The Role of RNF40 mediated H2B monoubiquitination in transcription

Dissertation

for the award of the degree
"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

Submitted by

Wanhua Xie

born in

Chongqing, China

Göttingen, 2016

Affidavit

I hereby declare that the PhD thesis entitled "The Role of RNF40 mediated H2B monoubiquitination in transcription" has been written independently and with no other sources and aids than quoted.

_____

Wanhua Xie

June, 2016

Göttingen

Table of Contents

Abstract

Active gene transcription is tightly correlated with the occupancy of histone H2B monoubiquitination (H2Bub1) in the transcribed region. Therefore, it is has been commonly assumed that H2Bub1 is an exclusively positively acting histone modification and that increased H2Bub1 occupancy correlates with its requirement for gene expression. In contrast, depletion of the H2B ubiquitin ligases RNF20 or RNF40 alters the expression of only a subset of genes. We show that genes occupied by low to moderate amounts of H2Bub1 are selectively regulated in response to *Rnf40* deletion whereas genes marked by high levels of H2Bub1 are mostly unaffected by *Rnf40* loss. Furthermore, RNF40-suppressed genes appear to play an equally important role compared to RNF40-dependent genes in the RNF40/H2Bub1 regulatory network. We show that decreased expression of RNF40-dependent genes is highly associated with widespread narrowing of H3K4me3 peaks following *Rnf40* deletion. Notably, genes upregulated following *Rnf40* deletion, including *Foxl2*, are enriched for H3K27me3, which is decreased following *Rnf40* deletion due to decreased expression of the *Ezh2* gene. As a consequence, increased expression of some RNF40-suppressed genes is associated with enhancer activation via FOXL2. Finally, consistent with our previous findings, H2Bub1 is essential for the activation of bivalent genes, whereby RNF40, presumably via H2Bub1, modulates bivalent gene expression via coordination of the active and repressive marks H3K4me3 and H3K27me3, respectively. Together these findings show the complexity and context-dependency whereby one histone modification can have divergent effects on gene transcription dependent upon the activity of other epigenetic regulatory proteins and histone modifications.

Acknowledgements

Prof. Dr. Hongsheng Ouyang to help me get the scholarship. Thanks a lot to Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences (GGNB) for providing lots of excellent courses.

Last but not least, I would like to thank my family for their love and support.

Abbreviations

APS	Ammonium persulfate
BGP	ß-Glycerophosphate
IAA	Iodacetamide
IBMX	Isobutyl-methyl-xanthine
NEM	N-ethylmaleimide
Bre1	BREfeldin A sensitivity
°C	degree Celsius / centrigrade
CDK7	Cyclin-Dependent Kinase 7
CDK9	Cyclin-Dependent Kinase 9
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
ChIP-seq	ChIP coupled with high-throughput sequencing
mRNA	Message ribonucleic acid
CO2	Carbon dioxide
COMPASS	Complex proteins associated with Set1p
CTD	Carboxy-terminal Domain
DEPC	Diethylpyrocarbonate
DMEM	Dulbecco/Vogt modified Eagle's minimal essential medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	deoxyribonucleotide
DSIF	DRB Sensitivity Inducing Factor
DUB	Deubiquitinating enzyme
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin-ligase
V	Voltage
EDTA	Ethylenediaminetetraacetic acid
et al.	and others
EtOH	Ethanol
FACT	Facilitates Active Chromatin Transcription
FDR	False Discovery Rate
H2A	Histone 2A
H2B	Histone 2B

H3	Histone 3
H3K4me1	Histone 3 monomethylated at lysine 4
H3K4me3	Histone 3 trimethylated at lysine 4
H3K27me3	Histone 3 trimethylated at lysine 27
H3K36me3	Histone 3 trimethylated at lysine 36
H3K79me3	Histone 3 trimethylated at lysine 79
H3K27ac	Histone 3 acetylation at lysine 27
H4	Histone 4
HAT	Histone acetyltransferase
hMSCs	Human Mesenchymal Stem Cells
HSC70	Heat shock 70kDa protein
TBP	TATA-binding protein
IgG	Immunoglobulin G
kDa	kilo Dalton
m	milli
μ	micro
n	nano
min	Minute
kb	Kilo base pairs
bp	Base pair
s	secend
h	hour
MLL	Myeloid/lymphoid or mixed-lineage
MEFs	Mouse embryonic fibroblasts
NELF	Negative elongation factor
NP-40	Nonidet P40
PBS	Phosphate Buffered Saline
mRNA	messenger RNA
cDNA	Complementary DNA
PCR	Polymerase Chain Reaction
qPCR	Quantitative real-time PCR
pH	Measurement of acidity or alkalinity of a solution
P-TEFb	Positive Transcription Elongation Factor beta
RNAP II	RNA Polymerase II
RNF20	Ring finger protein 20

RNF40	Ring finger protein 40
P300	E1A binding protein p300
CBP	CREB-binding protein
TF	Transcription factor
SAGA	Spt-Ada-Gcn5-Acetyltransferase
RT-PCR	Reverse Transcription PCR
SDS	Sodium dodecylsulfate
SDS-PAGE	Sodium dodecylsulfate polyacrylamide gel electrophoresis
siRNA	Small interfering RNA
Tris	Tris(hydroxymethyl)aminomethane
4-OHT	(Z)-4-Hydroxytamoxifen
U	Unit
UBE2A	Ubiquitin-conjugating Enzyme E2A
USP22	Ubiquitin Specific Peptidase 22
WAC	WW domain containing adaptor with coiled-coil
PcG	Polycomb-group proteins
EZH2	Enhancer of zeste homolog 2
PRC2	Polycomb Repressive Complex 2

List of Figures

1. Introduction

1.1 Nucleosome: How to organize the DNA double helix

In all organisms, from bacteria to human, the genetic instruction is carried by the DNA double helix. Within the nucleus of the eukaryotic cells, highly compacted DNA consists of a fundamental repeating chromatin unit called the nucleosome. Each nucleosome core particle is formed by the wrapping of approximately 147 base pairs of DNA around an octamer protein complex containing two of each of the highly conserved core histones (H2A, H2B, H3, and H4). In addition, the linker histone H1 as well as its isoforms bind to the linker region of DNA outside the nucleosome (Zhou et al. 1998). Through the interactions of the nucleosomes or other chromatin associated proteins like heterochromatin protein 1 (HP1), nucleosomes are further packaged into various compacted levels of chromatin structure from extended nucleosome arrays to higher-ordered chromatin fibers (Caterino and Hayes 2007).

Actually, eukaryotic cells utilize diverse mechanisms for DNA replication (Fragkos et al. 2015), transcription (Workman 2006), DNA damage repair (Ciccia and Elledge 2010), mRNA processing (Bentley 2014), etc. Those chromatin-associated molecular processes require regulatory machineries that control access to genome loci in a spatiotemporal manner. Moreover, it is becoming clear that the controlling of chromatin compaction and accessibility involves the modulation of histone tail – DNA interaction (Caterino and Hayes 2007). Numerous residues within histone proteins are subjected to different post-transitional modifications (PTMs) including phosphorylation, acetylation, methylation, ubiquitination, deamination (Kouzarides 2007), ADP ribosylation (Messner and Hottiger 2011), and formylation (Jiang et al. 2007; Wisniewski et al. 2008), etc. Until now, the best-understood modifications

focus on the N-terminal tail extending from the nucleosomal surface ([Fig. 1](#)). It suggested that this region is required for histone-DNA and nucleosome-nucleosome interactions (Iwasaki et al. 2013). Modifications to the histone N-terminal tail have profound influences on chromatin structure and accessibility. For example, in heterochromatin, H3 tri-methylation at Lysine 9 (H3K9me3) serves as an HP1 binding platform to compact chromatin (Fischle et al. 2005; Zinner et al. 2006). In euchromatin, acetylation of histone tails opens up chromatin via facilitating histone exchange and functioning as a transcription factor binding platform ,such as bromodomain proteins(Dhalluin et al. 1999).
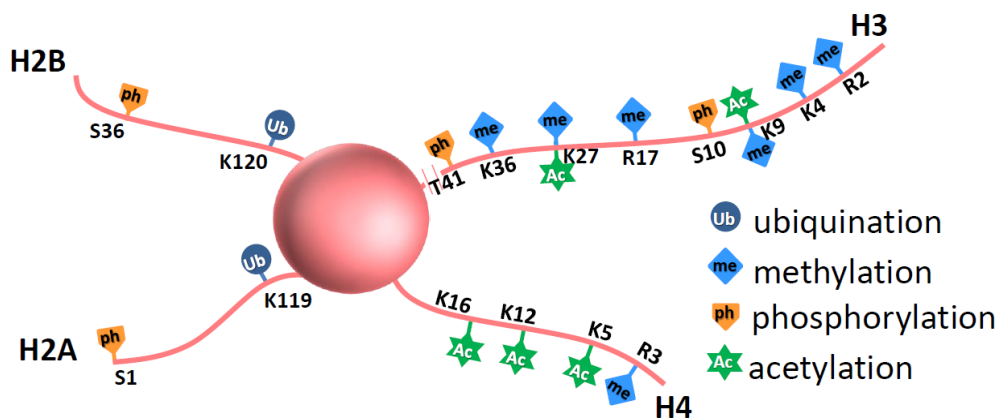
Fig. 1 Post-translational modifications in histone tails

Modifications at various amino acid residues are shown here (K: Lys, R: Arg, S: Ser, and T: Thr). Colors indicate different modifications   on each residue.

1.2 Histone modifications and gene expression

In multicellular organisms, each cell type has a unique epigenome to characterize specific transcriptional profile, which can be passed down to the daughter cells without changes in DNA sequences. In general, epigenome is defined as the combination of chemical modifications on DNA and histones (Schones and Zhao 2008). DNA modifications such as methylation have been frequently studied in the

past few decades. Unlike DNA, histones are subjected to various modifications, which play diverse functions in chromatin-associated processes such as transcription.

Transcription requires the access of multiple transcription factors to the specific gene locus to further recruit mRNA polymerase II to form the preinitiation complex near the transcription start sire (TSS). In the past few years, genome-scale studies on various histone modifications as well as their associated epigenetic factors have led to exciting advances in our knowledge of epigenetic regulation of gene expression, and put forward the 'histone code' hypothesis (Strahl and Allis 2000; Turner 2000), the 'charge-neutralization' hypothesis (Roth and Allis 1992), as well as the 'signaling-pathway' hypothesis (Schreiber and Bernstein 2002).

Genome-wide studies of histone modification maps on multiple eukaryotic cells have determined that each histone modification carried out a conserved function in transcription. For example, the best-studied H3K4me3, from yeast (Pokholok et al. 2005), fly (Schubeler et al. 2004), and mammal cells (Bernstein et al. 2005), is specifically localized near TSS, and facilitating transcription initiation (Guenther et al. 2007). However, monomethylation on H3 Lysine 4 (H3K4me1) occupies promoters and enhancers, and is required for enhancer-associated factors binding. Tri-methylation on H3 Lys 36 (H3K36me3) is selectively enriched on the transcribed regions of genes from the 5' to 3' region in a gradually increased manner, where it is associated with transcriptional elongation (de Almeida et al. 2011). In contrast, a subset of histone modifications exhibits a transcriptional repressive function. Tri-methylation of H3K9, H3K27, and H4K20 were found to be associated with heterochromatin and maintain genome silencing. Unlike H3K4me3 which selectively binds near transcription starting site, H3K27me3, H3K9me3, and H4K20me3 are

broadly enriched at the silenced non-coding and coding genomic regions (Roh et al. 2006; Magklara et al. 2011). However, di-methylation of H3K27 (H3K27me2) or monomethylation of H4K20 (H4K20me1) have been observed on less transcribed genes (Lee et al. 2015a; Svensson et al. 2015). Thus, histones can be diversely modified to comprehensively control gene expression on different levels.

