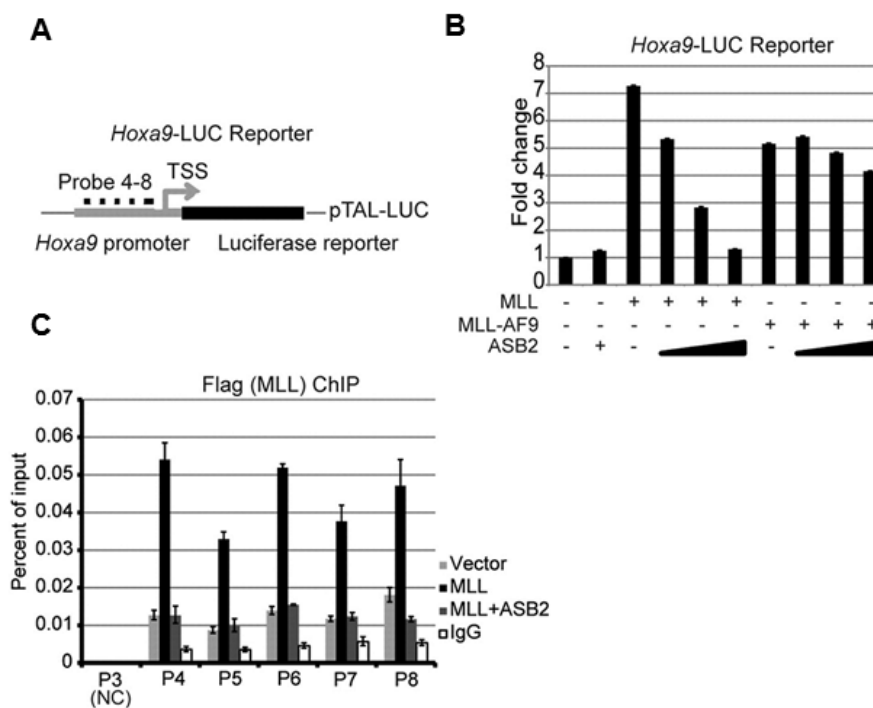
Figure 2.5. ASB2 expression leads to reduced MLL transactivation ability. (A) Schematic
diagram of the *Hoxa9*-LUC reporter. (B) Dual luciferase assay was performed in 293 cells with
Hoxa9-LUC reporter. Lane 3-6 shows expression of MLL with increasing dosage of ASB2, and
lanes 7-10 shows expression of MLL-AF9 with increasing dosage of ASB2. The ratios between
MLL or MLL-AF9 and ASB2 were the same as in (Figure 2.3 C). All changes are normalized to
lane 1, which includes *Hoxa9*-LUC and an empty expression vector. Error bars indicate

53

standard deviation. Results of one of more than three representative experiments performed are shown. (C) ChIP assay was performed in 293 cells transfected with *Hoxa9*-LUC, MLL and ASB2. The ratio between MLL and ASB2 was 6:4. Probes 4-8 cover the promoter region of *Hoxa9*-LUC. Probe3 recognizes a region that exists in the endogenous *Hoxa9* promoter but is not included in *Hoxa9*-LUC, and serves as a negative control. The nomenclature is consistent with Figure 2.21A, which shows the position of the probes on endogenous *Hoxa9* locus. Error bars indicate standard deviation. Results of one of more than three representative experiments performed are shown.

2.4.3   The bromodomain and 4th PHD finger of MLL mediate interaction with ASB2

The PHD/Bromo region of MLL contains four PHD fingers (PHD1-4) and an atypical

bromodomain between PHD3 and PHD4 (Figure 2.1A). To further characterize the

MLL-ASB2 interaction and identify which part of the PHD/Bromo region mediates the

interaction with ASB2, we made a series of deletion constructs spanning the

CxxC-PHD/Bromo region (Figure 2.6A), and performed co-immunoprecipitation

experiments with full-length ASB2. Deletion beyond the bromodomain led to

disruption of the MLL-ASB2 interaction (Figure 2.6B). A more robust interaction was

observed when PHD4 was included (Figure 2.6B compare lane 7 and 8) suggesting that

both the bromodomain and PHD4 are needed for efficient ASB2 binding.

Figure 2.6 The Bromodomain/PHD4 of MLL mediates interaction with ASB2
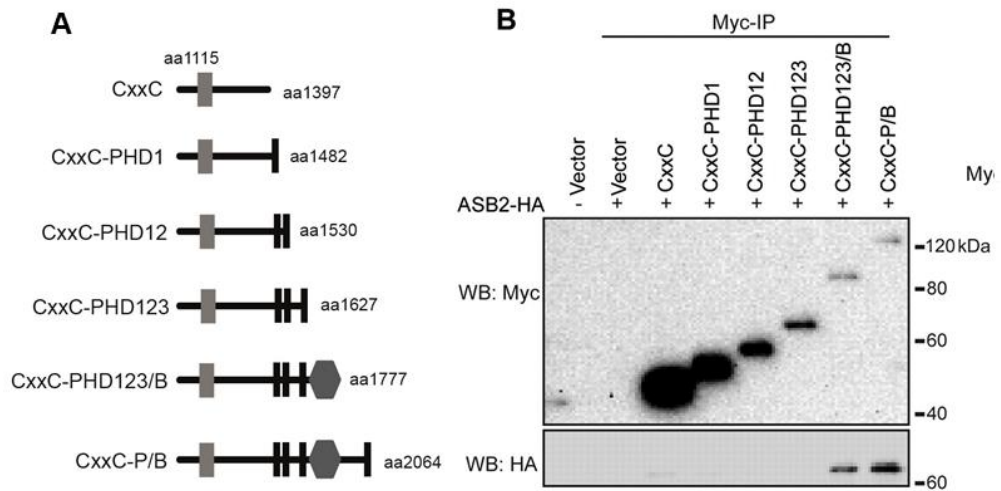
Figure 2.6. The Bromodomain/PHD4 of MLL mediates interaction with ASB2. (A) Schematic diagram of the CxxC-PHD/Bromo serial deletions. The first and last MLL amino acid retained in the constructs are indicated. (B) CxxC-PHD/Bromo deletion constructs described in (A) were cotransfected with ASB2 into 293 cells followed by anti-Myc immunoprecipitation. Binding of ASB2 was detected by western blotting with anti-HA antibody.

Consistent with this, MG132 treatment led to stabilization of the two deletion proteins

containing Bromo or Bromo-PHD4, supporting that this region is important for

ubiquitination (Figure 2.7A).

Figure 2.7 Bromodomain/PHD4 containing fragments are stabilized by MG132
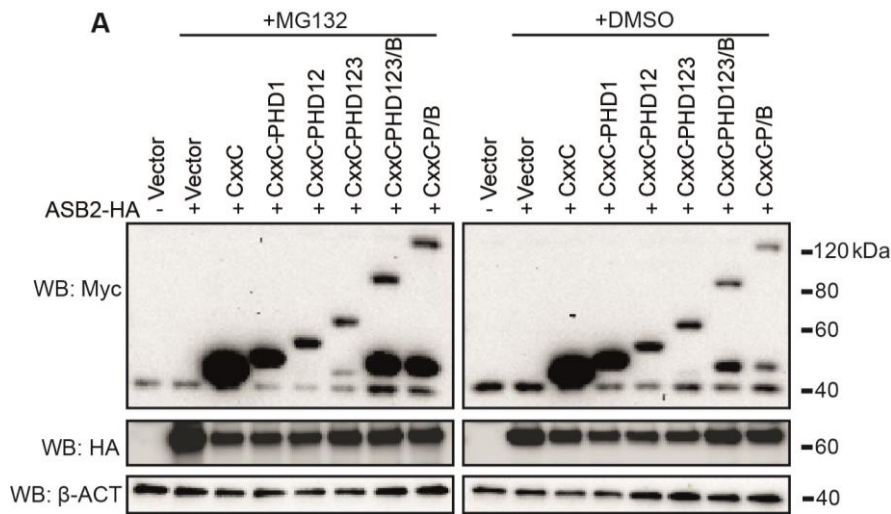
Figure 2.7. Bromodomain/PHD4 containing fragments are stabilized by MG132. (A). CxxC-PHD/Bromo deletion constructs described in (Figure 2.6A) were cotransfected with ASB2 into 293 cells followed by MG132 or DMSO treatment for 16hrs. Western blotting with whole cell lysate shows that CxxC-PHD123-Bromo and CxxC-P/B were stabilized by MG132 (lanes 7 and 8 in each blot).

To investigate whether MLL and ASB2 directly associate, *in vitro* binding assays were performed with bacterially expressed and purified His-GST tagged ASB2-HA and His-MBP tagged PHD3-Bromo-PHD4 (PHD3-4). ASB2 specifically bound to PHD3-4, but not to PHD3, suggesting a direct interaction (Figure 2.8A-B). In addition, to test if Bromo-PHD4 can serve as a substrate of ASB2, *in vitro* ubiquitination assays were performed with immunopurified ECS^{ASB2} complex from 293 cells and bacterially-purified His-MBP-PHD3-4. Purified ECS^{ASB2} complex showed ATP-dependent auto-ubiquitination activity *in vitro* (Lane 2, Figure 2.8C). An enhanced ubiquitin smear is observed upon addition of PHD3-4 but not PHD3 alone

indicating that ECS^{ASB2} requires the Bromo/PHD4 region of MLL for ubiquitination *in vitro* (Figure 2.8C).

Figure 2.8 ASB2 binds to and ubiquitinates PHD3-4 in vitro
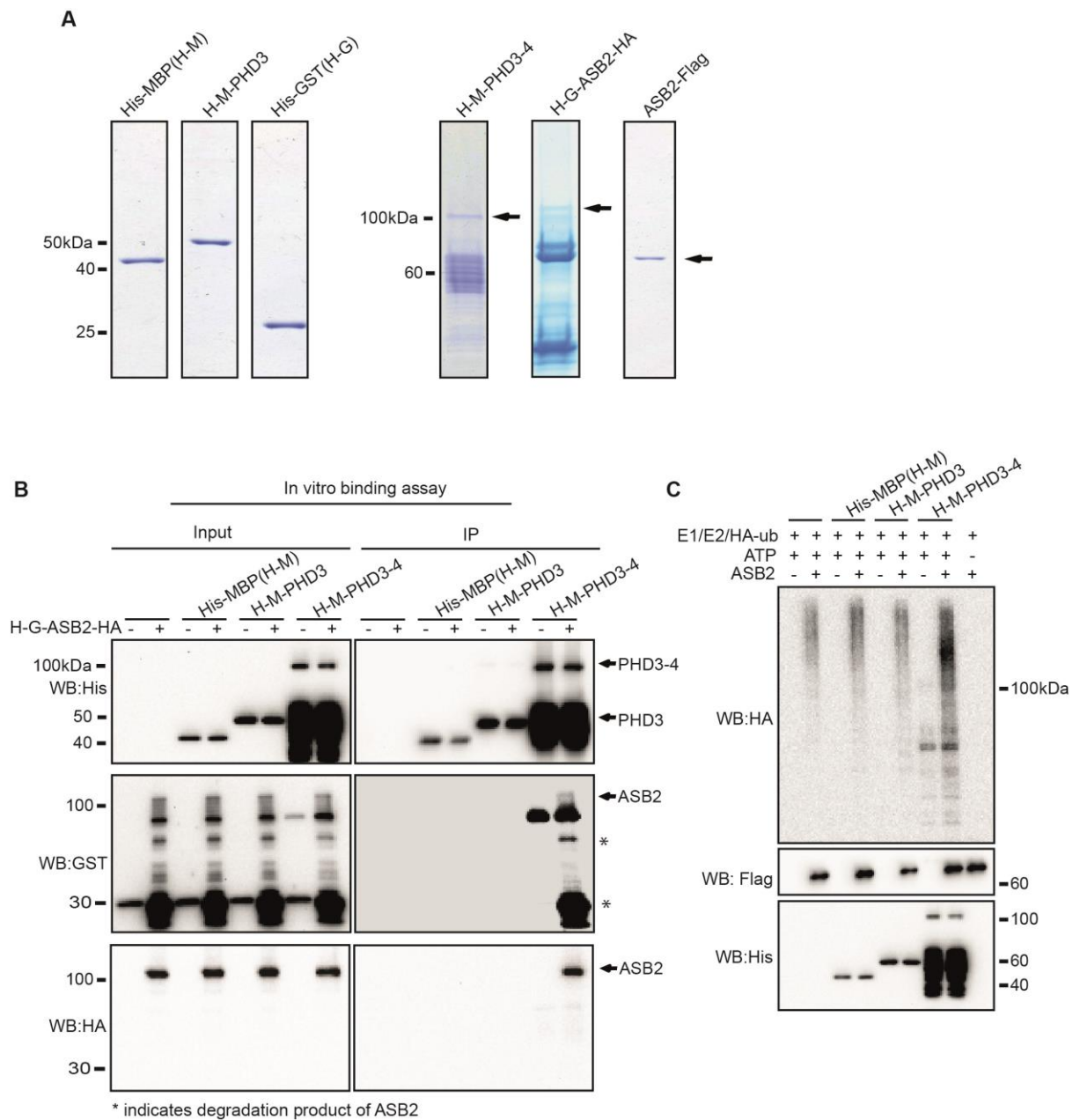
Figure 2.8. ASB2 binds to and ubiquitinates PHD3-4 in vitro. (A). Coomassie staining of purified His-MBP, His-MBP-PHD3, His-MBP-PHD3-4, His-GST, His-GST-ASB2 and ASB2-Flag. 1ug His-MBP, His-MBP-PHD3 and His-GST, 2ug of His-MBP-PHD3-4, His-GST-ASB2 and ASB2-Flag was loaded each lane. Arrow indicates purified proteins. (B).

Bacterially purified His-GST tagged ASB2 was incubated with His-MBP tagged PHD3 or PHD3-4, followed by pull-down using Amylose resin. Western blotting shows that ASB2 interacts with PHD3-4 but not with PHD3 alone. (C). Immunopurified ECS^{ASB2} complex from 293 cells was incubated with bacterially purified His-MBP tagged PHD3 or PHD3-4 in the presence of UBE1 (E1), UbcH5C (E2) and HA tagged ubiquitin. Western blotting using HA antibody shows that an ubiquitin smear was formed with PHD3-4 but not with PHD3

ASB2 contains 12 N-terminal ankyrin repeats which mediate protein-protein interactions, and a C-terminal SOCS box, which is necessary for the interaction with ELOC (Heuze et al., 2005). To map the interaction region on ASB2, we made serial deletions of ASB2 that contain different numbers of ankyrin repeats (Figure 2.9A). Co-immunoprecipitation experiments with Myc-CxxC-PHD/Bromo showed that the first five ankyrin repeats are sufficient for the interaction with MLL (Figure 2.9B).

Figure 2.9 The N terminal five ankyrin repeats of ASB2 mediate interaction with MLL
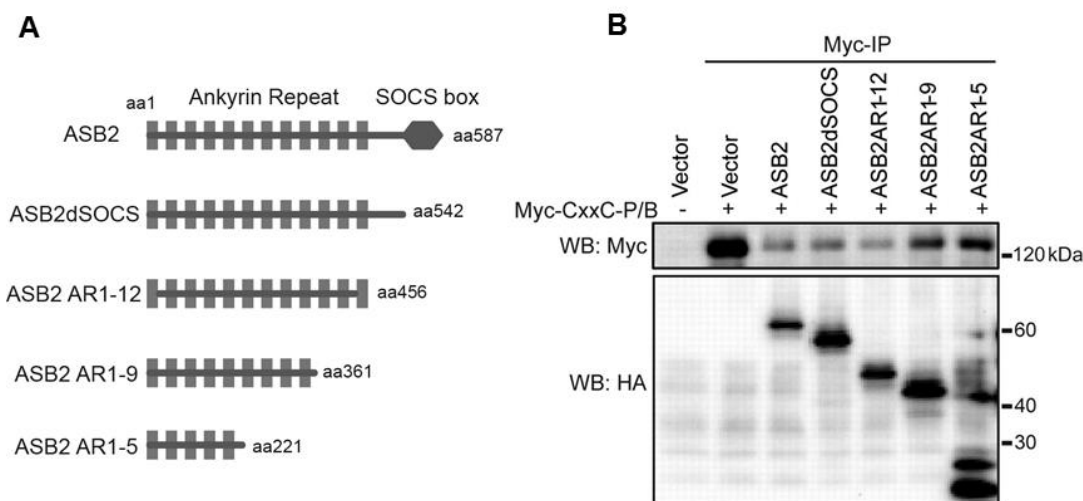
Figure 2.9. The N terminal five ankyrin repeats of ASB2 mediate interaction with MLL. (A) Schematic of HA tagged ASB2 serial deletions. The first and last ASB2 amino acid retained in the constructs are indicated. (B) ASB2 deletion constructs described in (A) were cotransfected with Myc-CxxC-PHD/Bromo in 293 cells followed by anti-Myc immunoprecipitation. Binding of ASB2 was detected by western blotting with anti-HA antibody.

2.4.4 Increased ASB2 expression during hematopoietic differentiation is associated with decreased MLL protein level and MLL target gene expression

Previous studies have shown that ASB2 expression is up-regulated during differentiation of myeloid leukemia cell lines and that ectopically expressed ASB2 induces myeloid growth arrest and differentiation (Guibal et al., 2002; Kohroki et al., 2001). To further explore the biological significance of MLL regulation by ASB2, we examined the levels of ASB2, MLL and MLL target genes during differentiation in both human and murine leukemia cell lines. First, the human leukemia NB4 cell line was treated with ATRA to induce cell differentiation. Consistent with previous publications (Kohroki et al., 2001). we observed a significant increase of *ASB2* expression upon ATRA treatment (Figure 2.10A). MLL protein levels decreased gradually and coordinately with *ASB2* up-regulation (Figure 2.10B). Notably, *MLL* transcription did not significantly change with differentiation as examined by RT-qPCR (Figure 2.10C), indicating the decrease in MLL protein is post-transcriptional. RT-qPCR was also performed to test the expression levels of *HOXA9*, a major target of MLL, which decreased approximately 40% after 72 hours' treatment (Figure 2.10D).

Figure 2.10 ASB2 expression leads to MLL degradation and MLL target gene

down-regulation during differentiation of NB4 cells
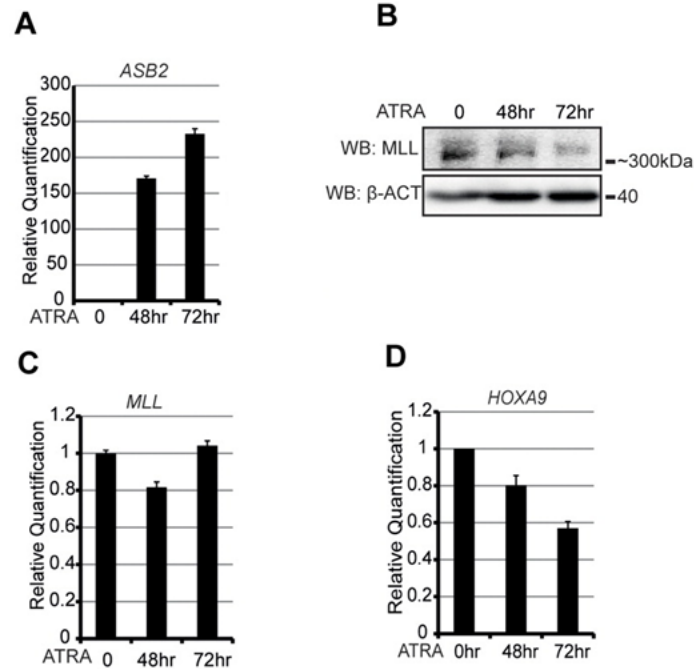
Figure 2.10. ASB2 expression leads to MLL degradation and MLL target gene down-regulation during differentiation of NB4 cells. (A-D) NB4 cells were treated with ATRA for 0, 48 and 72hrs to induce differentiation. The expression of ASB2 (A), the protein level (B) and transcription level (C) of MLL, and the expression of HOXA9 (D) were measured by RT-qPCR or western blotting. Transcription is shown relative to the level at 0 hr. Error bars indicate standard deviation. β-ACTIN blot indicates equal loading.

The expression of *ASB2*, MLL and another MLL target gene, *MEIS1,* was examined in ATRA treated K562 leukemia cells. Similar to NB4 cells, *ASB2* expression increased during ATRA induced cell differentiation, along with decreased MLL protein level and *MEIS1* expression, while the transcription of *MLL* did not change significantly (Figure 2.11 A-E).

Figure 2.11 ASB2 expression leads to MLL degradation and MLL target gene down-regulation during differentiation of K562 cells
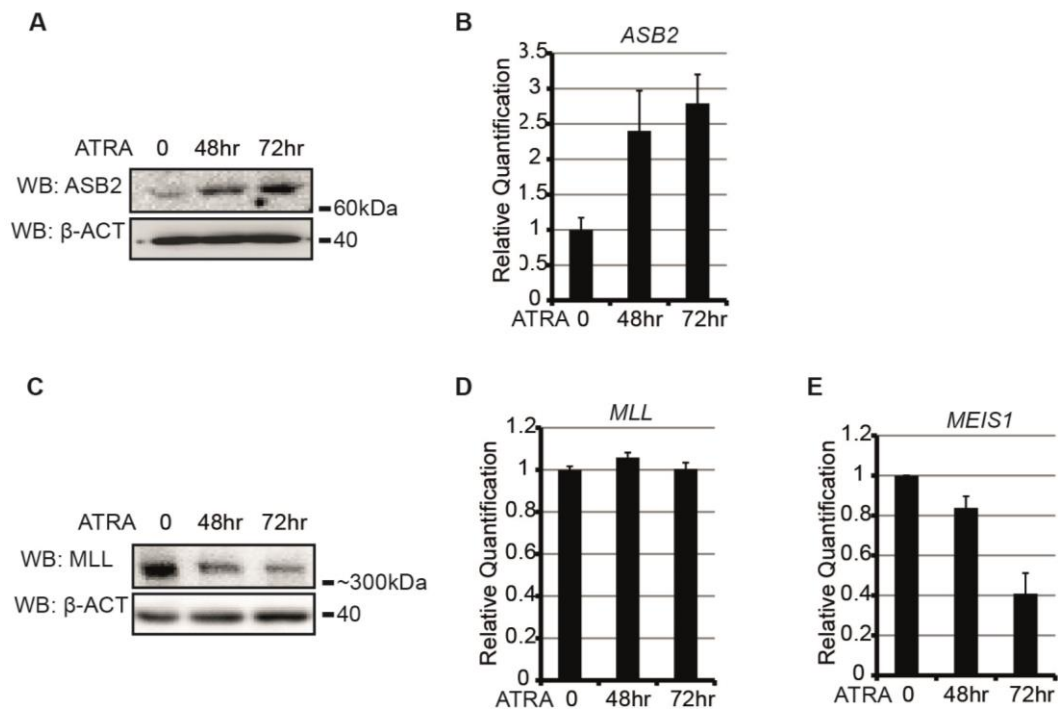
Figure 2.11. ASB2 expression leads to MLL degradation and MLL target gene down-regulation during differentiation of K562 cells. (A-E) K562 cells were treated with ATRA for 0, 48 and 72hrs to induce differentiation. The protein level and transcription level of ASB2 (A-B), MLL (C-D), and the expression of *MEIS1* (E) were measured by western blotting or RT-qPCR. Transcription is shown relative to the level at 0hr. Error bars indicate standard deviation. B-ACTIN indicates equal loading.

Because ATRA has a wide range of biological effects, we examined if the down-regulation of *HOXA9* and *MEIS1* is a direct consequence of ASB2 expression. To test this, ASB2 was transfected into NB4 cells and K562 cells by electroporation. ASB2 expression led to a significant decrease of *HOXA9* expression in NB4 cells (Figure 2.12A-B) and *MEIS1* expression in K562 cells (Figure 2.12C) suggesting the changes in gene expression are a direct result of ASB2 expression.

61

Figure 2.12 ASB2 expression down-regulates expression of *HOXA9* and *MEIS1*
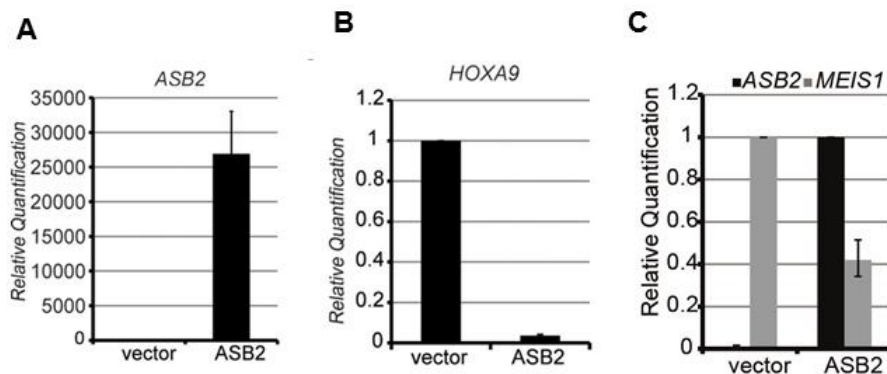
Figure 2.12. ASB2 expression down-regulates expression of *HOXA9* and *MEIS1*. (A-B) ASB2 expression vector was transfected into NB4 cells by electroporation, and the expression levels of *ASB2* (A) and *HOXA9* (B) were measured by RT-qPCR. Expression is shown relative to empty vector transfected control cells. Error bars indicate standard deviation. (C) ASB2 expression vector was transfected into K562 cells by electroporation, and the expression levels of *ASB2* (Black bar) and *MEIS1* (Grey bar) were measured by RT-qPCR. For visualization, expression of *ASB2* is shown relative to empty vector transfected control cells, and *MEIS1* expression is shown relative to ASB2 transfected cells. Error bars indicate standard deviation.

Asb2 expression was also examined in a differentiation system utilizing murine bone marrow cells transformed with an inducible Hoxa9-estrogen receptor (Hoxa9-ER) construct and maintained in 4-hydroxytamoxifen (4-OHT). Withdrawal of 4-OHT causes Hoxa9-ER translocation to the cytoplasm leading to cell differentiation and cell cycle arrest, which is largely complete within 5 days (Muntean et al., 2010). Microarray gene expression profiling was previously performed on Hoxa9-ER cells in the presence of 4-OHT and at 72, 96 and 120 hours post 4-OHT withdrawal (Muntean et al., 2010), and the 18 *Asb* family members were clustered according to their expression level changes. Of these, *Asb2* was the most highly up-regulated gene upon differentiation (Figure 2.13A and B). Along with *Asb2* up-regulation, we observed a decrease in Mll

protein level at both 72 and 120hrs post 4-OHT withdrawal compared to cells grown in

the presence of 4-OHT independent of changes in *Mll* transcription (Figure 2.13C and

D). Moreover, multiple Mll target genes, including *E2f1*, *Rbl*, *Meis1* and *Pbx3* were

down-regulated consistent with up-regulated *Asb2* degrading Mll protein (Figure

2.13E).

Figure 2.13 ASB2 expression leads to MLL degradation and MLL target gene

down-regulation during differentiation of Hoxa9-Meis1 transformed murine cells
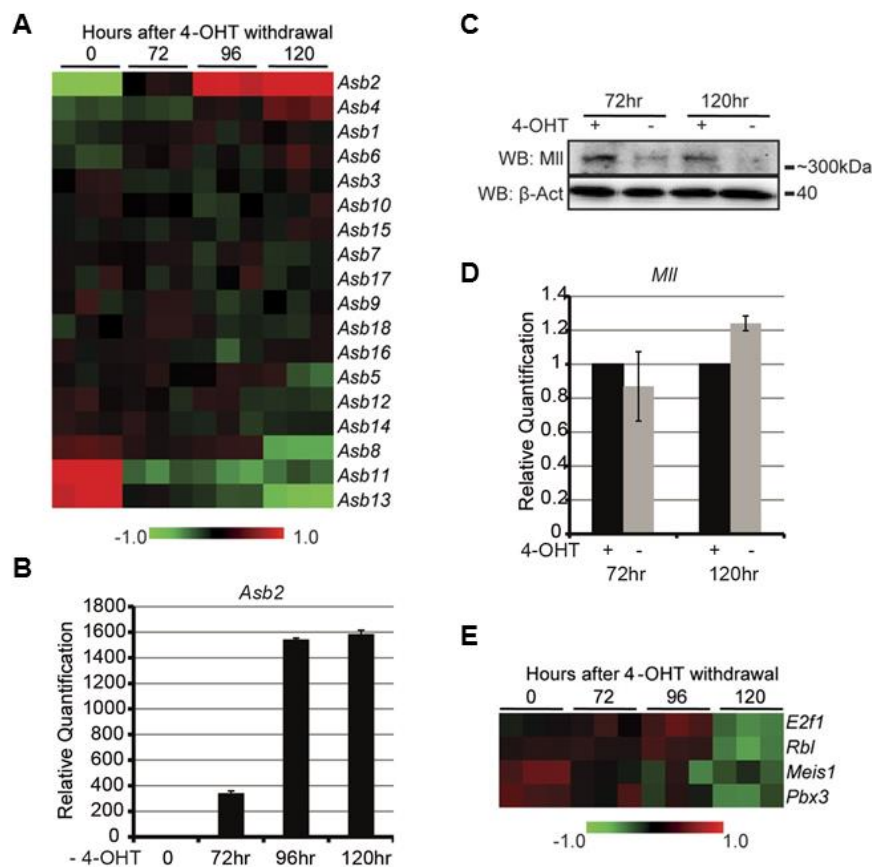
Figure 2.13. ASB2 expression leads to MLL degradation and MLL target gene down-regulation during differentiation of Hoxa9-Meis1 transformed murine cells. (A-E) Expression of *Asb2*, *Mll* and Mll target genes during 4-OHT withdrawal induced differentiation of the Hoxa9-ERHoxa9-ER cell line. (A) Heat map generated from gene expression array data

collected at different time points after 4-OHT withdrawal. The *Asb* genes were clustered according to their expression change. Data are shown in triplicate for each time point. (B) Expression of *Asb2* was measured by RT-qPCR at different time points during 4-OHT withdrawal induced differentiation of Hoxa9-ERHoxa9-ER cells, confirming the results obtained by microarray. Error bars indicate standard deviation. (C) Mll protein level in Hoxa9-ERHoxa9-ER cells was determined by western blotting at 72 and 120hrs with or without 4-OHT withdrawal. (D) Expression of *Mll* was measured by RT-qPCR at 72 and 120hr with or without 4-OHT withdrawal in Hoxa9-ERHoxa9-ER cells. Expression is shown relative to cells without 4-OHT withdrawal. No significant change was observed. Error bars indicate standard deviation. (E) Heat map of the Mll target gene expression generated from gene expression array data as in (A).

To further evaluate the role of ASB2 on MLL target gene expression and cell differentiation, we generated stable NB4 cells expressing an inducible pTRIPZ shRNA construct against ASB2 for knockdown studies. Following doxycycline treatment, shRNA expression was activated, which was reflected by expression of red fluorescent protein (RFP) on the pTRIPZ vector, leading to knockdown of ASB2 (Figure 2.14A-B).

Figure 2.14 Confirmation of shRNA mediated knockdown of ASB2
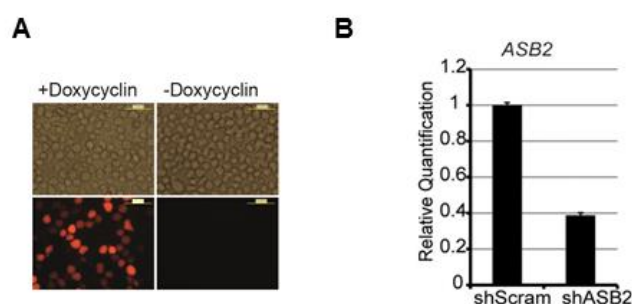
Figure 2.14. Confirmation of shRNA mediated knockdown of ASB2. (A) Image of NB4 cells transduced with pTRIPZ shRNA against ASB2 with (left) or without (right) doxycycline treatment for 24hrs, showing the activation of the shRNA. (B) RT-qPCR confirmation of *ASB2* knockdown in NB4 cells after doxycycline treatment for 24hrs. Expression is shown relative to scrambled shRNA-transduced control cells. Error bars indicate standard deviation.

With ASB2 knockdown, we observed an increase in MLL protein level, and a 60%

up-regulation of *HOXA9* expression (Figure 2.15A and C). Further, MLL turnover was

monitored using NB4 cells treated with cycloheximide. Prolonged MLL degradation

was observed in ASB2 knockdown cells compared to control scrambled

shRNA-transduced cells (Figure 2.15C).

Figure 2.15 ASB2 knockdown leads to stabilized MLL and increased expression
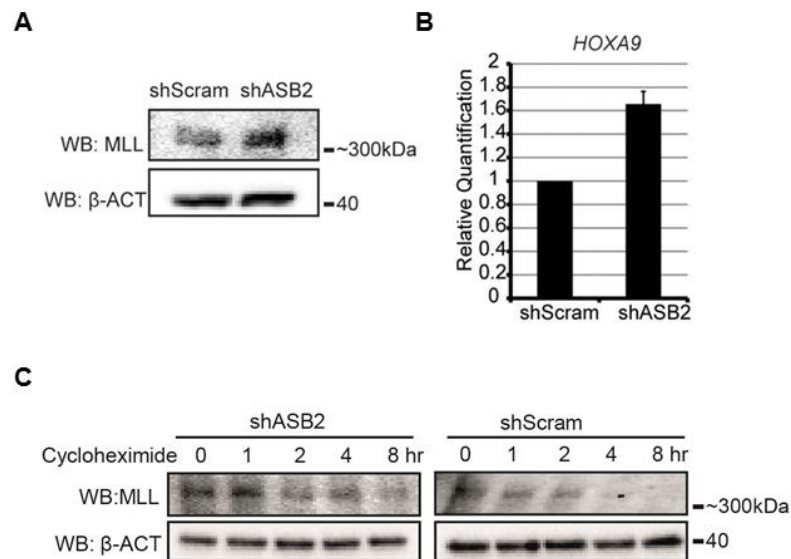
of *HOXA9*

Figure 2.15. ASB2 knockdown leads to stabilized MLL and increased expression of *HOXA9*.
(A) Protein level of MLL was detected by western blotting with whole cell lysate, which shows
an increase after doxycycline induced ASB2 knockdown. (B) Expression of *HOXA9* was
measured by RT-qPCR after doxycycline induced ASB2 knockdown. Expression is shown
relative to scrambled shRNA-transduced control cells. Error bars indicate standard deviation.
(C) NB4 cells were treated with cycloheximide for different time periods as indicated to block
translation. The turnover of MLL was monitored by western blotting using whole cell lysate.
MLL degradation was prolonged in ASB2 knockdown cells compared to control cells.