1.3 Histone H2B monoubiquitination

1.3.1 Biochemical pathway for the monoubiquitination of H2B

Monoubiquitination on H2B is sequentially catalyzed by an enzymatic cascade involving ATP-dependent ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and RING finger ubiquitin ligase (E3) which ligates the highly conserved ubiquitin containing of 76 amino acids to the C-terminal tail of histone H2B (at K123 of yeasts and K120 of mammals) (Robzyk et al. 2000; Hwang et al. 2003; Weake and Workman 2008). In this sequential process, ubiquitin is transferred to the ATP activated cysteine residue of E1; Afterwards, E2 is activated via ubiquitination by E1. Subsequently, the ubiquitin activated E2 is recruited to the lysine residue of target proteins by the substrate-specific binding E3 in an ATP independent manner. (Schulman and Harper 2009; van Wijk and Timmers 2010; Metzger et al. 2012).

The E2 and E3 enzymes responsible for H2B monoubiquitination were first identified in yeast. Rad6 with the E2 activity specifically interacts with Bre1 with the E3 activity to ubiquitinate H2B on lysine 123 (Robzyk et al. 2000; Hwang et al. 2003). E2 and E3 for H2B monoubiquitination are highly conserved, with two homologs in human cells. The human E2 homologs are called UBE2A and UBE2B (Kim et al. 2009). The E3 homologs RNF20 (hBre1A) and RNF40 (hBre1B) form a heterotetrameric

complex through their N-terminal region which together carries out H2B monoubiquitination at lysine 120 (Zhu et al. 2005; Kim et al. 2009). Depleting either RNF20 or RNF40 specifically results in the global loss of H2Bub1 (Fuchs et al. 2012; Karpiuk et al. 2012).

In contrast to poly-ubiquitination which commonly plays a role in targeting proteins for degradation via the proteasome, monoubiquitination on H2B plays multiple biological functions. It serves as a transcription coactivator for ongoing transcription elongation (Kim et al. 2005; Johnsen 2012; Fuchs et al. 2014). Additionally, it regulates replication-dependent histone mRNA 3'-end processing (Pirngruber et al. 2009). Furthermore, it plays a role in DNA double strand break (DSB) repair (Kari et al. 2011), DNA replication (Trujillo and Osley 2012), and maintains chromatin integrity (Ma et al. 2011; Sadeghi et al. 2014). Accordingly, H2B monoubiquitination plays a crucial role in stem cell differentiation and tumorigenesis (Chen et al. 2012; Johnsen 2012; Karpiuk et al. 2012; Cole et al. 2015).

1.3.2 H2Bub1 facilitates transcription elongation

Early studies in yeast (Robzyk et al. 2000; Sun and Allis 2002) and mammals (Zhu et al. 2005; Shema et al. 2008) suggest that H2Bub1 is tightly linked to gene transcription. Based on chromatin immunoprecipitation coupled with high throughout sequencing (ChIP-seq), H2Bub1 is selectively enriches transcribed regions in mammals and yeast (Jung et al. 2012; Sadeghi et al. 2014). Moreover, genome-wide mapping of H2Bub1 has uncovered that it is a better representative for the RNA polymerase II moving rate than H3K36me3 and H3K79me2 (Fuchs et al. 2014). Notably, RNF20-RNF40 ubiquitination complex is cooperatively recruited via

interacting to other transcription factors, involving PAF1 complex (Kim et al. 2009), WAC (Zhang and Yu 2011), TP53 (Liu et al. 2009), and RB1 (Wen and Ao 2000).

1.3.2.1 PAF1c regulates H2Bub1

In mammalian cells, the RNA polymerase II associated factor 1 complex (PAF1c) is composed of several subunits including PAF1, CTR9, CDC73, RTF1, SKI8, and LEO1 (Kim et al. 2010b). It is suggested that those components can directly interact with RNF20-RNF40 complex *in vitro* and mediate monoubiquitination of H2B at the transcribed regions (Fig. 2) (Laribee et al. 2005; Zhu et al. 2005; Kim et al. 2009). In the case of hyperparathyroidism-jaw tumor syndrome and Wilms tumor, it is suggested that low levels of H2Bub1 was related to the highly frequent mutation of the tumor suppressor CDC73 and CTR9. Disruption of PAF1 complex by knocking down *CDC73* leads to a global reduction of H2Bub1 levels in parathyroid cancer (Hahn et al. 2012; Hanks et al. 2014). In addition, PAF1c directly interacts with the histone chaperone complex Facilitates Chromatin Transcription (FACT) (Krogan et al. 2002), which releases the H2A-H2B dimer from the nucleosome to facilitate transcription elongation in an ATP-independent manner (Kireeva et al. 2002; Belotserkovskaya et al. 2003). Furthermore, H2Bub1 is able to increase the catalytic activity of FACT (Pavri et al. 2006). Accordingly, we can assume that PAF1c, FACT, and H2Bub1 forms a machinery to drive RNA polymerase II mediated transcription.

In addition, PAF1c shows functional diversity in transcription cycle. Depletion of *PAF1*, a subunit of PAF1c, leads the release of RNA polymerase II from the promoter-proximal pausing in an H2Bub1-independent manner (Chen et al. 2015a).

1.3.2.2 The CDK9-WAC-RNF20/40 regulatory pathway

The carboxy-terminal domain (CTD) of RNA polymerase II is comprised of the Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser7 heptapeptide repeats (Egloff and Murphy 2008). From yeast to human, the sequence of this repeat is highly conserved, while differing in number with 26 and 52 repeats, respectively (Eick and Geyer 2013). Dynamic phosphorylation of the different residues of the CTD is one of the most important events in the transcription cycle and is catalyzed by various cell cycle-dependent kinases and TFIIF-dependent phosphatases. Over the past decades, it was uncovered that the phosphorylation of CTD usually occurs at the Tyr1, Ser2, Thr4, Ser5, and Ser7 residues (Chapman et al. 2007; Hsin et al. 2011; Hintermair et al. 2012). Cyclin-dependent kinase-7 (CDK7), a component of the general transcription factor TFIIH, was initially identified to be responsible for phosphorylation of Ser5 (Ser5-P) in yeast and mammals (Feaver et al. 1991; Lu et al. 1992). Moreover, it was also suggested that TFIIH has Ser7 phosphorylation activity (Glover-Cutter et al. 2009). Early findings revealed that Ser2 is phosphorylated by the positive transcription elongation factor b (P-TEFb) composed of CDK9 and Cyclin T, which further facilitates transcription elongation by overcoming the promotor-proximal pausing of RNA polymerase II (Marshall and Price 1992). In addition, a recent study suggested that CDK12/CDK13, which has Ser2 phosphorylation activity, coupled with CDK9, coordinates transcription elongation (Blazek et al. 2011).

Due to the advances in the development of mono-antibody fo single phosphorylation at Tyr1, Ser2, Thr4, Ser5, or Ser7, ChIP-seq approaches were rendered possible and could reveal the dynamic changes of the CTD modifications through different transcription steps. Ser5-P and Ser7-P are highly enriched at the 5' ends (Kim et al. 2010a), where Ser5-P is involved in regulating mRNA capping via recruiting the

capping enzymes (Schwer and Shuman 2011). However, Ser2-P shows low occupancy at the 5' end, but gradually increases toward the 3' end, and is highly enriched surrounding transcription terminal site (TTS) (Bataille et al. 2012). Ser2-P modulates promoter-proximal pausing of RNA polymerase II, transcription elongation, and 3' RNA processing in transcription cycle.

During the generation of RNA transcripts, transcription elongation requires phosphorylation of Ser2 of CTD. Most of our understanding of the key machinery responsible for Ser2-P is P-TEFb involving CDK9. CDK9 is cooperatively recruited to transcription preinitiation complex (PIC), which directly or indirectly depends on Ser5-P (Eick and Geyer 2013). Ser5-P is specially catalyzed by CDK7, a component of the general transcription factor TFIIH (Lu et al. 1992). In the transcription initiation phase, CDK7 carries out a dual function via phosphorylating TFIIE and CDK9. At the beginning, the phosphorylated TFIIE is free from PIC, which provides a position for the combination of DRB sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF) thereby forming the promoter-proximal pausing of RNA polymerase II.

Stimulation of transcription elongation necessitates the activating of the two transcriptional elongation inhibitors by phosphorylation, which is catalyzed by the phosphorylated CDK9 (Fujinaga et al. 2004; Yamada et al. 2006). Usually, the kinase function of CDK9 is inhibited in complex with 7SK snRNA and HEXIM1, and dissociates in nucleus (Schulte et al. 2005). After recruiting near TSS under the control of other transcription factors like the Ser5-P dependent capping enzymes (Coudreuse and Nurse 2010), CDK9 is sequentially activated by CDK7-catalyzed T-loop phosphorylation (Larochelle et al. 2012).

H2Bub1 is globally dependent on Ser2-P as S2A mutation on CTD results in global loss of H2Bub1 (Pirngruber et al. 2009). In concordance, inhibiting kinase domain or depleting CDK9 abolished H2B monoubiquitination, while overexpression of CDK9 elevated H2Bub1 levels (Pirngruber et al. 2009). In transcription cycle, multiple molecular interactions are associated with the CDK9 kinase activity. Another important phosphorylated target is UBE2A, the E2 enzyme responsible for H2Bub1 in RNF20/RNF40 complex (Wood et al. 2005; Shchebet et al. 2012). Moreover, CDK9 can cooperatively facilitate H2B monoubiquitination via promoting the recruitment of PAFc, in which the phosphorylated SPT5 CTD by CDK9 serves as a binding domain for RTF1, a subunit of PAFc (Mbogning et al. 2013). Additionally, depletion of *PAF1* (a subunit of PAFc) increased CDK9 occupancy and Ser2-P levels and led to genome-wide loss of H2Bub1 (Chen et al. 2015a). This indicates that CDK9-faciated H2B monoubiquitination is dependent on PAFc.

The specific recruitment of RNF20/RNF40 complex to Ser2-P CTD is majorly mediated by the WW domain-containing adaptor with coiled-coil (WAC). WAC interacts with RNF20/RNF40 complex through its C-terminal coiled-coil region, and its N-terminal WW domain recognizes Ser2-P CTD of RNA polymerase II, thereby bridging RNF20/RNF40 complex to RNA polymerase II (Zhang and Yu 2011). Together, CDK9, WAC, and RNF20/RNF40 complex form a co-regulatory machinery to control tissue specific gene expression (Fig. 2) (Karpiuk et al. 2012).

Eventually, the recruitment of the key regulator CDK9 is crucial to regulate H2Bub1, which is mediated by various transcription factors involving BRD4. BRD4 containing two N-terminal bromodomains specifically recognizes acetylated histone tails, thereby bridging P-TEFb to the hyper-acetylated chromatin site through interacting

with BRD4 C-terminal domain PID (Bisgrove et al. 2007). In addition, BRD4 also promotes the release of the elongation negative factor NELF to active P-TEFb complex (Patel et al. 2013). Both overexpression and functional deletion experiments confirmed the P-TEFb mediated CTD phosphorylation is dependent on BRD4 (Bisgrove et al. 2007; Nagarajan et al. 2014). BRD4 genome-widely mediates the recruitment of CDK9 to transcribed regions, further stimulating coding and enhancer RNA transcription elongation. In turn, this process can be blocked by the bromodomain specific binding inhibitors (Kanno et al. 2014; Nagarajan et al. 2014). Unexpectedly, BRD4 partly serves the kinase activity to phosphorylate Ser2 of CTD in the absence of P-TEFb (Devaiah et al. 2012). Consistent with the role of Ser2-P on transcription elongation, BRD4 facilitates the release of RNA polymerase II into gene body (Kanno et al. 2014). Given the positive effect of Ser2-P on H2Bub1 (Karpiuk et al. 2012), our group further confirmed the role of BRD4 in H2Bub1 regulatory axis (Fig. 2), by which the depletion of BRD4 led to a decrease in H2Bub1 (Johnsen 2012; Nagarajan et al. 2014).

1.3.2.3 H2Bub1 facilitates nucleosome 'breathing' during transcription elongation

During transcription elongation, DNA replication, or DNA damage repair, the DNA template needs to be unwound from the nucleosome to enable epigenetic regulators to access the DNA template. Afterwards, DNA is repackaged within the nucleosome, and nucleosomal structure is restored. As other molecular processes, this nucleosome breathing requires the help of various chromatin regulators, which are classified into ATP-dependent nucleosome remodelers and ATP-independent histone chaperones.