The effect of ASB2 on cell differentiation was determined by monitoring the CD11b

surface marker expression and nitroblue tetrazoleum (NBT) assay after ATRA

treatment. ASB2 knockdown led to a delay in ATRA-induced NB4 cell differentiation,

as lower CD11b expression and a lower percentage of NBT positive cells was detected

in ASB2 knockdown cells at both 24 and 48 hours (Figure 2.15A and B). We also

observed that ASB2 knockdown resulted in faster growth compared to cells expressing

control shRNA. Moreover, ATRA treatment led to a rapid growth arrest in control cells,

while ASB2 knockdown cells retained the ability to grow, albeit at a slower rate (Figure

2.15C).

Figure 2.16 ASB2 knockdown inhibits cell differentiation
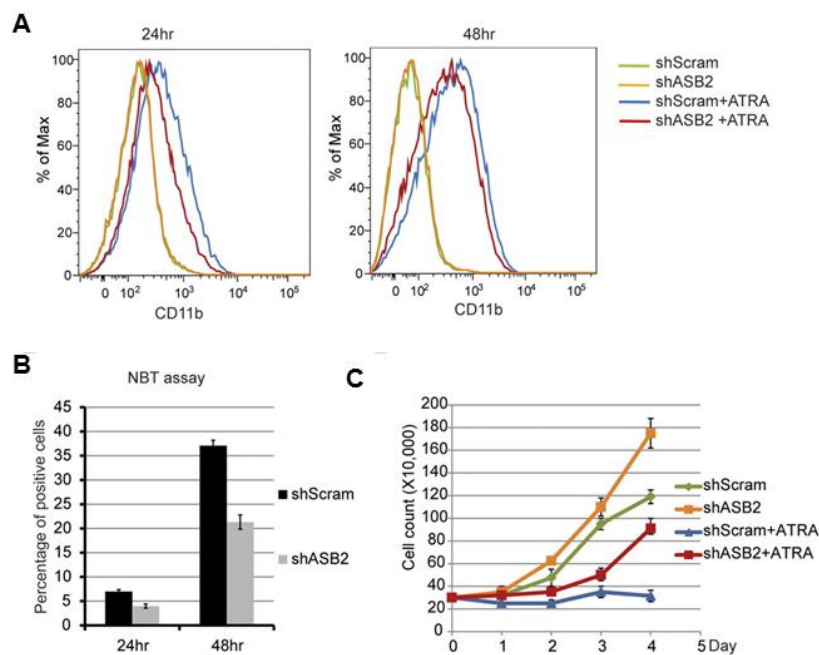
Figure 2.16. ASB2 knockdown inhibits cell differentiation. (A-C) NB4 cells were treated with doxycycline for 24hrs to induce ASB2 knockdown and then with ATRA to induce differentiation. (A) Expression of CD11b was determined by flow cytometry at 24 and 48hrs. (B) NBT assay was performed at 24 and 48hrs and the percentage of positive cells is shown. Error bars indicate standard deviation calculated from 3 independent experiments. (C) Control and ASB2 knockdown NB4 cells were grown in liquid culture with or without ATRA treatment. A proliferation advantage was observed from ASB2 knockdown cells in both conditions. Error bars indicate standard deviation.

As further confirmation, a constitutive pSM2 shRNA targeting another region of ASB2 was stably transduced into NB4 cells. Similar up-regulation of HOXA9 was observed with ASB2 knockdown (Figure 2.17A and B).

Figure 2.17 ASB2 knockdown with a second shRNA leads to increased expression of *HOXA9*
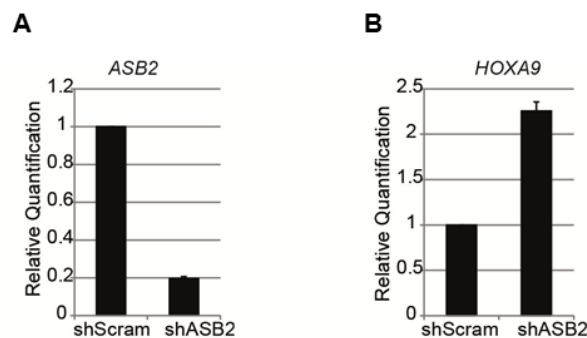
Figure 2.17. ASB2 knockdown leads to increased expression of *HOXA9*. (A) RT-qPCR confirmation of ASB2 knockdown with pSM2 shRNA against ASB2 in NB4 cells. Expression is shown relative to scrambled shRNA-transduced control cells. Error bars indicate standard deviation. (B) Expression of *HOXA9* was measured by RT-qPCR with ASB2 knockdown. Expression is shown relative to scrambled shRNA-transduced control cells. Error bars indicate standard deviation.

2.4.5   ASB2 reduced colony formation ability of MLL fusion transformed cells

Previous studies showed that expression of wild-type MLL is required for MLL fusion protein-mediated transformation (Thiel et al., 2010). As our data suggests that ASB2 leads to efficient degradation of MLL, we predicted that ASB2 expression would diminish MLL fusion mediated transformation through degradation of wild-type MLL. To test this, we established a murine cell line transformed with Flag tagged MLL-AF9. ASB2 was cloned into the MigR1 retroviral expression vector, and transduced into MLL-AF9 cells. Positively transduced cells were sorted by flow cytometry according

to green fluorescent protein (GFP) expression on the MigR1 vector, and colony

formation ability was examined by methylcellulose replating assays. ASB2 expression

led to a more than 60% decrease in colony number compared to empty

vector-transduced control cells (Figure 2.18A-C), as well as reduced cell proliferation

in liquid culture (Figure 2.18D). Consistent with this, a 40% decrease of *Hoxa9*

expression was observed in ASB2-transduced cells compared to vector control (Figure

2.18E), despite no detectable change in *Mll* transcription (Figure 2.18F).

Figure 2.18 ASB2 reduces colony formation of MLL-AF9 transformed cells
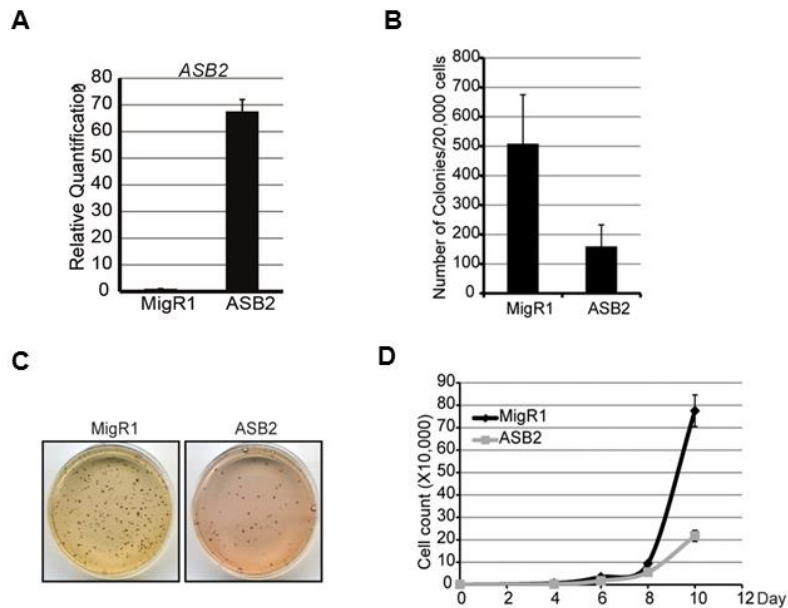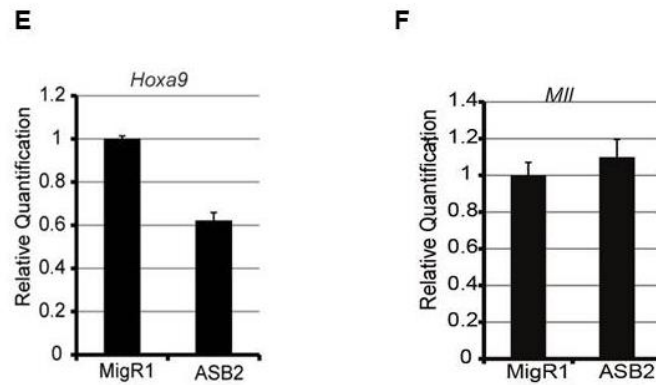
Figure 2.18. ASB2 reduces colony formation of MLL-AF9 transformed cells. (A) *ASB2* expression in MLL-AF9 cells was confirmed by RT-qPCR. Expression is shown relative to empty vector transduced control cells. Error bars indicate standard deviation. (B) Colony numbers of MLL-AF9 cells transduced with empty vector or ASB2 from methylcellulose replating assay. Error bars indicate standard deviation from three independent experiments. (C) P-iodonitro-tetrazolium violet staining of the colonies. (D) MLL-AF9 cells transduced with empty vector or ASB2 were grown in liquid culture. Significantly reduced cell proliferation was observed from ASB2-transduced cells. Error bars indicate standard deviation. (E) RT-qPCR was performed to measure the expression of *Hoxa9* in MLL-AF9 cells transduced with empty vector or ASB2. Expression is shown relative to empty vector-transduced control cells. Error bars indicate standard deviation. (F) Transcription of *Mll* was measured by RT-qPCR, which did not show significant change. Expression is shown relative to empty vector-transduced control cells. Error bars indicate standard deviation.

Notably, an ASB2 construct containing a deletion of the SOCS box (ASB2ΔSOCS)

still led to degradation of MLL, a phenomenon similar to that reported for ASB4 (Li et

al., 2007), and led to decreased colony formation when expressed in MLL-AF9 cells

(Figure 2.19).

Figure 2.19 ASB2ΔSOCS does not rescue colony formation of MLL-AF9 cells
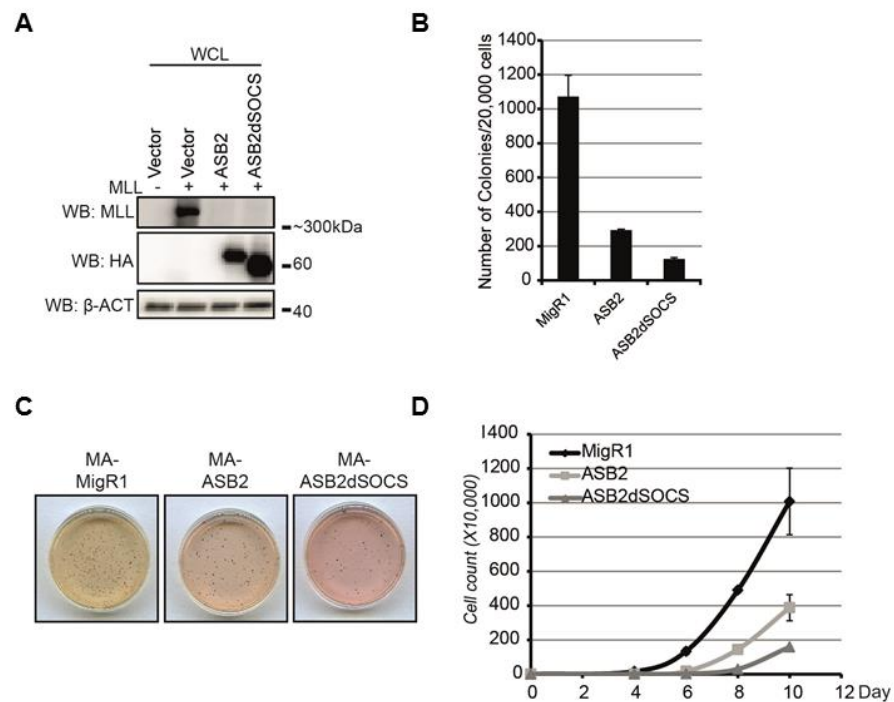
Figure 2.19. ASB2ΔSOCS does not rescue colony formation of MLL-AF9 cells. (A).
HA-tagged ASB2 or ASB2 with deletion of the SOCS box (ASB2dSOCS) was cotransfected
into 293 cells with MLL. Western blotting with whole cell lysate shows that both ASB2 and
ASB2dSOCS led to degradation of MLL to a similar extent. (B-D) Colony numbers (B), colony
staining (C) and proliferation (D) of MLL-AF9 cells transduced with empty vector, ASB2 or
ASB2dSOCS. Both ASB2 and ASB2dSOCS led to decreased colony formation and cell
proliferation. MA: MLL-AF9.

Decreased colony formation and cell proliferation was also observed using MLL-ENL

cells transduced with ASB2 (Figure 2.20A-C). Importantly, expression of ASB2 in

E2A-HLF cells, which transform in an MLL and *Hoxa9* independent manner, had no

effect on either colony formation or cell proliferation, indicating that the effect of

ASB2 is specific for MLL-fusion transformed cells (Figure 2.20D-F).

Figure 2.20 ASB2 reduces colony formation in MLL-ENL transformed cells but not in E2A-HLF transformed cells
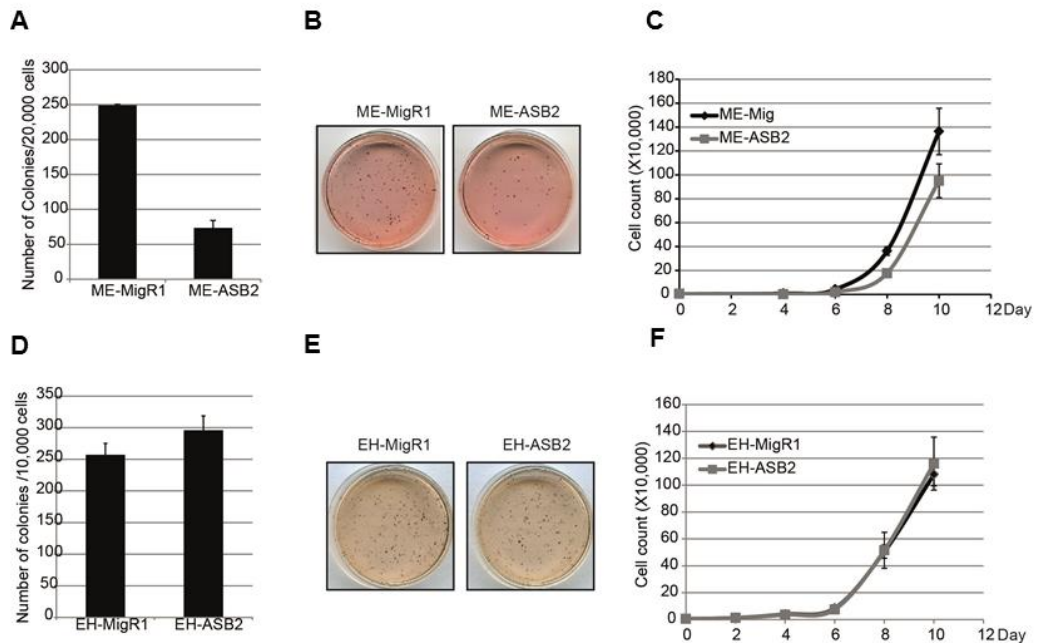
Figure 2.20. ASB2 reduces colony formation in MLL-ENL transformed cells but not in E2A-HLF transformed cells. (A) Colony numbers of MLL-ENL cells transduced with empty vector or ASB2 from methylcellulose replating assay. Error bars indicate standard deviation from duplicate experiments. Results of one of more than three representative experiments performed are shown. (B) P-iodonitro-tetrazolium violet staining of the colonies. (C) MLL-ENL cells transduced with empty vector or ASB2 were grown in liquid culture. Reduced cell proliferation was observed from ASB2 transduced cells. Error bars indicate standard deviation. (D-F) Colony numbers (D), colony staining (E), and proliferation (F) of E2A-HLF cells transduced with empty vector or ASB2. No decrease in colony formation or cell proliferation was observed. Results of one of more than three representative experiments performed are shown. Error bars indicate standard deviation. ME: MLL-ENL, EH: E2A-HLF.

To further confirm that the effect of ASB2 on colony formation is through wild-type Mll degradation, we examined the recruitment of Mll and MLL-AF9 to the *Hoxa9* locus. ChIP assays were performed with a series of probes spanning the *Hoxa9* promoter region and exons 1 and 2 (Figure 2.21A). The recruitment of endogenous Mll

was determined using an Mll antibody that recognizes a region that is lost in MLL-AF9

fusion protein. Mll showed a robust binding to the *Hoxa9* locus in control cells, while

the binding was significantly reduced upon ASB2 expression (Figure 2.21B).

Consistent with this, histone H3 lysine 4 tri-methylation (H3K4me3) is reduced in

ASB2-transduced cells compared to control cells (Figure 2.21C). In contrast,

recruitment of MLL-AF9, which was detected using Flag antibody, did not change

significantly with ASB2 expression (Figure 2.21D). These data indicate that ASB2

reduces the amount of wild-type Mll at target genes critical for leukemia, but does not

alter MLL-AF9 fusion protein binding.

Figure 2.21 ASB2 expression reduces recruitment of wild-type MLL but not

MLL-AF9 to target gene loci
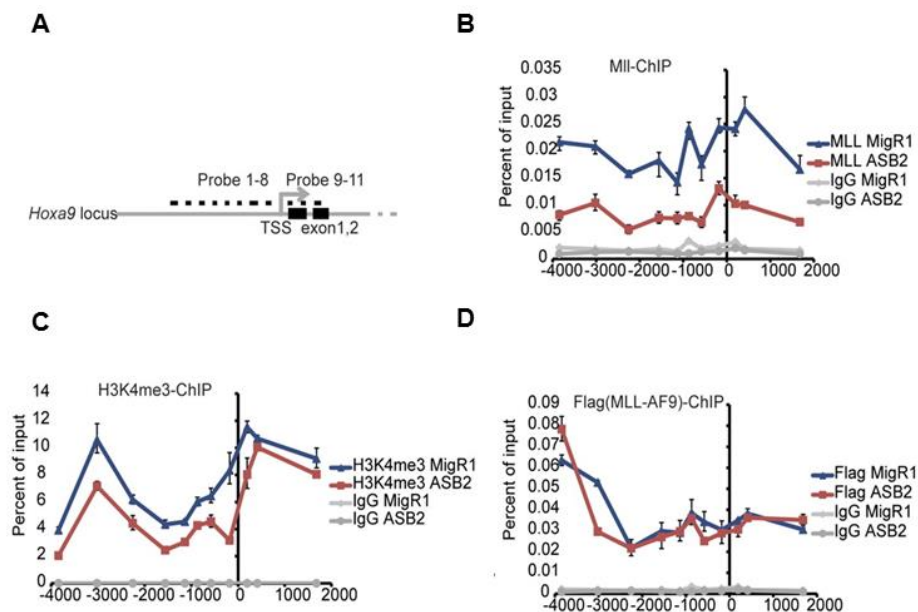
Figure 2.21. ASB2 expression reduces recruitment of wild-type MLL but not MLL-AF9 to target gene loci. (A) Schematic diagram showing the endogenous *Hoxa9* locus and probes used in the ChIP assay. TSS: transcription start site. (B-D) ChIP assays were performed to determine

the recruitment of Mll (B), the H3K4me3 level (C) and the recruitment of MLL-AF9 (D) on the *Hoxa9* locus with probes shown in (A). Antibodies against MLL, H3K4me3 and Flag were used respectively. Blue lines are data from MLL-AF9 cells transduced with empty vector; red lines are data from MLL-AF9 cells transduced with ASB2; grey lines are IgG controls for each cell line. Error bars indicate standard deviation.

2.5 Discussion

Post-translational modifications, such as phosphorylation, ubiquitination and acetylation, provide a rapid and flexible way to regulate protein activity in response to intrinsic and extrinsic cellular signals. As one of the major mechanisms of controlling protein abundance, ubiquitination has been shown to regulate a wide range of biological processes, including cell cycle progression, cell proliferation and differentiation. Recent studies have shown that several histone-modifying enzymes are regulated through ubiquitination. For example, the E3 ligase CUL4-DDB1 regulates the histone H4K20 monomethylase PR-Set7/Set8 during S phase, which is important for replication licensing, cell cycle progression and DNA damage response (Abbas et al., 2010; Centore et al., 2010; Oda et al., 2010; Tardat et al., 2010). Moreover, two components of the MLL complex, WDR5 and RBBP5, have been shown to be regulated by CUL4-DDB1, suggesting a strict regulation of H3K4 trimethylation by E3 ubiquitin ligases (Higa et al., 2006). Indeed, previous studies have shown that MLL stability is regulated during cell cycle progression by the E3 ligase SCF^{Skp2} and APC^{Cdc20}. Further work has shown that ATR mediated phosphorylation of MLL due to

genotoxic stress disrupts the interaction with SCF^{Skp2}, linking MLL to the mammalian

S phase checkpoint (Liu et al., 2007a; Liu et al., 2010).

In this study, we identified ASB2 as a novel regulator of MLL, and linked *HOX* gene

down-regulation during hematopoiesis to MLL ubiquitination and degradation through

the ECS^{ASB2} ubiquitin ligase complex. ASB2 expression leads to degradation of MLL

and reduces its transactivation ability (Figure 2.3 and 2.5). A significant increase in

ASB2 expression was observed during differentiation of both human leukemia (NB4

and K562) cell lines and murine leukemia (Hoxa9-ER) cell line, displaying a reciprocal

pattern to the MLL protein level and MLL target gene expression (Figure 2.10-2.13).

Furthermore, expression of ASB2 leads to *HOXA9* and *MEIS1* down-regulation, while

knockdown of ASB2 up-regulates *HOXA9* and renders the cells resistant to

differentiation stimuli (Figure 2.10-2.17). These findings support a model where MLL

is regulated post-translationally through ASB2 up-regulation during hematopoietic

differentiation leading to an active ECS^{ASB2} E3 ubiquitin ligase complex that rapidly

degrades MLL protein and down-regulates *HOX* gene (Figure 2.22A-B). In contrast,

loss of the PHD/Bromo region and escape from ASB2 mediated degradation may

contribute to MLL fusion mediated leukemogenesis (Figure 2.22C). Our study adds to

the list of regulators of MLL and suggests that SCF^{Skp2} and APC^{Cdc20} act coordinately to

control MLL protein turnover during cell cycle progression, while ECS^{ASB2} regulates

MLL during hematopoietic differentiation. Together, these data suggest that MLL

protein turnover is tightly controlled by several E3 ubiquitin ligase complexes during

various biological processes.

Figure 2.22 Model for the regulation of MLL degradation and *HOX* gene

expression by ASB2 during normal and malignant hematopoiesis
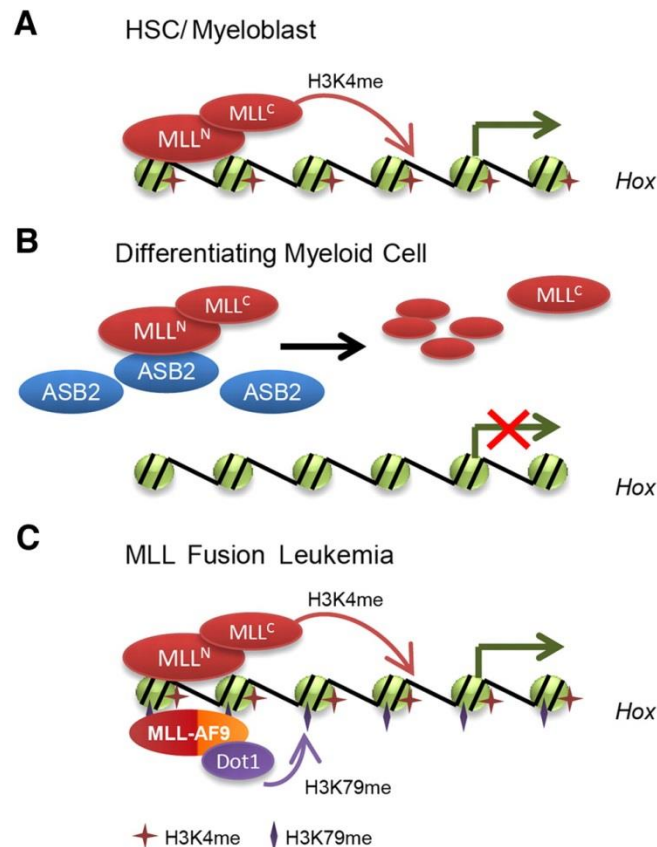
Figure 2.22. Model for the regulation of MLL degradation and *HOX* gene expression by ASB2 during normal and malignant hematopoiesis. (A) In hematopoietic stem cell and progenitor compartments, MLL maintains *HOX* gene expression. (B) During differentiation, ASB2 expression is induced, which leads to MLL ubiquitination and degradation. Less MLL is localized to *HOX* gene loci, resulting in inhibition of *HOX* gene transcription. (C) In MLL fusion protein-mediated leukemias, MLL fusion proteins lack the PHD/Bromo region that interacts with ASB2 and escape from ASB2-mediated degradation. Moreover, low expression of ASB2 leads to stabilization of wild-type MLL, which may also contribute to the constitutive activation of *HOX* genes and leukemogenesis.

Our experiments show that the interaction between ASB2 and MLL is mediated through the PHD/Bromo region (Figure 2.6 and 2.8). This region contains four PHD fingers and a bromodomain between PHD3 and PHD4. The third PHD finger of MLL has been reported to bind di/tri-methylated histone H3K4 as well as the cyclophilin Cyp33 simultaneously, switching MLL from an activator to a repressor (Chang et al., 2010b; Hom et al., 2010; Park et al., 2010; Wang et al., 2010). and loss of MLL PHD3 is required for MLL-ENL mediated transformation (Chen et al., 2008). The first and fourth PHD finger function together with the phenylalanine/tyrosine-rich (FYRN) domain to mediate the intramolecular interaction between MLL^N and MLL^C (Yokoyama et al., 2011). In our experiments, inclusion of the bromodomain is sufficient to mediate the interaction with ASB2, while the addition of PHD4 enhances binding (Figure 2.6). A number of bromodomains have been demonstrated to mediate protein-protein interaction through recognizing acetylated lysines (Sanchez and Zhou, 2009). Interestingly, although exhibiting a typical bromodomain fold, the bromodomain of MLL does not recognize acetylated histones, perhaps due to loss of a conserved asparagine found in acetyl-lysine binding bromodomains (Wang et al., 2010). Moreover, a sequence analysis of various human bromodomains defines the MLL bromodomain as an outlier, because of significant sequence variation (Sanchez et al., 2000; Sanchez and Zhou, 2009). Further work is needed to determine if the interaction between ASB2 and the MLL bromodomain is mediated through acetylation.

A recent study reports that the wild-type allele of *MLL* is required for MLL-AF9

mediated leukemogenesis. Wild-type Mll is recruited to the *Hoxa9* locus in MLL-AF9

transformed cells, and plays a crucial role for the maximal methylation of both H3K79

and H3K4, as well as the expression of *Hoxa9*. Knockdown or excision of wild-type

Mll in MLL-AF9 transduced cells inhibits cell proliferation, reduces colony formation

and inhibits the development of leukemia *in vivo* (Thiel et al., 2010). Based on these

observations, the authors propose that MLL fusion transformed leukemia cells may be

more sensitive to inhibition of wild-type MLL, as they have a reduced cellular level of

MLL, and wild-type MLL is a potential target for leukemia therapy (Thiel et al., 2010).

This is consistent with our findings, as ASB2 expression in MLL-AF9 transformed

cells phenocopies knockdown or excision of wild-type MLL, including decreased

H3K4 methylation, decreased *HoxA9* expression and reduced cell proliferation and

colony formation (Figure 2.18). Notably, ASB2 expression is rapidly induced by

ATRA (Guibal et al., 2002; Kohroki et al., 2001). This suggests that ATRA treatment,

which is already in use for treatment of promyelocytic leukemia, would have activity

against MLL associated leukemias. Indeed, two leukemia cell lines harboring the

MLL-AF9 fusion protein, THP-1 and MOLM-14 cells, have been shown to be sensitive

to ATRA induced differentiation (Matsushita et al., 2008).

In summary, our work has established a novel ubiquitination pathway for MLL degradation during hematopoietic differentiation by ECS^{ASB2}. This represents a mechanism for *HOX* gene down-regulation during blood cell maturation by degradation of MLL and provides a rationale for targeting the wild-type MLL protein for therapy of MLL associated leukemia.

Chapter3   A Subset of Mixed Lineage Leukemia Proteins Has PHD Domain-Mediated E3 Ligase Activity[23]

3.1   Abstract

The mixed lineage leukemia protein MLL1 (MLL) contains four highly conserved plant homeodomain (PHD) fingers, which are invariably deleted in oncogenic MLL1 fusion proteins in human leukemia. Here we show that the second PHD finger (PHD2) of MLL1 is an E3 ubiquitin ligase in the presence of the E2 conjugating enzyme CDC34. This activity is conserved in the second PHD finger of MLL4, the closest homolog to MLL1 but not in MLL2 or MLL3. Mutation of PHD2 leads to MLL1 stabilization, as well as increased transactivation ability and MLL1 recruitment to the target gene loci, suggesting that PHD2 negatively regulates MLL1 activity

3.2   Introduction

---

[2] This work has been published as Wang, J., Muntean, A.G., Wu, L., and Hess, J.L. (2012). A subset of mixed lineage leukemia proteins has PHD domain-mediated E3 ligase activity. J Biol Chem.

[3] In this chapter, the name MLL1 will be used for MLL for clarity.

The Mixed Lineage Leukemia protein MLL1 is a histone H3 lysine 4 (H3K4)

methyltransferase that positively regulates gene expression and plays a key role in

embryonic development and normal hematopoiesis (Hess et al., 1997). Chromosomal

translocations involving *MLL1* are associated with a variety of hematological

malignancies, accounting for up to 80% of infant leukemia and approximately 5%-10%

of adult leukemia overall (Aplan, 2006; Hess, 2004a). In most chromosomal

translocations, the N terminus of MLL1 is fused to the C terminus of a fusion partner,

thus forming a chimeric fusion protein with abnormal transactivation ability (Hess,

2004a). MLL1 contains multiple domains including a region of four PHD fingers on its

N terminus. Notably, the breakpoint cluster region within *MLL1* occurs before or

within the first PHD finger such that the PHD fingers are invariably lost in the resulting

MLL1 fusion proteins. Insertion of the PHD fingers into MLL1 fusion proteins

abolishes their transformation ability (Chen et al., 2008; Muntean et al., 2008).

In humans, the MLL protein family contains five members (MLL1 (MLL), MLL2 (chr.

12q13.12, previously referred to as MLL4), MLL3, MLL4 (chr. 19q13.1, previously

referred to as MLL2), and MLL5). MLL1 through MLL4 are homologous to yeast

methyltransferase Set1 (MLL5 is considered to be the homolog of the yeast protein

Set3/Set4) and share greater structural similarity, including multiple PHD fingers on

their N terminus. Although all the MLL family members have histone

methyltransferase activity, they regulate different biological and developmental

processes (Lee et al., 2009; Mohan et al., 2011; Shilatifard, 2012; Smith et al., 2011; Takahashi et al., 2011). Further, mutations and rearrangements of other MLL family member proteins have also been associated with malignancies and other disorders (Huntsman et al., 1999; Ng et al., 2010; Ruault et al., 2002; Saigo et al., 2008). Despite their biomedical importance, the mechanisms by which these proteins regulate transcription and their pathogenesis are not fully understood.

PHD fingers are zinc-binding motifs that exist in a wide range of proteins. Typical PHD fingers contain a Cys4-His-Cys3 consensus that coordinates two Zn^{2+} ions in a cross-brace topology (Bienz, 2006). The function of PHD fingers has been under intensive study due to their close association with various diseases. These studies have revealed that this domain has multiple functions, including recognition of different histone marks, DNA binding and mediating protein-protein interactions (Baker et al., 2008; Li and Li, 2012). Recent studies also discovered that PHD fingers are involved in ubiquitination, a post-translational modification implicated in proteasomal degradation as well as epigenetic regulation of transcription. Several PHD fingers have been found to have E3 ligase activity, including the viral proteins MIR1 and MIR2, as well as the eukaryotic proteins Msc1, MEKK1 and AIRE (Coscoy et al., 2001; Dul and Walworth, 2007; Lu et al., 2002; Uchida et al., 2004). Further, the PHD finger of KAP1 has been shown to have SUMO ligase activity, which mediates sumoylation, a process similar to ubiquitination (Ivanov et al., 2007).

In this study, we discovered that the second PHD finger of MLL1 has intrinsic E3 ligase activity that is conserved in at least one other closely related MLL family member, MLL4. This activity requires the E2 conjugating enzyme CDC34, which interacts with MLL1 in vitro and in vivo. Mutation of PHD2 in MLL1 is associated with increased transactivation ability and MLL1 recruitment to target gene promoters.

3.3   Experimental Procedures

3.3.1   Protein purification

cDNAs encoding the MLL PHD fingers were synthesized (GenScript) and cloned into the pMCSG9 vector which contains HIS-MBP tag. Expression plasmids were transformed into the AI bacteria strain and screened at the University of Michigan High-Throughput Protein Lab. Bacteria were grown at 37℃ in TB medium at 250 rpms with 50ug/ml spectinomycin and 100ug/ml ampicillin to an OD600 of approximately 0.7, and protein expression was induced with 200μM IPTG overnight at 20℃. Bacteria were lysed by sonication in sodium phosphate buffer (300 mM NaCl, 50 mM sodium phosphate buffer, 10 mM imidazole, pH 8.0, supplemented with 0.4 mg/ml lysozyme). Proteins were batch purified using Ni-NTA agarose (Qiagen) according to the manufacturer's instructions and visualized by Coomassie staining. The sequences of the PHD fingers purified are as below:

Protein	PHD finger	Position (Starting-ending amino acid)
MLL1	PHD1	I1393-F1481
	PHD2	K1480-G1544
	PHD3	H1551-P1629
	PHD4	D1914-L1980
MLL2	PHD1	E225-R280
	PHD2	E272-S330
	PHD3	T1359-S1437
	PHD4	G1421-S1484
	PHD5	A1502-V1560
	PHD6	T5090-G5140
MLL3	PHD5	S990-S1064
MLL4	PHD6	Q1239-N1313

3.3.2 In vitro ubiquitination assay

Purified HIS-MBP tagged PHD fingers (2ug) were incubated with 180ng UBE1 (E1), 560ng E2, 1ug HA-ubiquitin and 2mM ATP in 30ul reaction buffer (50mM Tris-HCl pH7.5, 5mM $MgCl_2$, 2mM NaF, 0.5mM DTT) at room temperature for 3 hours. Reactions were terminated by adding 30ul 2X SDS sample buffer (Invitrogen) and boiling at 95℃ for 5-10min. Western blotting was then performed to detect ubiquitin conjugates using HA antibody (Abcam). For the ubiquitination assays using histones as substrates, 1ug of recombinant histone H2A, H2B, H3 or H4 (Boston Biochem) were added to the reactions respectively. Human recombinant UBE1, CDC34, RAD6, UBCH5c, and HA tagged ubiquitin were purchased from Boston Biochem. Human recombinant UBCH13/Uev1a was kindly provided by Dr. Yali Dou (University of Michgan).

3.3.3 In vitro binding assay

Equal amounts of purified HIS-MBP/ HIS-MBP–PHD2 and recombinant CDC34 (1ug) were incubated in 500ul binding buffer (50mM Tris-HCl pH7.9, 300mM KCl, 0.05% NP40, 0.1% BSA) for 2 hours at 4℃. 20ul Amylose resin (New England Biolabs) was then added to the reaction and incubated for 1 hour at 4℃. Bound material was washed with wash buffer (50mM Tris-HCl pH 7.9, 300 mM KCl, 0.3% NP40) 3 times and eluted from beads by boiling in SDS sample buffer. Eluted material was detected by western blotting and Coomassie staining.

3.3.4 Immunoprecipitation and Western blotting

Immunoprecipitation and western blotting were performed as previously described (Chen et al., 2008). 293 cells were lysed with BC-300 lysis buffer (20mM Tris-HCl [pH 7.4], 10% glycerol, 300mM KCl, 0.1% NP-40) and incubated with protein-G agarose beads (Roche) and MLL1 antibody or normal IgG overnight at 4℃. Bound material was washed 3 times with BC-300 buffer. Proteins were eluted by boiling in SDS loading buffer, resolved by SDS-PAGE, and detected by western blotting. Primary antibodies against MLL1 and CDC34 were obtained from Bethyl and Upstate, and Boston Biochem respectively.

3.3.5 Vector construction and Dual Luciferase Assay

pCXN2-Flag-MLL1 have been previously described (Milne et al., 2002).
pCXN2-Flag-MLL1C1509A and pCXN2-Flag-MLL1ΔP2 are cloned from

pCXN2-Flag-MLL1 using the QuikChange Site-Directed Mutagenesis Kit (Agilent).

293 cells were transfected with MLL1 or MLL1 mutants, Renilla luciferase reporter

(serves as an internal control), and *Hoxa9*-LUC or *Myc*-LUC reporter with FuGene 6

according to the manufacturer's instructions. The amount of MLL1 and MLL1 mutant

constructs transfected was adjusted to yield equal protein levels (MLL1:

MLL1C1509A: MLL1ΔP2 =1:1.2:1). Vector DNA was added to ensure that the total

amount of DNA transfected was equal among all samples. Cells were serum starved in

0.5% FBS in OPTI-MEM media for 48 hours. Luciferase assays were performed using

the Dual Luciferase Assay Kit (Promega) according to the manufacturer's instructions.

Emission was detected using a Monolight 3010 Luminometer (BD Biosciences).

3.3.6 Chromatin immunoprecipitation (ChIP)

ChIP assays were performed as previously described (Milne et al., 2005). Antibody

against Flag tag was obtained from Sigma-Aldrich. qPCR was performed on the

precipitated DNAs with TaqMan primers and probes from Applied Biosystems.

Binding was quantitated as follows: $\Delta CT = CT(input) - CT(Chromatin\ IP)$, % total =

$2^{\Delta CT}$. Primer and probe sequences are described previously (Wang et al., 2011b).

3.4 Results

3.4.1 The second PHD finger of MLL1 has E3 ligase activity.

The MLL1 protein contains four PHD fingers (PHD1-4) with conserved Zn^{2+}

coordinating cysteine and histidine amino acids and variable intervening sequences

(Figure 3.1A).

Figure 3.1. Sequence alignment of the four PHD fingers of MLL1
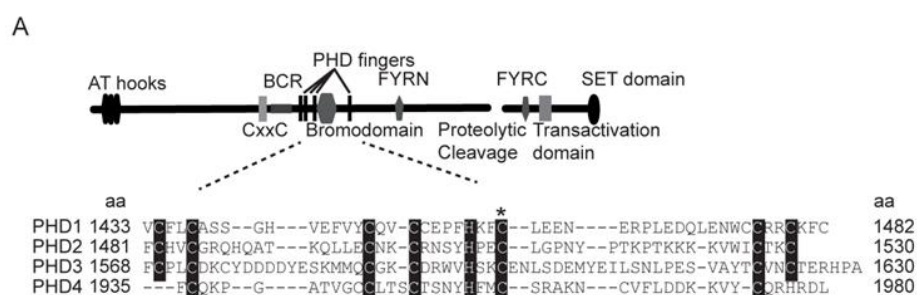
Figure 3.1. Sequence alignment of the four PHD fingers of MLL1. (A) (Upper) Schematic diagram of the structure of wild-type MLL1. (Lower) Sequence alignment of the four PHD fingers of MLL1. Black background indicates conserved cysteines and histidines. Grey background indicates other conserved amino acids. "*" indicates the cysteine in PHD2 that is mutated in this study.

Recent publications have shown that PHD fingers can function as an H3K4me3 "reader"

(Li et al., 2006). While the third PHD finger of MLL1 shares this high affinity

recognition specifically for H3K4me3, this activity is notably absent from PHD1, 2 and

4 of MLL1 (Chang et al., 2010b). As several PHD fingers function as E3 ligases, we

investigated whether any of the MLL1 PHD fingers have this activity. To this end, we

bacterially expressed and purified HIS-MBP-tagged individual PHD fingers of MLL1

(Figure 3.2A). In vitro ubiquitination assays were performed in the presence of E1

activating enzyme UBE1, E2 conjugating enzyme CDC34, HA-tagged Ubiquitin and

ATP. High molecular weight ubiquitin conjugates were observed only in the presence

86

of PHD2 (Figure 3.2B). This activity is ATP-dependent, as no ubiquitination was observed without ATP (Figure 3.2B, lane 6). To confirm this enzymatic activity, the sixth cysteine of PHD2 was mutated to alanine to disrupt the cross-brace structure of the domain. No ubiquitination activity was observed with the mutant, indicating that the observed E3 ubiquitin ligase activity is intrinsic to PHD2 (Figure 3.2C).

Figure 3.2 The second PHD finger of MLL1 has E3 ligase activity in vitro
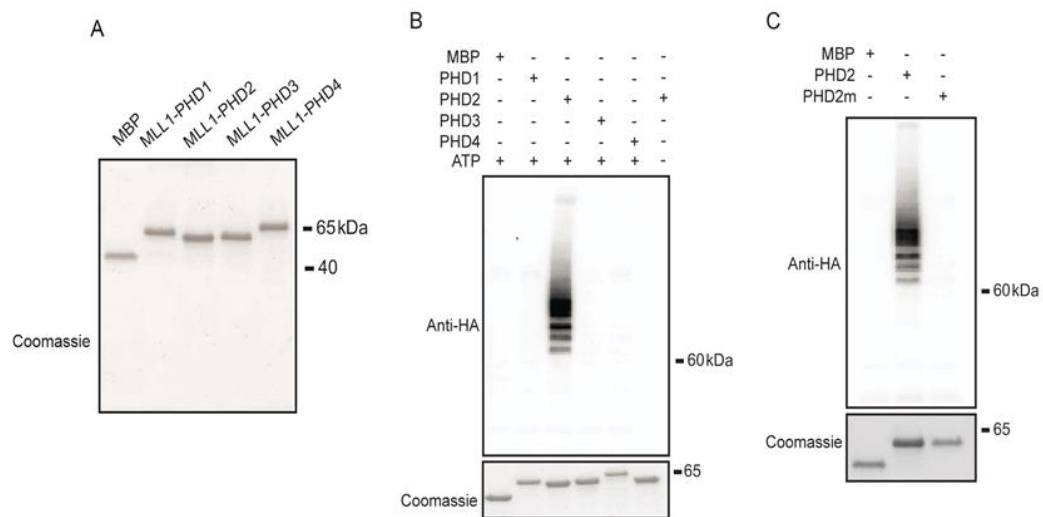
Figure 3.2. The second PHD finger of M LL1 has E3 ligase activity in vitro. (A) Coomassie staining of purified PHD fingers. (B) In vitro ubiquitination assays with the four PHD fingers of MLL1 followed by western blotting for HA-tagged ubiquitin shows that the second PHD finger promotes formation of poly-ubiquitin conjugates in the presence of ATP. Commassie staining indicates equal amount of PHD fingers used in the assay. (C) In vitro ubiquitination assays were performed with wild type or mutant PHD2 followed by HA western blotting showing mutation of PHD2 destroys ubiquitinating activity.

To test if PHD2 can ubiquitinate other substrates, in vitro ubiquitination assays were performed with free histones, including H2A, H2B, H3 and H4. PHD2 showed specific activity for histone H3 and H4, but not H2A or H2B (Figure 3.3A). To identify the sites on histone H3 and H4 that are ubiquitnated, mass spectrometry analysis was performed on these histones from the ubiquitination assays. Three sites were identified, including H3K79, H3K122 and H4K31.

Figure 3.3 PHD2 ubiquitinates histones in vitro.
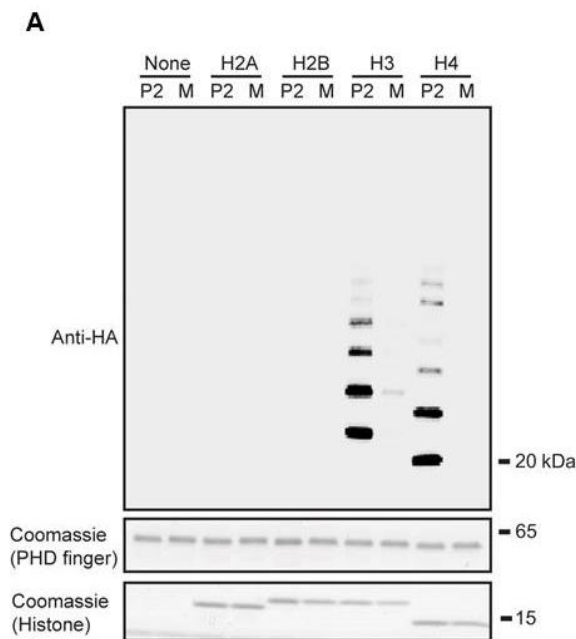
Figure 3.3. PHD2 ubiquitinates histones in vitro. (A) In vitro ubiquitination assays with PHD2 using histones as substrates. Western blotting shows that PHD2 leads to ubiquitination of histone H3 and H4. Commassie staining indicates equal amount of materials used in the assays. M = MBP, P2 = PHD2.

3.4.2   PHD E3 ligase activity is conserved in MLL4

The MLL protein family consists of five members (MLL1-MLL5), with MLL4 showing the greatest structural similarity to MLL1. To test if the E3 ligase activity is

conserved among the MLL family members, a sequence alignment of all predicted

PHD fingers in MLL1 through MLL4 was performed, which revealed the highest

homology between the second PHD finger (PHD2) of MLL4 and PHD2 of MLL1. The

fourth PHD finger (PHD4) of MLL2 and the fifth PHD finger (PHD5) of MLL3 are the

next most similar to PHD2 of MLL1 (Figure 3.4A and B). Further, the alignment shows

that each PHD finger in MLL4 clusters together with the corresponding PHD fingers in

MLL1, while five of the six PHD fingers in MLL2 cluster together with the

corresponding PHD fingers in MLL3 (Figure 3.4A). Notably, previous studies have

shown that MLL1 and MLL4 are functionally distinct from MLL2 and MLL3, which

form unique macromolecular protein complexes and target distinct genes (Figure 3.4C)

(Hughes et al., 2004; Mo et al., 2006; Shilatifard, 2012).

Figure 3.4 The second PHD finger (PHD2) of MLL4 shows highest similarity to PHD2 of MLL1
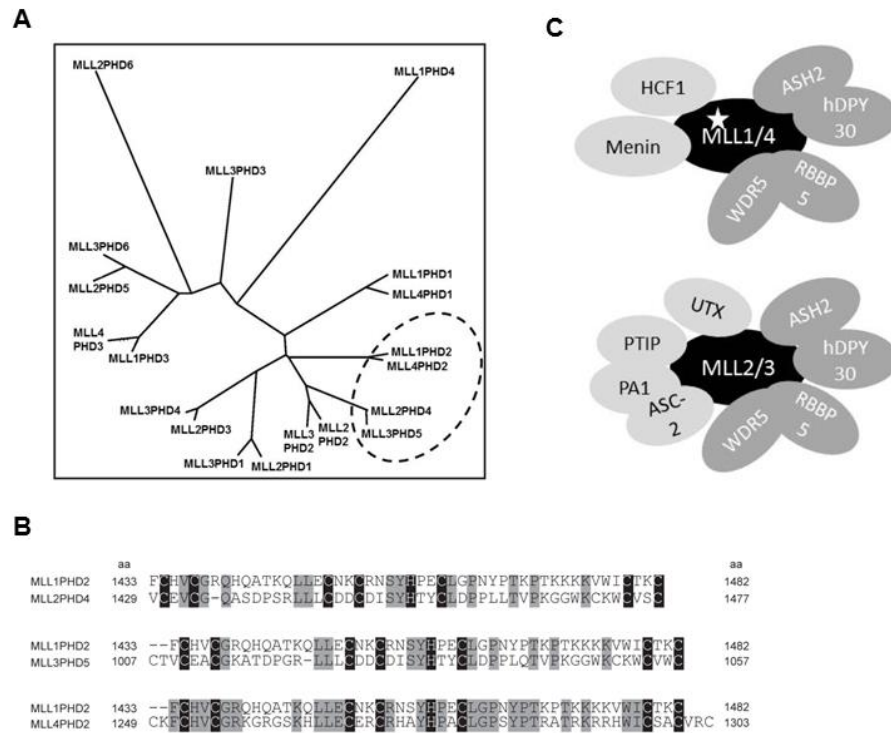
Figure 3.4. The second PHD finger (PHD2) of MLL4 shows highest similarity to PHD2 of MLL1. (A) Phylogenetic tree using Clustal W (Thompson et al., 1994) with neighbor joining method based on sequence alignment of the PHD fingers of MLL1 through MLL4. Dashed circle indicates the PHD fingers that are similar to PHD2 of MLL1. (B) Sequence alignment of PHD4 of MLL2, PHD5 of MLL3 and PHD2 of MLL4 with PHD2 of MLL1 respectively. Black background indicates conserved cysteines and histidines. Grey background indicates other conserved amino acids. (C) Schematic diagram showing the structure of MLL methyltransferase complexes. Black indicates MLL proteins. Dark grey indicates common subunits. Light grey indicates unique subunits. Star indicates E3 ligase activity.

To test for conservation of E3 ligase activity, HIS-MBP-tagged PHD2 of MLL4, PHD5 of MLL3 and all PHD fingers of MLL2 were bacterially expressed, purified and tested in in vitro ubiquitination assays (Figure 3.5A-B). These showed that only the most homologous PHD finger (PHD2 of MLL4) has conserved E3 ligase activity, while no

activity was observed in other MLL family member PHD fingers under the same

conditions (Figure 3.5C-D). These data demonstrate a novel E3 ubiquitin ligase activity

associated with PHD2 of MLL1 that is evolutionarily conserved in PHD2 of MLL4.

Figure 3.5 The E3 ligase activity is conserved in PHD2 of MLL4
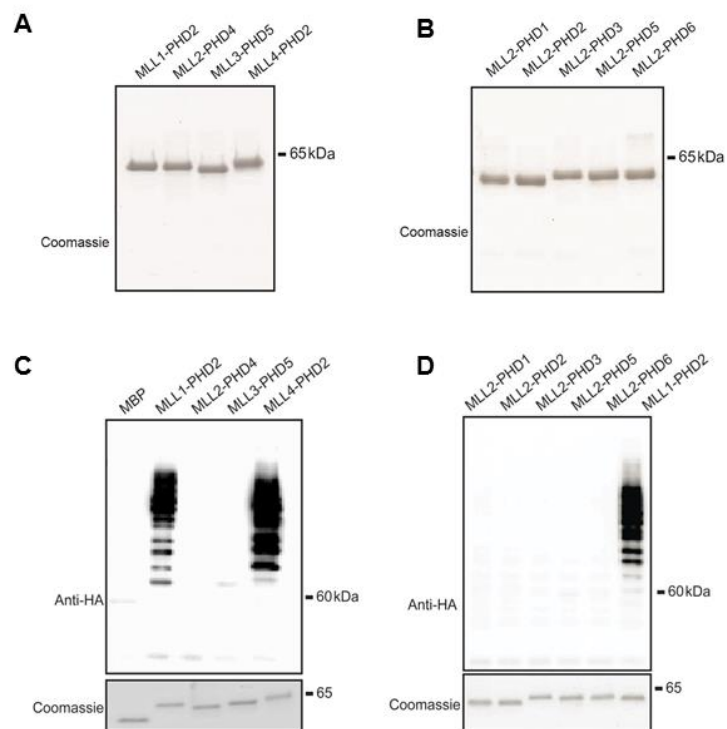
Figure 3.5. The E3 ligase activity is conserved in PHD2 of MLL4. (A-B) Commassie staining

of bacteria expressed and purified PHD fingers. (C) In vitro ubiquitination assays with PHD2 of

MLL1, PHD4 of MLL2, PHD5 of MLL3 and PHD2 of MLL4. Western blotting shows that the

PHD finger of MLL1 and MLL4 contain ubiquitinating activity. (D) In vitro ubiquitination

assays with PHD2 of MLL1 and the PHD fingers of MLL2. None of the MLL2 PHD fingers

show ubiquitinating activity.

3.4.3   CDC34 facilitates the E3 ligase activity of PHD2 and interacts with MLL1

The interaction between E2 conjugating enzymes and E3 ligases plays a central role in ubiquitination as it provides specificity to the reaction. We sought to identify the E2 conjugating enzyme required for PHD2 activity by performing ubiquitination assays with a panel of E2 enzymes, including CDC34, RAD6, UBCH5c and UBCH13. Ubiquitin conjugates were only observed in the presence of CDC34, demonstrating specificity for CDC34 in PHD2-mediated ubiquitination (Figure 3.6A). To test for a direct interaction, in vitro binding assays were performed using purified HIS-MBP-PHD2 and recombinant CDC34. Immunoprecipitation of PHD2 but not the tag control leads to co-precipitation of CDC34 (Figure 3.6B). We then tested for an endogenous interaction between full length MLL1 and CDC34 in 293 cells. Endogenous MLL1 was immunoprecipitated using anti-MLL1 antibody and western blotting using anti-CDC34 antibody showed that endogenous CDC34 co-precipitates with full length MLL1 (Figure 3.6C).

Figure 3.6 CDC34 facilitates the E3 ligase activity of PHD2 and interacts with MLL1
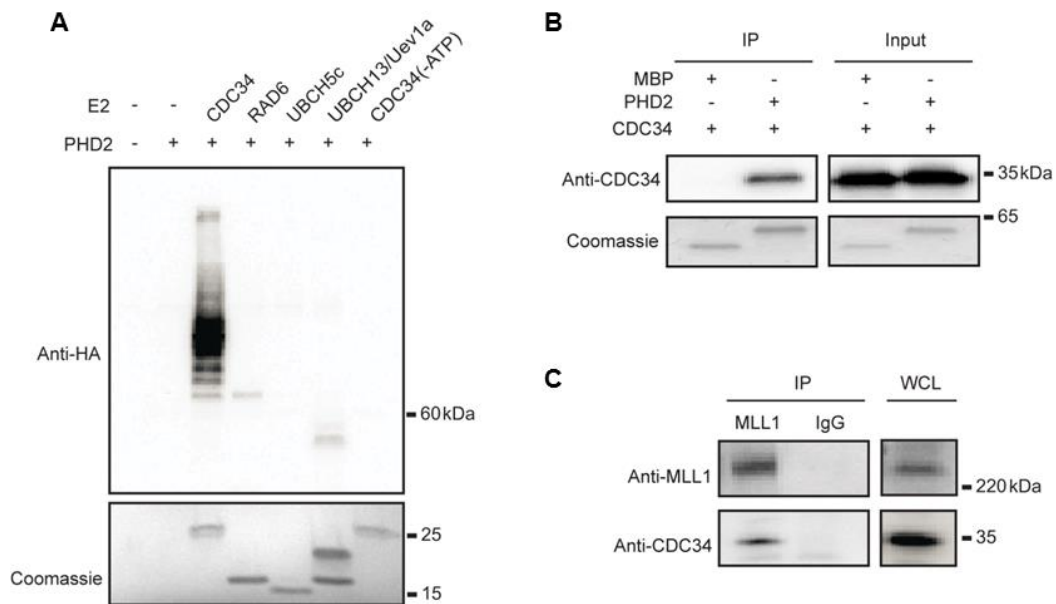
Figure 3.6. CDC34 facilitates the E3 ligase activity of PHD2 and interacts with MLL1.

(A) In vitro ubiquitination assays of PHD2 with different E2 enzymes showing a specific

requirement for CDC34. Commassie staining indicates the amount of E2 enzymes for each

assay. The two bands in lane 6 represent UBCH13 (lower) and Uev2a (upper). (B) In vitro

binding assays with purified PHD2 and recombinant CDC34 indicates that CDC34

co-precipitates with PHD2 but not the HIS-MBP tag control. (C) Immunoprecipitation of

endogenous MLL1 co-precipitates endogenous CDC34.

3.4.4 Transcriptional activity of MLL1 is augmented by mutation of PHD2

MLL1 is a positive regulator of gene expression. To test the effect of PHD2 on MLL1

transactivation ability, two mutants were made in the context of full length MLL1.

MLL1C1509A replaces the sixth cysteine of PHD2 with alanine, which resembles the

PHD2 mutant used in the in vitro ubiquitination assay; and MLL1ΔP2 encodes full

length MLL1 with a deletion of the entire PHD2. Wild type MLL1 or MLL1 mutants

were expressed in 293 cells with a luciferase reporter driven by the promoter of *Hoxa9*

(*Hoxa9*-LUC), an important target gene of MLL. Dual Luciferase Assays were

performed to test the activation of the reporter, which reflects the transactivation ability

of MLL1 (Figure 3.7A). Both MLL1C1509A and MLL1ΔP2 showed higher

transactivation compared to wild type MLL1 (Figure 3.7B and C). Similar results were

observed when using a luciferase reporter driven by the *Myc* promoter, another target of

MLL1 (Figure 3.7D).

Figure 3.7 Mutation of PHD2 in MLL1 leads to increased MLL1 transactivation
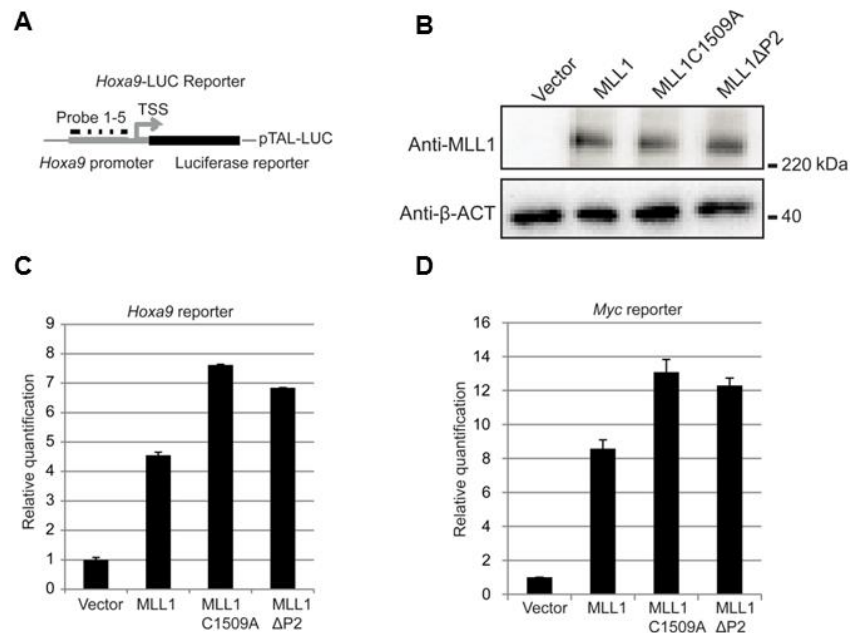
ability

Figure 3.7. Mutation of PHD2 in MLL1 leads to increased MLL1 transactivation ability. (A)

Schematic diagram of the *Hoxa9*-LUC reporter. Black bars indicate the position of the probes

used in ChIP-qPCR experiments. (B) Expression of wild type MLL1 and MLL1 mutants

in 293 cells. β-ACTIN blot shows equal loading. (C) Dual Luciferase Assays were performed in 293 cells with *Hoxa9*-LUC reporter showing that mutations of PHD2 in full length MLL1 lead to higher activation of *Hoxa9* reporter compared to wild type MLL1. All changes are normalized to lane 1, which contains *Hoxa9*-LUC and an empty expression vector. Error bars indicate SD. (D) Dual luciferase assay was performed in 293 cells with *Myc*-LUC reporter.

Further, chromatin immunoprecipitation (ChIP)–qPCR experiments were performed to test the recruitment of MLL1 or MLL1 mutants to the promoter region of the *Hoxa9* reporter. 293 cells were transfected with MLL1 or MLL1 mutant expression vectors and *Hoxa9*-LUC reporter at the same ratio as used in the Dual Luciferase Assays. This yielded comparable levels of MLL1 and MLL1 mutant expression (data not shown). ChIP-qPCR revealed a more robust recruitment of both MLL1 mutants compared to wild type MLL1 (Figure 3.8).

Figure 3.8 Mutation of PHD2 in MLL1 leads to increased MLL1 recruitment to target gene loci
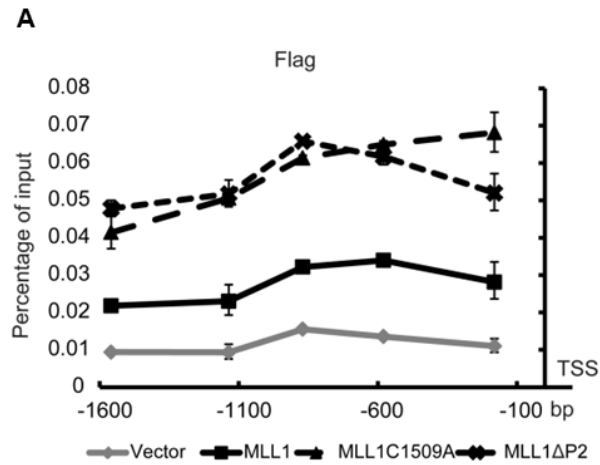
Figure 3.8. Mutation of PHD2 in MLL1 leads to increased MLL1 recruitment to target gene loci. (A) ChIP-qPCR assays were performed in 293 cells transfected with *Hoxa9*-LUC, MLL1 or MLL1 mutants. X axis labels the distance of the probes relative to the transcription start site. Grey line: vector control. Solid line with squares: wild-type MLL1. Long dashed line with triangles: MLL1C1509A. Short dashed line with crossings: MLL1ΔP2.

To test the effect of PHD2 on MLL1 stability, the turnover of MLL1 and MLL1C1509A was monitored following treatment with cycloheximide. In these experiments, MLL1C1509A displayed prolonged degradation and greater stability than wild type MLL1 (Figure 3.9).

Figure 3.9 Mutation of PHD2 in MLL1 leads to increased MLL1 stability
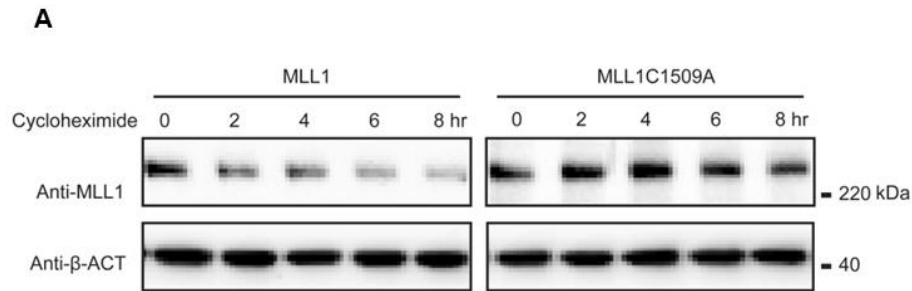
Figure 3.9. Mutation of PHD2 in MLL1 leads to increased MLL1 stability. (A) 293 cells

transfected with MLL1 or MLL1C1509A were treated with cycloheximide for different time

periods as indicated to block translation. The turnover of MLL1 or MLL1C1509A was

monitored by western blotting using whole cell lysate. MLL1C1509A shows prolonged protein

degradation compared to wild type MLL1.

3.5 Discussion

MLL1 contains four PHD fingers, which are invariably deleted in leukemogenic MLL1

fusion proteins. Moreover, insertion of the PHD fingers into MLL1 fusion proteins

abolishes their transformation ability (Chen et al., 2008; Muntean et al., 2008).

Sequence variation between the MLL1 PHD fingers suggests that they have different

functions. Indeed, the third PHD finger of MLL1 has been reported to bind

di/tri-methylated histone H3K4 and the cyclophilin Cyp33 simultaneously while the

first and fourth PHD fingers are involved in mediating intramolecular interactions

between the N-terminal and C-terminal fragments of MLL1 (Chang et al., 2010b; Park

et al., 2010; Yokoyama et al., 2011). Here we report that MLL1 contains a second enzymatic activity in addition to the methyltransferase activity intrinsic to the SET domain. The second PHD finger functions as an E3 ligase in conjugation with the E2 enzyme CDC34. Mutation of PHD2 leads to increased transactivation ability of MLL1 and its recruitment to target genes, which is likely due to increased protein stability (Figure 3.8-3.9). However, PHD2 may also ubiquitinate other substrates that contribute to its inhibitory effect. Interestingly, we observed that PHD2 ubiquitinates histone H3 and H4 in vitro, and three ubiquitination sites were identified, including H3K79, H3K122 and H4K31. Notably, all these sites have been found to be ubiquitinated in vivo (Kim et al., 2011; Wagner et al., 2011), while the E3 ligases that mediate these modifications and their biological functions remain to be determined. Further, a sequence analysis was reported that characterizes the surrounding sequences of each lysine residue in human histones H2A, H2B, H3 and H4. In this analysis, clustering was performed with each lysine and its twelve flanking residues. H3K79, H3K122 and H4K31 cluster together (Basu et al., 2009), indicating that these sites share sequence similarities that are preferentially recognized and ubiquitinated by PHD2. Further experiments are required to determine if histone H3 and H4 are *bona fide* substrates in vivo. Three E3 ligases have been shown to ubiquitinate MLL1, including SCF[Skp2], APC[Cdc20] and ECS[ASB2] (Liu et al., 2007a; Wang et al., 2011b). Our findings indicate that the second PHD finger also plays a role in MLL1 ubiquitination. CDC34 is an

important regulator of cell cycle progression (King et al., 1996), raising the possibility that PHD2 plays a role in regulating MLL1 degradation during cell cycle.

The MLL protein family contains five members, with MLL1 and MLL4 being functionally distinct from MLL2 and MLL3 (Lee et al., 2009; Mohan et al., 2011; Shilatifard, 2012; Smith et al., 2011; Takahashi et al., 2011). This is reflected from the sequence alignment of the MLL PHD fingers, as the PHD fingers of MLL1 and MLL4 cluster together while the PHD fingers of MLL2 and MLL3 cluster together (Figure 3.4 A). Further, the PHD2 of MLL4 also has E3 ligase activity (Figure 3.5). Like *MLL1*, *MLL4* also undergoes cancer associated gene rearrangements. In one study, intronic insertions of hepatitis B virus (HBV) genomic sequences or chromosomal translocations of *MLL4* were found in 26 out of 42 human hepatocellular carcinoma (HCC) cases suggesting MLL4 overexpression or alteration plays a key role to the pathogenesis of HCC (Saigo et al., 2008). Given the central role of MLL family members and their frequent rearrangements and mutations in both leukemia and solid tumors, additional studies appear warranted to define the role of PHD finger mediated E3 ligase activity in normal gene regulation and carcinogenesis.

Chapter4   The Functional Interactions between HOXA9, C/EBPA and PU.1

4.1   Abstract

HOXA9 is a homedomain-containing transcription factor that plays a crucial rule in embryonic development and normal hematopoiesis. Constitutive activation of HOXA9 is associated with a variety of hematopoietic malignancies, including MLL fusion mediated leukemia, and plays a key role in their pathogenesis. Using motif analysis of genome-wide HOXA9 binding and mass spectrometry, we previously found that C/EBPA and PU.1 are potential HOXA9 cofactors. In this study, we characterized the co-binding between Hoxa9, C/ebpα and Pu.1 on the genome and its association with Hoxa9 regulated gene expression using next-generation DNA sequencing coupled with chromatin immunoprecipitation (ChIP-seq) and whole transcriptome sequencing (RNA-seq). Different ChIP-seq and RNA-seq analysis methods were tested and compared to optimize the analysis pipeline. Based on these data, we found that genome-wide Hoxa9 binding overlaps with C/ebpα and Pu.1 binding. Co-bound regions are enriched with the enhancer mark H3K4me1, and are enriched for hematopoiesis related pathways. Further, co-binding of C/ebpα and Pu.1 associates with Hoxa9 regulated genes but does not predict activation or repression activity of Hoxa9.

4.2 Introduction

HOXA9 is a member of the *HOX* transcription factor family and is one of the best-studied MLL target genes. It maintains the self-renewal ability of hematopoietic stem cells and progenitors and functions as a master regulator of hematopoiesis (Argiropoulos and Humphries, 2007). Constitutive activation of Hoxa9 is associated with a variety of hematopoietic malignancies, including MLL translocation-mediated leukemia, NPM1 mutation associated AML and CDX2/CDX4 mediated leukemia, and plays a key role in their pathogenesis (Bansal et al., 2006; Mullighan et al., 2007; Rawat et al., 2008). In fact, HOXA9 up-regulation is present in approximately 50% of all AML cases and it was identified as the most highly correlated gene for poor prognosis in AML (Golub et al., 1999; Lawrence et al., 1999). Further, HOXA9-NUP98 fusion resulting from chromosome translocation leads to aberrant gene expression profiling and is associated with AML, trilineage MDS and chronic myelomonocytic leukemia (Borrow et al., 1996; Hatano et al., 1999; Inaba et al., 1996; Kroon et al., 2001; Nakamura et al., 1996; Wong et al., 1999).