To date, several ATP-dependent chromatin remodeling enzymes have been implicated in transcription activation such as SWI/SNF and NuRD/Mi-2/CHD complex. In agreement with the transcription coactive role of H2Bub1, H2Bub1 directs SWI/SNF complex recruitment to target genes via functioning as a binding platform for chromatin remodeling complex such as BAF155, a subunit of SWI/SNF complex (Shema-Yaacoby et al. 2013). The correlation between H2Bub1 and SWI/SNF could be confirmed in gene expression due to the fact that disruption of SWI/SNF complex resulted in significant reduction of H2Bub1-dependent genes. Unexpectedly, one early finding suggested that BAF250, another subunit of SWI/SNF complex, partly acts as an E3 ubiquitin ligase for H2Bub1 (Li et al. 2010).

Unlike chromatin remodelers, histone chaperones mediated nucleosome breathing doesn't require ATP. One of the best-studied chaperones is Facilitates Chromatin Transcription (FACT) complex containing two conserved subunits: SPT16H and SSRP1. FACT facilitates transcription by releasing H2A-H2B dimer from the core nucleosome, and directly interacts with H3-H4 dimer to allow RNA polymerase II to efficiently overcome the nucleosome barrier (Belotserkovskaya et al. 2003). On the structural basis of a recent study, FACT priority displaces H2A-H2B dimer from nucleosome, which requires the interaction between the novel 'U-turn' motif in Spt16M domain and the α1 helix of H2B to weaken the binding between H2A-H2B dimer and DNA (Hondele et al. 2013). FACT has the capability of a dual function as it carries out the nucleosome reassembly after transcription. Furthermore, it was proposed that H2Bub1 is capable of increasing FACT enzymatic activity to promote H2A-H2B dimer displacing. FACT can in turn facilitate H2B monoubiquitination through SPT16H recruits RNF20/RNF40 to Ser2-P CTD of RNA polymerase II (Endoh et al. 2004). Taking together, this cooperative interaction between FACT and

H2Bub1 enhances nucleosome breathing ([Fig. 2](#)), Subsequently, H2Bub1-facilitated transcription elongation shows FACT dependency (Pavri et al. 2006).

1.3.2.4 Deubiquitination of H2Bub1 by SAGA complex

Monoubiquitination of H2B is dynamically controlled by ubiquitination and deubiquitination. Ubiquitination of H2B is catalyzed by RNF20/RNF40 complex in mammals and Rad6/Bre1 in yeast, as described above. Deubiquitination of H2B is carried out by an evolutionary conserved multifunctional machinery-SAGA (Spt–Ada–Gcn5 acetyltransferase). In yeast, architecture of SAGA is composed of several modules: TATA-binding protein (TBP), histone acetyltransferase (HAT), and deubiquitinase (DUB) (Rodriguez-Navarro 2009; Samara et al. 2012; Morgan et al. 2016). DUB is comprised of several subunits involving ubiquitin-specific protease, Sgf73, Sgf11, and Sus1.

On the functional and structural basis of early studies, DUB is independently folded into a sub-complex and crosslinks to the core SAGA complex via the C-terminal region of Sgf73 (Han et al. 2014), or the human homolog ATXN7. The disruption of DUB modules by knockdown of ATXN7 leads to a significantly global increase in H2Bub1 (Bonnet et al. 2014). The conserved zinc finger domain of Sgf11 (human, ATXN7L3) bridges DUBs to H2B, and the N-terminal zinc finger–ubiquitin binding (ZnF-UBP) domain of Ubp8 (human, USP22) and is essential to deubiquitination (Samara et al. 2010; Samara et al. 2012; Morgan et al. 2016).

The DUB activity is mainly carried out by Ubp8 (Henry et al. 2003) and Ubp10 (Emre et al. 2005) in yeast. Genome-wide studies uncovered that Ubp8 and Ubp10 deubiquitinated H2BK123 in an epigenomic context dependent manner (Schulze et al. 2011). Ubp8 is suggested to serve as a transcription coactivator (Daniel and

Grant 2007), while Ubp10 is associated with telomere silencing (Emre et al. 2005). In human, multiple H2B ubiquitin-specific proteases have been identified such as USP3 (Nicassio et al. 2007), USP7 (Sarkari et al. 2009), USP12 (Joo et al. 2011), USP46 (Joo et al. 2011), USP44 (Fuchs et al. 2012), USP22 (Zhao et al. 2008), and two potential deubiquitinases USP27X and USP51 (Johnsen 2012). Possibly, as in yeast, the diverse DUBs in human control H2Bub1 levels in a context-dependent manner (Johnsen 2012). Notably, SAGA is required for a subset of gene transcription, depletion of ATXN7L3 leads to significant reduction of gene transcription regardless of H2Bub1 level (Bonnet et al. 2014). The question is how this contradictory phenomenon occur: 1. H2Bub1 stabilizes nucleosome in yeast (Chandrasekharan et al. 2009), which characterizes a potential transcription repressive function; 2. H2Bub1 serves as a transcription coactivator; 3, Loss of DUBs with H2Bub1 increasing is harmful for transcription; 4, FACT enriched transcribed region is sensitive to H2Bub1 depletion. One possible explanation is a dynamic balance between ubiquitination and deubiquitination that could increase the FACT-mediated nucleosome 'breathing' rate, and promotes transcription elongation. Thus, tipping the balance by loss of any of those functional opposing machineries (RNF20/RNF40 complex and UBDs) could decrease FACT-facilitated transcription elongation rate.
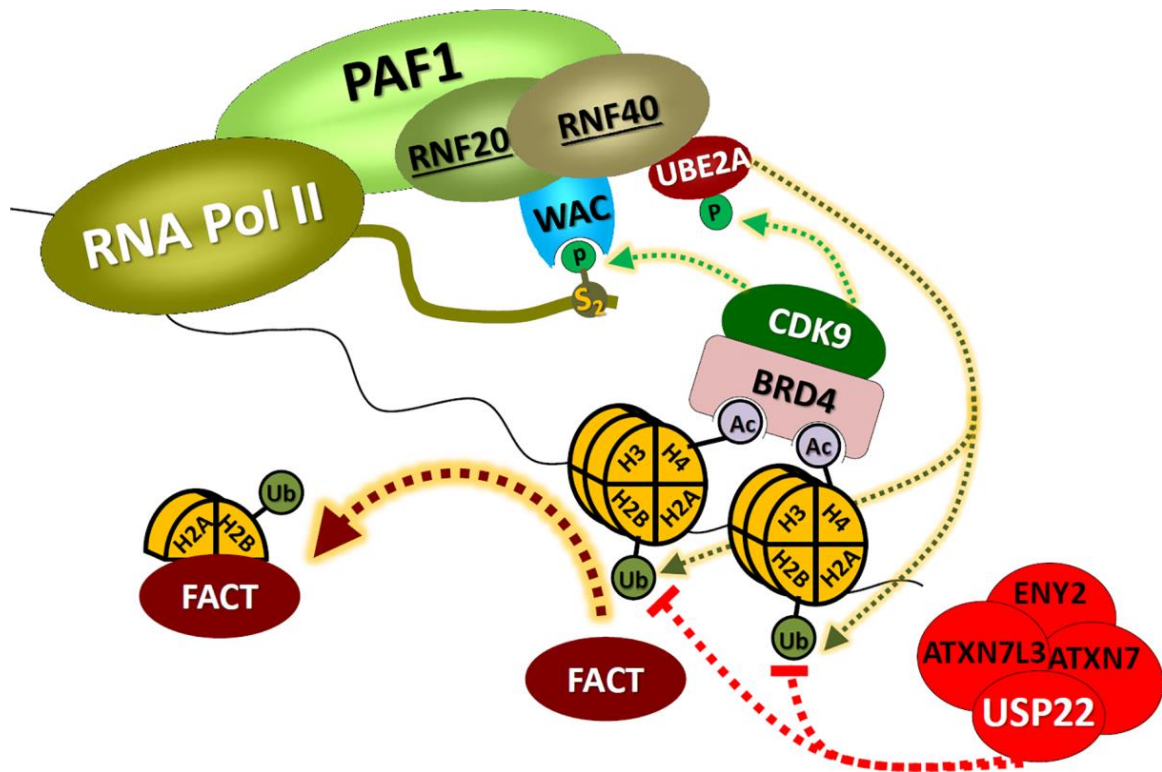
Fig. 2 H2Bub1 associated machinery for transcription elongation modified from (Johnsen 2012).

Monoubiquitination of H2B begins with the recruitment of CDK9 mediated by BRD4, which contains two bromodomain regions which bind to acetylated histone H4. CDK9 phosphorylates Ser2 residue on the CTD tail of polymerase II to create a binding domain for WAC, further recruiting RNF20-RNF40 complex. Meanwhile, CDK9 activates the E2 (UBE2A) by phosphorylation. Notably, PAF1 complex also provides binding domains to recruit RNF20 and RNF40. Those components work together to ubiquitinate H2B on Lys 120. H2B monoubiquitinated nucleosome further recruits FACT, which removes H2A-H2B dimer from the core nucleosome to facilitate transcription elongation. The single ubiquitin on H2B can be specifically removed by SAGA complex involving ubiquitin-specific protease-22 (USP22).

1.3.3 H2Bub1 also functions as a transcription suppressor

H2Bub1 is also proposed to play a repressive role in transcription both in yeast and human cells. Depletion of RNF20 led to a subset of H2Bub1-targeted genes upregulation (Espinosa 2008; Shema et al. 2008). In yeast, the transcription repressive function of H2Bub1 is associated with the occupancy of H2Bub1 on promoters. The nucleosome at TSS is highly occupied with H2Bub1 compared to transcribed regions in H2Bub1-suppressed genes, where H2Bub1 inhibits transcription factor binding such as TBP via stabilizing the nucleosome (Gardner et al. 2005; Fleming et al. 2008). Collectively, Espinosa put forward a hypothesis that H2Bub1-mediated transcription suppression is related to promoter-proximal pausing of RNA polymerase II as deubiquitination of H2Bub1 by DUB such as USP22 could facilitate the release of RNA polymerase II from the promoter-proximal pausing (Espinosa 2008). However, genome-wide analysis of H2Bub1 in human cells failed to show additional occupancy of H2Bub1 on promoters (Shema et al. 2008). Moreover, loss of H2Bub1 by depleting hBRE1A was not able to promote the release of RNA polymerase II from promoter-proximal pausing (Chen et al. 2015a). In addition, another study proposed that RNF20 acted as a tumor suppressor by inhibiting pro-oncogenic gene transcription via mechanically blocking the cooperative binding of transcription elongation factor S-II (TFIIS) to PAFc and RNA polymerase II (Shema et al. 2011). Moreover, another genome-wide study suggested that H2Bub1-suppressed genes show low transcription level and are unmarked by H2Bub1 (Jung et al. 2012). In addition, the H2Bub1 ubiquitinating complex and DUB have opposing function in the same genes in yeast and human cells. For example, the *IRF1* is upregulated following RNF20 knockdown, while downregulated in USP22 depleted cells (Johnsen 2012). But it still less known how the two enzymes, which play these

opposing functions on H2Bub1, together modulate gene transcription. It is proposed that H2Bub1-suppressed genes display epigenetic context dependency (Buro et al. 2010; Chipumuro and Henriksen 2012).

1.3.4 H2Bub1 crosstalk with H3 methylation

Studies in both yeast and human have shown that H2Bub1-mediated chromatin processes such as DNA damage repair and transcription are correlated to the histone trans-tail crosstalk between H2Bub1 and H3 methylation at lysine 4 or 79. Similar to the occupancy pattern of H2Bub1, genome-wide mapping of H3K79me3 indicates that H3K79me3 is preferentially enriched on transcribed gene both in yeast and human cells (Schulze et al. 2011; Jung et al. 2012). Early studies uncovered that H2Bub1 directly stimulates DOT1-like (DOT1L) methyltransferase activity through allosteric mechanisms to facilitate H3K79 methylation (Ng et al. 2002; McGinty et al. 2008).

Unlike H3K79me3 and H2Bub1, H3K4me3 is preferentially enriched on promoters, where it modulates the recruitment of transcription initiation factors (Guenther et al. 2007). Methylation of H3K4 is catalyzed by the COMPASS/Set1 complex containing Set1, Cps35, and other six subunits in yeast. Compared to yeast, multiple complexes are responsible for H3K4me3 in mammals, including six major players in the MLL/Set1 complex: SET1A and SET1B (the homologue of *Drosophila* Set1), MLL1 and MLL2 (the homologue of *Drosophila* Trithorax), as well as MLL3 and MLL4 (the homologue of *Drosophila* Trithorax-related) (Smith et al. 2011). Two hypotheses about H2Bub1 and H3K4 methylation crosstalk were put forward in yeast. In the 'bridge' model, the H3K4 methyltransferase Set1/COMPASS complex is directed to transcribed genes via the interaction between the accessory COMPASS subunit

yCps35 or hWDR82 and H2Bub1 (Lee et al. 2007; Vitaliano-Prunier et al. 2008; Wu et al. 2008). Notably, H2Bub1 is also able to mediate the interaction between Cps35 and Dot1 to facilitate H2Bub1-dependent H3K79 methylation (Lee et al. 2007). Another study suggested that Rad6/Bre1 is capable of activating Cps35 via adding a ubiquitin at lysine 68 and lysine 69. The ubiquitinated Cps35 further promotes the recruitment of Spp1, a required COMPASS subunit for H3K4me3 (Vitaliano-Prunier et al. 2008). In the 'wedge' model, H2Bub1 acts as a chromatin opener which enhances the access of H3K4 methyltransferases to active regions. However, most of the studies were performed in yeast. Considering that H3K4 methylation machineries are more complex in mammals than COMPASS/Set1 complex in yeast, exhaustive study is direly needed to resolve the mechanism of crosstalk between H2Bub1 and H3K4me3 in mammals. While the COMPASS/Set1 complex is responsible for the bulk of H3K4me3 in yeast, the MLL complexes in mammals are more likely to be catalyzed in a gene-specific manner (Shilatifard 2012). The question is if the crosstalk between H2Bub1 and H3K4me3 occurs in a gene-specific manner, or if H3K4me3 is globally modulated by H2Bub1 on mammals.

Several studies have shown that H2Bub1 is coupled to H3 methylation (H3K4me3 and H3K79me3) to modulate transcription of certain genes such as homeobox (*HOX*) genes, which are required for anatomical structures development in various organisms (Zhu et al. 2005; Kim et al. 2009). Two recent studies suggested that broad H3K4me3 was able to facilitate transcription elongation and had additional enhancer activity (Benayoun et al. 2014; Dincer et al. 2015). While less is known about the factors driving H3K4me3 spreading, given the correlations between H2Bub1 and H3K4me3, we propose that H2Bub1 could facilitate H3K4me3 broadening.

1.4 Polycomb group proteins (PcG)

PcG genes were first identified as suppressor of *Hox* genes during anterior-posterior development in *Drosophila*. In various organisms, multiple PcG complexes are recognized as chromatin modifiers which are capable of mediation of H2A monoubiquitination (H2Aub1), H3K27 methylation, as well as DNA methylation to silence gene expression, and play crucial roles in cell fate, development, and diseases.

1.4.1 PcG in gene regulation

1.4.1.1 The core components of PcG complex

Two main PcG complexes, the **P**olycomb **R**epressive **C**omplex 1 and 2 (PRC1 and PRC2) have been identified in mammals. The core components of canonical PRC1 includes one of each of RING1 (RING1a/b), chromobox-domain (CBX) protein (CBX2, 4, 6, 7, and 8), polycomb group ring finger family (PCGF1-8), and PHC (PHC1-3) (Di Croce and Helin 2013). These components work together to monoubiquitinate H2A on Lys119 through the E3 ligase RING1a and RING1b activity (Wang et al. 2004; Cao et al. 2005). The core components of PRC2 contain Enhancer of zeste homolog 2 or 1 (EZH2 and EZH1), embryonic ectoderm development (EED), and suppressor of zeste 12 (SUZ12). EZH2 or EZH1 catalyzes additional methyl groups on H3 Lys 27 (mono-, di-, and tri- methylation) (Cao et al. 2002). Moreover, EZH2 is capable of catalyzing DNA methylation (Vire et al. 2006). The function of PRC1 and PRC2 in genome silencing is believed to be associated with nucleosome compaction(Sexton et al. 2012).

1.4.1.2 Different PcG complexes mediate specific genome locus silencing

During embryogenesis and tumorigenesis, different genome loci are spatiotemporally activated or repressed. In *Drosophila* and mammals, diverse epigenetic factors are capable of interacting with core PcG components to form different PcG complexes, which selectively meditate genome silencing at the right time. For example, the silencing of a subset of tissue-specific genes in ESC cells requires Jarid2 (a member of Jumonji C (JmjC) domain protein family) to direct the core components of PRC2 to the promoters of these genes (Peng et al. 2009; Shen et al. 2009). KMT1D and KMT1C (H3K9me1 and H3K9me2 methyltransferases) could be purified from E2F6-dependent PRC2 complex, which mediates silencing of the E2F6 bound promoters and Myc-response elements through the interaction with E2F6 (Ogawa et al. 2002). Another PRC2-associated protein PHF1 is suggested to mediate *HOX* gene repression (Sarma et al. 2008). Recently, genome-wide study of yeast and embryonic stem cells uncovered that more than 50% of nucleosomes were modified by H3K27 methylation (mono-methylation, di-methylation, and tri-methylation) (Voigt et al. 2012; Lee et al. 2015b). This indicates that methylation of H3K27 may have very broad effects on chromatin.

In the case of PRC1, approximately 180 PRC1 complexes comprised of the core components with other epigenetic factors have been postulated in mammals (Di Croce and Helin 2013). Usually, the recruitment of PRC1 complex is mediated by the interaction between CBX family (subunits of PRC1) and H3K27me3. In mouse embryonic stem cells, CBX7-mediated PRC1 recruitment is required to silence differentiation genes and maintain pluripotency, while CBX7 is replaced by CBX2 and CBX4 to mediate pluripotent genes silencing during differentiation (Morey et al.

2012; O'Loghlen et al. 2012). More PcG complex-associated proteins and their functions are reviewed by Di Croce (Di Croce and Helin 2013).

Overall, the interaction of the core components of PcG complex with other epigenetic factors mediates the specific silencing of genome loci and characterizes gene regulation in a context-dependent manner.

1.4.2 H2Bub1 and Bivalency

Due to the advancement in ChIP-sequencing, early studies in ESC cells found that approximately 22% of promoters are marked with the transcription coactivator H3K4me3 and the repressive H3K27me3, which is referred to as bivalency (Bernstein et al. 2006; Mikkelsen et al. 2007). Strikingly, most of those bivalent domains are occupied by RNA polymerase II while being transcribed at a low degree (Min et al. 2011). Moreover, most development-associated genes are located in bivalent domains such as HOX clusters (Bernstein et al. 2006). Thus, it is believed that bivalent domains maintain transcription silencing in ESCs while allowing to be rapidly activated following developmental stimuli.

Bivalent domains are evolutionarily conserved and present in mammals and zebrafish (Lindeman et al. 2011). In addition to ESCs, bivalency is also found in various cell lines such as MEFs, B cells, neural progenitors, mesenchymal stem cells (MSC). Furthermore, bivalency is also studied in cancer cells, and increasing evidence show the dynamic changes in specifc bivalent domains in cancer cells and after therapeutic treatment (Bapat et al. 2010; Ntziachristos et al. 2014). Given that drug-induced DNA demethylation in colon cancer cells forms new bivalent regions (McGarvey et al. 2008), it can be speculated that the subsequent silencing of some tumor suppressors might be the origin of the establishment of bivalency.

To date, multiple factors are proposed to form bivalent domains (Voigt et al. 2013), including CpG island dependent recruitment, DNA methylation status, histone modification, and noncoding RNAs (ncRNA),etc. The key step is controlling the access of the central players responsible for H3K4me3 and H3K27me3, namely MLL and PRC2 complex, to genomic loci. Changes in either H3K4me3 or H3K27me3 can affect the expression of the bivalent genes resulting in either their full activation or repression (Agger et al. 2007; Wang et al. 2009; Agarwal and Jothi 2012). Given the importance of H2Bub1 on facilitating H3K4me3, the co-regulation of Hox bivalent genes by H2Bub1 and H3K4me3 was proposed (Zhu et al. 2005). H2Bub1 is also suggested to decrease H3K27me3 on some individual tissue-specific bivalent genes during MSC differentiation (Karpiuk et al. 2012). However the role of H2Bub1 in bivalency is still unclear. Genome-wide studies can uncover the effect of H2Bub1 on bivalent domains via utilizing ChIP-sequencing approaches.

1.4.3 Another poised chromatin signature: the dynamic equilibrium between H3K27me3 and H3K27ac

The lysine 27 of H3 is not only subjected to methylation, but also modified with acetylation. In general, histone acetylation is catalyzed by histone acetyltransferases (HATs) via transferring an acetyl group from acetyl CoA to form ε-N-acetyl lysine. H3K27ac shows an opposing effect on gene transcription compared to H3K27me3. Multiple HAT families have been identified involving Gcn5-related N-acetyltransferases (GNATs), p300/CBP complex, SRC family, as well as MYST families. One of the best-known HATs is p300/CBP complex, which is believed to bind to active cis-regulatory elements.

Surprisingly, p300/CBP is also found at hypoacetylated promoters and enhancers (Rada-Iglesias et al. 2011; Zentner et al. 2011). In most cases, those p300/CBP bound regions are occupied by H3K27me3, and their annotated genes are not highly transcribed. Generally, this epigenetic signature endows those regions with a 'poised' feature like bivalent domains, which allows their annotated genes to be rapidly activated (Rada-Iglesias et al. 2011). Acetylation of histone is required not only in the recruitment but also in the activation of the enzymatic activity of p300/CBP (Holmqvist and Mannervik 2013). In order to prevent pre-acetylation by p300/CBP on those poised region, antagonistic mechanisms must exist. H3K27me3 is typically considered to decrease chromatin accessibility. While the recruitment of p300/CBP complex is not affected by the presence of H3K27me3, H3K27me3 blocks the p300 enzymatic activity on those unique chromatin regions (Rada-Iglesias et al. 2011; Calo and Wysocka 2013; Holmqvist and Mannervik 2013).

The opposing functions of H3K27me3 and H3K27ac on gene transcription indicate that p300/CBP serves as a transcription switch on the 'poised' regions. The H3K27-specific demethylase UTX can directly bind to CBP (Tie et al. 2012) thus elevating H3K27ac by overexpressing CBP which antagonizes PcG-mediated H3K27me3. Additionally, knockdown of CBP results in antagonistic changes in H3K27me3 in *Drosophila* (Tie et al. 2009). Furthermore, inhibiting EZH2 methyltransferase domain leads to a global increasing of H3K27ac (Johnson et al. 2015). Given the similar 'poised' function of these unique regions and bivalent domains in gene transcription, the bivalent domain might be occupied by p300/CBP. However, factors affecting the dynamic equilibrium between H3K27me3 and H3K27ac are still necessary to be identified.

1.5 Enhancers and transcription

The human genome is as long as approximately 6 billion base pairs, while less than 2% of this DNA sequence (coding DNA) can be transcribed into mRNA (Hawes and O'Brien 2008). The spatiotemporal expression of genes is modulated by the interaction of epigenetic factors and cis-regulatory elements. In the past decades, some functionally diversified cis-regulatory elements (such as promoters, enhancers, silencers, and insulators) within the noncoding genome regions have been well studied. Unlike promoters which usually control the transcription of their nearby genes, enhancers are able to activate genes at remote distances, ranging from several to hundreds kilo base pairs. Even enhancers can activate genes at different chromatins (Lomvardas et al. 2006). Owing to the recent advances in high throughout epigenomic profiling technologies, enhancers are proved to play a crucial role in regulating cell type-specific genes during embryogenesis and tumorigenesis.