Although being established for a long time as the key regulator of hematopoiesis and important mediator of leukemia, the mechanisms of HOXA9 regulated gene expression remains poorly understood. Specifically, how does HOXA9 recognize its target genes, and how is the gene activation or repression activity determined? To solve these questions, the concept of "Hoxasome" was brought up, meaning that

transcription factors and chromatin modifiers (HOXA9 cofactors) form a complex

with HOXA9 and bind to DNA in a cooperative pattern, thus increasing the binding

specificity, affinity or transactivation ability of HOXA9 (Mann et al., 2009). One

example of HOXA9 cofactor is the non-Hox homeodomain protein MEIS1. MEIS1

can dimerize with HOXA9, or form a triple complex with HOXA9 and PBX proteins

(PBX1-3), another class of HOXA9 cofactors that contain a homeodomain with ~40%

identity to that of MEIS1. These interactions stabilize the HOXA9 DNA binding

complex. Also, addition of MEIS1 binding consensus (TGACAG) and PBX binding

consensus (TGAT) increases the specificity of HOXA9 recognition on target loci

(Shen et al., 1997; Shen et al., 1999).

To better understand the mechanisms of HOXA9 regulated gene expression and to

discover other HOXA9 cofactors, the Hess lab performed ChIP-seq experiment to

define the genome-wide binding pattern of Hoxa9 using a murine cell line

transformed with Hoxa9 and Meis1. This identified ~400 binding peaks for Hoxa9

with more than 50% of them overlapping with Meis1 binding, consistent with the role

of Meis1 as Hoxa9 cofactor. Hoxa9 binding sites show high levels of histone

H3K4me1 and CBP/P300 binding characteristic of enhancers, suggesting a role of

Hoxa9 on the enhancer elements. Interestingly, de novo motif analysis reveals that

Hoxa9 peaks are enriched with binding motifs of several transcription factors and

chromatin modifiers, such as PU.1, C/EBP, RUNX1 and STAT5 (Huang et al., 2012).

Further, mass spectrometry analysis identified that C/ebp and Stat5 interact with

Hoxa9, which was also confirmed by in vitro binding assays and/or

co-immunoprecipitation. These findings support that these proteins are potential

HOXA9 cofactors.

In this study, we focus on the functional interaction between HOXA9, the

CCAAT/enhancer-binding protein alpha (C/EBPA) and the ETS transcription factor

PU.1 (also called SPI1). C/EBPA and PU.1 are both key regulators of hematopoiesis,

especially myeloid lineage development. *PU.1* is expressed in lymphoid-myeloid

progenitor (LMP) cells and their progeny cells, and is required for the generation of

both common lymphoid progenitor (CLP) and granulocyte-macrophage progenitor

(GMP) cells. Knockout of *Pu.1* in mice leads to significant decrease of CLP and GMP

populations, thus absence of B cells and monocytes and greatly reduced neutrophils

(Friedman, 2007; Iwasaki et al., 2005). *C/EBPA* is expressed in HSC, common

myeloid progenitor (CMP) and GMP cells. *C/ebpα* knockout in mice blocks the

transition from CMP to GMP cells. Consistent with this, these mice lack neutrophils

and eosinophils and have reduced monocytes in their fetal and newborn livers

(Friedman, 2007; Zhang et al., 2004).

4.3   Experimental Procedures

4.3.1   Cell culture (Performed by Cailin Collins)

Hoxa9, Meis1 transformed murine cell line was established as described in (Huang et

al., 2012; Muntean et al., 2010). Mouse bone marrow cells were harvested from

5-fluorouracil–treated female 6- to 8-week-old C57BL/6 mice and transformed into

myeloblastic cell lines with a murine stem cell virus –based retrovirus expressing

HA-tagged Hoxa9 fused to a modified estrogen receptor ligand-binding domain

(Hoxa9-ER) and Meis1. Cells were cultured in IMDM with 15% FBS (StemCell

Technologies) with penicillin/streptomycin, IL-3 (10ng/ml, R&D Systems) and

4-hydroxytamoxifen (4-OHT; 100nM, Sigma-Aldrich). Double Hoxa9/Meis1

transductants were selected by FACS (bicistronic Meis1+GFP expression using MigR1

vector; a gift from Dr Warren S. Pear, University of Pennsylvania).

4.3.2   ChIP-sequencing (Performed by Cailin Collins)

Chromatin Immunoprecipitation (ChIP) was performed as described previously

(Huang et al., 2012). Thirty million cells were used per IP and the following antibodies

are used: HA (Abcam, Ab9110), C/ebpα (Santa Cruz, 14AA SC-61), Pu.1 (Santa Cruz,

sc352 T21), H3K4me1 (Abcam, ab8895), H3K27me3 (Millipore, 07-449). For

ChIP-seq analysis, 10 ng of ChIPed DNA was processed for library generation using

the ChIP-seq Library Preparation Kit (Illumina) following the manufacturer's protocol.

The sequencing was performed on Illumina™ HiSeq2000 at University of Michigan

DNA sequencing core and raw ChIP-seq data were processed using the Illumina

software pipeline.

4.3.3   RNA-sequencing (Performed by Cailin Collins)

For RNA sequencing, one million cells in culture with 4-OHT or at 72hrs upon withdrawal of 4-OHT were used per sample. RNA was extracted using Qiagen RNeasy kit following the manufacturer's protocol. Sequencing was performed on Illumina™ HiSeq2000 at University of Michigan DNA sequencing core and raw RNA-seq data were processed using the Illumina software pipeline.

4.3.4 Peak Calling

Sequenced reads were pre-processed to remove contamination of adaptor sequences and then aligned to mouse reference genome (mm9) using BWA (version 0.6.2) (Li and Durbin, 2009). Three peak calling programs were tested and compared, including Model-based Analysis for ChIP-Seq (MACS) (Zhang et al., 2008), CisGenome (Ji et al., 2008) and ChIPseeqer (Giannopoulou and Elemento, 2011). The parameters used are as below (default values are used if not specified):

MACS: format=BED -g mm --nomodel --shiftsize 75 -w –S

ChIPseeqer : -t 15 -fold_t 2 -readlen 52 -fraglen 150

CisGenome: default parameters are used

Distribution of peaks in the promoter (-1kb to +100bp of TSS), exon, intron, and intergenic regions (outside of the defined regions above) was estimated using HOMER (Heinz et al., 2010). Peaks are annotated to their nearest gene using CisGenome (Ji et al., 2008).

Histograms for the distance between Hoxa9 peaks and the transcription start site (TSS)

of their nearest gene are plotted using R program with bin width set to

max(X)-min(X))/(1+log2(length(X)) according to the Sturges' bin rule. Densities of the

peak-TSS distance for Hoxa9, C/ebpα and Pu.1 were estimated using Kernel Density

Estimation using Gaussian distribution as the kernel.

4.3.5   Functional and statistical analysis of ChIP-seq data

Peak overlap was performed using HOMER with the criteria that there is at least 1

basepair overlap between tested peaks (Heinz et al., 2010). The overlap was also

confirmed using the R package ChIPpeakAnno, which generated similar results. The

significance of peak overlap was calculated using hypergeometric test from the R

package ChIPpeakAnno. The background (total number of tests) was set to 159,029 as

an estimation of total transcription factor binding sites obtained from K526 leukemia

cells (Yip et al., 2012). A smaller background of 100,189 was also used, which is the

peak number of H3K4me1, representing the total number of enhancer elements

specifically in our murine cell line. This is a more conservative background as the three

proteins can also bind to non-enhancer regions. Both backgrounds generated consistent

results. The p-values from the former one are presented in the text.

Quantification of histone mark binding signals was calculated using HOMER.

Sequenced tags of H3K4me1 and H3K27me3 around Hoxa9 peaks were first

normalized to 10 million and then normalized that the resulting units are per bp per peak. Heatmap was made with Cluster 3 (de Hoon et al., 2004), and visualized using Treeview (Page, 1996). Histone mark distributions are plotted in excel (Heinz et al., 2010).

De novo motif analysis was performed using MEME-ChIP (Machanick and Bailey, 2011). Hoxa9 binding regions were trimmed to a centered peak core sequence of 200 bp for analysis. A motif is considered to be statistically significantly enriched if the number of sequences in which the motif is found to be present is significantly higher than its expected whole-genome occurrences according to E score (comparable to p-value) <0.05 generated by the MEME and DREME algorithms. The de novo motif discovery module of HOMER and ChIPseeqer were also used which generated similar results to MEME-ChIP. Comprehensive search of known motifs was performed against 865 transcription factor motifs included in Genomatix proprietary Mat Base Matrix Family Library (Version 8.2, January 2010). A motif is considered to be statistically significantly enriched according to standard z-test (z-score >2.81; p-value < 0.005), using the whole genome as background.

Pathway analysis was performed using GREAT software based on Binomial test p-value <0.05 (McLean et al., 2010). Other analysis was performed using R programming.

4.3.6   Differential expression analysis

Sequenced reads were aligned to mouse reference genome (mm9) using Bowtie and

Tophat (version 2.0.3) (Kim et al., 2013). Two programs were tested and compared for

differential expression analysis, Cuffdiff (Trapnell et al., 2010) and EdgeR (Robinson

et al., 2010), which generated similar results. Pathway analysis was performed using

Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005).

4.3.7   Incorporation of ChIP-seq and RNA-seq data

Incorporation of ChIP-seq and RNA-seq data was performed using R programming.

Genes are divided into classes according to binding patterns of Hoxa9, C/ebpα and Pu.1,

and their expression levels are extracted from RNA-seq data.

To compare the percentages of genes with significant expression change between

classes of genes with different binding patterns, and the percentages of down- (up-)

regulated genes in genes, Fisher Exact test was performed for pairwise comparison, and

Chi square test was performed for multiple group comparison. P-value <=0.05 is

considered to be statistically significant.

To compare the magnitude of gene expression change, fold change for each gene was

transformed by log2. Nonparametric Wilcoxon Rank Sum test was performed for

pairwise comparison and Kruskal-Wallis test was performed for multiple group

comparison. P-value <=0.05 is considered to be statistically significant. Parametric

t-test and analysis of variance (ANOVA) were also used for pairwise and multiple

group comparison, respectively, which generate consistent results with nonparametric tests.

4.4 Results

4.4.1 Characterization of genome-wide binding of Hoxa9, C/ebpα and Pu.1

To study the genome-wide binding patterns of Hoxa9, C/ebpα and Pu.1, we used a murine cell line in which bone marrow cells were transduced with HA-tagged Hoxa9 fused to a modified estrogen receptor ligand-binding domain (HA-Hoxa9-ER) and Flag tagged Meis1 expression vectors, and maintained in the presence of 4-hydroxytamoxifen (4-OHT). The presence of HA tag enables efficient immunoprecipitation of Hoxa9, as we were not able to obtain a ChIP-seq qualified antibody against endogenous Hoxa9 (Muntean et al., 2010). Western blotting showed that the expression levels of Hoxa9 and Meis1 in these cells were comparable to MLL-AF9 leukemia cells (Huang et al., 2012). Illumina based ChIP-sequencing libraries were prepared using antibodies against HA, C/ebpα and Pu.1. As previous ChIP-on-chip experiments show that factors associated with enhancers are present on Hoxa9 binding regions, such as p300, CBP and H3K4me1, we also included antibodies against enhancer associated histone marks, including H3K4me1, which marks general enhancers, H3K27ac, which marks active enhancers, and H3K27me3, which marks poised enhancers (Rada-Iglesias et al., 2011). Unfortunately, the H3K27ace ChIP-seq library did not pass the quality control, thus was excluded from current analysis.

ChIP-seq samples were sequenced on Illumina™ HiSeq2000 system. Sequenced reads

were then aligned to the mouse mm9 reference genome. Overall, for each sample we

obtained more than 20 million reads aligned to the reference genome (Table 4.1).

Table 4-1 Summary of the number of sequenced and aligned reads

Sample	Total No. of sequences	No. of sequences aligned	Percentage of alignment
Input	54,768,760	40,932,499	74.7%
HA (Hoxa9)	43,378,206	33,630,123	77.5%
C/ebpα	40,609,786	29,679,136	73.1%
Pu.1	34,363,360	23,702,997	69%
H3K27me3	55,631,183	40,191,253	72.2%
H3K4me1	38,787,174	20,961,523	54%

Three peak calling programs were tested for their performance using Hoxa9 ChIP-seq

data, including Model-based Analysis for ChIP-Seq (MACS) (Zhang et al., 2008),

CisGenome (Ji et al., 2008) and ChIPseeqer (Giannopoulou and Elemento, 2011).

Overall, MACS detected the largest number of peaks, consistent with previous reports

that MACS has high sensitivity, while ChIPseeqer detected the smallest number of

peaks (Figure 4.1B). However, results from the three programs showed significant

overlap (hypergeometric test p-values =0 for pairwise comparisons of MACS *vs.*

CisGenome, CisGenome *vs* ChIPseeqer, and MACS *vs.* ChIPseeqer), with 653 peaks

detected by all three, representing 75% of ChIPseeqer detected peaks, 18% of

CisGenome detected peaks, and 10% of MACS detected peaks.

Results of the three peak calling programs were further evaluated based on seven gene

loci that we previously discovered and confirmed using ChIP-qPCR for Hoxa9 binding,

including two sites with high level binding (>30), three sites with medium level binding

(10 < IP signal/IgG signal < 30) and two sites with low level binding (5 < IP signal/IgG

< 10) (Figure 4.1A). All of the three programs detected the two sites with high Hoxa9

binding. However, only MACS was able to detect the three sites with medium level

binding, while CisGenome and ChIPseeqer missed one and two sites, respectively.

None of the programs was able to detect the two sites with low Hoxa9 binding,

supporting the stringent setting of their parameters to reduce false positive detection

(Figure 4.1A, C). So we decided that MACS best captures Hoxa9 binding signal, and

this program was used for later analysis. Default p-value cutoff of 1e-5 was used, as a

more stringent cutoff (1e-6) failed to detect all the median level sites (Figure 4.1C

column 5 and 6). Further, detected peaks also significantly overlap with previous report

of genome-wide Hoxa9 binding (Hypergeometric test p-value=0) (Huang et al., 2012).

Figure 4.1 Comparison of three peak calling programs, MACS, CisGenome and
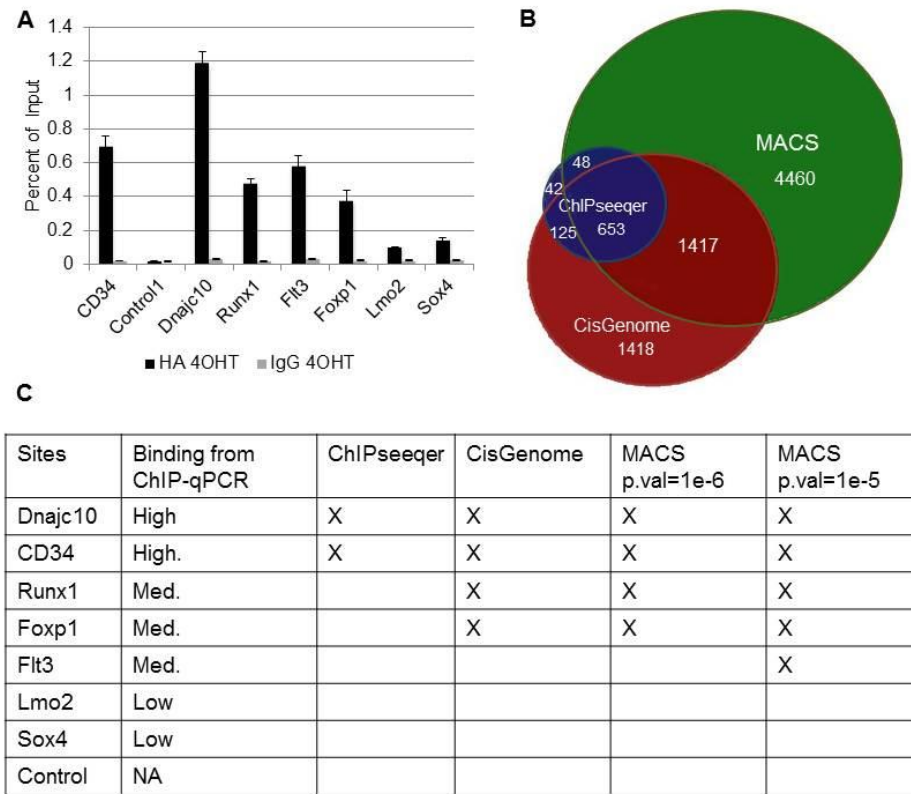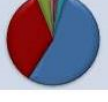
ChIPseeqer

Figure 4.1. Comparison of three peak calling programs, MACS, CisGenome and ChIPseeqer. (A) ChIP-qPCR of seven gene loci with different levels of Hoxa9 binding. (B) Venn Diagram showing the overlap of peaks detected by the three programs. (C) Summary of the Hoxa9 binding levels of the seven gene loci and whether they are detected by each peak calling program. For MACS, two p-value cutoffs were tested. p-value =1e-5 is the default setting of this program.

Overall, 6578 Hoxa9 binding peaks, 19093 Pu.1 binding peaks, 26446 C/ebpα binding peaks, 100189 H3K4me1 binding peaks and 33769 H3K27me3 binding peaks were detected with false discovery rate (FDR) < 0.05 (Figure 4.2). The peaks were then annotated with information on genomic regions including promoter (defined as -1kb to +100bp from transcription start site (TSS)), intergenic, intronic regions, etc. For Hoxa9,

~48% of the peaks reside in intergenic regions; ~46% reside in intronic regions; while

only ~2% of the peaks reside in promoter regions. This distribution is similar to

previous report, and the lower percentage of promoter associated peaks is due to a more

stringent definition of promoters (compared to promoter defined as -2kb to +1kp of

TSS in (Huang et al., 2012)) (Figure 4.2A). The Hoxa9 peaks were then annotated to

the nearest gene, and the distance between each peak and the gene transcription start

site (TSS) was calculated and plotted (Figure 4.2B, histogram and blue line). The

average distance is ~40kb, and ~80% of the peaks reside in between 5kb and 250kb

regions from the TSS. Pu.1 and C/ebpα also have majority of their peaks residing in

intergenic and intronic regions, although these two proteins have a higher percentage of

promoter associated peaks (Figure 4.2A). The distribution of the peak-TSS distance for

C/ebpα and Pu.1 are highly similar, and are also similar to that of Hoxa9 (Figure 4.2B).

Figure 4.2 Summary of Hoxa9, C/ebpα, Pu.1, H3K4me3 and H3K27me3 binding peaks

A

Sample	No. of Peaks	FDR	Peak distribution
Hoxa9	6578	0.004865	
Pu.1	19093	0.000995	
Cebpa	26446	0.002609	
H3K4me1	100189	0.00504	
H3K27me3	33769	0.013296	

Legend:
- 3'UTR
- 5'UTR
- Exon
- Intergenic
- Intron
- non-coding
- Promoter
- TTS

B

Distribution of Hoxa9/Cebpa/ Pu.1 peaks to TSS
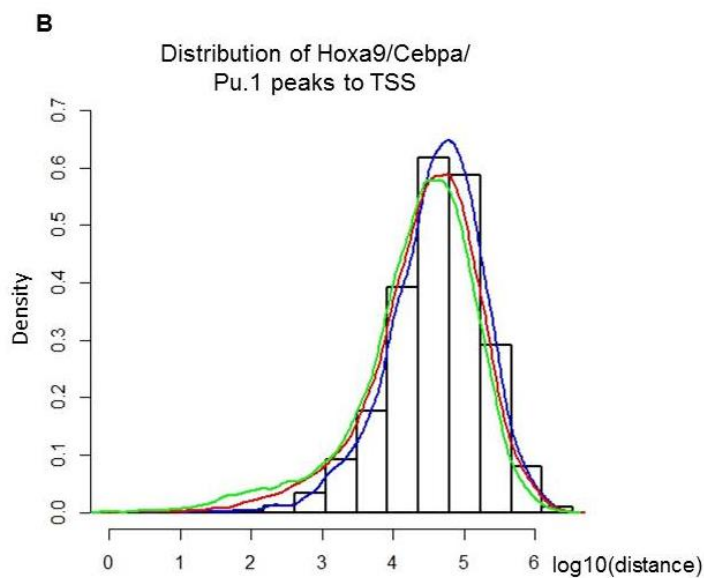
Figure 4.2. (A) Summary of Hoxa9, C/ebpα, Pu.1, H3K4me3 and H3K27me3 binding peaks, including total number of peaks (column2), false discovery rate (FDR) (column 3) and distribution of the peaks (column4). Red indicates intronic regions. Blue indicates intergenic regions. Green indicates promoter regions. TTS: transcription termination site. (B) Distribution

114

of the distance between Hoxa9, C/ebpα and Pu.1 peaks to the transcription start site of the nearest genes. Histogram of Hoxa9 peak distribution was plotted. Lines indicate estimated densities of peak-TSS distance for Hoxa9 (blue), C/ebpα (red) and Pu.1 (green). X axis labels the log10 transformation of the distance to TSS in basepair. Y axis labels the density of the distribution.

To characterize the distribution of histone marks around Hoxa9 binding sites, heatmap was generated displaying the binding signals of H3K4me1 and H3K27me3 -3kb to +3kb from the center of each Hoxa9 peak. Hoxa9 binding regions are largely enriched with the enhancer mark H3K4me1. Importantly, H3K27me3, which generally marks poised enhancers and are associated with inactive gene expression, was depleted from Hoxa9 binding regions. This data together with the distribution of Hoxa9 peaks in intergenic and intronic regions support an active role of Hoxa9 on the enhancer elements(Figure 4.3).

Figure 4.3 Distribution of histone marks around Hoxa9 binding peaks
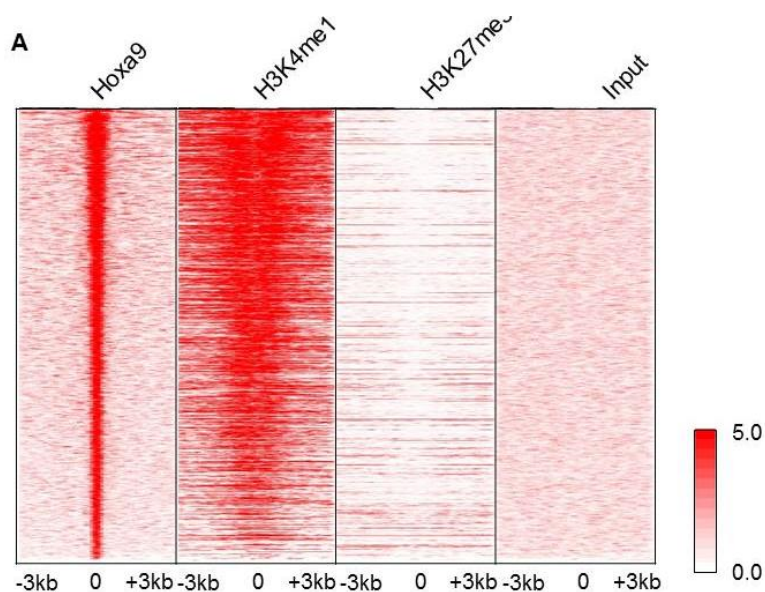
Figure 4.3. Distribution of histone marks around Hoxa9 binding peaks. (A) Heatmap showing the density of H3K4me1, H3K27me3 and input control -3kb to 3kb from center (0) of Hoxa9

binding peaks. Each row indicates a Hoxa9 peak, ranked from high to low according to average Hoxa9 density.

4.4.2 Hoxa9 co-localizes with C/ebpα and Pu.1 on the genome

The similar distribution of Hoxa9, C/ebpα and Pu.1 binding peaks (Figure 4.2) indicates that the three proteins may co-localize on the genome. To test this, we calculated the overlap between Hoxa9, C/ebpα and Pu.1 binding peaks. Among the 6578 Hoxa9 binding peaks, 3584 (~54%) overlap with C/ebpα binding peaks (Hypergeometric test p-value=0), 1851 (28%) overlap with Pu.1 binding peaks (Hypergeometric test p-value=1.917467e-271), while 1307 (~20%) overlap with both C/ebpα and Pu.1 (Figure 4.4).

Figure 4.4 Overlap of Hoxa9, C/ebpα and Pu.1 binding peaks
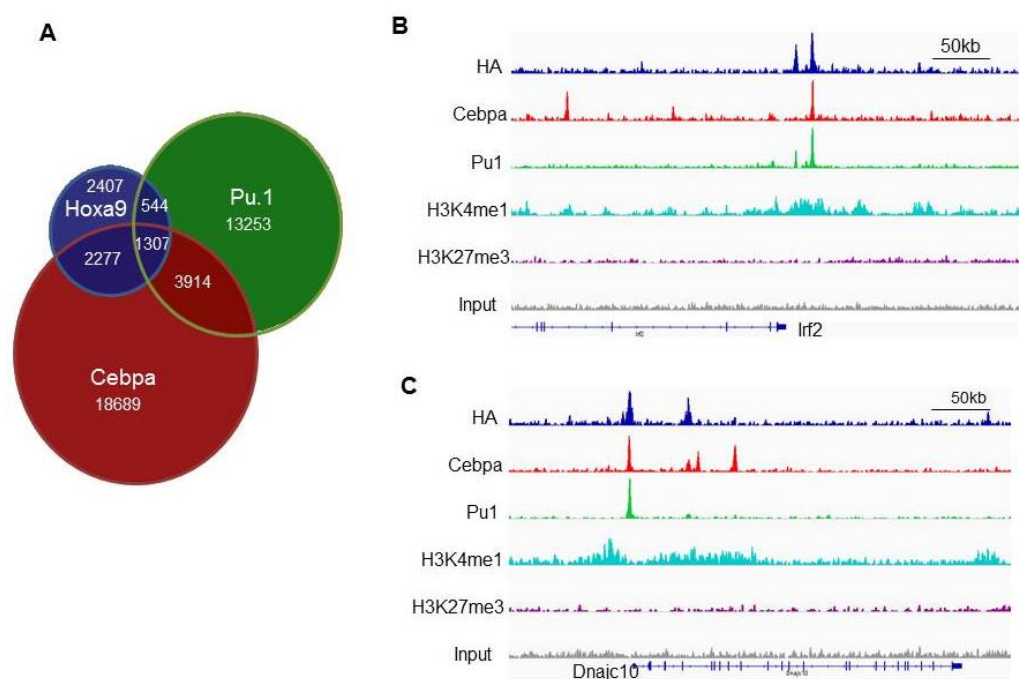
Figure 4.4. Overlap of Hoxa9, C/ebpα and Pu.1 binding peaks. (A) Venn Diagram showing the overlap of Hoxa9, C/ebpα and Pu.1 binding peaks. (B) Gene tracks of two representative gene

116

loci (Irf2, Dnajc10) co-bound by C/ebpα and Pu.1. Density plots were generated from raw read data and loaded into the UCSC Genome Browser as custom tracks. Blue indicates Hoxa9 binding signal. Red indicates C/ebpα binding signal. Green indicates Pu.1 binding signal. Light blue indicates H3K4me1 binding signal. Purple indicates H3K27me3 binding signal and gray indicates input signal which serves as control.

This is also supported by motif analysis. First, the Hoxa9 binding peaks were scanned for known transcription factor binding motifs against MatBase (Genomatix), which contains 865 motifs. 91% of Hoxa9 peaks contain HOXA9 motif (Z score=10.54, Z score >2 or <-2 corresponds to p-value <0.05) and 60% contain HOX-MEIS1 combinatorial binding motif (Z score=13.39), confirming the quality of the ChIP-seq. Notably, 95% of Hoxa9 peaks contain ETS (PU.1) binding motif (Z score=26.49) and 74% contain C/EBP binding motif (Z score=29.21) (Figure 4.5A). Further, de novo motif discovery using MEME-ChIP discovered that ETS and C/EBP motifs are enriched in Hoxa9 binding regions (E value= 1.1e-343, 2.6e-32, respectively. E value < 0.001 is considered to be significant) (Machanick and Bailey, 2011) (Figure 4.5B). Two other motifs discovered by this approach are STAT and RUNX motifs (Figure 4.5B) (Machanick and Bailey, 2011).

Figure 4.5 Motif analysis of Hoxa9 binding peaks
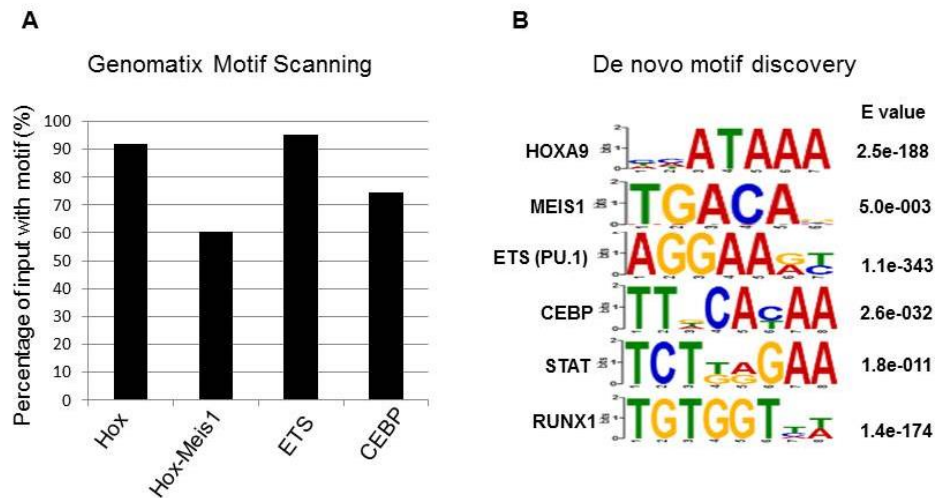
Figure 4.5 Motif analysis of Hoxa9 binding peaks. (A). Motif scanning of Hoxa9 binding peaks was performed by Genomatix against MatBase. Bar graph shows the percentages of Hoxa9 binding peaks that contain the indicated motifs. For all the four motifs, Z score >10. (B). STAMP logo of the six motifs discovered by de novo motif analysis that have significant enrichment in Hoxa9 binding peaks. E values indicate the significance (comparable to p-values).

4.4.3 Hoxa9, C/ebpα and Pu.1 cobound regions are enriched with enhancer marks

As Hoxa9 binding regions are enriched with the enhancer mark H3K4me1, and Pu.1 has been reported to promote H3K4me1 and enhancer establishment, we would like to determine if co-binding of C/ebpα and Pu.1 affects the level of histone marks. To test this, the Hoxa9 binding peaks were divided into four classes, those only bound by Hoxa9, co-bound by Hoxa9 and C/ebpα, co-bound by Hoxa9 and Pu.1 and co-bound by all three. The read densities of H3K4me1 and H3K27me3 3kb up- and downstream of the center of Hoxa9 peaks were normalized and calculated for each of the four classes and plotted (Heinz et al., 2010). All the four classes of Hoxa9 peaks have enriched

H3K4me1. Notably, the triple-bound peaks have the highest level of enrichment;

double-bound peaks of Hoxa9 and C/ebpα or Hoxa9 and Pu.1 show similar levels of

enrichment, which is lower than the triple-bound peaks, while peaks only bound by

Hoxa9 have the lowest level of enrichment (Figure 4.6A). Further, H3K27me3 mark

was depleted from these peaks, with triple-bound regions having the highest level of

depletion, and single-bound regions having the lowest level of depletion (Figure 4.6B).

These data support that co-binding of C/ebpα and Pu.1 associates with enriched active

enhancer marks.

Figure 4.6 Distribution of histone marks around Hoxa9 binding peaks with

different co-bindings of C/ebpα and Pu.1
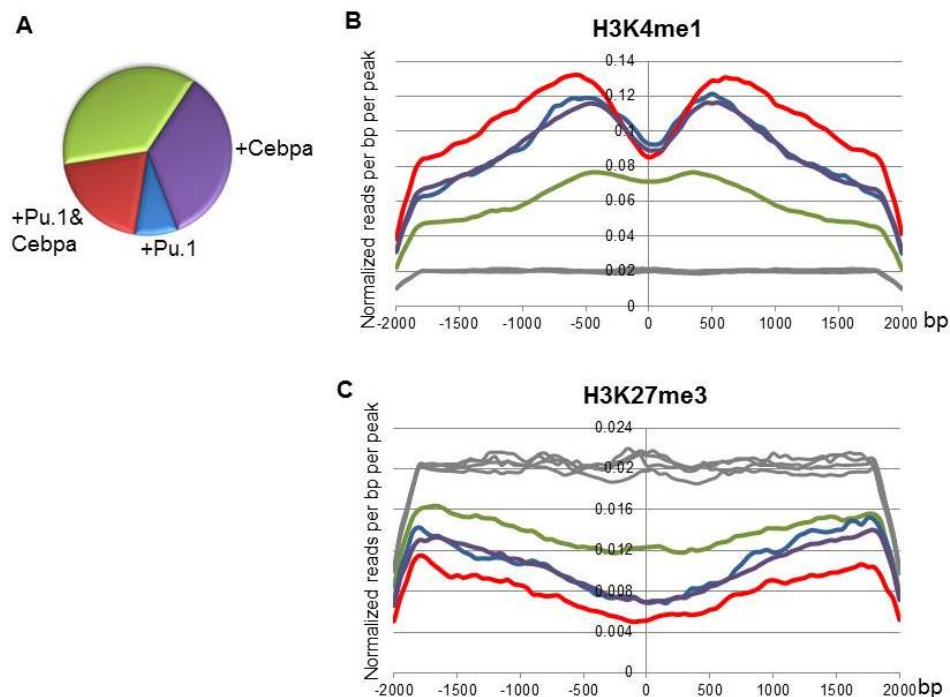
Figure 4.6. Distribution of histone marks around Hoxa9 binding peaks with different
co-bindings of C/ebpα and Pu.1. (A). Distribution of Hoxa9 peaks co-bound by C/ebpα or Pu.1.
Green indicates peaks that are only bound by Hoxa9. Purple indicates peaks co-bound by

119

Hoxa9 and C/ebpα. Blue indicates peaks co-bound by Hoxa9 and Pu.1. Red indicates peaks co-bound by Hoxa9, C/ebpα and Pu.1. (B-C). Density of H3K4me1 (B) and H3K27me3 (C) around Hoxa9 binding peaks with different co-binding patterns. The X-axis shows distance from the peaks. 0 indicates center of the peaks, and negative and positive values indicate upstream and downstream, respectively. Y-axis shows normalized average number of reads for each binding pattern. Colors corresponds to binding patterns the same way as in (A), with green indicating peaks that are only bound by Hoxa9, purple indicating peaks co-bound by Hoxa9 and C/ebpα, blue indicating peaks co-bound by Hoxa9 and Pu.1 and red indicating peaks co-bound by Hoxa9, C/ebpα and Pu.1. Grey indicates signals of the input.

Further, to characterize the genes associated with the Hoxa9 peaks with or without co-binding of C/ebpα and Pu.1, pathway analysis on single- and triple-bound regions was performed using GREAT (McLean et al., 2010). Indeed, genes associated with triple-bound regions are highly enriched for hematopoiesis related pathways, such as immunity and immune cell physiology (p-value<0.001). Although the single-bound regions are also enriched with a few hematopoiesis related pathways, the significance scores are a lot lower than those of the triple-bound regions (Figure 4.7).

Figure 4.7 Hoxa9, C/ebpα and Pu.1 co-bound regions associate with genes involved in hematopoiesis
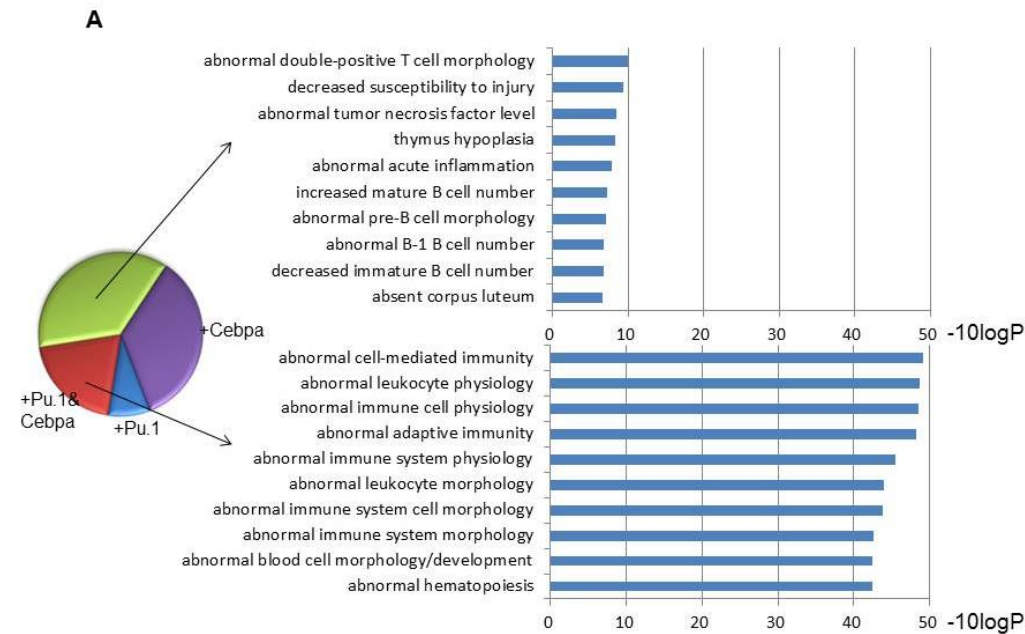
Figure 4.7. Hoxa9, C/ebpα and Pu.1 co-bound regions associate with genes involved in hematopoiesis. Pathway analysis of regions bound only by Hoxa9 (upper panel) and regions co-bound by Hoxa9, C/ebpα and Pu1 (lower panel). Analysis was performed using GREAT against the mouse phenotype database, y axis indicates top 10 enriched pathways, x axis indicates -10*log (p-value).

4.4.4   C/ebpα and Pu.1 associate with Hoxa9-responsive genes

The presence of C/ebpα and Pu.1 on Hoxa9 binding sites suggest that they are Hoxa9 cofactors and cooperate with Hoxa9 to regulate gene expression. To test this, we would like to answer two questions: Does co-binding of C/ebpα and Pu.1 affect the responsiveness of the genes to Hoxa9? And does it specify the activation or repression activity of Hoxa9? To this end, we first sought to define the genome-wide target genes of Hoxa9. Withdrawal of 4-OHT blocks translocation of Hoxa9-ER into the nucleus, thus enables us to monitor gene expression with or without functional Hoxa9. RNA was

121

extracted from cells cultured in the presence of 4-OHT and at 72hrs upon withdrawal of

4-OHT, which is approximately the earliest time point that depletion of Hoxa9 on

representative binding sites (those used in Figure 4.1A) was observed by ChIP-qPCR

experiments. RNA-sequencing was then performed with duplicated samples on

Illumina™ HiSeq2000 system. More than 20 million reads were obtained for each

sample and aligned to the reference genome. For differential expression analysis, two

programs were tested and compared, the Tophat-Cuffdiff module (Kim et al., 2013;

Trapnell et al., 2010) and EdgeR package (Robinson et al., 2010). The two methods

generated similar results and Tophat-Cuffdiff was used for later analysis (Figure 4.8A).

Upon 4-OHT withdrawal, 774 genes were up-regulated and 420 genes were

down-regulated (FDR<0.05), consistent with previous report that Hoxa9 both activates

and represses gene expression (Huang et al., 2012). Gene Set Enrichment Analysis

(GSEA) was performed (Subramanian et al., 2005), showing that genes up-regulated

with 4-OHT withdrawal are enriched with genes repressed in hematopoietic stem cells

(False discovery rate (FDR)=0, Normalized enrichment score (NES)=1.99) and

leukemia stem cells (FDR=0, NES=1.94), indicating that loss of Hoxa9 leads to loss of

stem cell signature, supporting the role of Hoxa9 in maintaining hematopoietic stem

cell identity and confirming the quality of the RNA-seq (Figure 4.8A, B).

Figure 4.8 Hoxa9 regulates stem cell signature genes
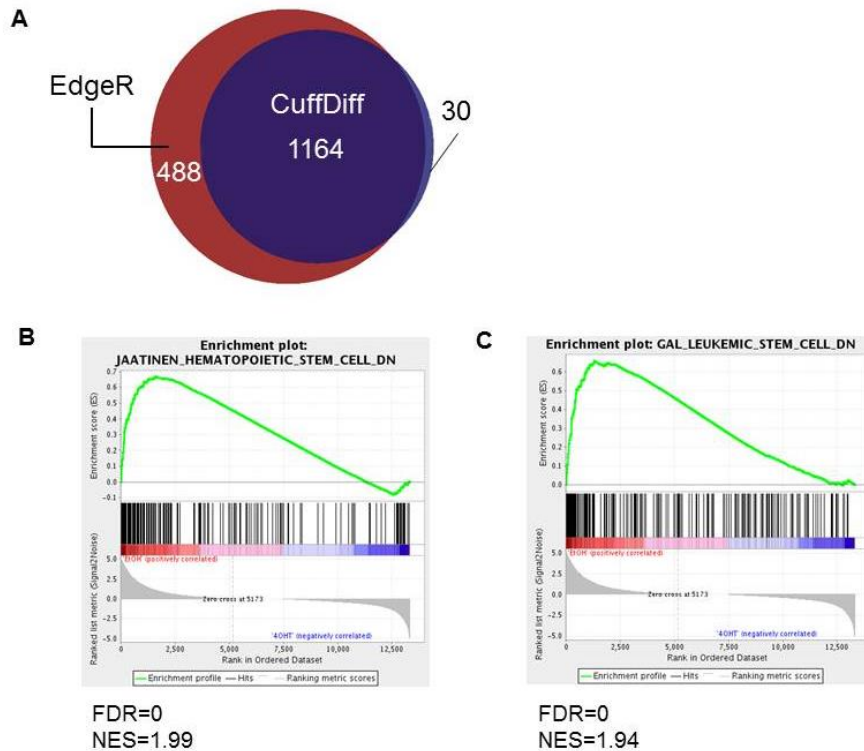
Figure 4.8. Hoxa9 regulates stem cell signature genes. (A) Venn diagram showing the number of genes with significant expression change detected by CuffDiff (blue) and EdgeR (red). (B-C) Gene set enrichment analysis discovered that hematopoietic stem cell signature (B) and leukemic stem cell signature (C) were reversed upon 4-OHT withdrawal. FDR: false discovery rate. NES: normalized enrichment score.

Genes with changed expression may be directly or indirectly regulated by Hoxa9, thus we define "Hoxa9 core target genes" as genes that are both bound by Hoxa9 and have expression change for more than 1.5 fold upon 4-OHT withdrawal. To this end, we mapped the 6578 Hoxa9 binding peaks to the nearest gene, and obtained 3602 unique genes (15.6% of the total number of genes detected in RNA-seq) that are nearest to one or more Hoxa9 peaks. Among these 3602 genes, 739 genes meet the criteria (Hoxa9 core target genes, Appendix. A).

These 739 genes represent 20.51% (out of 3602) of the Hoxa9-bound genes. The

percentage is significantly higher than that of non-Hoxa9 bound genes, as only 6.72%

(1292 out of 19214) of those showed significant change in expression (one sided

Fisher exact test p-value<2.2e-16) (Figure 4.9A). To further test if the responsiveness

of the genes to Hoxa9 activity is affected by co-binding of C/ebpα and Pu.1, genes

bound by Hoxa9 were further divided into four classes: those only bound by Hoxa9

(985 genes), co-bound by Hoxa9 and C/ebpα (1294 genes) or Pu.1 (364 genes) and

co-bound by all three of them (1048 genes). Notably, 14.51% (143 out of 985) of the

genes only bound by Hoxa9 showed significant change; the percentages are higher for

double-bound genes by Hoxa9 and C/ebpα (20.94%, 271 out of 1294) or Pu.1

(21.42%, 78 out of 364); while the triple-bound genes have the highest percentage,

which is 25.85% (271 out of 1048, Appendix. B) (Chi square test of the five classes

p-value < 2.2e-16, and the pairwise Fisher exact one-sided test p-values are shown in

Figure 4.9C). These data support that C/ebpα and Pu.1 associate with genes that are

more responsive to Hoxa9.

Figure 4.9 Co-binding of C/ebpα and Pu.1 associates with Hoxa9 responsive genes

p-value	No Hoxa9	Hoxa9 alone	Hoxa9+Cebpa	Hoxa9+Pu.1	Hoxa9+Cebpa+Pu.1
No Hoxa9		< 2.2e-16*	< 2.2e-16*	< 2.2e-16*	< 2.2e-16*
Hoxa9 alone			4.487e-05*	0.001834*	1.172e-10*
Hoxa9+Cebpa				0.446	0.002969*
Hoxa9+Pu.1					0.05179
Hoxa9+Cebpa+Pu.1					
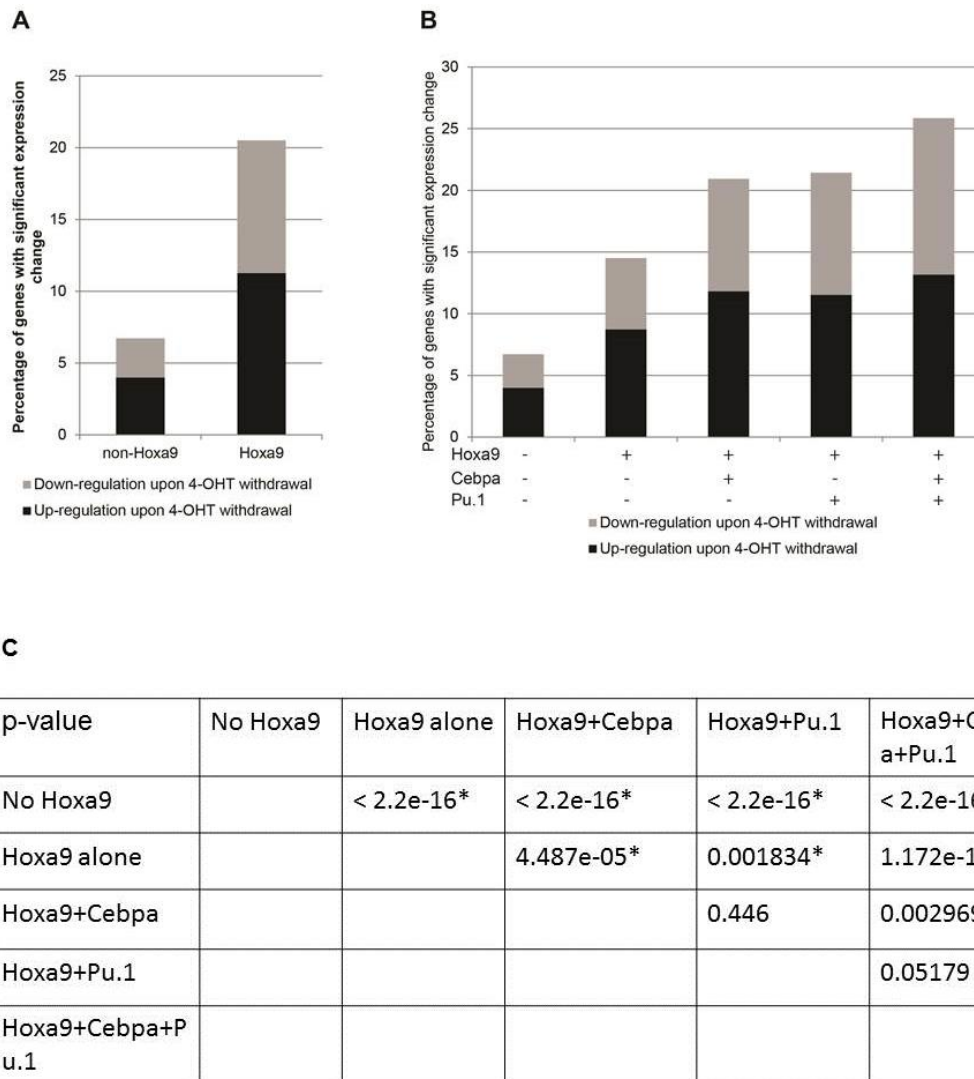
Figure 4.9. Co-binding of C/ebpα and Pu.1 associates with Hoxa9 responsive genes. (A) Bar graph showing the percentage of genes with significant expression change upon 4-OHT withdrawal in none-Hoxa9 and Hoxa9 bound genes. Black indicates up-regulated genes upon 4-OHT withdrawal. Grey indicates down-regulated genes upon 4-OHT withdrawal. (B) Bar graph showing the percentage of genes with significant expression change upon 4-OHT withdrawal in genes that are not bound by Hoxa9, bound only by Hoxa9, co-bound by Hoxa9 and C/ebpα or Pu.1 and co-bound by Hoxa9, C/ebpα and Pu.1. Black indicates up-regulated genes upon 4-OHT withdrawal. Grey indicates down-regulated genes upon 4-OHT withdrawal. (C). Matrix showing pairwise p-values of Fisher Exact test comparing the percentages of significant expression change for genes with different binding patterns in (B). * indicates p-value<0.05.

However, binding of Hoxa9 does not associate with the magnitude of gene expression change, as for both up- and down-regulation upon 4-OHT withdrawal, the global fold change are comparable between non-Hoxa9 bound genes and Hoxa9 bound genes (Wilcoxon Rank Sum two-sided test p-value= 0.09817 for up-regulated genes, p-value = 0.06788 for down-regulated genes) (Figure 4.10A and B). Fold change was also compared among genes with different co-bindings of C/ebpα and Pu.1. No difference was detected for up-regulated genes (Kruskal-Wallis test p-value= 0.06152) (Figure 4.10C). For down-regulated genes, only those with co-binding of Hoxa9 and Pu.1 show slightly higher fold change compared to other groups (Kruskal-Wallis test p-value= 0.02459).

Figure 4.10 Co-binding of C/ebpα and Pu1 does not associate with specific

activity of Hoxa9
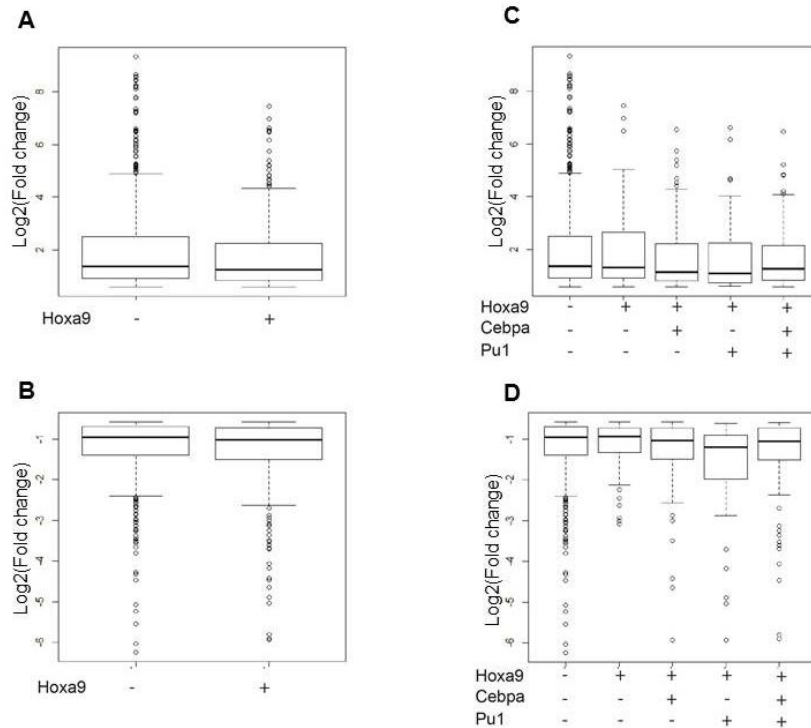
Figure 4.10. Co-binding of C/ebpα and Pu1 does not associate with the magnitude of gene expression change. (A) Boxplot showing the log2 transformed fold change of genes with significant up-regulation upon 4-OHT withdrawal in none-Hoxa9 and Hoxa9 bound genes. (B) Boxplot showing the log2 transformed fold change of genes with significant up-regulation upon 4-OHT withdrawal with different binding patterns. (C) Boxplot showing the log2 transformed fold change of genes with significant down-regulation upon 4-OHT withdrawal in none-Hoxa9 and Hoxa9 bound genes.   (D) Boxplot showing the log2 transformed fold change of genes with significant down-regulation upon 4-OHT withdrawal with different binding patterns. log2 transformed fold change for genes with co-binding of Hoxa9 and Pu.1 is higher compared to those with no Hoxa9, Hoxa9 alone, Hoxa9 and C/ebpα and Hoxa9, C/ebpα and Pu.1, with pairwise Wilcoxon Rank Sum two-sided test p-values = 0.001513, 0.01255, 0.01662, 0.02069, respectively).

Finally, we tested if co-binding of C/ebpα and Pu.1 is associated with activation or

repression activity of Hoxa9. For genes with significant expression change upon

4-OHT withdrawal and without Hoxa9 binding, 40.63% were down-regulated

(activated by Hoxa9) （Figure 4.9A）. This percentage is comparable to those bound

by Hoxa9, which is 45.06% (Fisher Exact test two-sided p-value = 0.05561). Further,

no difference was detected when comparing genes with different co-bindings of

C/ebpα and Pu.1 (Chi square test p-value= 0.1089) (Figure 4.9B). These data indicate

that co-binding of C/ebpα and Pu.1 does not specify the activation or repression

activity of Hoxa9.

4.5   Discussion

The development of high throughput sequencing technologies, such as ChIP-seq and

RNA-seq, has greatly improved our ability to study protein-DNA interaction or gene

transcription in the genome-wide scale, and revolutionized the way for biomedical

research. Despite the advantage of these techniques as being fast, economic and

highly reproducible, the generation of large amount of data brings challenge for the

development of analysis algorithms. For ChIP-seq analysis, peak calling is the key

step. Currently, there are more than thirty programs that use different ways to

generate reads density, model peak distribution, adjust for background and assign

peaks to genomic regions (Wilbanks and Facciotti, 2010), so they work differently in

balancing sensitivity, which is the ability to "detect more peaks", and specificity,

which is the ability to "detect real peaks", and may give very different results. So

choosing the proper method for data analysis according to the need of study is critical.

In this work, we tested and compared three different peak calling programs, MACS,

CisGenome and ChIPseeqer. MACS models the distribution of reads on the genome using poisson distribution with dynamic parameter λ estimated from nearby genomic regions. Also, it estimates the length of the sequencing fragments and shifts peaks accordingly, which increases the accuracy in estimating the center of peaks (Zhang et al., 2008). CisGenome is one of the earliest developed peak calling programs and gains its popularity by having a graphical user interface. Cisgenome uses conditional binomial model to estimate the distribution of reads (Ji et al., 2008). ChIPseeqer, as MACS, also uses a poisson model to estimate the reads distribution, but adjusts the length of fragment length by extending reads to the average length of the DNA fragment in the ChIP-seq library (Giannopoulou and Elemento, 2011). For Hoxa9 ChIP-seq data, the results of the three programs show significant overlap, although the total numbers of peaks detected are quite different. MACS detected the largest number of peaks, while ChIPseeqer detected the smallest number of peaks (Figure 4.1). Comparing to several gene loci that have been confirmed by ChIP-qPCR assays to have Hoxa9 binding, we found that although ChIPseeqer and CisGenome both detected the regions with high level of Hoxa9 binding, they missed some regions with median level of Hoxa9 binding, while MACS has higher sensitivity and detected all the regions with high and median levels Hoxa9 binding. Notably, the high sensitivity of MACS does not associate with an obvious loss of specificity, as although we cannot obtain the "true Hoxa9 binding sites", more than 90% of MACS peaks are detected to have HOX motif, supporting the reliability of the peaks. Thus in our

specific experiment, we conclude that MACS performs best in detecting Hoxa9

binding.

Similar to ChIP-seq, several methods have been developed for differential expression

analysis of RNA-seq data. In this study, we tested and compared two methods, EdgeR

package, which uses overdispersed poisson model to estimate the variation in the

sequencing read counts (Robinson et al., 2010) and empirical Bayes method to

estimate the degree of overdispersion, and Cuffdiff from Cufflinks RNA-Seq analysis

tools, which uses beta negative binomial model to estimate variability across

replicates and variability caused by ambiguously mapped reads from gene isoforms

(Trapnell et al., 2010). Using the same alignment method (Tophat), the two methods

give very similar results (Figure 4.8), indicating the consistency of these algorithms.

Cuffdiff detects slightly fewer differentially expressed genes, possibly because taking

into consideration of the isoform information decreases the false discovery rate.

In this study, we characterized the genome-wide binding for Hoxa9, C/ebpα and Pu.1,

and found that all these three transcription factors predominantly bind to intergenic

and intronic regions, supporting their role in enhancer function. Hoxa9, C/ebpα and

Pu.1 have significant overlap on the genome, and the co-bound regions are enriched

with the enhancer marker H3K4me1 (Figure 4.4). Specifically, triple-bound regions

by Hoxa9, C/ebpα and Pu.1 have the highest levels of H3K4me1, double-bound

regions have lower enrichment and single-bound regions only by Hoxa9 have the

lowest level of enrichment (Figure 4.6), supporting the role of C/ebpα and Pu.1 in

promoting H3K4me1 and enhancer establishment. Notably, the depletion of

H3K27me3, the poised enhancer marker, indicates that these enhancers play active

roles in transcription regulation (Figure 4.6). Consistent with this, co-binding of

C/ebpα and Pu.1 associates with genes whose expression are responsive to Hoxa9

activity. These data support that C/ebpα and Pu.1 are Hoxa9 cofactors and cooperate

with Hoxa9 to regulate gene expression. Indeed, our lab observed that depletion of

C/ebpα or Pu.1 in Hoxa9-Meis1 transformed cells blocks transformation, reflected by

cell differentiation and decreased cell proliferation (unpublished data). Important

questions remain to be answered. Especially, how does Hoxa9 cooperate with C/ebpα

and Pu.1? Does binding of one factor affect the recruitment of the other two? Also,

although Hoxa9, C/ebpα and Pu.1 co-bound regions associate with higher percentage

of genes response to Hoxa9 activity (25%), the majority of these genes do not respond,

and C/ebpα and Pu.1 co-binding does not specify the activation or repression activity

of Hoxa9 (Figure 4.9). Are there other co-factors, either on the enhancers or

promoters, also involved in this transcriptional regulation? These questions will be the

focus of Hess lab for the future study.

Chapter5   Concluding Remarks and Future Directions

The MLL-HOXA9 axis plays a critical role in the regulation of development and hematopoiesis. Chromosome translocations of MLL are closely associated with human leukemia and constitutive activation of HOXA9 is required for the leukemogenesis. Although being studied for decades, important questions remain to be answered in this field. For example, how is the activity of MLL regulated during hematopoiesis? What are the mechanisms of MLL fusion-mediated transformation? And how does HOXA9 regulate gene expression and mediate leukemogenesis? The work described in this dissertation touches every aspect of the above questions.

In Chapter 2 and Chapter 3, I addressed the first two questions by studying the function of the PHD/Bromo region of MLL. This region is invariably disrupted or deleted in oncogenic MLL fusion proteins and is incompatible to MLL fusion mediated transformation. We found that this region mediates MLL ubiquitination in multiple ways. The bromodomain recruits ECS^{ASB2} E3 ligase complex through interaction with ASB2. ECS^{ASB2} mediates ubiquitination and degradation of MLL during hematopoietic differentiation, thus contributing to the down-regulation of MLL target genes, including *HOXA9* and *MEIS1*. This is important for the normal progression of hematopoiesis, as knockdown of ASB2 stabilizes MLL and blocks cell differentiation. In addition to the bromodomain, the second PHD finger (PHD2) of

MLL has intrinsic E3 ligase activity in the presence of the E2 conjugating enzyme CDC34. This activity requires the intact structure of the PHD finger and is conserved in the second PHD finger of MLL4. Further, mutation of PHD2 affects MLL stability and transactivation ability. These data support a model in which the activity and stability of MLL is regulated by ubiquitination in multiple ways, possibly depending on different biological context (Figure 5.1). Notably, the oncogenic MLL fusion proteins do not have the PHD/Bromo region, and are resistant to both degradation pathways. The increased protein stability may contribute to MLL fusion mediated transformation.

Figure 5.1 Model for the function of the PHD/Bromodomain in MLL ubiquitination
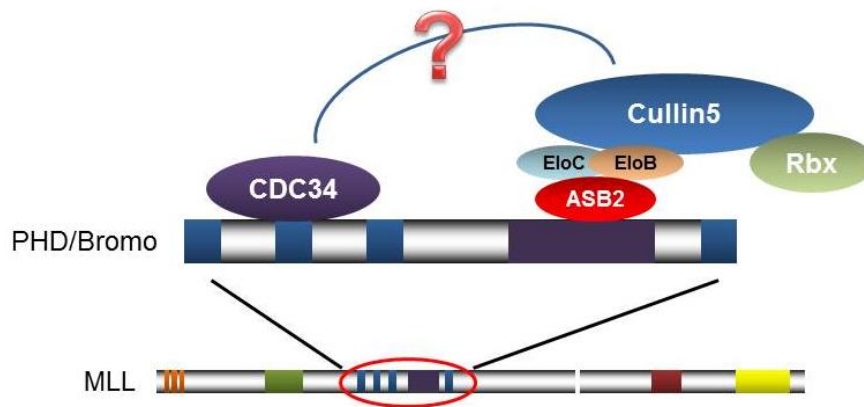
Figure 5.1 Model for the function of the PHD/Bromodomain in MLL ubiquitination. The bromodomain interacts with ASB2 and recruits the ECS[ASB2] E3 ligase complex, while the second PHD finger interacts with CDC34 and has intrinsic E3 ligase activity. It remains to be determined if there is any crosstalk between ECS[ASB2] mediated MLL ubiquitination and PHD2 mediated ubiquitination.

An interesting and important question that arises from examining the PHD/Bromo

region of MLL is what are the biological functions of ASB2. The rapid expression of

ASB2 upon cell differentiation and its effect on MLL degradation supports its role as

a positive regulator of hematopoietic differentiation. It will be interesting to establish

a knockout mouse model of Asb2 and systematically assess its functions in embryonic

development and hematopoiesis, specifically, whether depletion of Asb2 leads to

hyperproliferation of hematopoietic cells, and sensitizes these cells to leukemogenic

transformation. Also, besides MLL, FilaminA and FilaminB, ASB2 may have other

substrates that remain to be discovered. Among the MLL family proteins, MLL4

shares the greatest sequence similarity to MLL, raising the possibility that ASB2 can

also ubiquitinate MLL4. Indeed, blast of MLL bromodomain and flanking sequences

(aa1631-1934) reveals a region in MLL4 (1401-1642) that shares 40% sequence

identity. Experiments are needed to determine if ASB2 is a specific E3 ligase of MLL,

or it also regulates MLL4 and other MLL family proteins. Further, ASB2 contains

twelve ankyrin repeats which are implicated to mediate the interaction with its

substrates. We mapped the interaction between ASB2 and MLL to the first five

ankyrin repeats, which may not be the minimal region required for this interaction.

Co-localization of ASB2 and FilaminA requires the first ten ankyrin repeats but the

direct interaction may be mediated only by the N-terminal twenty amino acid residues

that are structurally homologous to filamin-binding motifs (Lamsoul et al., 2011). So

it is possible that other ankyrin repeats of ASB2 interact with additional substrates.

Identifying these substrates, for example, by mass spectrometry, and studying their functions in hematopoiesis and other biological processes will provide insights for the functions of ASB2.

A recent study reports that the wild-type allele of *MLL* is required for MLL-AF9 mediated leukemogenesis, as knockdown or excision of wild-type Mll in MLL-AF9 transduced cells inhibits cell proliferation, reduces colony formation and inhibits the development of leukemia *in vivo* (Thiel et al., 2010). Based on these observations, the authors propose that MLL fusion protein-transformed leukemia cells may be more sensitive to inhibition of wild-type MLL, as they have a reduced cellular level of MLL, and wild-type MLL is a potential target for leukemia therapy (Thiel et al., 2010). This is consistent with our findings, as ASB2 expression in MLL-AF9 transformed cells phenocopies knockdown or excision of wild-type MLL, including decreased H3K4 methylation, decreased *HoxA9* expression and reduced cell proliferation and colony formation. Notably, ASB2 expression is rapidly induced by ATRA (Guibal et al., 2002; Kohroki et al., 2001). This suggests that ATRA treatment, which is already in use for treatment of promyelocytic leukemia, would have activity against MLL associated leukemias. Indeed, two leukemia cell lines harboring the MLL-AF9 fusion protein, THP-1 and MOLM-14 cells, have been shown to be sensitive to ATRA induced differentiation (Matsushita et al., 2008). More MLL associated leukemia cell lines need to be tested to confirm this application.

Other than ASB2, several members of the ASB family have potential roles in hematopoiesis as well. The expression profile of these *Asb* genes during differentiation of Hoxa9-ER/Meis1 transformed cells reveals that besides *Asb2*, which has increased expression upon cell differentiation, three other *Asb* genes, *Asb8*, *Asb11* and *Asb13* show decreased expression in this process, indicating that they may have opposite functions in hematopoiesis to *Asb2* (Figure 2.13A). Notably, ASB11 has been shown in zebrafish to be required for proper neurogenesis by activating the Notch signaling pathway, which is also a critical signaling pathway for self-renewal in hematopoietic stem cells (Diks et al., 2008; Sartori da Silva et al., 2010). The function of ASB8 and ASB13 is still unknown. It will be interesting to study the function of these ASB proteins in hematopoiesis, which may reveal novel regulators of this process.

The biological significance of the MLL PHD2 E3 ligase activity also remains to be determined. Considering the role of CDC34 as a master regulator of cell cycle (King et al., 1996), it is possible that PHD2 regulates MLL degradation during cell cycle progression. It will be interesting to establish a cell line with MLL PHD2 mutants, for example, in Mll$^{-/-}$ MEF cells, and test its effect on cell cycle progression and Mll target gene expression.

Further, it will be interesting to test if PHD2 has other substrates besides auto-ubiquitination. In vitro ubiquitination assays show that PHD2 specifically

ubiquitinates histone H3 and H4, but not H2A and H2B. The ubiquitination sites were further identified to be H3K79, H3K122 and H4K31. All these sites have been shown to be ubiquitinated in vivo using proteomics approaches, while their functions remain unknown (Kim et al., 2011; Wagner et al., 2011). Further, sequence analysis reveals that the flanking sequences of these sites share similarity, which may be preferentially recognized and ubiquitinated by PHD2 (Basu et al., 2009). It remains to be determined if histone H3 and H4 are *bona fide* substrates of MLL PHD2 in vivo, and if so, what are the biological functions. These three sites on histone H3 and H4 can be modified by other post-translational modifications. Especially, H3K79 is methylated by DOT1L, which is a cofactor of MLL fusion proteins and is required for MLL fusion mediated transformation (Okada et al., 2005). H3K122 can be acetylated, which promotes transcription (Tropberger et al., 2013). Ubiquitination may compete with these modifications and represents a second way that MLL regulates transcription besides H3K4 methylation. Further, it has been shown that MLL mediated histone H3K4 methylation depends on ubiquitination of histone H2B (Kim et al., 2009). It will be interesting to test if PHD2 mediated ubiquitination affects H3K4 methylation, which may represent another layer of histone crosstalk.

Besides ubiquitin, proteins are conjugated with other ubiquitin-like molecules, which regulate protein stability and activity (Herrmann et al., 2007). The PHD finger of KAP1 has been shown to mediate sumoylation of the adjacent bromodomain (Ivanov

et al., 2007). This activity is critical for the recruitment of NuRD, as both SETDB1

and CHD3, the subunits of NuRD complex, contain a SUMO-interacting motif and

interact with KAP1 through sumo binding (Capili et al., 2001; Ivanov et al., 2007;

Zeng et al., 2008). In MLL, PHD2 and the bromodomain are also adjacent to each

other. It is worth testing if PHD2 has sumo ligase activity, for example, to interact

with the SUMO E2 conjugating enzyme UBC9, and mediate sumoylation of the

bromodomain or other substrates. Further, as the bromodomain mediates interaction

with ASB2, this modification may affect the interaction between MLL and ASB2, and

ASB2 mediated MLL ubiquitination. Besides SUMO protein, NEDD8 is another

ubiquitin-like protein. NEDD8 has been shown to be conjugated to all Cullin proteins.

This neddylation results in conformational rearrangements of the ECS complex,

which are necessary for ubiquitin transfer to a substrate (Soucy et al., 2010). It

remains to be determined if PHD2 can mediate conjugation of NEDD8 on Cullin5 and

regulates its activity, which may represent another way of crosstalk between PHD2

and ECS^{ASB2} mediated MLL ubiquitination.