1.5.1 Enhancer-promoter looping activates transcription

In general, activation of enhancers begins with the binding of multiple transcription factors such as "pioneer" factors, which are able to recognize specific DNA sequences (Cirillo et al. 2002). In fact, most transcription factors can only bind to nucleosome-free DNA, the pioneer factors have the additional ability to directly bind to the nucleosomal DNA via a cooperative interaction mechanism (Zaret and Carroll 2011). For example, the forkhead box (FOX) proteins, FOXA1, FOXA2, and FOXA3 pioneers bind to sequence-specific nucleosomal DNA through a conserved 80-100 amino acid formed motif (McPherson et al. 1993; Zaret and Carroll 2011; Spitz and Furlong 2012). Genome-wide mapping further discovered that multiple FOX proteins directly modulate enhancers' activity (Spitz and Furlong 2012; Georges et al. 2014). In fact, the ability of enhancers to facilitate gene transcription from a great distance is

mediated by the interaction with coactivators, including histone modifiers (e.g., p300/CBP), chromatin remodelers (e.g., CHD7), and mediators (Weake and Workman 2010). The pioneer factors are able to stimulate an open chromatin state via repositioning nucleosomes and decompacting chromatin. Additionally, it facilitates the binding of other transcription factors via recruiting the chromatin remodelers such as SWI/SNF complex (McPherson et al. 1993; Li et al. 2012). Notably, enhancers are also bound by general transcription factors (e.g., TFIID) and mRNA polymerase II (Malik and Roeder 2010). Based on the chromosome conformation capture technology (such as 3C, 4C, 5C and Hi−C), diversity of enhancer-promoter communications have been observed (Ong and Corces 2011). In order to activate transcription, enhancer-bound transcription factors loop out the intervening sequences and interact with the promoter regions (Fig. 3).
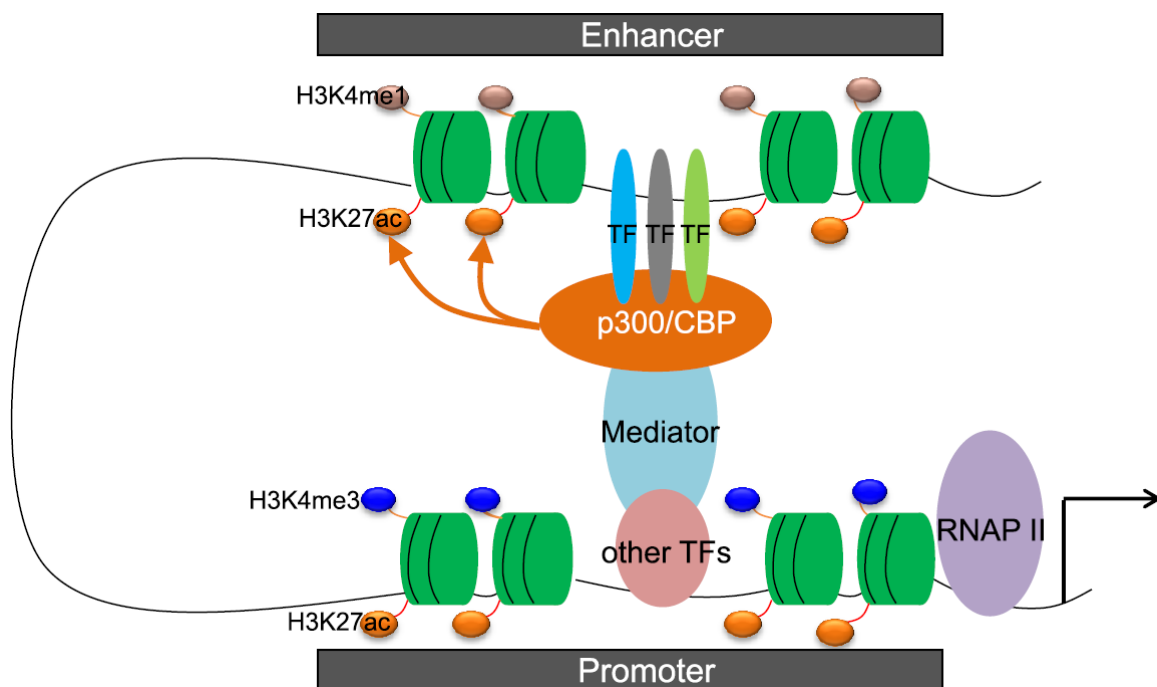
Fig.3   the model of enhancer-promoter loop

TF, transcription factors, such as FOXA1, FOXA2, and FOXA3.

1.5.2 Enhancer associated histone modifications

1.5.2.1 H3K4 monomethylation at enhancers

The central feature of enhancers is their ability to bind to multiple transcription factors to activate gene transcription from a great distance In order for the transcription factors to have access to the enhancer, they must overcome the nucleosome barrier. Similar to pioneer factors, histone modifications have the capability of activating enhancers though various mechanisms, such as rearranging nucleosomes, serving as an integrated transcription factor binding platform, and promoting histone exchange. H3K4me1 was the first identified enhancer associated histone signature (Heintzman et al. 2007). Notably, H3K4me1 is not only a unique signature for enhancers, but also broadly occupies the 5' region of transcribed genes. The pre-presence of H3K4me1 on intergenic regions is utilized to identify cell type-specific enhancers, which opens a window to study enhancers' role in cell type-specific gene transcription during embryogenesis and tumorigenesis (Akhtar-Zaidi et al. 2012; Herz et al. 2012). During embryonic stem cell differentiation, the activation of tissue-specific genes is associated with their enhancer activity (Bogdanovic et al. 2012). In addition to being tightly coupled to enhancer activity, H3K4me1 also marks pre-activated enhancers which lack histone acetylation (Creyghton et al. 2010). Additional genome-wide studies have shown that p300/CBP pre-occupies those unactivated enhancers which are also enriched by H3K27me3 (Rada-Iglesias et al. 2011). This unique chromatin region is referred to by poised enhancers as described **1.4.3**. Thus, H3K4me1 is a unique epigenetic signature that can broadly be used to define potential enhancers.

1.5.2.2 An over-activated enhancer marker: H3K4me3

Although H3K4me3 typically presents on activated promoter, genome-wide mapping of H3K4me3 observe a detectable level of H3K4me3 on enhancers (Clouaire et al. 2012). Recent studies suggested that broad H3K4me3 domain at 5' regions additionally serves as an enhancer that activates tumor suppressors and tissue-specific genes (Benayoun et al. 2014; Clouaire et al. 2014; Chen et al. 2015b). Moreover, loss of the H3K4me3-specific demethylase KDM5C leads the overactivation of enhancers characterized by the elevation of H3K4me3, which is associated with overexpression of oncogenes (Shen et al. 2016). H3K4me2 is also able to occupy both promoters and enhancers (Pekowska et al. 2010). Studies in *Drosophila* and human cells confirmed that the bulk of H3K4me2 and H3K4me3 are majorly catalyzed by the SET1A/B complex (Wu et al. 2008; Ardehali et al. 2011). However, the CxxC domain-containing CFP1, one specific subunit of SET1a/b complex, directs H3K4me3 on nonmethylated CpG-islands enriched promoters (Lee and Skalnik 2005). Interestingly, disruption of SET1 complex by depletion of WDR82, SET1, and CFP1 decreased H3K4me3 occupancy on promoters, while increased enhancer activity proved by increased H3K4me3 occupancy (Clouaire et al. 2012; Clouaire et al. 2014). Given that H2Bub1 increases H3K4me3 mainly via modulating the activity of SET1 complex, we speculate that the gene-body specific H2Bub1 modulates enhancer activity at a great distance.

1.6 Aims of this project

In the past decades, H2Bub1 has been believed to be tightly associated with embryogenesis and tumorigenesis. Although active gene transcription is tightly correlated to the occupancy of H2Bub1 in the transcribed region, depletion of the H2B ubiquitin ligases RNF20 or RNF40 alters (including increasing and decreasing)

the expression of only a subset of genes. In order to investigate the role of H2Bub1 in gene transcription, we utilized the next generation sequencing technology to study the genome-wide occupancy of H2Bub1, H3K4me3, H3K27me3, and H3K27ac occupancy in inducible *Rnf40* knockout mouse embryo fibroblasts (MEF). We observed that low and moderate levels of H2Bub1 are particularly associated with RNF40-dependent gene expression changes. Interestingly, the downregulation of RNF40-dependent genes was related to the narrowing of broad H3K4me3 peaks in H2Bub1-deficient MEFs while the upregulation of gene expression was dependent upon a loss of *Ezh2* transcription and decreased H3K27me3 near TSS, resulting in the up-regulation of many H3K27me3-targeted bivalent genes. Moreover, many upregulated genes are highly associated with the activation of FOXL2-bound enhancers. Together these findings uncover a previously unknown function of H2Bub1 and RNF40 in the indirect repression of gene transcription via the maintenance/activation of PRC2 and indirect repression of *Foxl2* transcription and provide further insight into the context-dependent intricacies of epigenetic regulation.

2. Materials

2.1 Equipment

2.1.1 Technical equipment

Equipment	Source
Pipettes "Research" Series	Eppendorf AG, Hamburg
Qubit® 2.0 Fluorometer	Invitrogen GmbH, Karlsruhe
Nano Drop® ND-1000	Peqlab Biotechnology GmbH, Erlangen
Bioruptor	Diagenode SA, Liège, Belgium
Agarose gel chamber	Harnischmacher Labortechnik, Kassel
Centrifuge 4 °C	Eppendorf AG, Hamburg
Balance	Sartorius AG, Göttingen
MiniVE	GE Healthcare Europe GmbH, München
CFX96™ Optical Reaction Module	Bio-Rad Laboratories GmbH, München
C1000TM Thermal Cycler	Bio-Rad Laboratories GmbH, München
Counting chamber (Neubauer)	Brand GmbH & Co. KG, Wertheim
Biological Safety Cabinet	Thermo Fisher Scientific, Waltham, USA
Freezer -20 °C	Liebherr GmbH, Biberach
Freezer -80 °C "Hera freeze"	Thermo Fisher Scientific, Waltham, USA
5100 Cryo 1 °C Freezing Container	Thermo Fisher Scientific, Waltham, USA
Microwave	Clatronic International GmbH, Kempen
Incubator (cell culture)	Thermo Fisher Scientific, Waltham, USA
Pipettes "Research" Series	Eppendorf AG, Hamburg
Thermomixer comfort	Eppendorf AG, Hamburg

Ultrapure Water System "Aquintus"	MembraPure GmbH, Bodenheim
Gel Imager "Gel iX imager"	Intas Science Imaging GmbH, Göttingen
Magnet stirrer "MR3001"	Heidolph GmbH & Co. KG, Schwabach
Microscope Axio Scope A1	Carl Zeiss MicroImaging GmbH, Göttingen
Vacuum pump	Integra Bioscienc. AG, Zizers, Switzerland
Bandelin Sonoplus Sonicator	Bandelin electr. GmbH & Co. KG, Berlin
Microwave	Clatronic International GmbH, Kempen
pH meter	inoLab® WTW GmbH, Weilheim
Repeat Pipette	Eppendorf AG, Hamburg
Table centrifuge (GMC-060)	LMS Co., Ltd., Tokyo, Japan
Scanner (CanoScan 8600F)	Canon GmbH, Krefeld
Pipette Aid® portable XP	Drummond Scientific Co., Broomall, USA
Vortex mixer	Scientific Industries, Inc., Bohemia, USA
Shaker "Rocky"	Schütt Labortechnik GmbH, Göttingen
Test tube rotator	Schütt Labortechnik GmbH, Göttingen
Pestle	Sartorius AG, Göttingen
Qubit 2.0 fluorometer	Life Technologies, USA
Water bath "TW 20"	JULABO Labortechnik GmbH, Seelbach
Centrifuge (Megafuge 1.OR)	Thermo Fisher Scientific, Waltham, USA
Incubator (bacteria culture)	Infors AG, Bottmingen
Incubator (bacteria)	Memmert GmbH & Co. KG, Schwabach
ChemiDoc™ MP Imaging System	Bio-Rad Laboratories GmbH, München
DynaMag™-96 Side Magnet	Thermo Fisher Scientific, Waltham, USA
MagnaRack™ Magnetic Separation Rack	Thermo Fisher Scientific, Waltham, USA

Agilent 2100 Bioanalyzer	Agilent Technologies, California, USA
Freezer -150 °C (MDF-C2156VAN)	Panasonic, Japan
DELL Optiplex 7010	DELL, Texas, USA

2.1.2 Bioinformatic software

Name	Source
Galaxy Cistrome	http://cistrome.org/ap/root
Galaxy	http://galaxyproject.org/
Galaxy deepTools	http://deeptools.ie-freiburg.mpg.de/
Bowtie 2.0	
Ruby Script	https://github.com/judofyr/rubyscript
DESeq	
Gene Set Enrichment Analysis (GSEA)	http://www.broadinstitute.org/gsea/index.jsp
DAVID analysis	https://david.ncifcrf.gov/
Primer designing tool NCBI/Primer-BLAST	www.ncbi.nlm.nih.gov/tools/primer-blast/
Integrative Genomics Viewer (IGV)	https://www.broadinstitute.org/igv/
Reduce Visualize Gene Ontology (REViGO)	http://revigo.irb.hr/
Genomic Regions Enrichment of Annotations Tool (GREAT)	http://bejerano.stanford.edu/great/public/html/
R statistical software	https://www.r-project.org/

2.2 Consumable materials

Name	Source
Syringe filter, Ca-membrane, 0,20 m	Sartorius AG, Göttingen
Protan® Nitrocellulose transfer membrane	Whatman GmbH, Dassel

Pipette filter tips	Sarstedt AG & Co., Nümbrecht
Pipette tips	Greiner Bio-One GmbH, Frickenhausen
Petri dish 92×16 mm	Sarstedt AG & Co., Nümbrecht
Parafilm® "M"	Pechiney Plastic Packaging, Chicago, USA
Microtube 0,5 ml, 1.5 ml, 2 ml	Sarstedt AG & Co., Nümbrecht
Eppendorf® LoBind microcentrifuge tubes	Eppendorf AG, Hamburg
NORM-JECT Syringes	Henke Sass Wolf GmbH, Tuttlingen
96 Multiply® PCR plate	Bio-Rad Laboratories GmbH, München
Microtube 1.5 ml, conical	VWR International GmbH, Darmstadt
Gel blotting paper (Whatman paper)	Sartorius AG, Göttingen
Cryo TubeTM Vial (1.8 ml)	Thermo Fisher Scientific, Waltham, USA
Cellstar tissue culture dish 100×20 mm	Greiner Bio-One GmbH, Frickenhausen
Cell scraper (16 cm, 25 cm)	Sarstedt AG & Co., Nümbrecht
Cellstar tissue culture dish 145×20 mm	Greiner Bio-One GmbH, Frickenhausen
Cellstar 6- and 12-well cell culture plate	Greiner Bio-One GmbH, Frickenhausen
Cellstar PP-tube 15 and 50 ml	Greiner Bio-One GmbH, Frickenhausen

2.3 Chemicals

Name	Source
Albumin Fraction V	Carl Roth GmbH & Co. KG, Karlsruhe
Aprotinin	Carl Roth GmbH & Co. KG, Karlsruhe

Ammonium persulfate	Carl Roth GmbH & Co. KG, Karlsruhe
Adefodur WB fixing concentrate	Adefo-Chemie GmbH, Dietzenbach
Adefodur WB developing concentrate	Adefo-Chemie GmbH, Dietzenbach
Agarose	Biozym Scientific GmbH, Oldendorf
Acetic acid	Carl Roth GmbH & Co. KG, Karlsruhe
Ammonium sulfate	Carl Roth GmbH & Co. KG, Karlsruhe
Bromophenol blue	Sigma-Aldrich Co., St. Louis, USA
Chloroform	Carl Roth GmbH & Co. KG, Karlsruhe
Calcium Chloride	Carl Roth GmbH & Co. KG, Karlsruhe
Charcoal Dextran treated FBS	Thermo Scientific HyClone, Logan, USA
Co-precipitant Pink	Bioline, Luckenwalde
Diethylpyrocarbonate (DEPC)	Carl Roth GmbH & Co. KG, Karlsruhe
Dimethyl sulfoxide (DMSO)	AppliChem GmbH, Darmstadt
GlutaMAX™ DMEM	GIBCO®, Invitrogen GmbH, Darmstadt
dNTPs	Carl Roth GmbH & Co. KG, Karlsruhe
Ethidium bromide	Carl Roth GmbH & Co. KG, Karlsruhe
Ethanol absolute	Th. Geyer GmbH & Co. KG, Renningen
Ethylenediaminetetraacetic acid (EDTA)	Carl Roth GmbH & Co. KG, Karlsruhe

EPZ 6438,EZH2 Inhibitor	Absource Diagnostics GmbH, München
Fetal Bovine Serum (FBS)	Thermo Scientific HyClone, Logan, USA
FBS Superior	Biochrom GmbH, Berlin, Germany
Formaldehyde	Sigma-Aldrich Co., St. Louis, USA
Glycine	Carl Roth GmbH & Co. KG, Karlsruhe
ß-Glycerolphosphate (BGP)	Sigma-Aldrich Co., St. Louis, USA
Hydrochloric acid (HCl)	Carl Roth GmbH & Co. KG, Karlsruhe
Isopropanol	Carl Roth GmbH & Co. KG, Karlsruhe
Lithium chloride (LiCl), 8M	Sigma-Aldrich Co., St. Louis, USA
Leupeptin	Carl Roth GmbH & Co. KG, Karlsruhe
lipofectamine ^{TM} 2000	Invitrogen GmbH, Karlsruhe
QIAzol™ Lysis Reagent	QIAGEN, Hilden
Methanol	M. Baker B.V., Deventer, Netherlands
Magnesium chloride (MgCl2)	Carl Roth GmbH & Co. KG, Karlsruhe
M-MLV Reverse Transcriptase	New England Biolabs, Frankfurt am Main
N-ethylmaleimide (NEM)	Sigma-Aldrich Co., St. Louis, USA
NonidetTM P40 (NP-40)	Sigma-Aldrich Co., St. Louis, USA
Non-Essential Amino Acid (NEAA)	GIBCO®, Invitrogen GmbH, Darmstadt

Opti-MEM	GIBCO®, Invitrogen GmbH, Darmstadt
PBS tablets	GIBCO®, Invitrogen GmbH, Darmstadt
Penicillin-Streptomycin solution	Sigma-Aldrich Co., St. Louis, USA
Pefabloc SC Protease Inhibitor	Carl Roth GmbH & Co. KG, Karlsruhe
Potassium chloride (KCl)	AppliChem GmbH, Darmstadt
Protein-A Sepharose CL-4B	GE Healthcare, Uppsala, Sweden
Protein-G Sepharose 4 Fast Flow	GE Healthcare, Uppsala, Sweden
Lipofectamine® RNAiMAX Reagent	Invitrogen GmbH, Karlsruhe
RNase inhibitor	New England Biolabs, Frankfurt am Main
Rotiphorese® Gel 30	Carl Roth GmbH & Co. KG, Karlsruhe
Roti®-Phenol	Carl Roth GmbH & Co. KG, Karlsruhe
Rotipuran® Chloroform	Carl Roth GmbH & Co. KG, Karlsruhe
Rotipuran® Isoamylalcohol	Carl Roth GmbH & Co. KG, Karlsruhe
SepharoseTM CL-4B	GE Healthcare, Uppsala, Sweden
Sodium chloride	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium acetate	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium aside	AppliChem GmbH, Darmstadt
Skim milk powder	Carl Roth GmbH & Co. KG, Karlsruhe

Sodium deoxycholate	AppliChem GmbH, Darmstadt
Sodium fluoride	AppliChem GmbH, Darmstadt
Sodium dodecylsulfate	Carl Roth GmbH & Co. KG, Karlsruhe
Sodium hydroxide	Carl Roth GmbH & Co. KG, Karlsruhe
SYBR Green I	Roche Diagnostics GmbH, Mannheim
TEMED	Carl Roth GmbH & Co. KG, Karlsruhe
Tris	Carl Roth GmbH & Co. KG, Karlsruhe
Trypsin-EDTA (0.05%)	GIBCO®, Invitrogen GmbH, Darmstadt
Tween-20	AppliChem GmbH, Darmstadt
Triton X-100	AppliChem GmbH, Darmstadt
(Z)-4-Hydroxytamoxifen (4-OHT)	Sigma-Aldrich Co., St. Louis, USA

2.4 Kits and reagents

Names	Source
Qubit dsDNA HS assay	Invitrogen GmbH, Karlsruhe
NucleoBond® Xtra Midi	MACHEREY-NAGEL GmbH & Co. KG, Düren
NucleoSpin® Gel and PCR Clean-up	MACHEREY-NAGEL GmbH & Co. KG, Düren

innuPREP Plasmid Mini Kit Plus	AJ Innuscreen GmbH, Berlin
NEBNext® Ultra™ RNA Library Prep Kit for Illumina®	New England Biolabs, Frankfurt am Main
NEBNext® ChIP-Seq Library Prep Master Mix Set for Illumina®	New England Biolabs, Frankfurt am Main
NEXTflex™ RNA-Seq Kit	Bioo Scientific, Austin
NEBNext® Poly(A) mRNA Magnetic Isolation Module	New England Biolabs, Frankfurt am Main
SuperSignal® West Dura	Thermo Fisher Scientific, Waltham, USA
Immobilon Western Chemiluminescent HRP Substrate	Millipore, Billerica, USA
Bioanalyzer DNA High sensitivity kit	Agilent, Santa Clara, USA

2.5 Nucleic acids

2.5.1 Primers for PCR

All primers in this study were designed using the NCBI primer designing tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/), and ordered from Sigma Aldrich, Hamburg. Reverse Transcription primers were used 9-mer random primer.

qRT-PCR primers

Gene	Forward Primer (5' to 3')	Reverse Primer (5' to 3')
Rplp0	TTGGCCAATAAGGTGCCAGC	CTCGGGTCCTAGACCAGTGT
Ezh2	TCCATGCAACACCCAACACA	AACTCCTTAGCTCCCTCCAGAT

Eed	AAGAACCTGGAGGGAGGCG	TGACAGCATCGTCATTTTCGTC
Ezh1	GCAAGTCCCCCAACTTCCAA	ACATACAGAGCCTTTGCTCCC
Suz12	AGCATCAAAAGCTTGTCTGCAC	ACTTTCACAAGCAGGACTTCCA
Psrc1	AATTCGAACGAAGCTCCGCC	CGGCTTAGACCTCGCTTCAC
Myl9	GCGCCGAGGACTTTTCTTCT	TCTTGGCCTTGGCTCTCTTG
Loxl3	TTGTGCCTAGTCGAAGTGCCC	GGGCAGCAATACCACACACT
Hoxc6	ATGAATTCGCACAGTGGGGT	GCCGAGTTAGGTAGCGGTTG
Hoxc8	GATGAGACCCCACGCTCCT	CTTCAATCCGGCGCTTTCTG
Hoxc10	CGGATAACGAAGCTAAAGAGGAGA	TCCAATTCCAGCGTCTGGTG
Nat8l	GCCCTGCGCTACTACTACAG	CCCAGAAACAGGAACCAGGTG
Kcnc3	CCATCCGAAAAGCCACTGGT	GCTCGTCCACTAGGGGGATA
Tgfa	CTGCTAGCGCTGGGTATCCT	CTGAGTGTGGGAATCTGGGC
Chd5	TGGACCCTGACTACTGGGAG	TCATCCTGCCACTCCTGGTC
Foxl2	TGCAACCGAGTTCTCATCCC	TAAAGACCTGGCCTGGAGGT
Foxl2os	AGCAAGCTGGTCTAACGCTC	AGAGCCAAAAGGTACCTGCC
Hoxc13	CCCTGTTGAAGGCTACCAGC	AACCACGTCTGGGAAGGGAG
Esr2	GAAAGCTGCTGGATGGAGGT	CCTCATCCCTGTCCAGAACGA
Efna5	TGCAATCCCAGACAACGGAA	TGGCTCGGCTGACTCATGTA

ChIP-qPCR primers		
Gene	**Forward Primer (5' to 3')**	**Reverse Primer (5' to 3')**
Ezh2- TSS	CCGGAATCCACAGTTCACTCG	GACAGCTTTCTGAGCGGTCG
Ezh2- Gene body	TGAAGGTTTTGGGAGGGTGG	AGGCAGGTAAGCAGTTTGGG
Hoxc13- TSS	GCCGGAGAGCCTTATGTACG	CCAATACAGGGTGCGGGAG
Hoxc6-TSS	CCACCGCCTATGATCCAGTG	GGGAGTCGAGTAGATCCGGT

2.5.2 Plasmid and primers for cloning

Vector pSG5-HA-ERT2-P2A-Hyg was used to overexpress wildtype and mutated Ezh2.