Finally, although the structure for the binding of PHD fingers to histone modifications

is well studied, the structure and sequence basis of the PHD fingers with E3 ligase

activity remains unknown. It will be important to characterize the interaction between

these PHD fingers and the E2 conjugating enzymes, and determine the key foldings

and residues involved. This will help classify and predict the specific functions of

PHD fingers, and also provide insights for development of PHD finger inhibitors for therapeutic applications.

In Chapter 4, I explored the mechanisms of HOXA9 regulated gene expression by studying the cooperation between Hoxa9 and its potential cofactors C/ebpα and Pu.1 using high throughput sequencing technologies. Analysis of ChIP-seq and RNA-seq data reveals that genome-wide binding of C/ebpα and Pu.1 significantly overlaps with the binding of Hoxa9. Further, the co-bound regions have enriched enhancer marker H3K4me1 and are associated with hematopoiesis related pathways and Hoxa9-responsive genes.

Experiments are ongoing to study the biological significance of C/ebpα and Pu.1 for the function of Hoxa9 in normal and malignant hematopoiesis. Indeed, our preliminary data show that depletion of C/ebpα or Pu.1 in Hoxa9/Meis1 transformed cells blocks transformation, reflected by cell differentiation and decreased cell proliferation (unpublished data), supporting that these two proteins are required for Hoxa9/Meis1 mediated transformation. Along with this, it is important to map the interaction regions on Hoxa9, C/ebpα and Pu.1, as disruption of these interactions may have therapeutic values for Hoxa9 mediated leukemia. The mechanisms of the recruitment of the three proteins remained to be determined. It has been reported that C/ebpα and Pu.1 establish the enhancer elements and prime recruitment of the NF-κB p65 subunit (Heinz et al., 2010; Jin et al., 2011).   Experiments are needed to test if

the same mechanism works for Hoxa9. The lab has obtained conditional knockout

mice for *C/ebpα* and *Pu.1*, which enables us to deplete one or both of these proteins

and monitor the recruitment of Hoxa9, and the establishment of the enhancer marks.

Further, the ChIP-seq and RNA-seq data provide a valuable resource to study other

aspects of Hoxa9 regulated gene expression. For example, does the location of Hoxa9

binding on different genomic regions (promoter/intergenic/intronic regions) or with

different distances from the transcription start sites correlate with gene expression?

Also, combining the ChIP-seq and RNA-seq data, we obtained the genes that are

bound by Hoxa9 (and its cofactors) and have their expression regulated by Hoxa9. It

will be important to find the key genes and pathways for Hoxa9 regulated

hematopoiesis and leukemogenesis. Finally, co-binding of C/ebpα and Pu.1 does not

predict the activation or repression activity of Hoxa9, so the direction of Hoxa9's

activity still remains a mystery. Other possibilities can be explored. For example,

separate motif analysis of Hoxa9 binding regions associated with activated or

repressed genes may reveal additional Hoxa9 cofactors involved in this process. Or

the histone mark on gene promoters may determine the direction of gene transcription

together with the enhancer mark. Deciphering the functional interaction between

Hoxa9 and its cofactors will provide not only insight for Hoxa9 regulated gene

expression, but also potential therapeutic targets for Hoxa9 associated leukemia and

other malignancies.

Appendix A. Core Hoxa9 Target Genes

Gene name	Locus	Gene name	Locus
1110002B05Rik	chr12:55746360-5575755 9	Klf5	chr14:99697909-99712 628
1110059E24Rik	chr19:21671802-2172728 1	Klf6	chr13:5860734-586963 9
1600012F09Rik	chr13:18078473-1808518 5	Klf7	chr1:64082246-641679 63
1600014C10Rik	chr7:38968235-3898258 4	Klhl2	chr8:67218658-673737 16
1700012B09Rik	chr9:14562637-1457547 4	Klhl5	chr5:65522469-655593 81
1700020L24Rik	chr11:83251195-832547 34	Krt80	chr15:101180001-1012 00556
1810033B17Rik	chr8:3665761-3668905	Lair1	chr7:3958674-4014806
2310037I24Rik	chr15:98349531-983646 52	Lama5	chr2:179911077-17996 0564
4930550C14Rik	chr9:53213390-5324090 8	Lass4	chr8:4493404-4526079
4931406C07Rik	chr9:15087780-1516223 2	Lbp	chr2:158132228-15815 8588
4932438H23Rik	chr16:91054179-910693 90	Ldlrad3	chr2:101790357-10202 6534
5031414D18Rik	chr14:75415833-754523 39	Lhfpl2	chr13:94827750-94965 364
5730469M10Rik	chr14:41807028-418270 64	Lilrb3	chr7:3664106-3671984
Abca13	chr11:9091944-9584262	Lima1	chr15:99608898-99705 887
Abcd2	chr15:90976301-910222 38	Lims1	chr10:57786213-57887 439
Abcg3	chr5:105364075-105411 736	Litaf	chr16:10959365-10993 214
Ablim1	chr19:57107753-572905 22	Lmo2	chr2:103798151-10382 2035
Abr	chr11:76230235-763912 29	Lpgat1	chr1:193541902-19360 8134

Acad8	chr9:26781725-2680713 4	Lrrc8c	chr5:105948489-10603 7973
Acot7	chr4:151552208-151645 964	Lrrfip2	chr9:111020614-11112 8172
Acoxl	chr2:127680363-127949 628	Luzp1	chr4:136025675-13610 5233
Acpl2	chr9:96723761-9678984 1	Ly6g5b	chr17:35250889-35252 345
Acsf2	chr11:94418415-944631 00	Ly6i	chr15:74810241-74813 860
Actr3b	chr5:25265843-2535615 9	Lypd6	chr2:49921948-500490 89
Adam17	chr12:21322252-213794 52	2310010M24Ri k	chr2:49643205-498043 66
Adam8	chr7:147164839-147178 387	Lyst	chr13:13682675-13869 707
Adamts6	chr13:105077952-10528 4843	Maml3	chr3:51491534-519089 28
Adamtsl4	chr3:95480126-9549178 1	Man1a	chr10:53625838-53795 602
Add3	chr19:53214934-533218 89	Map3k14	chr11:103081077-1031 28715
Adrb2	chr18:62337469-623396 13	Map3k8	chr18:4331324-435295 1
Adssl1	chr12:113858257-11387 9566	Map4k4	chr1:39957757-400831 55
Aff4	chr11:53164268-532353 32	Mapk4	chr18:74088140-74224 603
Centg2	chr1:91351385-9179185 7	Matk	chr10:80720289-80725 726
Agl	chr3:116442916-116511 084	Mbd2	chr18:70727945-70785 785
Ahr	chr12:36182650-362196 61	Mboat1	chr13:30228358-30338 563
Aim1	chr10:43670112-437246 52	Mcph1	chr8:18595172-188031 88
Ak2	chr4:128670508-128688 773	Mdga1	chr17:29964902-30024 827
Akap2	chr4:57858119-5790985 6	Mdm1	chr10:117578842-1176 06055
Akr1c18	chr13:4131872-4149877	Mdn1	chr4:32744093-328621 92

Akr1e1	chr13:4591735-4608410	Med13l	chr5:119010727-11921 5446
Alas1	chr9:106135785-106150 285	Megf9	chr4:70092960-701959 62
Alcam	chr16:52249108-524531 10	Meis1	chr11:18780430-18918 972
Aldh1a3	chr7:73535778-7357236 3	Met	chr6:17413956-175239 80
Aldh3b1	chr19:3913490-3929716	Metrnl	chr11:121563740-1215 78703
Alox5	chr6:116360088-116411 196	2610205E22Ri k	chr2:30663496-306785 34
Als2	chr1:59177778-5929407 5	Mgat4b	chr11:50038836-50048 605
Ammecr1	chrX:139288016-13940 1271	Mgat5	chr1:129101562-12937 9549
Anapc10	chr8:82235718-8230122 0	Mgl2	chr11:69943858-69951 044
Angpt1	chr15:42256270-425083 41	Mgll	chr6:88674405-887783 54
Ralgps1	chr2:32992492-3322699 8	Cbara1	chr10:59165382-59326 870
Ank	chr15:27396431-275246 62	Mid1	chrX:166123178-16642 8729
Ankrd22	chr19:34175432-342405 31	Minpp1	chr19:32560258-32589 860
Ankrd28	chr14:32513200-326436 01	Mmp12	chr9:7347373-7360461
Ankrd9	chr12:112213562-11221 7231	Mmp13	chr9:7272513-7283333
Tmem16k	chr9:122084996-122203 492	Mmp8	chr9:7558428-7568486
Anpep	chr7:86966688-8698723 8	Mocs1	chr17:49567688-49594 755
Anxa1	chr19:20447923-204651 61	Mphosph9	chr5:124700967-12479 1853
Apbb1ip	chr2:22629846-2273117 3	Mppe1	chr18:67247989-67405 484
Apobec3	chr15:79719961-797388 59	Mrc1	chr2:14151040-142536 50
Aqp9	chr9:70958467-7101109 6	Mreg	chr1:72205806-722588 81

Arhgap5	chr12:53617063-5366883 8	Ms4a2	chr19:11691682-11698 133
Arhgap6	chrX:165233030-16574 2367	Msi2	chr11:88152883-88579 581
Arhgef3	chr14:28051224-2821709 0	Msra	chr14:64741457-65074 740
Arid3a	chr10:79389816-7941775 7	Mtap	chr4:88783273-888269 94
Arid3b	chr9:57638315-5768204 1	Mtap6	chr7:106415956-10648 5647
Arid5b	chr10:67558340-6774147 4	Mtap7	chr10:19868725-20001 396
Arl1	chr10:88194158-8820594 7	Mthfs	chr9:89106027-891350 63
Arl11	chr14:61928589-6193077 3	Mtm1	chrX:68463941-685684 70
Arl15	chr13:114584715-11494 7669	Muc13	chr16:33794122-33820 013
Arl4c	chr1:90594800-9059876 6	Muc20	chr16:32735971-32797 521
Arsb	chr13:94541633-947129 71	Mycn	chr12:12942898-12948 642
Ddef1	chr15:63920100-6421448 1	Myh9	chr15:77591018-77672 545
Asprv1	chr6:86578167-8657969 8	Myl4	chr11:104411976-1044 48533
Asrgl1	chr19:9186208-9210056	Mylk	chr16:34785035-35002 520
Atf3	chr1:192994175-193007 212	Myo1d	chr11:80295628-80593 527
Atg7	chr6:114593118-114810 632	Myo1e	chr9:70055156-702478 74
Atp10a	chr7:65913571-6608479 6	Myst4	chr14:22318991-22491 700
Atp11a	chr8:12757015-1286872 8	1200013P24Rik	chr16:3884618-390478 1
Atp1b1	chr1:166237805-166388 486	Ncf1	chr5:134696129-13470 5495
Atp6v0d1	chr8:108048369-108089 940	Ncf4	chr15:78075240-78093 010
Atp6v0d2	chr4:19803984-1984971 3	Neb	chr2:51992166-521943 18

Atp8a1	chr5:68009379-6823867 0	Nebl	chr2:17268360-176526 95
Auts2	chr5:131913551-133018 213	Nedd9	chr13:41405284-41582 689
Axl	chr7:26541518-2657375 2	Neil3	chr8:54672220-547244 19
Azgp1	chr5:138422748-138431 460	Nek6	chr2:38367216-384430 10
B230118H07Rik	chr2:101400937-101472 685	Nfe2l2	chr2:75513575-755426 98
B3galt5	chr16:96457407-965414 65	Nfil3	chr13:53062577-53076 408
B3gnt5	chr16:19760326-197728 46	Nfkbia	chr12:56590395-56593 634
B4galt4	chr16:38742376-387691 58	Nfxl1	chr5:72904542-729508 84
Basp1	chr15:25293031-253435 19	Nlrp3	chr11:59356187-59380 458
Baz2b	chr2:59737419-5996379 7	Noc3l	chr19:38862617-38893 727
BC017647	chr11:77908174-779708 41	Nod1	chr6:54873935-549226 06
Bcar3	chr3:122122697-122233 100	Nos1ap	chr1:172247626-17251 9980
Bcas1	chr2:170172490-170253 345	Notch2	chr3:97817460-979542 90
Bcat1	chr6:144942354-145024 677	Nov	chr15:54577482-54585 316
Bcl2	chr1:108434755-108610 867	Nr4a3	chr4:48064119-480962 24
Bcl2l11	chr2:127951773-127988 283	Nrg1	chr8:32928499-330286 75
Bcl6	chr16:23965137-239886 98	Nrgn	chr9:37352077-373603 30
Bdh1	chr16:31422382-314589 87	Nrp1	chr8:130882972-13102 9375
Bmp2k	chr5:97426707-9754061 5	Nt5e	chr9:88222446-882669 27
Bmpr1a	chr14:35224253-353157 32	Nucb2	chr7:123647887-12368 4102
Bms1	chr6:118333398-118369 435	Nudt7	chr8:116657478-11667 6205

Bpil2	chr10:85422435-8547460 5	Numb	chr12:85136388-85183 293
Bst1	chr5:44210131-4423470 7	Nupr1	chr7:133766759-13376 8984
Btd	chr14:32454242-324813 83	Obfc2a	chr1:51526331-515352 43
Btla	chr16:45224449-452530 08	Ocm	chr5:144780675-14481 1478
Bzw2	chr12:36818421-368834 12	Odc1	chr12:17551678-17558 308
C3ar1	chr6:122797157-122806 175	Olfml2b	chr1:172574662-17261 2920
C5ar1	chr7:16832091-1684488 9	Orai2	chr5:136623331-13664 6526
Camk2d	chr3:126299890-126547 972	Osbpl5	chr7:150874666-15092 7868
Cast	chr13:74831733-749453 69	1810007P19Rik	chr4:125735808-12576 6578
Cbwd1	chr19:24994405-250361 06	P2rx3	chr2:84836708-848759 91
Ccdc101	chr7:133792822-133816 293	P2rx7	chr5:123093919-12314 1441
Ccdc109a	chr10:58909731-590794 40	P2ry1	chr3:60806716-608129 01
Ccdc109b	chr3:129617877-129673 124	P2ry10	chrX:104284673-10430 0309
Ccdc85a	chr11:28285691-284842 96	P2ry14	chr3:58810899-591481 78
Ccdc88c	chr12:102150909-10226 7193	Pan3	chr5:148242155-14836 0077
Ccl3	chr11:83461344-834628 80	Paqr5	chr9:61801544-618745 97
Ccl4	chr11:83476085-834781 85	Parp8	chr13:117643630-1178 14323
Ccl6	chr11:83395562-834371 95	Parvb	chr15:84062472-84146 039
Ccl9	chr11:83386418-833921 38	Pax8	chr2:24276079-243310 86
Ccne1	chr7:38883002-3889250 9	Pcnx	chr12:82961016-83101 911
Ccr1	chr9:123876958-123883 525	Pcp4l1	chr1:173103394-17312 6399

Ccr2	chr9:124016921-124023879	Pcsk9	chr4:106114938-10613 6930
Ccr3	chr9:123936702-123946421	Pcyt1a	chr16:32431006-32475 151
Ccr5	chr9:124036281-124041922	Pdcd4	chr19:53966720-54107 768
Cd200r3	chr16:44943790-44981493	Pde11a	chr2:75829188-761767 10
Cd226	chr18:89366818-89439719	Pde4b	chr4:101927346-10227 9867
Cd27	chr6:125174229-125189952	Pde7a	chr3:19125917-192113 22
Cd28	chr1:60803231-60830203	Pde9a	chr17:31523178-31613 254
Cd2ap	chr17:42929899-43013373	Pdgfrb	chr18:61204803-61244 721
Cd33	chr7:50782825-50788541	Pdpn	chr4:142857324-14288 9410
Cd34	chr1:196765014-196803153	Peli2	chr14:48740543-48880 558
Cd37	chr7:52489001-52494209	Per2	chr1:93312558-933559 05
Cd38	chr5:44260065-44303613	Pex26	chr6:121133684-12114 6203
Cd44	chr2:102651299-102741822	Pfkfb4	chr9:108893946-10893 4739
Cd48	chr1:173612185-173635388	Pfn2	chr3:57645816-576516 79
Cd53	chr3:106561778-106593067	Phb	chr11:95528270-95542 087
Cd63	chr10:128345974-128349874	Phf15	chr11:51626957-51670 983
Cd93	chr2:148262386-148269271	Phlda1	chr10:110943341-1109 45705
Cdc42ep3	chr17:79733364-79754431	Phldb2	chr16:45746343-45844 491
Cdc42ep4	chr11:113588163-113613129	Phtf2	chr5:20264481-203879 42
Cdh1	chr8:109127267-109194146	Piga	chrX:160857718-16087 1847
Cdh17	chr4:11685303-11745052	Pigu	chr2:155103987-15518 3160

Pctk2	chr10:92623620-9270408 7	Pip5k1b	chr19:24369285-24630 317
Cdk6	chr5:3344311-3522225	Pira1	chr7:3788411-3867103
Cdkn2b	chr4:88952197-8895694 1	Pkd1l2	chr8:119519578-11960 6349
Cdyl2	chr8:119092623-119256 891	Pkib	chr10:57351786-57460 918
Ch25h	chr19:34548273-345496 25	Pla2g4a	chr1:151676751-15180 8414
Chn2	chr6:53989925-5438021 5	Pla2g4c	chr7:13911015-139460 17
Chsy1	chr7:73254400-7331868 4	Plcb4	chr2:135567565-13583 8804
Clec12a	chr6:129300261-129315 321	Plcg1	chr2:160557045-16069 8726
Clec4d	chr6:123212124-123225 286	Pld1	chr3:27837601-280322 84
Clec4e	chr6:123231806-123239 889	Pld4	chr12:113998865-1140 07197
Clec4n	chr6:123179860-123197 042	Plekha5	chr6:140372619-14054 3427
Clec5a	chr6:40524896-4053580 4	Plekha6	chr1:135142673-13520 0012
Clec7a	chr6:129411608-129422 795	Plekha7	chr7:123267631-12333 3355
Clip1	chr5:124029078-124134 300	9430067K14Ri k	chr1:64835694-650033 98
Cln5	chr14:103469432-10347 6847	Plxdc2	chr2:16277948-166737 42
Clnk	chr5:39097698-3926793 2	Plxnd1	chr6:115904828-11594 5023
Cma1	chr14:56560287-565634 98	Pnpla2	chr7:148641086-14864 6642
Cmah	chr13:24419288-245691 54	Pparg	chr6:115311238-11544 0419
Cnga1	chr5:72994935-7303399 1	Ppm1k	chr6:57456495-574854 20
Cnksr3	chr10:3134303-3227479	Ppm1m	chr9:106097283-10610 1564
Cnnm2	chr19:46836098-469530 70	Ppp1cb	chr5:32761342-327960 85
Cobl	chr11:12136678-123649	Pqlc1	chr18:80451983-80489

	63		463
Cobll1	chr2:64926395-6507668 3	Prickle1	chr15:93329544-93426 322
Cpa3	chr3:20115514-2014208 0	Prkar2b	chr12:32643343-32746 144
Cpd	chr11:76590709-766605 10	Prkcb1	chr7:129432638-12977 7915
Cradd	chr10:94637379-947867 31	Prkcd	chr14:31408539-31439 394
Creb5	chr6:53523367-5364582 6	Prkch	chr12:74686027-74879 171
Crebl2	chr6:134780216-134807 901	Prkcq	chr2:11094008-112228 53
Crhbp	chr13:96201330-962147 86	Prmt5	chr14:55126018-55136 307
Crim1	chr17:78599587-787759 32	Prrc1	chr18:57514386-57552 373
Crisp1	chr17:40430706-404561 56	Prss28	chr17:25445590-25448 821
Crispld2	chr8:122516369-122576 693	Prss34	chr17:25435338-25437 106
Crtap	chr9:114277441-114299 830	Prtn3	chr10:79342411-79345 917
Cryba4	chr5:112675552-112681 522	Psat1	chr19:15979612-15999 549
Csf1	chr3:107543965-107563 387	Psd3	chr8:70212980-704984 73
Ctla2a	chr13:61035515-610379 86	Pstpip1	chr9:55937782-559766 97
Ctps	chr4:120212472-120242 881	Pstpip2	chr18:78033288-78121 618
Ctsg	chr14:56718717-567214 11	Ptafr	chr4:132119981-13213 6781
Cxcl2	chr5:91332924-9133496 4	Ptger2	chr14:45607785-45623 495
Cxcl3	chr5:91215128-9121711 6	Ptger3	chr3:157229855-15730 7722
Cxcr4	chr1:130484775-130488 876	Ptgs2	chr1:151947253-15195 5142
Cyfip2	chr11:46007350-461263 61	Ptk2b	chr14:66772093-66899 889
Cyp2b10	chr7:26682682-2671164	Ptplad2	chr4:88058833-880848

	3		30
Cysltr1	chrX:103771855-103905896	Ptplb	chr16:35022506-3510961
Pscdbp	chr2:57981549-58012533	Ptpn12	chr5:20492462-2056161 5
D10Bwg1379e	chr10:18307816-1846364	Ptprc	chr1:139959437-1400718 82
D630023F18Rik	chr1:65153886-6516976 7	Ptprg	chr14:12386066-1307453 553
Dab2ip	chr2:35405053-3558651 4	Pygl	chr12:71291801-7132867 670
Dach1	chr14:98186065-9856876 62	Qrsl1	chr10:43593995-4362154 542
Dapp1	chr3:137593969-137644513	Rab40b	chr11:121217434-1212149565
Ddc	chr11:11714103-1179814 47	Rab8b	chr9:66691470-6676751 12
Depdc1b	chr13:109106530-1091799751	Rad17	chr13:101387118-101421014
Dgat2	chr7:106302172-1063311223	Rai14	chr15:10498732-1064438 386
Dgki	chr6:36796021-3724997 6	Ralgps2	chr1:158734296-1588669757
Dhrs1	chr14:56357856-5636452 521	Ranbp17	chr11:33111793-33413746
Diablo	chr5:123960468-123974173	Rap1b	chr10:117251652-117283030
Dio2	chr12:91962991-9197687 78	Rap2a	chr14:120877682-1209 06414
Dip2a	chr10:75725793-75808036	Rasa2	chr9:96439718-9653191 22
Diras2	chr13:52599743-5262620 05	Rasgrp1	chr2:117105728-1171688613
Dirc2	chr16:35694988-3576944 42	Rb1	chr14:73595308-7372559 598
Dlc1	chr8:37630792-3801549 6	Rbm19	chr5:120566521-120648980
Dnahc8	chr17:30763880-3101220 09	Rbms1	chr2:60590009-6080126 61
Dnajc10	chr2:80155622-8019421 2	Rbpms	chr8:34893115-3504031 13
Dnajc6	chr4:101169252-101315	Rcbtb2	chr14:73542316-73583

150

	404		861
Dock7	chr4:98603355-9878760 6	Rcl1	chr19:29175864-29218 333
Dync2h1	chr9:6928502-7177046	Rell1	chr5:64300136-643601 36
E2f8	chr7:56121798-5613641 1	Reps2	chrX:158849885-15908 1533
Ednra	chr8:80186927-8024835 1	Rere	chr4:149656024-14999 6075
Ehf	chr2:103103589-103143 353	Rftn1	chr17:50132631-50329 822
Eif4e3	chr6:99575130-9961676 5	Rfxdc2	chr9:72380046-724707 56
Ela2	chr10:79349056-793509 61	Rgl1	chr1:154364659-15447 2241
Elk3	chr10:92710160-927739 04	Rgs1	chr1:146091798-14609 6234
Elovl6	chr3:129235303-129341 411	Rgs10	chr7:135517138-13556 1686
Elovl7	chr13:109004597-10907 7301	Rgs12	chr5:35292096-353762 42
Emr1	chr17:57498108-576229 52	Rhbdf2	chr11:116459478-1164 88333
Entpd1	chr19:40734283-408160 92	Rhob	chr12:8504564-850679 1
Epsti1	chr14:78304045-784024 63	Rhoj	chr12:76409299-76502 442
D5Wsu178e	chr5:30559157-3059897 6	Rin3	chr12:103521283-1036 29064
Ern1	chr11:106258933-10634 9110	Riok2	chr17:17511295-17531 863
Errfi1	chr4:150229199-150242 989	Rnd1	chr15:98499635-98507 892
Esr1	chr10:5342779-5734495	Rora	chr9:68501608-692268 35
Exoc6	chr19:37624907-377577 39	Rpn1	chr6:88034466-880552 98
Ext1	chr15:52899817-531777 38	Rpp40	chr13:35986972-35998 216
Ext2	chr2:93535787-9366272 5	Rps6ka1	chr4:133403205-13344 3714
Extl3	chr14:65670895-657169	Rrp12	chr19:41937340-41970

	43		643
Eya2	chr2:165480797-165597131	Rspry1	chr8:97125840-9718476
F13a1	chr13:36959046-37142113	Rtn4r	chr16:18127798-18152501
Fabp5	chr3:10012605-10016610	Satb1	chr17:51875511-51972615
5730446C15Rik	chr19:21727798-21760127	Sc5d	chr9:42062259-42072383
1810015C04Rik	chr15:25773018-25903442	Scp2	chr4:107716434-107791152
AI847670	chr5:20930720-20987942	Sec14l1	chr11:116976485-117020582
Fastkd1	chr2:69524880-69550663	Sec22c	chr9:121589162-121614147
Fbxo17	chr7:29501808-29523159	Serinc5	chr13:93381092-93481901
Fbxo27	chr7:29477867-29484356	Serpine2	chr1:79790895-79855240
Fbxo9	chr9:77929305-77956860	Sgms1	chr19:32197216-32462944
Fcgrt	chr7:52348362-52359192	Sgsh	chr11:119176100-119216824
Fes	chr7:87522643-87532832	Sh2d3c	chr2:32576574-32610527
Fgd4	chr16:16422070-16560289	Sh3kbp1	chrX:156065339-156413849
Fgd6	chr10:93498745-93608084	Sh3tc1	chr5:36039828-36071925
Flt3	chr5:148142316-148212065	Shc1	chr3:89222472-89233951
Fmnl2	chr2:52716901-52993236	Siglecg	chr7:50663649-50673719
Fmo5	chr3:97432726-97459210	Sipa1l2	chr8:127941962-128016610
Fn1	chr1:71632096-71699745	Sirpa	chr2:129418574-129457964
Fnbp1l	chr3:122241636-122322585	Tg	chr15:66502331-66682282
Fndc3a	chr14:72937759-73109810	Slc13a2	chr11:78207986-78235687
Fndc3b	chr3:27315083-2760936	Slc16a10	chr10:39753340-39862

	1		060
Fnip1	chr11:54251680-54331743	Slc17a5	chr9:78384315-78435834
Foxp1	chr6:98875335-99385339	Slc18a1	chr8:71561606-71637616
Frat2	chr19:41920464-41922622	Slc1a4	chr11:20202182-20232716
Frmd5	chr2:121371264-121632793	3110004L20Rik	chr13:34271026-34437051
Furin	chr7:87534079-87550317	Slc22a3	chr17:12612839-12700570
Fut8	chr12:78339090-78576983	Slc23a2	chr2:131878231-131970844
Fyb	chr15:6529870-6613312	Slc25a33	chr4:149118144-149148376
Fyn	chr10:39089604-39285180	Slc26a3	chr12:32123165-32158584
Fzd1	chr5:4753838-4758216	Slc28a2	chr2:122252212-122286866
Gab3	chrX:72233886-72330244	Slc2a3	chr6:122677826-122692763
Gal	chr19:3409916-3414457	Slc35a5	chr16:45139685-45158786
Galnt2	chr8:126755293-126869622	Slc37a3	chr6:39284769-39327706
Galnt3	chr2:65920822-65962850	Slc41a2	chr10:82693883-82800562
Galntl4	chr7:118615174-118923491	Slc4a7	chr14:15535538-15632457
Gata1	chrX:7536385-7544941	Slc7a11	chr3:50168857-50247535
Gch1	chr14:47773569-47809077	Slc9a9	chr9:94570310-95130871
Gcnt2	chr13:40955136-41056261	Smad1	chr8:81862297-81923367
Gda	chr19:21465796-21547151	Sms	chrX:153881885-153929978
Gem	chr4:11631593-11642140	Snx24	chr18:53405315-53550479
Gfi1b	chr2:28464969-28477502	Soat1	chr1:158358238-158404459
Gfod1	chr13:43290887-433995	Socs3	chr11:117827400-1178

	41		30680
A030007L17Rik	chr6:54935088-54942861	Sod1	chr16:90220986-90226569
Gimap1	chr6:48689045-48693794	Sorcs2	chr5:36359829-36740788
Gimap6	chr6:48651581-48658243	Sorl1	chr9:41776571-41932372
OTTMUSG00000000971	chr11:83517557-83519771	Sox4	chr13:29042598-29045551
Gm885	chr11:106612581-106628882	Sp100	chr1:87546624-87606023
Gm962	chr19:5568073-5571261	Specc1	chr11:61890598-62036515
LOC751864	chr3:15296550-15332302	Spn	chr7:134276977-134281331
Gmpr	chr13:45602837-45641750	Spns3	chr11:72311657-72363748
Gng2	chr14:20691780-20796471	Spp1	chr5:104864129-104870072
Gnl2	chr4:124707257-124742901	Spred1	chr2:116947185-117005072
Gnptab	chr10:87842156-87910074	Spry1	chr3:37538871-37543521
Got2	chr8:98388036-98412265	Srpk2	chr5:22940246-23122389
Gpam	chr19:55144223-55173937	Ssbp2	chr13:91600701-91925753
Gpr174	chrX:104451216-104492108	Ssh2	chr11:77029926-77273721
Gpr179	chr11:97193422-97213387	St3gal1	chr15:66934436-67008444
Cask	chrX:13094207-13423667	St7	chr6:17699215-17893022
Gpr56	chr8:97501008-97538108	Stam	chr2:13995738-14069957
Gpr77	chr7:16819933-16829503	Stard10	chr7:108469832-108494826
Gpr84	chr15:103138665-103140869	Rage	chr12:112046007-112079149
Gpsm2	chr3:108456830-108525217	Stxbp5	chr10:9475344-9620838
Gpt2	chr8:88016515-8805145	Stxbp6	chr12:45953472-46175

	7		470
Gramd3	chr18:56591785-5666346	Susd5	chr9:113966471-11400 7851
Gria3	chrX:38754480-3903178	Syk	chr13:52678805-5274416 1
Grlf1	chr7:17079821-17200342	Syne1	chr10:5151833-5326342
Gsr	chr8:34763709-3480863 4	Syngr1	chr15:79921763-79949 931
Gtf2ird2	chr5:134659907-134694 013	Tagap	chr17_random:39676-4 8574
Gtf3a	chr5:147760232-147775 373	Tbc1d10b	chr7:134340973-13435 1982
Gulp1	chr1:44608515-4485368 1	Tbc1d14	chr5:36833252-369288 75
Glt8d3	chr15:93070172-9310555 15	Tbc1d2	chr4:46617261-466630 71
Gypa	chr8:83017943-8303468 4	Tbc1d22b	chr17:29686746-29743 753
H2afy2	chr10:61178010-612466 12	Tbc1d8	chr1:39424695-395355 92
H2-Ob	chr17:34375849-343828 53	Tcf12	chr9:71692058-719596 26
Haao	chr17:84230693-842461 30	Tcf7	chr11:52066105-52095 752
Hacl1	chr14:32420411-324541 51	Tec	chr5:73146961-732596 87
Hao1	chr2:134323096-134380 088	Tfpi	chr2:84273016-843169 32
Havcr2	chr11:46268432-462947 56	Tgfbr2	chr9:115996812-11608 4481
Hbb-b2	chr7:110961037-110962 437	Tgm3	chr2:129838109-12987 6135
Hck	chr2:152934203-152977 177	Thbs1	chr2:117937657-11795 2869
Hdac4	chr1:93829331-9404497 0	Thy1	chr9:43851466-438566 62
Hdac9	chr12:35056361-352137 54	Tjp1	chr7:72441050-725161 30
Hdhd2	chr18:77182853-772109 10	Tle1	chr4:71778175-718619 53
Heatr2	chr5:139626176-139662	Tle3	chr9:61220172-612663

Gene	Location	Gene	Location
	459		04
Heg1	chr16:33684551-33768281	Tlr1	chr5:65316072-65323938
Hgf	chr5:16059367-16125257	Tlr13	chrX:103338613-103355832
Hiatl1	chr13:65166337-65214290	Tlr2	chr3:83640193-83645530
Hip1	chr5:135882387-136020992	Tmco7	chr8:109206967-109375339
Hivep2	chr10:13686184-13871184	Tmed3	chr9:89594040-89599881
Hlx	chr1:186551023-186556372	Tmem154	chr3:84470113-84508497
Hmgn2	chr4:133520653-133523906	Tmem176b	chr6:48783810-48791373
Hnf4a	chr2:163372923-163398643	Tmem2	chr19:21852831-21932817
Hs3st1	chr5:40005173-40035870	Tmem50b	chr16:91574752-91597925
Hsdl2	chr4:59594434-59631566	Tmem51	chr4:141586907-141644016
Hspa9	chr18:35097068-35114005	Tmprss7	chr16:45656429-45693771
1500019G21Rik	chr7:4612122-4636565	Tmsb4x	chrX:163645025-163647150
Iars	chr13:49777498-49829636	Tmtc2	chr10:104624719-105011535
Ibtk	chr9:85580966-85642941	Txndc13	chr2:134420241-134469857
Id2	chr12:25778663-25780957	Tnfaip3	chr10:18720716-18735216
Ide	chr19:37343230-37405103	Tnfrsf1b	chr4:144802270-144836773
Ier3	chr17:35958657-35959856	Tnfrsf26	chr7:150793589-150813750
Ifngr1	chr10:19311763-19330031	Tnfrsf9	chr4:150294263-150320211
Igf1	chr10:87321800-87399792	Tnfsf11	chr14:78677252-78707850
Igf1r	chr7:75097142-75378553	Tnfsf13b	chr8:10006632-10035999
Igf2r	chr17:12875271-130610	Tox	chr4:6614532-6917870

Gene	Coordinates	Gene	Coordinates
Mettl9	chr7:128177958-128220349	Tpst1	chr5:130549195-130611602
Ikzf1	chr11:11586215-11672929	Traf1	chr2:34798777-34817292
Ikzf3	chr11:98326006-98407345	Traf3	chr12:112404758-112505359
Il10ra	chr9:45061921-45077229	Trib1	chr15:59480208-59488654
Il10rb	chr16:91406479-91426079	Trpm2	chr10:77370466-77432617
Il18r1	chr1:40522396-40557699	Tsc22d1	chr14:76815627-76907573
Il1r2	chr1:40133056-40182070	Ttc21a	chr9:119846723-119876911
Il27	chr7:133732808-133738424	Ttc21b	chr2:66022385-66094674
Il2ra	chr2:11564418-11614821	Ttc27	chr17:75117089-75262910
Il4ra	chr7:132695795-132722988	1810054D07Rik	chr4:82866204-82970093
Il6	chr5:30339700-30346508	Txk	chr5:73087216-73144012
Inhba	chr13:16106307-16119044	Txndc16	chr14:45754122-45839069
Inpp4a	chr1:37356737-37467580	Ubac2	chr14:122277827-122420257
Inpp4b	chr8:84239098-84646460	Ugcg	chr4:59202421-59235705
Inpp5a	chr7:146575007-146768696	Upp1	chr11:9018010-9036173
Ipmk	chr10:70810540-70848633	Usp10	chr8:122434751-122481457
Iqgap2	chr13:96397131-96661877	Usp12	chr5:147546387-147606532
Irf8	chr8:123260275-123280592	Usp40	chr1:89841695-89905126
Irs2	chr8:10986963-11054541	Vcl	chr14:21748654-21852895
Isoc1	chr18:58819135-58839224	Vegfc	chr8:55162885-55271808
Itga4	chr2:79095582-79269144	Vipr2	chr12:117316195-1173

	5		84734
Itga6	chr2:71625139-7169481 5	Vit	chr17:78907402-79026 749
Itgav	chr2:83564553-8364707 3	Vldlr	chr19:27291509-27328 721
Itgax	chr7:135273081-135294 171	Tmem49	chr11:86397366-86497 324
Itgb5	chr16:33829750-339494 24	Vps13c	chr9:67688202-678434 41
Itih5	chr2:10075169-1017815 6	Wdfy2	chr14:63456526-63575 723
Itk	chr11:46138651-462030 17	Wipi1	chr11:109434834-1094 72703
Itpr2	chr6:146060001-146450 434	Wwc2	chr8:48912963-490759 05
Itsn1	chr16:91729615-919208 24	Xdh	chr17:74233247-74299 522
Jazf1	chr6:52718061-5301861 8	Zc3h7b	chr15:81575277-81626 699
Jhdm1d	chr6:39068471-3915677 2	Zc3hc1	chr6:30316387-303410 10
Jun	chr4:94715726-9471891 3	Zcchc4	chr5:53174305-532110 04
Kcnj2	chr11:110927477-11093 8139	Zfp36l1	chr12:81208746-81214 000
Kcnk5	chr14:20959279-210010 04	Zfp566	chr7:30862355-308755 29
Kctd12	chr14:103375797-10338 1854	Zfp703	chr8:28087807-280919 34
Ndg1	chr1:21339757-2134228 0	Zfp760	chr17:21844707-21862 603
Kifap3	chr1:165709809-165847 231	Zkscan1	chr5:138526311-13854 9050
Klf13	chr7:71031236-7108380 1		

Appendix B. Core Hoxa9 Target Genes with Co-binding of C/ebpα and Pu.1

Gene name	Locus	Gene name	Locus
1700012B09Rik	chr9:14562637-14575474	Lims1	chr10:57786213-57887439
1810033B17Rik	chr8:3665761-3668905	Litaf	chr16:10959365-10993214
5031414D18Rik	chr14:75415833-75452339	Lmo2	chr2:103798151-103822035
Abcg3	chr5:105364075-105411736	2310010M24Rik	chr2:49643205-49804366
Abr	chr11:76230235-76391229	Maml3	chr3:51491534-51908928
Acot7	chr4:151552208-151645964	Man1a	chr10:53625838-53795602
Acpl2	chr9:96723761-96789841	Map4k4	chr1:39957757-40083155
Adamtsl4	chr3:95480126-95491781	Mapk4	chr18:74088140-74224603
Add3	chr19:53214934-53321889	Mbd2	chr18:70727945-70785785
Ak2	chr4:128670508-128688773	Mboat1	chr13:30228358-30338563
Akr1c18	chr13:4131872-4149877	Mcph1	chr8:18595172-18803188
Alas1	chr9:106135785-106150285	Mdga1	chr17:29964902-30024827
Angpt1	chr15:42256270-42508341	Mdm1	chr10:117578842-117606055
Ank	chr15:27396431-27524662	Mdn1	chr4:32744093-32862192
Tmem16k	chr9:122084996-122203492	Med13l	chr5:119010727-119215446
Anxa1	chr19:20447923-20465161	Megf9	chr4:70092960-70195962
Apbb1ip	chr2:22629846-22731173	Meis1	chr11:18780430-18918972

Aqp9	chr9:70958467-7101109 6	Met	chr6:17413956-1752398 0
Arid5b	chr10:67558340-677414 74	Mgat5	chr1:129101562-129379 549
Arl1	chr10:88194158-882059 47	Mid1	chrX:166123178-16642 8729
Asprv1	chr6:86578167-8657969 8	Mmp12	chr9:7347373-7360461
Atf3	chr1:192994175-193007 212	Mmp13	chr9:7272513-7283333
Atp10a	chr7:65913571-6608479 6	Mmp8	chr9:7558428-7568486
Atp1b1	chr1:166237805-166388 486	Mreg	chr1:72205806-7225888 1
Atp6v0d1	chr8:108048369-108089 940	Ms4a2	chr19:11691682-116981 33
B3galt5	chr16:96457407-965414 65	Mtap	chr4:88783273-8882699 4
B3gnt5	chr16:19760326-197728 46	Mtap7	chr10:19868725-200013 96
B4galt4	chr16:38742376-387691 58	Mthfs	chr9:89106027-8913506 3
Baz2b	chr2:59737419-5996379 7	Myh9	chr15:77591018-776725 45
Bcas1	chr2:170172490-170253 345	Myl4	chr11:104411976-10444 8533
Bcat1	chr6:144942354-145024 677	Myo1d	chr11:80295628-805935 27
Bmp2k	chr5:97426707-9754061 5	Myst4	chr14:22318991-224917 00
Bst1	chr5:44210131-4423470 7	Ncf1	chr5:134696129-134705 495
C5ar1	chr7:16832091-1684488 9	Ncf4	chr15:78075240-780930 10
Ccdc101	chr7:133792822-133816 293	Neb	chr2:51992166-5219431 8
Ccdc109a	chr10:58909731-590794 40	Nebl	chr2:17268360-1765269 5
Ccl3	chr11:83461344-834628 80	Nedd9	chr13:41405284-415826 89
Ccl9	chr11:83386418-833921 38	Nek6	chr2:38367216-3844301 0

160

Ccr3	chr9:123936702-123946421	Nlrp3	chr11:59356187-59380458
Ccr5	chr9:124036281-124041922	Nrg1	chr8:32928499-33028675
Cd200r3	chr16:44943790-44981493	Numb	chr12:85136388-85183293
Cd28	chr1:60803231-60830203	P2rx3	chr2:84836708-84875991
Cd33	chr7:50782825-50788541	Pan3	chr5:148242155-148360077
Cd34	chr1:196765014-196803153	Parp8	chr13:117643630-117814323
Cd53	chr3:106561778-106593067	Parvb	chr15:84062472-84146039
Cd63	chr10:128345974-128349874	Pax8	chr2:24276079-24331086
Cd93	chr2:148262386-148269271	Pcnx	chr12:82961016-83101911
Cdk6	chr5:3344311-3522225	Pcp4l1	chr1:173103394-173126399
Cdkn2b	chr4:88952197-88956941	Pcyt1a	chr16:32431006-32475151
Chsy1	chr7:73254400-73318684	Pde11a	chr2:75829188-76176710
Clec12a	chr6:129300261-129315321	Pde4b	chr4:101927346-102279867
Clec4d	chr6:123212124-123225286	Pde7a	chr3:19125917-19211322
Clec4e	chr6:123231806-123239889	Phf15	chr11:51626957-51670983
Clec4n	chr6:123179860-123197042	Pira1	chr7:3788411-3867103
Clec5a	chr6:40524896-40535804	Pla2g4a	chr1:151676751-151808414
Cln5	chr14:103469432-103476847	Pld4	chr12:113998865-114007197
Cmah	chr13:24419288-24569154	Pnpla2	chr7:148641086-148646642
Cradd	chr10:94637379-94786731	Ppm1m	chr9:106097283-106101564
Creb5	chr6:53523367-53645826	Prickle1	chr15:93329544-93426322

Crhbp	chr13:96201330-9621478 6	Prkar2b	chr12:32643343-3274614 4
Crispld2	chr8:122516369-1225766 93	Prkch	chr12:74686027-7487917 1
Cxcl3	chr5:91215128-9121711 6	Prtn3	chr10:79342411-7934591 7
Cxcr4	chr1:130484775-1304888 76	Ptafr	chr4:132119981-1321367 81
D630023F18Rik	chr1:65153886-6516976 7	Ptger2	chr14:45607785-4562349 5
Dab2ip	chr2:35405053-3558651 4	Ptger3	chr3:157229855-1573077 22
Dapp1	chr3:137593969-137644 513	Ptgs2	chr1:151947253-1519551 42
Dgki	chr6:36796021-3724997 6	Ptplad2	chr4:88058833-8808483 0
Dnajc10	chr2:80155622-8019421 2	Ptpn12	chr5:20492462-2056161 5
E2f8	chr7:56121798-5613641 1	Ptprc	chr1:139959437-140071 882
Ehf	chr2:103103589-103143 353	Qrsl1	chr10:43593995-436215 42
Eif4e3	chr6:99575130-9961676 5	Rbm19	chr5:120566521-120648 980
Elovl7	chr13:109004597-10907 7301	Rbpms	chr8:34893115-3504031 3
Emr1	chr17:57498108-576229 52	Rere	chr4:149656024-149996 075
Ern1	chr11:106258933-10634 9110	Rgl1	chr1:154364659-154472 241
Errfi1	chr4:150229199-150242 989	Rgs1	chr1:146091798-146096 234
Esr1	chr10:5342779-5734495	Rgs12	chr5:35292096-3537624 2
Exoc6	chr19:37624907-377577 39	Rora	chr9:68501608-6922683 5
Ext1	chr15:52899817-531777 38	Rpn1	chr6:88034466-8805529 8
Extl3	chr14:65670895-657169 43	Rrp12	chr19:41937340-419706 43
F13a1	chr13:36959046-371421 13	Rtn4r	chr16:18127798-181525 01

Fabp5	chr3:10012605-1001661 0	Scp2	chr4:107716434-107791 152
1810015C04Rik	chr15:25773018-259034 42	Sipa1l2	chr8:127941962-128016 610
AI847670	chr5:20930720-2098794 2	3110004L20Ri k	chr13:34271026-344370 51
Fbxo9	chr9:77929305-7795686 0	Slc23a2	chr2:131878231-131970 844
Fes	chr7:87522643-8753283 2	Slc25a33	chr4:149118144-149148 376
Fgd4	chr16:16422070-165602 89	Slc28a2	chr2:122252212-122286 866
Fmnl2	chr2:52716901-5299323 6	Slc2a3	chr6:122677826-122692 763
Fndc3a	chr14:72937759-731098 10	Slc35a5	chr16:45139685-451587 86
Fndc3b	chr3:27315083-2760936 1	Slc37a3	chr6:39284769-3932770 6
Foxp1	chr6:98875335-9938533 9	Slc7a11	chr3:50168857-5024753 5
Fut8	chr12:78339090-785769 83	Snx24	chr18:53405315-535504 79
Galnt2	chr8:126755293-126869 622	Sod1	chr16:90220986-902265 69
Gcnt2	chr13:40955136-410562 61	Sorcs2	chr5:36359829-3674078 8
Gda	chr19:21465796-215471 51	Specc1	chr11:61890598-620365 15
Gem	chr4:11631593-1164214 0	Spred1	chr2:116947185-117005 072
Gfod1	chr13:43290887-433995 41	Spry1	chr3:37538871-3754352 1
A030007L17Rik	chr6:54935088-5494286 1	Ssbp2	chr13:91600701-919257 53
Gimap1	chr6:48689045-4869379 4	Ssh2	chr11:77029926-772737 21
Gimap6	chr6:48651581-4865824 3	St3gal1	chr15:66934436-670084 44
OTTMUSG00000 000971	chr11:83517557-835197 71	St7	chr6:17699215-1789302 2
LOC751864	chr3:15296550-1533230 2	Stam	chr2:13995738-1406995 7

Gpam	chr19:55144223-5517393 7	Rage	chr12:112046007-11207 9149
Gpr174	chrX:104451216-10449 2108	Stxbp5	chr10:9475344-9620838
Cask	chrX:13094207-134236 67	Syngr1	chr15:79921763-799499 31
Gpt2	chr8:88016515-8805145 7	Tagap	chr17_random:39676-4 8574
Gria3	chrX:38754480-390317 78	Tbc1d14	chr5:36833252-3692887 5
Grlf1	chr7:17079821-1720034 2	Tbc1d22b	chr17:29686746-297437 53
Gsr	chr8:34763709-3480863 4	Tbc1d8	chr1:39424695-3953559 2
Glt8d3	chr15:93070172-931055 15	Tcf12	chr9:71692058-7195962 6
Gypa	chr8:83017943-8303468 4	Tec	chr5:73146961-7325968 7
H2-Ob	chr17:34375849-343828 53	Tgfbr2	chr9:115996812-116084 481
Hdac4	chr1:93829331-9404497 0	Thbs1	chr2:117937657-117952 869
Hdhd2	chr18:77182853-772109 10	Tle1	chr4:71778175-7186195 3
Hlx	chr1:186551023-186556 372	Tle3	chr9:61220172-6126630 4
Hs3st1	chr5:40005173-4003587 0	Tmed3	chr9:89594040-8959988 1
Hspa9	chr18:35097068-351140 05	Tmem2	chr19:21852831-219328 17
Iars	chr13:49777498-498296 36	Tmtc2	chr10:104624719-10501 1535
Ibtk	chr9:85580966-8564294 1	Tnfsf13b	chr8:10006632-1003599 9
Ifngr1	chr10:19311763-193300 31	Tox	chr4:6614532-6917870
Igf1r	chr7:75097142-7537855 3	Tpst1	chr5:130549195-130611 602
Il10ra	chr9:45061921-4507722 9	Traf1	chr2:34798777-3481729 2
Il27	chr7:133732808-133738 424	Trpm2	chr10:77370466-774326 17

164

Il4ra	chr7:132695795-132722988	Tsc22d1	chr14:76815627-76907573
Inhba	chr13:16106307-16119044	Ttc27	chr17:75117089-75262910
Inpp5a	chr7:146575007-146768696	Txndc16	chr14:45754122-45839069
Iqgap2	chr13:96397131-96661877	Ubac2	chr14:122277827-122420257
Itgav	chr2:83564553-83647073	Ugcg	chr4:59202421-59235705
Itgb5	chr16:33829750-33949424	Usp12	chr5:147546387-147606532
Itpr2	chr6:146060001-146450434	Vcl	chr14:21748654-21852895
Jazf1	chr6:52718061-53018618	Vipr2	chr12:117316195-117384734
Kcnj2	chr11:110927477-110938139	Tmem49	chr11:86397366-86497324
Klf13	chr7:71031236-71083801	Vps13c	chr9:67688202-67843441
Klf6	chr13:5860734-5869639	Wdfy2	chr14:63456526-63575723
Klhl2	chr8:67218658-67373716	Wwc2	chr8:48912963-49075905
Lbp	chr2:158132228-158158588	Xdh	chr17:74233247-74299522
Lbp	chr2:158132228-158158588	Zfp36l1	chr12:81208746-81214000
Lhfpl2	chr13:94827750-94965364		

Reference

Abbas, T., Shibata, E., Park, J., Jha, S., Karnani, N., and Dutta, A. (2010). CRL4(Cdt2) regulates cell proliferation and histone gene expression by targeting PR-Set7/Set8 for degradation. Mol Cell *40*, 9-21.

Agricola, E., Randall, R.A., Gaarenstroom, T., Dupont, S., and Hill, C.S. (2011). Recruitment of TIF1gamma to chromatin via its PHD finger-bromodomain activates its ubiquitin ligase and transcriptional repressor activities. Mol Cell *43*, 85-96.

Aplan, P.D. (2006). Chromosomal translocations involving the MLL gene: molecular mechanisms. DNA Repair (Amst) *5*, 1265-1272.

Arai, S., Yoshimi, A., Shimabe, M., Ichikawa, M., Nakagawa, M., Imai, Y., Goyama, S., and Kurokawa, M. (2011). Evi-1 is a transcriptional target of mixed-lineage leukemia oncoproteins in hematopoietic stem cells. Blood *117*, 6304-6314.

Argiropoulos, B., and Humphries, R.K. (2007). Hox genes in hematopoiesis and leukemogenesis. Oncogene *26*, 6766-6776.

Armstrong, S.A., Staunton, J.E., Silverman, L.B., Pieters, R., den Boer, M.L., Minden, M.D., Sallan, S.E., Lander, E.S., Golub, T.R., and Korsmeyer, S.J. (2002). MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia. Nat Genet *30*, 41-47.

Ayton, P.M., and Cleary, M.L. (2003). Transformation of myeloid progenitors by MLL oncoproteins is dependent on Hoxa7 and Hoxa9. Genes Dev *17*, 2298-2307.

Baker, L.A., Allis, C.D., and Wang, G.G. (2008). PHD fingers in human diseases: disorders arising from misinterpreting epigenetic marks. Mutat Res *647*, 3-12.

Bansal, D., Scholl, C., Frohling, S., McDowell, E., Lee, B.H., Dohner, K., Ernst, P., Davidson, A.J., Daley, G.Q., Zon, L.I.*, et al.* (2006). 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 *103*, 16924-16929.

Basu, A., Rose, K.L., Zhang, J., Beavis, R.C., Ueberheide, B., Garcia, B.A., Chait, B., Zhao, Y., Hunt, D.F., Segal, E.*, et al.* (2009). Proteome-wide prediction of acetylation substrates. Proc Natl Acad Sci U S A *106*, 13785-13790.

Bello, N.F., Lamsoul, I., Heuze, M.L., Metais, A., Moreaux, G., Calderwood, D.A., Duprez, D., Moog-Lutz, C., and Lutz, P.G. (2009). The E3 ubiquitin ligase specificity subunit ASB2beta is a novel regulator of muscle differentiation that targets filamin B to proteasomal degradation. Cell Death Differ *16*, 921-932.

Bernard, O.A., Romana, S.P., Schichman, S.A., Mauchauffe, M., Jonveaux, P., and Berger, R. (1995). Partial duplication of HRX in acute leukemia with trisomy 11. Leukemia *9*, 1487-1490.

Bienz, M. (2006). The PHD finger, a nuclear protein-interaction domain. Trends Biochem Sci *31*, 35-40.

Birke, M., Schreiner, S., Garcia-Cuellar, M.P., Mahr, K., Titgemeyer, F., and Slany, R.K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res *30*, 958-965.

Borrow, J., Shearman, A.M., Stanton, V.P., Jr., Becher, R., Collins, T., Williams, A.J., Dube, I., Katz, F., Kwong, Y.L., Morris, C.*, et al.* (1996). The t(7;11)(p15;p15) translocation in acute myeloid leukaemia fuses the genes for nucleoporin NUP98 and class I homeoprotein HOXA9. Nat Genet *12*, 159-167.

Bursen, A., Schwabe, K., Ruster, B., Henschler, R., Ruthardt, M., Dingermann, T., and Marschalek, R. (2010). The AF4.MLL fusion protein is capable of inducing ALL in mice without requirement of MLL.AF4. Blood *115*, 3570-3579.

Caligiuri, M.A., Strout, M.P., Lawrence, D., Arthur, D.C., Baer, M.R., Yu, F., Knuutila, S., Mrozek, K., Oberkircher, A.R., Marcucci, G.*, et al.* (1998). Rearrangement of ALL1 (MLL) in acute myeloid leukemia with normal cytogenetics. Cancer Res *58*, 55-59.

Caligiuri, M.A., Strout, M.P., Schichman, S.A., Mrozek, K., Arthur, D.C., Herzig, G.P., Baer, M.R., Schiffer, C.A., Heinonen, K., Knuutila, S.*, et al.* (1996). Partial tandem duplication of ALL1 as a recurrent molecular defect in acute myeloid leukemia with trisomy 11. Cancer Res *56*, 1418-1425.

Capili, A.D., Schultz, D.C., Rauscher, I.F., and Borden, K.L. (2001). Solution structure of the PHD domain from the KAP-1 corepressor: structural determinants for PHD, RING and LIM zinc-binding domains. EMBO J *20*, 165-177.

Caslini, C., Yang, Z., El-Osta, M., Milne, T.A., Slany, R.K., and Hess, J.L. (2007). Interaction of MLL amino terminal sequences with menin is required for transformation. Cancer Res *67*, 7275-7283.

Centore, R.C., Havens, C.G., Manning, A.L., Li, J.M., Flynn, R.L., Tse, A., Jin, J., Dyson, N.J., Walter, J.C., and Zou, L. (2010). CRL4(Cdt2)-mediated destruction of the histone methyltransferase Set8 prevents premature chromatin compaction in S phase. Mol Cell *40*, 22-33.

Chakravarty, S., Zeng, L., and Zhou, M.M. (2009). Structure and site-specific recognition of histone H3 by the PHD finger of human autoimmune regulator. Structure *17*, 670-679.

Champagne, K.S., and Kutateladze, T.G. (2009). Structural insight into histone recognition by the ING PHD fingers. Curr Drug Targets *10*, 432-441.

Champagne, K.S., Saksouk, N., Pena, P.V., Johnson, K., Ullah, M., Yang, X.J., Cote, J., and Kutateladze, T.G. (2008). The crystal structure of the ING5 PHD finger in complex with an H3K4me3 histone peptide. Proteins *72*, 1371-1376.

Chang, M.J., Wu, H., Achille, N.J., Reisenauer, M.R., Chou, C.W., Zeleznik-Le, N.J., Hemenway, C.S., and Zhang, W. (2010a). Histone H3 lysine 79 methyltransferase Dot1 is required for immortalization by MLL oncogenes. Cancer Res *70*, 10234-10242.

Chang, P.Y., Hom, R.A., Musselman, C.A., Zhu, L., Kuo, A., Gozani, O., Kutateladze, T.G., and Cleary, M.L. (2010b). Binding of the MLL PHD3 finger to histone H3K4me3 is required for MLL-dependent gene transcription. J Mol Biol *400*, 137-144.

Chen, J., Santillan, D.A., Koonce, M., Wei, W., Luo, R., Thirman, M.J., Zeleznik-Le, N.J., and Diaz, M.O. (2008). Loss of MLL PHD finger 3 is necessary for MLL-ENL-induced hematopoietic stem cell immortalization. Cancer Res *68*, 6199-6207.

Chignola, F., Gaetani, M., Rebane, A., Org, T., Mollica, L., Zucchelli, C., Spitaleri, A., Mannella, V., Peterson, P., and Musco, G. (2009). The solution structure of the first PHD finger of autoimmune regulator in complex with non-modified histone H3 tail reveals the antagonistic role of H3R2 methylation. Nucleic Acids Res *37*, 2951-2961.

Chung, A.S., Guan, Y.J., Yuan, Z.L., Albina, J.E., and Chin, Y.E. (2005). Ankyrin repeat and SOCS box 3 (ASB3) mediates ubiquitination and degradation of tumor necrosis factor receptor II. Mol Cell Biol *25*, 4716-4726.

Cierpicki, T., Risner, L.E., Grembecka, J., Lukasik, S.M., Popovic, R., Omonkowska, M., Shultis, D.D., Zeleznik-Le, N.J., and Bushweller, J.H. (2010). Structure of the MLL CXXC domain-DNA complex and its functional role in MLL-AF9 leukemia. Nat Struct Mol Biol *17*, 62-68.

Coscoy, L., Sanchez, D.J., and Ganem, D. (2001). A novel class of herpesvirus-encoded membrane-bound E3 ubiquitin ligases regulates endocytosis of proteins involved in immune recognition. J Cell Biol *155*, 1265-1273.

Daigle, S.R., Olhava, E.J., Therkelsen, C.A., Majer, C.R., Sneeringer, C.J., Song, J., Johnston, L.D., Scott, M.P., Smith, J.J., Xiao, Y.*, et al.* (2011). Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. Cancer Cell *20*, 53-65.

Dawson, M.A., Prinjha, R.K., Dittmann, A., Giotopoulos, G., Bantscheff, M., Chan, W.I., Robson, S.C., Chung, C.W., Hopf, C., Savitski, M.M.*, et al.* (2011). Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature *478*, 529-533.

de Hoon, M.J., Imoto, S., Nolan, J., and Miyano, S. (2004). Open source clustering software. Bioinformatics *20*, 1453-1454.

Debernardi, S., Lillington, D.M., Chaplin, T., Tomlinson, S., Amess, J., Rohatiner, A., Lister, T.A., and Young, B.D. (2003). Genome-wide analysis of acute myeloid leukemia with normal karyotype reveals a unique pattern of homeobox gene expression distinct from those with translocation-mediated fusion events. Genes Chromosomes Cancer *37*, 149-158.

Deshpande, A.J., Chen, L., Fazio, M., Sinha, A.U., Bernt, K.M., Banka, D., Dias, S., Chang, J., Olhava, E.J., Daigle, S.R.*, et al.* (2013). Leukemic transformation by the MLL-AF6 fusion oncogene requires the H3K79 methyltransferase Dot1l. Blood *121*, 2533-2541.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K., and Zhou, M.M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. Nature *399*, 491-496.

Dhar, S.S., Lee, S.H., Kan, P.Y., Voigt, P., Ma, L., Shi, X., Reinberg, D., and Lee, M.G. (2012). Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. Genes Dev *26*, 2749-2762.

Diks, S.H., Sartori da Silva, M.A., Hillebrands, J.L., Bink, R.J., Versteeg, H.H., van Rooijen, C., Brouwers, A., Chitnis, A.B., Peppelenbosch, M.P., and Zivkovic, D. (2008). d-Asb11 is an essential mediator of canonical Delta-Notch signalling. Nat Cell Biol *10*, 1190-1198.

Dul, B.E., and Walworth, N.C. (2007). The plant homeodomain fingers of fission yeast Msc1 exhibit E3 ubiquitin ligase activity. J Biol Chem *282*, 18397-18406.

Elliott, B., Richardson, C., and Jasin, M. (2005). Chromosomal translocation mechanisms at intronic alu elements in mammalian cells. Mol Cell *17*, 885-894.

Ernst, P., Mabon, M., Davidson, A.J., Zon, L.I., and Korsmeyer, S.J. (2004). An Mll-dependent Hox program drives hematopoietic progenitor expansion. Curr Biol *14*, 2063-2069.

Ernst, P., Wang, J., Huang, M., Goodman, R.H., and Korsmeyer, S.J. (2001). MLL and CREB bind cooperatively to the nuclear coactivator CREB-binding protein. Mol Cell Biol *21*, 2249-2258.

Fiedler, M., Sanchez-Barrena, M.J., Nekrasov, M., Mieszczanek, J., Rybin, V., Muller, J., Evans, P., and Bienz, M. (2008). Decoding of methylated histone H3 tail by the Pygo-BCL9 Wnt signaling complex. Mol Cell *30*, 507-518.

French, C.A., Miyoshi, I., Kubonishi, I., Grier, H.E., Perez-Atayde, A.R., and Fletcher, J.A. (2003). BRD4-NUT fusion oncogene: a novel mechanism in aggressive carcinoma. Cancer Res *63*, 304-307.

Friedman, A.D. (2007). Transcriptional control of granulocyte and monocyte development. Oncogene *26*, 6816-6828.

Giannopoulou, E.G., and Elemento, O. (2011). An integrated ChIP-seq analysis platform with customizable workflows. BMC Bioinformatics *12*, 277.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A.*, et al.* (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531-537.

Gomez, C.A., Ptaszek, L.M., Villa, A., Bozzi, F., Sobacchi, C., Brooks, E.G., Notarangelo, L.D., Spanopoulou, E., Pan, Z.Q., Vezzoni, P.*, et al.* (2000). Mutations in conserved regions of the predicted RAG2 kelch repeats block initiation of V(D)J

recombination and result in primary immunodeficiencies. Mol Cell Biol *20*, 5653-5664.

Gozani, O., Karuman, P., Jones, D.R., Ivanov, D., Cha, J., Lugovskoy, A.A., Baird, C.L., Zhu, H., Field, S.J., Lessnick, S.L.*, et al.* (2003). The PHD finger of the chromatin-associated protein ING2 functions as a nuclear phosphoinositide receptor. Cell *114*, 99-111.

Gregory, G.D., Vakoc, C.R., Rozovskaia, T., Zheng, X., Patel, S., Nakamura, T., Canaani, E., and Blobel, G.A. (2007). Mammalian ASH1L is a histone methyltransferase that occupies the transcribed region of active genes. Mol Cell Biol *27*, 8466-8479.

Grembecka, J., Belcher, A.M., Hartley, T., and Cierpicki, T. (2010). Molecular basis of the mixed lineage leukemia-menin interaction: implications for targeting mixed lineage leukemias. J Biol Chem *285*, 40690-40698.

Grembecka, J., He, S., Shi, A., Purohit, T., Muntean, A.G., Sorenson, R.J., Showalter, H.D., Murai, M.J., Belcher, A.M., Hartley, T.*, et al.* (2012). Menin-MLL inhibitors reverse oncogenic activity of MLL fusion proteins in leukemia. Nat Chem Biol *8*, 277-284.

Gu, Y., Cimino, G., Alder, H., Nakamura, T., Prasad, R., Canaani, O., Moir, D.T., Jones, C., Nowell, P.C., Croce, C.M.*, et al.* (1992). The (4;11)(q21;q23) chromosome translocations in acute leukemias involve the VDJ recombinase. Proc Natl Acad Sci U S A *89*, 10464-10468.

Guibal, F.C., Moog-Lutz, C., Smolewski, P., Di Gioia, Y., Darzynkiewicz, Z., Lutz, P.G., and Cayre, Y.E. (2002). ASB-2 inhibits growth and promotes commitment in myeloid leukemia cells. J Biol Chem *277*, 218-224.

Hatano, Y., Miura, I., Nakamura, T., Yamazaki, Y., Takahashi, N., and Miura, A.B. (1999). Molecular heterogeneity of the NUP98/HOXA9 fusion transcript in myelodysplastic syndromes associated with t(7;11)(p15;p15). Br J Haematol *107*, 600-604.

Hay, R.T. (2005). SUMO: a history of modification. Mol Cell *18*, 1-12.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Herrmann, J., Lerman, L.O., and Lerman, A. (2007). Ubiquitin and ubiquitin-like proteins in protein regulation. Circ Res *100*, 1276-1291.

Hess, J.L. (2004a). Mechanisms of transformation by MLL. Crit Rev Eukaryot Gene Expr *14*, 235-254.

Hess, J.L. (2004b). MLL: a histone methyltransferase disrupted in leukemia. Trends Mol Med *10*, 500-507.

Hess, J.L., Yu, B.D., Li, B., Hanson, R., and Korsmeyer, S.J. (1997). Defects in yolk sac hematopoiesis in Mll-null embryos. Blood *90*, 1799-1806.

Heuze, M.L., Guibal, F.C., Banks, C.A., Conaway, J.W., Conaway, R.C., Cayre, Y.E., Benecke, A., and Lutz, P.G. (2005). ASB2 is an Elongin BC-interacting protein that can assemble with Cullin 5 and Rbx1 to reconstitute an E3 ubiquitin ligase complex. J Biol Chem *280*, 5468-5474.

Heuze, M.L., Lamsoul, I., Baldassarre, M., Lad, Y., Leveque, S., Razinia, Z., Moog-Lutz, C., Calderwood, D.A., and Lutz, P.G. (2008). ASB2 targets filamins A and B to proteasomal degradation. Blood *112*, 5130-5140.

Higa, L.A., Wu, M., Ye, T., Kobayashi, R., Sun, H., and Zhang, H. (2006). CUL4-DDB1 ubiquitin ligase interacts with multiple WD40-repeat proteins and regulates histone methylation. Nat Cell Biol *8*, 1277-1283.

Hom, R.A., Chang, P.Y., Roy, S., Musselman, C.A., Glass, K.C., Selezneva, A.I., Gozani, O., Ismagilov, R.F., Cleary, M.L., and Kutateladze, T.G. (2010). Molecular mechanism of MLL PHD3 and RNA recognition by the Cyp33 RRM domain. J Mol Biol *400*, 145-154.

Hsieh, J.J., Cheng, E.H., and Korsmeyer, S.J. (2003a). Taspase1: a threonine aspartase required for cleavage of MLL and proper HOX gene expression. Cell *115*, 293-303.

Hsieh, J.J., Ernst, P., Erdjument-Bromage, H., Tempst, P., and Korsmeyer, S.J. (2003b). Proteolytic cleavage of MLL generates a complex of N- and C-terminal fragments that confers protein stability and subnuclear localization. Mol Cell Biol *23*, 186-194.

Huang, Y., Sitwala, K., Bronstein, J., Sanders, D., Dandekar, M., Collins, C., Robertson, G., MacDonald, J., Cezard, T., Bilenky, M.*, et al.* (2012). Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood *119*, 388-398.

Hudson, B.P., Martinez-Yamout, M.A., Dyson, H.J., and Wright, P.E. (2000). Solution structure and acetyl-lysine binding activity of the GCN5 bromodomain. J Mol Biol *304*, 355-370.

Hughes, C.M., Rozenblatt-Rosen, O., Milne, T.A., Copeland, T.D., Levine, S.S., Lee, J.C., Hayes, D.N., Shanmugam, K.S., Bhattacharjee, A., Biondi, C.A.*, et al.* (2004). Menin associates with a trithorax family histone methyltransferase complex and with the hoxc8 locus. Mol Cell *13*, 587-597.

Huntsman, D.G., Chin, S.F., Muleris, M., Batley, S.J., Collins, V.P., Wiedemann, L.M., Aparicio, S., and Caldas, C. (1999). MLL2, the second human homolog of the Drosophila trithorax gene, maps to 19q13.1 and is amplified in solid tumor cell lines. Oncogene *18*, 7975-7984.

Inaba, T., Shimazaki, C., Yoneyama, S., Hirai, H., Kikuta, T., Sumikuma, T., Sudo, Y., Yamagata, N., Ashihara, E., Goto, H.*, et al.* (1996). t(7;11) and trilineage myelodysplasia in acute myelomonocytic leukemia. Cancer Genet Cytogenet *86*, 72-75.