	Name	Primer (5' to 3')
Wildtype *Ezh2*	BamHI-*Ezh2*-For	GCTGACGGATCCATGGGCCAGACTGGGAAGAAATCTG
	SpeI-*Ezh2*-Rev	GCTGACACTAGTAGGGATTTCCATTTCTCGTTCGATG
Mutated *Ezh2*	H689A-For	TGCTAATGCTTCAGTAAATCCAAAC
	H689A-Rev	GCATAGCAGTTTGGATTTACTGAAGCATTAGCA

2.5.3 Primers for genome typing mouse embryos

Detected target	Name	Primer (5' to 3')
LoxP site	mRnf40_3LoxP_13254F	TGGGCCCAGGTGGATGCCTGAA
	mRnf40_3LoxP_15985R	AGGCCACAGCAGGGACCATCA
Cre^{ERT2}	ERT2-For	AAAGTCGCTCTGAGTTGTTAT
	ERT2-Rev1	GGAGCGGGAGAAATGGATATG
	ERT2-Rev2	CCTGATCCTGGCAATTTCG3

2.5.4 siRNA Oligonucleotides

All siRNA Oligonucleotides were ordered from Dharmacon, Lafayette, CO, USA.

Target gene	Sequence (5' to 3')
siGENOME Nontargeting siRNA pool #5	---
Foxl2 SMARTpool siRNAs	5'-GCGCAGUCAAAGAGGCCGA-3'
	5'-ACUCGUACGUGGCGCUCAU-3'
	5'-UAGCCAAGUUCCCGUUCUA-3'
	5'-CGGGACAACACCGGAGAAA-3'

2.6 Proteins

2.6.1 Molecular weight standards

GeneRuler™ 1 Kb Plus DNA Ladder	Thermo Fisher Scientific, Waltham, USA
PageRuler™ Plus Prestained Protein Ladder (10 to 250 kDa)	Thermo Fisher Scientific, Waltham, USA

2.6.2 Antibodies

Primary antibodies used for western blot (WB) and ChIP

Target	Clone	Cat.No.	WB	ChIP	Source
H3	-	ab10799	1:5000	-	Abcam
HSC70	B-6	sc-7298	1:50,000	-	Santa Cruz
H2B	53H3	2934	1:5000	-	Cell signaling
RNF40	-	15621-1-AP	1:500	-	Acris
EZH1	-	ab13665	1:1000	-	Abcam
SUZ12	D39F6	3737	1:1000	-	Cell signaling
EZH2	-	4905	1:1000	-	Cell signaling
H2Bub1	7B4	-	1:10	-	Hybridoma (Prenzel et al., 2011)
H2Bub1	D11	5546	-	1.5 µl	Cell signaling
H3K27me3	-	pAb-195-050	1:1000	1 µg	Diagenode
H3K27ac	-	pAb-196-050	1:1000	1 µg	Diagenode
H3K4me3	-	pAb-003-050	1:1000	1 µg	Diagenode
IgG	-	Ab37415	-	1 µg	Abcam

Secondary antibodies for western blot		
Name	**Cat.No.**	**Source**
Goat anti-rabbite IgG HRP	sc-2004	Santa Cruz
Goat anti-mouse IgG HRP	sc-2005	Santa Cruz

2.6.3 Enzymes

Name	**Source**
Proteinase K	Invitrogen GmbH, Karlsruhe
RNase A	Qiagen GmbH, Hilden
Phusion DNA Polymerases	Thermo Fisher Scientific, Waltham, USA
M-MLV Reverse Transcriptase	New England Biolabs, Frankfurt am Main
Taq DNA Polymerase	Prime Tech, Mink, Belarus
T4 DNA Ligase	New England Biolabs, Frankfurt am Main
BamHI	Thermo Fisher Scientific, Waltham, USA
SpeI	Thermo Fisher Scientific, Waltham, USA

2.7 Animal and Cells

2.7.1 Bacterial Cells

Escherichia coli DH10BTM was got from Invitrogen GmbH, Karlsruhe.

2.7.2 Mice

Rosa26-Cre^{ERT2}, *Rnf40*$^{loxP/wt}$ mice was generated from Johnsen's lab.

2.7.3 Mouse embryonic fibroblast cells (MEFs)

Rnf40 inducible knockout MEFs were isolated from 13.5 postcoitum mouse embryos.

2.8 Buffers and cell culture medium

2.8.1 Buffers for PCR

10X PCR buffer (store at room temperature up to 1 month)

Stock	Final concentration
1.5 M Tris-HCl (pH 8.8)	750 mM
1 M $(NH_4)_2SO_4$	200 mM
10% Tween-20	0.1%

PCR Master Mix (store at -20°C up to 4 months)

Stock	Final concentration
10X PCR buffer	1X
25 mM $MgCl_2$	3 mM
SYBR Green (1:100)	1:80000
20 mM dNTPs	0.2 mM
5 U/µl Taq DNA Polymerase	20 U/ml
10% Triton X-100	0.25%
1 M Trehalose	300 mM

2.8.2 Buffers for western blot

10X PBS (store at room temperature up to 1 month)

Stock	Final concentration
NaCl	0.73 M
KCl	0.027 M
NaH2PO4 * 7H2O	14.3 mM
KH2PO4	14.7 mM

RIPA Lysis Buffer (store at 4°C up to 1 month)

Stock	Final concentration
5 M NaCl	150 mM
0.5 M EDTA (pH 8.0)	5 mM
1 M Tris (pH 8.0)	50 mM
10% NP-40	1.0%

10% Sodium deoxycholate	0.5%
10% SDS	0.1%

10X TBS-T (pH 7.6) (store at 4°C up to 1 month)

Stock	Final concentration
Tris	0.1 M
NaCl	1.5 M
Tween-20	1%

10X Western salts (store at 4°C up to 1 month)

Stock	Final concentration
Tris	0.25 M
Glycine	0.86 M
SDS	0.7 mM

6X Laemmli buffer (store at -20°C up to 4 months)

Stock	Final concentration
1M Tris-Cl (pH 6.8)	0.375 M
SDS	12%
Glycerol	60%
DTT	0.6M
Bromophenol blue	0.06%

Transfer Buffer (store at 4°C, can be used for two or three times)

Tris	0.048M
glycine	0.039M
methanol	20%
SDS	0.00375%

Blocking solution (store at 4°C up to 1 month)

TBST	1X
Milk	5%

2.8.3 Buffers for ChIP

Nuclear preparation buffer (store at 4°C up to 1 month)

Stock	Final concentration
5M NaCl	150 mM
0.5 M EDTA (pH 8.0)	20 mM
1 M Tris-HCl (pH 7.5)	50 mM
10% NP-40	0.5%
10% Triton X-100	1%
0.5 M NaF	20 mM

Sonication buffer-1 (freshly used)

Stock	Final concentration
1 M Tris-HCl (pH 8.0)	50 mM
0.5 M EDTA (pH 8.0)	10 mM
10% SDS	1%

Sonication buffer-2 (store at 4°C up to 1 month)

Stock	Final concentration
0.5 M EDTA (pH 8.0)	20 mM
1 M Tris-HCl (pH 8.0)	50 mM
5M NaCl	150 mM
10% NP-40	1%
0.5 M NaF	20 mM

Dilution buffer (store at 4°C up to 1 month)

Stock	Final concentration
0.5 M EDTA (pH 8.0)	20 mM
1 M Tris-HCl (pH 8.0)	50 mM
5M NaCl	150 mM
10% NP-40	1%
0.5 M NaF	20 mM
10% (w/v) Sodium deoxycholate	0.5%

IP Buffer (store at 4°C up to 1 month)

Stock	Final concentration
0.5 M EDTA (pH 8.0)	20 mM
1 M Tris-HCl (pH 8.0)	50 mM
5M NaCl	150 mM
10% NP-40	1%
0.5 M NaF	20 mM
10% (w/v) Sodium deoxycholate	0.5%
10% (w/v) SDS	0.1%

Wash buffer (store at 4°C up to 1 month)

Stock	Final concentration
8 M LiCl	0.5 M (add it freshly before using this buffer)
10% NP-40	1%
10% (w/v) Sodium deoxycholate	0.5%
0.5 M EDTA (pH 8.0)	20 mM
1 M Tris-HCl (pH 8.5)	10 mM
0.5 M NaF	20 mM

TE buffer (store at 4°C up to 1 month)

Stock	Final concentration
0.5 M EDTA (pH 8.0)	1 mM
1 M Tris-HCl (pH 8.0)	10 mM

2.8.4 Cell culture medium

LB-medium (1 L) (store at 4°C up to 1 month)

Yeast extract	5 g
Tryptone	10 g
NaCl	10 g

MEF growth medium (store at 4°C up to 1 month)

GlutaMAX™ DMEM, high-glucose	---
FBS Superior	10%
Non-Essential Amino Acid (100X)	1X
Penicillin	100 U/ml
Streptomycin	100 µg/ml

3. Methods

3.1 Conditional *Rnf40* knockout mouse model

All animal work was performed in agreement with the Institutional Animal Care and Use Committee and the Institutional Guidelines for Humane Use of Animals in Research. Conditional *Rnf40* knockout mice were generated using a construct containing two loxP sites flanking exons 3 and 4 of the *Rnf40* gene (Fig.4) and a neomycin selection cassette was surrounded by two short flippase recognition target (FRT) sites. The targeting construct was -transfected in MPI II ES cells by electroporation and targeted clones were identified by quantitative and long-range PCR. Following the generation of chimeras and verification of germline transmission, the neomycin cassette was removed to generate *Rnf40*^{loxP} mice by crossing to a transgenic mouse line expressing the FLP recombinase in all tissues (Farley et al. 2000). The *Rnf40*^{loxP} mice were next crossed to a transgenic line expressing a tamoxifen-inducible Cre recombinase (Cre^{ERT2}) inserted into the ubiquitously expressed *Rosa26* locus (Hameyer et al. 2007).
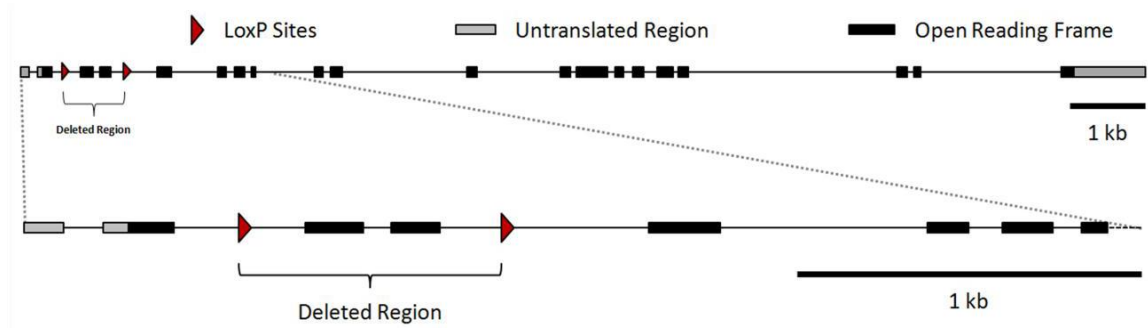
Fig. 4 Conditional *Rnf40* knockout construct model

3.2 Cell culture

3.2.1 Isolation and culturing of primary MEFs

The inducible *Rnf40* knockout MEFs were obtained by intercrossing *Rosa26*-Cre[ERT2], *Rnf40*[loxP/wt] mice. MEFs were isolated from 13.5 postcoitum mouse embryos as previously described (Takahashi and Yamanaka 2006). In brief, the legs, brain and dark red organs were removed the rest of the cleaned tissue was made into single cells sequentially via cutting into small pieces and trypsinizing with 0.25% of trypsin-EDTA. The single cells from each embryo were separately plated in 15 cm tissue culture dish and cultured in growth medium (GlutaMAX™ DMEM supplemented with 1x NEAA, 10% FBS Superior, 100 U/ml penicillin, and 100 µg/ml streptomycin) at 37°C and 5% CO_2 conditions. The primary cells were frozen at 70% confluency in MEF freezing media (DMEM high glucose supplemented with 8% DMSO and 50% FBS) and stored in 150°C freezer.