Ishii, E., Eguchi, M., Eguchi-Ishimae, M., Yoshida, N., Oda, M., Zaitsu, M., Fujita, I., Miyazaki, S., Hamasaki, Y., and Mizutani, S. (2002). In vitro cleavage of the MLL gene by topoisomerase II inhibitor (etoposide) in normal cord and peripheral blood mononuclear cells. Int J Hematol *76*, 74-79.

Ivanov, A.V., Peng, H., Yurchenko, V., Yap, K.L., Negorev, D.G., Schultz, D.C., Psulkowski, E., Fredericks, W.J., White, D.E., Maul, G.G.*, et al.* (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. Mol Cell *28*, 823-837.

Iwasaki, H., Somoza, C., Shigematsu, H., Duprez, E.A., Iwasaki-Arai, J., Mizuno, S., Arinobu, Y., Geary, K., Zhang, P., Dayaram, T.*, et al.* (2005). Distinctive and indispensable roles of PU.1 in maintenance of hematopoietic stem cells and their differentiation. Blood *106*, 1590-1600.

Jacobson, R.H., Ladurner, A.G., King, D.S., and Tjian, R. (2000). Structure and function of a human TAFII250 double bromodomain module. Science *288*, 1422-1425.

Jang, M.K., Mochizuki, K., Zhou, M., Jeong, H.S., Brady, J.N., and Ozato, K. (2005). The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell *19*, 523-534.

Jeanmougin, F., Wurtz, J.M., Le Douarin, B., Chambon, P., and Losson, R. (1997). The bromodomain revisited. Trends Biochem Sci *22*, 151-153.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. Nat Biotechnol *26*, 1293-1300.

Jin, F., Li, Y., Ren, B., and Natarajan, R. (2011). 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 *108*, 5290-5295.

Jude, C.D., Climer, L., Xu, D., Artinger, E., Fisher, J.K., and Ernst, P. (2007). Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell Stem Cell *1*, 324-337.

Kaltenbach, S., Soler, G., Barin, C., Gervais, C., Bernard, O.A., Penard-Lacronique, V., and Romana, S.P. (2010). NUP98-MLL fusion in human acute myeloblastic leukemia. Blood *116*, 2332-2335.

Kile, B.T., Metcalf, D., Mifsud, S., DiRago, L., Nicola, N.A., Hilton, D.J., and Alexander, W.S. (2001). Functional analysis of Asb-1 using genetic modification in mice. Mol Cell Biol *21*, 6189-6197.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol *14*, R36.

Kim, J., Guermah, M., McGinty, R.K., Lee, J.S., Tang, Z., Milne, T.A., Shilatifard, A., Muir, T.W., and Roeder, R.G. (2009). RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell *137*, 459-471.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J.*, et al.* (2011). Systematic and quantitative assessment of the ubiquitin-modified proteome. Mol Cell *44*, 325-340.

King, R.W., Deshaies, R.J., Peters, J.M., and Kirschner, M.W. (1996). How proteolysis drives the cell cycle. Science *274*, 1652-1659.

Kohroki, J., Fujita, S., Itoh, N., Yamada, Y., Imai, H., Yumoto, N., Nakanishi, T., and Tanaka, K. (2001). ATRA-regulated Asb-2 gene induced in differentiation of HL-60 leukemia cells. FEBS Lett *505*, 223-228.

Kohroki, J., Nishiyama, T., Nakamura, T., and Masuho, Y. (2005). ASB proteins interact with Cullin5 and Rbx2 to form E3 ubiquitin ligase complexes. FEBS Lett *579*, 6796-6802.

Kroon, E., Thorsteinsdottir, U., Mayotte, N., Nakamura, T., and Sauvageau, G. (2001). NUP98-HOXA9 expression in hemopoietic stem cells induces chronic and acute myeloid leukemias in mice. EMBO J *20*, 350-361.

Lallous, N., Legrand, P., McEwen, A.G., Ramon-Maiques, S., Samama, J.P., and Birck, C. (2011). The PHD finger of human UHRF1 reveals a new subgroup of unmethylated histone H3 tail readers. PLoS One *6*, e27599.

Lamsoul, I., Burande, C.F., Razinia, Z., Houles, T.C., Menoret, D., Baldassarre, M., Erard, M., Moog-Lutz, C., Calderwood, D.A., and Lutz, P.G. (2011). Functional and structural insights into ASB2alpha, a novel regulator of integrin-dependent adhesion of hematopoietic cells. J Biol Chem *286*, 30571-30581.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Laughon, A. (1991). DNA binding specificity of homeodomains. Biochemistry *30*, 11357-11367.

Lawrence, H.J., Rozenfeld, S., Cruz, C., Matsukuma, K., Kwong, A., Komuves, L., Buchberg, A.M., and Largman, C. (1999). Frequent co-expression of the HOXA9 and MEIS1 homeobox genes in human myeloid leukemias. Leukemia *13*, 1993-1999.

Lee, S., Roeder, R.G., and Lee, J.W. (2009). Roles of histone H3-lysine 4 methyltransferase complexes in NR-mediated gene transcription. Prog Mol Biol Transl Sci *87*, 343-382.

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

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. Nature *442*, 91-95.

Li, J.Y., Chai, B.X., Zhang, W., Liu, Y.Q., Ammori, J.B., and Mulholland, M.W. (2007). Ankyrin repeat and SOCS box containing protein 4 (Asb-4) interacts with GPS1 (CSN1) and inhibits c-Jun NH2-terminal kinase activity. Cell Signal *19*, 1185-1192.

Li, Y., and Li, H. (2012). Many keys to push: diversifying the 'readership' of plant homeodomain fingers. Acta Biochim Biophys Sin (Shanghai) *44*, 28-39.

Lin, C., Smith, E.R., Takahashi, H., Lai, K.C., Martin-Brown, S., Florens, L., Washburn, M.P., Conaway, J.W., Conaway, R.C., and Shilatifard, A. (2010). AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. Mol Cell *37*, 429-437.

Liu, H., Cheng, E.H., and Hsieh, J.J. (2007a). Bimodal degradation of MLL by SCFSkp2 and APCCdc20 assures cell cycle execution: a critical regulatory circuit lost in leukemogenic MLL fusions. Genes Dev *21*, 2385-2398.

Liu, H., Takeda, S., Kumar, R., Westergard, T.D., Brown, E.J., Pandita, T.K., Cheng, E.H., and Hsieh, J.J. (2010). Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint. Nature *467*, 343-346.

Liu, L., Qin, S., Zhang, J., Ji, P., Shi, Y., and Wu, J. (2012). Solution structure of an atypical PHD finger in BRPF2 and its interaction with DNA. J Struct Biol *180*, 165-173.

Liu, Y., Subrahmanyam, R., Chakraborty, T., Sen, R., and Desiderio, S. (2007b). A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. Immunity *27*, 561-571.

Lovett, B.D., Strumberg, D., Blair, I.A., Pang, S., Burden, D.A., Megonigal, M.D., Rappaport, E.F., Rebbeck, T.R., Osheroff, N., Pommier, Y.G.*, et al.* (2001). Etoposide metabolites enhance DNA topoisomerase II cleavage near leukemia-associated MLL translocation breakpoints. Biochemistry *40*, 1159-1170.

Lu, Z., Xu, S., Joazeiro, C., Cobb, M.H., and Hunter, T. (2002). The PHD domain of MEKK1 acts as an E3 ubiquitin ligase and mediates ubiquitination and degradation of ERK1/2. Mol Cell *9*, 945-956.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics *27*, 1696-1697.

Mann, R.S., Lelli, K.M., and Joshi, R. (2009). Hox specificity unique roles for cofactors and collaborators. Curr Top Dev Biol *88*, 63-101.

Mansfield, R.E., Musselman, C.A., Kwan, A.H., Oliver, S.S., Garske, A.L., Davrazou, F., Denu, J.M., Kutateladze, T.G., and Mackay, J.P. (2011). Plant homeodomain (PHD) fingers of CHD4 are histone H3-binding modules with preference for unmodified H3K4 and methylated H3K9. J Biol Chem *286*, 11779-11791.

Marschalek, R. (2011). Mechanisms of leukemogenesis by MLL fusion proteins. Br J Haematol *152*, 141-154.

Martin, D.G., Baetz, K., Shi, X., Walter, K.L., MacDonald, V.E., Wlodarski, M.J., Gozani, O., Hieter, P., and Howe, L. (2006). The Yng1p plant homeodomain finger is a methyl-histone binding module that recognizes lysine 4-methylated histone H3. Mol Cell Biol *26*, 7871-7879.

Martin, M.E., Milne, T.A., Bloyer, S., Galoian, K., Shen, W., Gibbs, D., Brock, H.W., Slany, R., and Hess, J.L. (2003). Dimerization of MLL fusion proteins immortalizes hematopoietic cells. Cancer Cell *4*, 197-207.

Matsushita, H., Nakajima, H., Nakamura, Y., Tsukamoto, H., Tanaka, Y., Jin, G., Yabe, M., Asai, S., Ono, R., Nosaka, T.*, et al.* (2008). C/EBPalpha and C/EBPvarepsilon induce the monocytic differentiation of myelomonocytic cells with the MLL-chimeric fusion gene. Oncogene *27*, 6749-6760.

Matthews, A.G., Kuo, A.J., Ramon-Maiques, S., Han, S., Champagne, K.S., Ivanov, D., Gallardo, M., Carney, D., Cheung, P., Ciccone, D.N.*, et al.* (2007). RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. Nature *450*, 1106-1110.

Matthews, J.M., Bhati, M., Lehtomaki, E., Mansfield, R.E., Cubeddu, L., and Mackay, J.P. (2009). It takes two to tango: the structure and function of LIM, RING, PHD and MYND domains. Curr Pharm Des *15*, 3681-3696.

McDaneld, T.G., and Spurlock, D.M. (2008). Ankyrin repeat and suppressor of cytokine signaling (SOCS) box-containing protein (ASB) 15 alters differentiation of mouse C2C12 myoblasts and phosphorylation of mitogen-activated protein kinase and Akt. J Anim Sci *86*, 2897-2902.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol *28*, 495-501.

McMahon, K.A., Hiew, S.Y., Hadjur, S., Veiga-Fernandes, H., Menzel, U., Price, A.J., Kioussis, D., Williams, O., and Brady, H.J. (2007). Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. Cell Stem Cell *1*, 338-345.

Meyer, C., Kowarz, E., Hofmann, J., Renneville, A., Zuna, J., Trka, J., Ben Abdelali, R., Macintyre, E., De Braekeleer, E., De Braekeleer, M.*, et al.* (2009). New insights to the MLL recombinome of acute leukemias. Leukemia *23*, 1490-1499.

Meyer, C., Schneider, B., Jakob, S., Strehl, S., Attarbaschi, A., Schnittger, S., Schoch, C., Jansen, M.W., van Dongen, J.J., den Boer, M.L.*, et al.* (2006). The MLL recombinome of acute leukemias. Leukemia *20*, 777-784.

Miller, T.C., Rutherford, T.J., Johnson, C.M., Fiedler, M., and Bienz, M. (2010). Allosteric remodelling of the histone H3 binding pocket in the Pygo2 PHD finger triggered by its binding to the B9L/BCL9 co-factor. J Mol Biol *401*, 969-984.

Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D., and Hess, J.L. (2002). MLL targets SET domain methyltransferase activity to Hox gene promoters. Mol Cell *10*, 1107-1117.

Milne, T.A., Dou, Y., Martin, M.E., Brock, H.W., Roeder, R.G., and Hess, J.L. (2005). MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A *102*, 14765-14770.

Mo, R., Rao, S.M., and Zhu, Y.J. (2006). Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. J Biol Chem *281*, 15714-15720.

Mohan, M., Herz, H.M., Smith, E.R., Zhang, Y., Jackson, J., Washburn, M.P., Florens, L., Eissenberg, J.C., and Shilatifard, A. (2011). The COMPASS family of H3K4 methylases in Drosophila. Mol Cell Biol *31*, 4310-4318.

Monroe, S.C., Jo, S.Y., Sanders, D.S., Basrur, V., Elenitoba-Johnson, K.S., Slany, R.K., and Hess, J.L. (2010). MLL-AF9 and MLL-ENL alter the dynamic association of transcriptional regulators with genes critical for leukemia. Exp Hematol.

Monroe, S.C., Jo, S.Y., Sanders, D.S., Basrur, V., Elenitoba-Johnson, K.S., Slany, R.K., and Hess, J.L. (2011). MLL-AF9 and MLL-ENL alter the dynamic association of transcriptional regulators with genes critical for leukemia. Exp Hematol *39*, 77-86 e71-75.

Moore, M.A., Chung, K.Y., Plasilova, M., Schuringa, J.J., Shieh, J.H., Zhou, P., and Morrone, G. (2007). NUP98 dysregulation in myeloid leukemogenesis. Ann N Y Acad Sci *1106*, 114-142.

Mujtaba, S., He, Y., Zeng, L., Farooq, A., Carlson, J.E., Ott, M., Verdin, E., and Zhou, M.M. (2002). Structural basis of lysine-acetylated HIV-1 Tat recognition by PCAF bromodomain. Mol Cell *9*, 575-586.

Mujtaba, S., He, Y., Zeng, L., Yan, S., Plotnikova, O., Sachchidanand, Sanchez, R., Zeleznik-Le, N.J., Ronai, Z., and Zhou, M.M. (2004). Structural mechanism of the bromodomain of the coactivator CBP in p53 transcriptional activation. Mol Cell *13*, 251-263.

Mujtaba, S., Zeng, L., and Zhou, M.M. (2007). Structure and acetyl-lysine recognition of the bromodomain. Oncogene *26*, 5521-5527.

Mullighan, C.G., Kennedy, A., Zhou, X., Radtke, I., Phillips, L.A., Shurtleff, S.A., and Downing, J.R. (2007). 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 *21*, 2000-2009.

Muntean, A.G., Giannola, D., Udager, A.M., and Hess, J.L. (2008). The PHD fingers of MLL block MLL fusion protein-mediated transformation. Blood *112*, 4690-4693.
Muntean, A.G., and Hess, J.L. (2012). The pathogenesis of mixed-lineage leukemia. Annu Rev Pathol *7*, 283-301.

Muntean, A.G., Tan, J., Sitwala, K., Huang, Y., Bronstein, J., Connelly, J.A., Basrur, V., Elenitoba-Johnson, K.S., and Hess, J.L. (2010). The PAF complex synergizes with MLL fusion proteins at HOX loci to promote leukemogenesis. Cancer Cell *17*, 609-621.

Nakamura, T., Largaespada, D.A., Lee, M.P., Johnson, L.A., Ohyashiki, K., Toyama, K., Chen, S.J., Willman, C.L., Chen, I.M., Feinberg, A.P.*, et al.* (1996). Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. Nat Genet *12*, 154-158.

Nakamura, T., Mori, T., Tada, S., Krajewski, W., Rozovskaia, T., Wassell, R., Dubois, G., Mazo, A., Croce, C.M., and Canaani, E. (2002). ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell *10*, 1119-1128.

Nakamura, Y., Umehara, T., Nakano, K., Jang, M.K., Shirouzu, M., Morita, S., Uda-Tochio, H., Hamana, H., Terada, T., Adachi, N.*, et al.* (2007). Crystal structure of the human BRD2 bromodomain: insights into dimerization and recognition of acetylated histone H4. J Biol Chem *282*, 4193-4201.

Nakanishi, H., Nakamura, T., Canaani, E., and Croce, C.M. (2007). ALL1 fusion proteins induce deregulation of EphA7 and ERK phosphorylation in human acute leukemias. Proc Natl Acad Sci U S A *104*, 14442-14447.

Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem *75*, 4646-4658.

Ng, S.B., Bigham, A.W., Buckingham, K.J., Hannibal, M.C., McMillin, M.J., Gildersleeve, H.I., Beck, A.E., Tabor, H.K., Cooper, G.M., Mefford, H.C.*, et al.* (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nat Genet *42*, 790-793.

Nie, L., Zhao, Y., Wu, W., Yang, Y.Z., Wang, H.C., and Sun, X.H. (2011). Notch-induced Asb2 expression promotes protein ubiquitination by forming non-canonical E3 ligase complexes. Cell Res *21*, 754-769.

Nitiss, J.L. (2009). Targeting DNA topoisomerase II in cancer chemotherapy. Nat Rev Cancer *9*, 338-350.

Noordzij, J.G., de Bruin-Versteeg, S., Verkaik, N.S., Vossen, J.M., de Groot, R., Bernatowska, E., Langerak, A.W., van Gent, D.C., and van Dongen, J.J. (2002). The immunophenotypic and immunogenotypic B-cell differentiation arrest in bone marrow of RAG-deficient SCID patients corresponds to residual recombination activities of mutated RAG proteins. Blood *100*, 2145-2152.

O'Connell, S., Wang, L., Robert, S., Jones, C.A., Saint, R., and Jones, R.S. (2001). Polycomblike PHD fingers mediate conserved interaction with enhancer of zeste protein. J Biol Chem *276*, 43065-43073.

Oda, H., Hubner, M.R., Beck, D.B., Vermeulen, M., Hurwitz, J., Spector, D.L., and Reinberg, D. (2010). Regulation of the histone H4 monomethylase PR-Set7 by CRL4 Cdt2-mediated PCNA-dependent degradation during DNA damage. Mol Cell *40*, 364-376.

Oettinger, M.A., Schatz, D.G., Gorka, C., and Baltimore, D. (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science *248*, 1517-1523.

Okada, Y., Feng, Q., Lin, Y., Jiang, Q., Li, Y., Coffield, V.M., Su, L., Xu, G., and Zhang, Y. (2005). hDOT1L links histone methylation to leukemogenesis. Cell *121*, 167-178.

Org, T., Chignola, F., Hetenyi, C., Gaetani, M., Rebane, A., Liiv, I., Maran, U., Mollica, L., Bottomley, M.J., Musco, G.*, et al.* (2008). The autoimmune regulator PHD finger binds to non-methylated histone H3K4 to activate gene expression. EMBO Rep *9*, 370-376.

Owen, D.J., Ornaghi, P., Yang, J.C., Lowe, N., Evans, P.R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, A.A. (2000). The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. EMBO J *19*, 6141-6149.

Page, R.D. (1996). TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci *12*, 357-358.

Park, S., Osmers, U., Raman, G., Schwantes, R.H., Diaz, M.O., and Bushweller, J.H. (2010). The PHD3 domain of MLL acts as a CYP33-regulated switch between MLL-mediated activation and repression. Biochemistry *49*, 6576-6586.

Pena, P.V., Davrazou, F., Shi, X., Walter, K.L., Verkhusha, V.V., Gozani, O., Zhao, R., and Kutateladze, T.G. (2006). Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature *442*, 100-103.

Pena, P.V., Musselman, C.A., Kuo, A.J., Gozani, O., and Kutateladze, T.G. (2009). NMR assignments and histone specificity of the ING2 PHD finger. Magn Reson Chem *47*, 352-358.

Pickart, C.M. (2001). Mechanisms underlying ubiquitination. Annu Rev Biochem *70*, 503-533.

Pineault, N., Helgason, C.D., Lawrence, H.J., and Humphries, R.K. (2002). Differential expression of Hox, Meis1, and Pbx1 genes in primitive cells throughout murine hematopoietic ontogeny. Exp Hematol *30*, 49-57.

Qiu, Y., Liu, L., Zhao, C., Han, C., Li, F., Zhang, J., Wang, Y., Li, G., Mei, Y., Wu, M.*, et al.* (2012). Combinatorial readout of unmodified H3R2 and acetylated H3K14 by the tandem PHD finger of MOZ reveals a regulatory mechanism for HOXA9 transcription. Genes Dev *26*, 1376-1391.

Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S.A., Flynn, R.A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature *470*, 279-283.

Rajakumara, E., Wang, Z., Ma, H., Hu, L., Chen, H., Lin, Y., Guo, R., Wu, F., Li, H., Lan, F.*, et al.* (2011). PHD finger recognition of unmodified histone H3R2 links UHRF1 to regulation of euchromatic gene expression. Mol Cell *43*, 275-284.

Ramon-Maiques, S., Kuo, A.J., Carney, D., Matthews, A.G., Oettinger, M.A., Gozani, O., and Yang, W. (2007). The plant homeodomain finger of RAG2 recognizes histone H3 methylated at both lysine-4 and arginine-2. Proc Natl Acad Sci U S A *104*, 18993-18998.

Rawat, V.P., Thoene, S., Naidu, V.M., Arseni, N., Heilmeier, B., Metzeler, K., Petropoulos, K., Deshpande, A., Quintanilla-Martinez, L., Bohlander, S.K.*, et al.* (2008). Overexpression of CDX2 perturbs HOX gene expression in murine progenitors depending on its N-terminal domain and is closely correlated with deregulated HOX gene expression in human acute myeloid leukemia. Blood *111*, 309-319.

Reichel, M., Gillert, E., Angermuller, S., Hensel, J.P., Heidel, F., Lode, M., Leis, T., Biondi, A., Haas, O.A., Strehl, S.*, et al.* (2001). Biased distribution of chromosomal breakpoints involving the MLL gene in infants versus children and adults with t(4;11) ALL. Oncogene *20*, 2900-2907.

Reynoird, N., Schwartz, B.E., Delvecchio, M., Sadoul, K., Meyers, D., Mukherjee, C., Caron, C., Kimura, H., Rousseaux, S., Cole, P.A.*, et al.* (2010). Oncogenesis by sequestration of CBP/p300 in transcriptionally inactive hyperacetylated chromatin domains. EMBO J *29*, 2943-2952.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140.

Ruault, M., Brun, M.E., Ventura, M., Roizes, G., and De Sario, A. (2002). MLL3, a new human member of the TRX/MLL gene family, maps to 7q36, a chromosome region frequently deleted in myeloid leukaemia. Gene *284*, 73-81.

Saigo, K., Yoshida, K., Ikeda, R., Sakamoto, Y., Murakami, Y., Urashima, T., Asano, T., Kenmochi, T., and Inoue, I. (2008). Integration of hepatitis B virus DNA into the myeloid/lymphoid or mixed-lineage leukemia (MLL4) gene and rearrangements of MLL4 in human hepatocellular carcinoma. Hum Mutat *29*, 703-708.

Sanchez, R., Pieper, U., Melo, F., Eswar, N., Marti-Renom, M.A., Madhusudhan, M.S., Mirkovic, N., and Sali, A. (2000). Protein structure modeling for structural genomics. Nat Struct Biol *7 Suppl*, 986-990.

Sanchez, R., and Zhou, M.M. (2009). The role of human bromodomains in chromatin biology and gene transcription. Curr Opin Drug Discov Devel *12*, 659-665.

Sartori da Silva, M.A., Tee, J.M., Paridaen, J., Brouwers, A., Runtuwene, V., Zivkovic, D., Diks, S.H., Guardavaccaro, D., and Peppelenbosch, M.P. (2010). Essential Role for the d-Asb11 cul5 Box Domain for Proper Notch Signaling and Neural Cell Fate Decisions In Vivo. PLoS One *5*, e14023.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev *16*, 919-932.

Schultz, D.C., Friedman, J.R., and Rauscher, F.J., 3rd (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev *15*, 428-443.

Schwarz, K., Gauss, G.H., Ludwig, L., Pannicke, U., Li, Z., Lindner, D., Friedrich, W., Seger, R.A., Hansen-Hagge, T.E., Desiderio, S.*, et al.* (1996). RAG mutations in human B cell-negative SCID. Science *274*, 97-99.

Shamay, M., Barak, O., Doitsh, G., Ben-Dor, I., and Shaul, Y. (2002). Hepatitis B virus pX interacts with HBXAP, a PHD finger protein to coactivate transcription. J Biol Chem *277*, 9982-9988.

Shen, W., Xu, C., Huang, W., Zhang, J., Carlson, J.E., Tu, X., Wu, J., and Shi, Y. (2007). Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails. Biochemistry *46*, 2100-2110.

Shen, W.F., Montgomery, J.C., Rozenfeld, S., Moskow, J.J., Lawrence, H.J., Buchberg, A.M., and Largman, C. (1997). AbdB-like Hox proteins stabilize DNA binding by the Meis1 homeodomain proteins. Mol Cell Biol *17*, 6448-6458.

Shen, W.F., Rozenfeld, S., Kwong, A., Kom ves, L.G., Lawrence, H.J., and Largman, C. (1999). HOXA9 forms triple complexes with PBX2 and MEIS1 in myeloid cells. Mol Cell Biol *19*, 3051-3061.

Shi, A., Murai, M.J., He, S., Lund, G., Hartley, T., Purohit, T., Reddy, G., Chruszcz, M., Grembecka, J., and Cierpicki, T. (2012). Structural insights into inhibition of the bivalent menin-MLL interaction by small molecules in leukemia. Blood *120*, 4461-4469.

Shi, X., Hong, T., Walter, K.L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Pena, P., Lan, F., Kaadige, M.R.*, et al.* (2006). ING2 PHD domain links histone H3 lysine 4 methylation to active gene repression. Nature *442*, 96-99.

Shi, X., Kachirskaia, I., Walter, K.L., Kuo, J.H., Lake, A., Davrazou, F., Chan, S.M., Martin, D.G., Fingerman, I.M., Briggs, S.D.*, et al.* (2007). Proteome-wide analysis in Saccharomyces cerevisiae identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. J Biol Chem *282*, 2450-2455.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annu Rev Biochem *81*, 65-95.

Shinkai, Y., Rathbun, G., Lam, K.P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charron, J., Datta, M., Young, F., Stall, A.M.*, et al.* (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell *68*, 855-867.

Slany, R.K. (2009). The molecular biology of mixed lineage leukemia. Haematologica *94*, 984-993.

Smith, E., Lin, C., and Shilatifard, A. (2011). The super elongation complex (SEC) and MLL in development and disease. Genes Dev *25*, 661-672.

So, C.W., Lin, M., Ayton, P.M., Chen, E.H., and Cleary, M.L. (2003). Dimerization contributes to oncogenic activation of MLL chimeras in acute leukemias. Cancer Cell *4*, 99-110.

Soucy, T.A., Dick, L.R., Smith, P.G., Milhollen, M.A., and Brownell, J.E. (2010). The NEDD8 Conjugation Pathway and Its Relevance in Cancer Biology and Therapy. Genes Cancer *1*, 708-716.

Strout, M.P., Marcucci, G., Bloomfield, C.D., and Caligiuri, M.A. (1998). The partial tandem duplication of ALL1 (MLL) is consistently generated by Alu-mediated homologous recombination in acute myeloid leukemia. Proc Natl Acad Sci U S A *95*, 2390-2395.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Super, H.G., Strissel, P.L., Sobulo, O.M., Burian, D., Reshmi, S.C., Roe, B., Zeleznik-Le, N.J., Diaz, M.O., and Rowley, J.D. (1997). Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia. Genes Chromosomes Cancer *20*, 185-195.

Takahashi, Y.H., Westfield, G.H., Oleskie, A.N., Trievel, R.C., Shilatifard, A., and Skiniotis, G. (2011). Structural analysis of the core COMPASS family of histone H3K4 methylases from yeast to human. Proc Natl Acad Sci U S A *108*, 20526-20531.

Tardat, M., Brustel, J., Kirsh, O., Lefevbre, C., Callanan, M., Sardet, C., and Julien, E. (2010). The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. Nat Cell Biol *12*, 1086-1093.

Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., Chait, B.T.*, et al.* (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. Mol Cell *24*, 785-796.

Terranova, R., Agherbi, H., Boned, A., Meresse, S., and Djabali, M. (2006). Histone and DNA methylation defects at Hox genes in mice expressing a SET domain-truncated form of Mll. Proc Natl Acad Sci U S A *103*, 6629-6634.

Thiel, A.T., Blessington, P., Zou, T., Feather, D., Wu, X., Yan, J., Zhang, H., Liu, Z., Ernst, P., Koretzky, G.A.*, et al.* (2010). MLL-AF9-induced leukemogenesis requires coexpression of the wild-type Mll allele. Cancer Cell *17*, 148-159.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res *22*, 4673-4680.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol *28*, 511-515.

Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E.T., Buhler, M.*, et al.* (2013). Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. Cell *152*, 859-872.

Uchida, D., Hatakeyama, S., Matsushima, A., Han, H., Ishido, S., Hotta, H., Kudoh, J., Shimizu, N., Doucas, V., Nakayama, K.I.*, et al.* (2004). AIRE functions as an E3 ubiquitin ligase. J Exp Med *199*, 167-172.

Wagner, S.A., Beli, P., Weinert, B.T., Nielsen, M.L., Cox, J., Mann, M., and Choudhary, C. (2011). A proteome-wide, quantitative survey of in vivo ubiquitylation sites reveals widespread regulatory roles. Mol Cell Proteomics *10*, M111 013284.

Wang, C., Shen, J., Yang, Z., Chen, P., Zhao, B., Hu, W., Lan, W., Tong, X., Wu, H., Li, G.*, et al.* (2011a). Structural basis for site-specific reading of unmodified R2 of histone H3 tail by UHRF1 PHD finger. Cell Res *21*, 1379-1382.

Wang, G.G., Song, J., Wang, Z., Dormann, H.L., Casadio, F., Li, H., Luo, J.L., Patel, D.J., and Allis, C.D. (2009a). Haematopoietic malignancies caused by dysregulation of a chromatin-binding PHD finger. Nature *459*, 847-851.

Wang, J., Muntean, A.G., and Hess, J.L. (2011b). ECSASB2 mediates MLL degradation during hematopoietic differentiation. Blood.

Wang, J., Muntean, A.G., Wu, L., and Hess, J.L. (2012). A subset of mixed lineage leukemia proteins has plant homeodomain (PHD)-mediated E3 ligase activity. J Biol Chem *287*, 43410-43416.

Wang, P., Lin, C., Smith, E.R., Guo, H., Sanderson, B.W., Wu, M., Gogol, M., Alexander, T., Seidel, C., Wiedemann, L.M.*, et al.* (2009b). 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 *29*, 6074-6085.

Wang, Z., Song, J., Milne, T.A., Wang, G.G., Li, H., Allis, C.D., and Patel, D.J. (2010). Pro isomerization in MLL1 PHD3-bromo cassette connects H3K4me readout to CyP33 and HDAC-mediated repression. Cell *141*, 1183-1194.

Whitmarsh, R.J., Saginario, C., Zhuo, Y., Hilgenfeld, E., Rappaport, E.F., Megonigal, M.D., Carroll, M., Liu, M., Osheroff, N., Cheung, N.K.*, et al.* (2003). Reciprocal DNA topoisomerase II cleavage events at 5'-TATTA-3' sequences in MLL and AF-9 create homologous single-stranded overhangs that anneal to form der(11) and der(9) genomic

breakpoint junctions in treatment-related AML without further processing. Oncogene *22*, 8448-8459.

Wilbanks, E.G., and Facciotti, M.T. (2010). Evaluation of algorithm performance in ChIP-seq peak detection. PLoS One *5*, e11471.

Wong, K.F., So, C.C., and Kwong, Y.L. (1999). Chronic myelomonocytic leukemia with t(7;11)(p15;p15) and NUP98/HOXA9 fusion. Cancer Genet Cytogenet *115*, 70-72.

Wong, P., Iwasaki, M., Somervaille, T.C., So, C.W., and Cleary, M.L. (2007). Meis1 is an essential and rate-limiting regulator of MLL leukemia stem cell potential. Genes Dev *21*, 2762-2774.

Wu, W., and Sun, X.H. (2011). A mechanism underlying notch-induced and ubiquitin-mediated Jak3 degradation. J Biol Chem.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859-872.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Yagi, H., Deguchi, K., Aono, A., Tani, Y., Kishimoto, T., and Komori, T. (1998). Growth disturbance in fetal liver hematopoiesis of Mll-mutant mice. Blood *92*, 108-117.

Yip, K.Y., Cheng, C., Bhardwaj, N., Brown, J.B., Leng, J., Kundaje, A., Rozowsky, J., Birney, E., Bickel, P., Snyder, M., *et al.* (2012). Classification of human genomic regions based on experimentally determined binding sites of more than 100 transcription-related factors. Genome Biol *13*, R48.

Yokoyama, A., and Cleary, M.L. (2008). Menin critically links MLL proteins with LEDGF on cancer-associated target genes. Cancer Cell *14*, 36-46.

Yokoyama, A., Ficara, F., Murphy, M.J., Meisel, C., Naresh, A., Kitabayashi, I., and Cleary, M.L. (2011). Proteolytically cleaved MLL subunits are susceptible to distinct degradation pathways. J Cell Sci *124*, 2208-2219.

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

Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. Nature *378*, 505-508.

Yu, M., Honoki, K., Andersen, J., Paietta, E., Nam, D.K., and Yunis, J.J. (1996). MLL tandem duplication and multiple splicing in adult acute myeloid leukemia with normal karyotype. Leukemia *10*, 774-780.

Zeisig, B.B., Milne, T., Garcia-Cuellar, M.P., Schreiner, S., Martin, M.E., Fuchs, U., Borkhardt, A., Chanda, S.K., Walker, J., Soden, R.*, et al.* (2004). Hoxa9 and Meis1 are key targets for MLL-ENL-mediated cellular immortalization. Mol Cell Biol *24*, 617-628.

Zeng, L., Yap, K.L., Ivanov, A.V., Wang, X., Mujtaba, S., Plotnikova, O., Rauscher, F.J., 3rd, and Zhou, M.M. (2008). Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. Nat Struct Mol Biol *15*, 626-633.

Zeng, L., Zhang, Q., Li, S., Plotnikov, A.N., Walsh, M.J., and Zhou, M.M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature *466*, 258-262.

Zhang, P., Iwasaki-Arai, J., Iwasaki, H., Fenyus, M.L., Dayaram, T., Owens, B.M., Shigematsu, H., Levantini, E., Huettner, C.S., Lekstrom-Himes, J.A.*, et al.* (2004). Enhancement of hematopoietic stem cell repopulating capacity and self-renewal in the absence of the transcription factor C/EBP alpha. Immunity *21*, 853-863.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9*, R137.

Zhang, Y., and Rowley, J.D. (2006). Chromatin structural elements and chromosomal translocations in leukemia. DNA Repair (Amst) *5*, 1282-1297.

Zorko, N.A., Bernot, K.M., Whitman, S.P., Siebenaler, R.F., Ahmed, E.H., Marcucci, G.G., Yanes, D.A., McConnell, K.K., Mao, C., Kalu, C.*, et al.* (2012). Mll partial tandem duplication and Flt3 internal tandem duplication in a double knock-in mouse recapitulates features of counterpart human acute myeloid leukemias. Blood *120*, 1130-1136.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M.*, et al.* (2011). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature *478*, 524-528.