3.2.2 Inducing *Rnf40* knockout and inhibiting EZH2 enzymatic activity in MEFs

For deletion of the conditional *Rnf40* allele, 10^5 MEFs were passaged in growth medium supplemented with 250nM of (Z)-4-Hydroxytamoxifen (4-OHT), while the control cells were treated with 0.5 µl of ethanol. After 5 days, cells were grown for another 3 days in the absence 4-OHT. Cells were passaged every three days.

For inhibition of EZH2 methyltransferase activity in *Rnf40* wild type (*Rnf40*[+/+]) MEFs, 10^5 cells were cultured in growth medium supplemented with 1µM of an EZH2 small molecular inhibitor (EPZ-6438) for 2 days, while the control cells were treated with 1µl DMSO.

3.2.3 siRNA transfection in *Rnf40$^{+/+}$*, *Rnf40$^{-/-}$*, and EZH2 inhibited MEFs

Forward and reverse siRNA transfection were performed in *Rnf40$^{+/+}$*, *Rnf40$^{-/-}$*, and EZH2 inhibited (EZH2i) MEFs using lipofectamine® RNAiMAX Reagent according to the manufacturer's protocol. For knockdown of *Foxl2* gene, targeted mouse *Foxl2* SMARTpool siRNAs were transfected, while cells transfected with non-targeting siRNA were used as a negative control. These experiments were performed in 6-well plates. 30 pmol of siRNA, 500 µl of Opti-MEM medium, and 5 µl of lipofectamine® RNAiMAX Reagent were pre-mixed in a 1.5 ml tube, and incubated for 20 min at room temperature. Meanwhile, cells were trypsinized into single cells and diluted into 50,000 cells/ml using penicillin/streptomycin-free MEF growth medium. 2 ml of diluted MEFs and 0.5 ml of siRNA-lipofectamine® RNAiMAX complex was added to each well of the 6-well plate, and then mixed properly. After approximately 16 hours, the medium was changed into fresh MEF growth medium containing penicillin/streptomycin, and EZH2i group cells were additionally treated with 1 µM of EPZ-6438. Cells were harvested after 48 hours.

3.3 Molecular biology

3.3.1 Genome typing for mouse embryos

Tissues from legs of each embryo were lysed overnight at 56°C in 500 µl of Lysis buffer (100 mM Tris-HCl pH 8.5, 5 mM EDTA, 0.2% SDS, and 200 mM NaCl) supplemented with 100 µg of proteinase K. After centrifugation at full speed for 10 min, the supernatant was vortexed with 500 µl isopropanol. DNA was further pelleted by centrifugation at full speed for 10 min at room temperature, and washed with 70% EtOH. After brief air-drying, DNA pellet was redissolved in 50 µl water. DNA concentration was detected using Nano Drop® ND-1000. 300 ng of each DNA

sample was used for genotyping. PCR primers for detection of Cre-ER^{T2} or the sequence containing loxP are listed in **2.5.3**.

A total 50 µl volume of PCR reaction for detection of Cre-ER^{T2} or loxP sit:

	Cre-ER^{T2} (µl)	loxP site (µl)
Waster	13.8	14.8
10× Hot fire buffer	2.5	2.5
dNTP (2 mM)	2.5	2.5
MgCl$_2$ (25 mM)	2	2
	1	1
primers	1	1
	1	
Hot Fire DNA polymerase	0.2	0.2

PCR reaction was performed in C1000TM Thermal Cycler:

loxP site	
95°C	15 min
95°C	30 sec
60°C	30 sec
72°C	1 min
72°C	10 min
4°C	Infinite hold

35 cycles (95°C 30 sec, 60°C 30 sec, 72°C 1 min)

Cre-ER^{T2}	
95°C	15 min
95°C	30 sec
55°C	45 sec
72°C	1 min
72°C	10 min
4°C	Infinite hold

35 cycles (95°C 30 sec, 55°C 45 sec, 72°C 1 min)

The PCR produces were separated in 0.8% agarose gel, and the detected in Gel Imager.

3.3.2 RNA isolation

Total RNA from cultured cells were isolated using QIAzol™ Lysis Reagent according to the manufacturer's protocol. First of all, cells in 6-well plates were washed twice using PBS and lysed by adding 500 µl of QIAzol™ Lysis Reagent. Cells were then

harvested into 1.5 ml tubes using scrapers and kept on ice. Each sample was mixed with 100 µl of chloroform by vortexing for 15 sec and then centrifuged at 12,000 x g for 15 min at 4 °C. The upper, aqueous phase (approximately 250 µl) was collected in a new 1.5 ml tube, and mixed with 300 µl of isopropanol by vortexing for 15 sec. All samples were placed at -20 °C overnight. Next day, RNA pellets were collected by centrifuging at 12,000 x g for 30 min at 4°C, and washed twice by 70% of ethanol (prepared with DEPC water). After brief air-drying the RNA pellet was redissolved in 40 µl of DEPC water.

3.3.3 First-Strand Synthesis of cDNA

Before cDNA synthesis, the concertation of total RNA was measured using Nano Drop® ND-1000. 1 µg of total RNA from each sample, 2 µl of 15 µM random primers, and 4 µl of 2.5 µM dNTP were pre-mixed and incubated at 70°C for 5 min. After that, 2 µl of M-MLV 10× reaction buffer, 10 U of RNase inhibitor, 25 U of M-MLV Reverse Transcriptase, and 1.625 µl of DEPC water were mixed with each sample. First-strand cDNA was synthesized by incubating at 42°C for 1 hour followed by 95°C for 5 min. Finally, each sample was diluted in 30 µl of DEPC water.

3.3.4 High throughout RNA sequencing (RNA-Seq)

Transcript profiles in *Rnf40*$^{+/+}$ and *Rnf40*$^{-/-}$ MEFs at passage 3 were investigated using RNA-Seq. After measuring the concentration of total RNA in the three replicates of each condition, the integrity of each sample was detected using Bioanalyzer 2100. Libraries were prepared from 1 µg of total RNA using the NEXTflex™ Rapid Directional RNA-Seq Kit according to the manufacturer's protocol. In brief, Libraries were performed by a series of steps as follows: mRNA purification using Oligo d(T)25 magnetic beads, mRNA fragmentation, first-strand cDNA synthesis, cDNA amplification, adaptor extension and size selection, and

amplification of size-selected fragments. Each library was quantified by Qubit 3.0, and then the fragment size (approximately 300 bp) was determined by using the DNA 1000 chip on the Bioanalyzer 2100.

Finally, cDNA library sequencing was performed by using the cBot and HiSeq2000 from Illumina (SR; 1x50 bp; 6 GB ca. 30-35 million reads per sample) at the Transcriptome Analysis Laboratory (TAL), University of Göttingen.

3.3.5 Chromatin immunoprecipitation (ChIP)

Cross linking and cell harvest

First of all, approximately 2×10^6 MEFs in 15 cm plates were crosslinked for 10 min by adding 1% formaldehyde at room temperature. The formaldehyde was quenched by adding 125 mM glycine. After washing twice in ice-cold PBS, the nuclear pellets were harvested in cold nuclear preparation buffer supplemented multiple proteinase inhibitors (1 mM N-ethylmaleimide, 10 mM β-glycerophosphate, 1 ng/µl Aprotinin/Leupeptin, 1 mM Pefabloc, 10µM iodo acetamide and 1 mM nickel chloride). The pellets were further washed once with nuclear preparation buffer, frozen in liquid nitrogen and stored at -80°C.

Sonication

The nuclear pellets were re-suspended and lysed in 300 µl of sonication buffer-1for 15 min at 4°C. The sample was diluted using 100 µl of sonication buffer-2 and aliquoted in two 1.5 ml tubes, and sonicated for 30 cycles in the biorupter with 30sec on/off setting. The soluble chromatin fragments were cleared by centrifugation at 12,000 g for 10 min at 8°C, and diluted in 600 µl dilution buffer.

Preclearing

Chromatin fragments were pre-cleared by incubating with 100 μl of 50% slurry of sepharose beads for 1 hour at 4°C. After centrifugation at 12,000 g for 2 min at 4°C, the cleared supernatant was aliquoted in three 1.5 ml tubes (each around 300 μl) and stored at -80°C. 30 μl of supernatants was kept as input.

Immunoprecipitation

300 μl of each sample was diluted in 600 μl of IP buffer and incubated overnight at 4°C with protein specific targeting primary antibodies listed in 2.6.2, or the background binding non-specific control IgG antibody. The antibody bound chromatin fragments were precipitated by adding 30 μl of 50% slurry of protein A-sepharose. The sepharose beads were collected by centrifuging at 2,000 g for 2 min at 4°C. Finally, the immunoglobulin complexes were washed with several buffers: ice-cold IP buffer twice, wash buffer twice, IP buffer twice again, and TE buffer twice.

DNA isolation

RNAs in the samples were degraded by incubating for 30 min at 37°C with 50 μl of RNAse-A solution. In addition, proteins were eliminated by adding 50 μl of 2× Weinmann lysis buffer (100 mM Tris-HCl at pH 8, 20 mM EDTA, and 2% SDS) supplemented with 1 μg proteinase K and incubated overnight at 65°C. After centrifugation at 2,000 g for 2 min, 100 μl of 10 mM Tris-HCl (pH8) solution was added to the supernatant, and incubated 10 min at 65°C.This was followed by adding 10 μl 8M of LiCl, 4 μl colorless coprecipitant, and 200 μl of Phenol/chloroform/Isoamylic alcohol extraction (25:24:1) and vortexing. The aqueous phase containing DNA was collected in 1.5 ml LoBind microcentrifuge tubes after centrifugation at full speed for 2 min. In addition, the phenolic phase was vortexed with 200 μl of solution containing 10 mM of Tris-HCl (pH8) and 0.4 M LiCl, and then

the aqueous phase was harvested by centrifugation at full speed for 2 min. 400 µl of the aqueous phase containing DNA was mixed with 1 ml of EtOH and incubated for 2 hours at -80°C. DNA was pelleted via centrifugation at 15,000 g for 5 min at 4°C. After brief air-drying, the DNA pellet was redissolved in 40 µl of nuclease-free water.

3.3.6 Quantitative real-time PCR (qPCR)

Before starting qPCR, standard samples were prepared from all cDNA samples or input DNA of ChIP samples in different dilutions (1:1, 1:4, 1:16, 1:64, 1:256, and water). Each qPCR reaction involved 1 µl cDNA or ChIP DNA, 30 nM primers, 8.5 µl water, and 14 µl of PCR Master Mix (75 mM Tris-HCl (pH 8.8), 20 mM $(NH4)_2SO_4$, 0.01% Tween-20, 3 mM MgCl2, 1:80000 SYBR Green, 0.2 mM dNTPs, 20 U/ml Taq DNA Polymerase, 0.25% Triton X-100, and 300 mM Trehalose).

PCR reaction was performed in CFX96TM Optical Reaction Module using two-step protocol:

$$
\begin{array}{ll}
95°C & 2\ min \\
\left\{ \begin{array}{ll} 95\ °C & 15\ sec \\ 60\ °C & 1\ min \end{array} \right\} & 40\ cycles
\end{array}
$$

The PCR reaction was followed by a melting curve analysis from 60°C to 95°C with reads every 0.5°C.

3.3.7 ChIP-Seq library preparation

In order to investigate the genome-wide H2Bub1, H3K4me3, H3K27ac, and H3K27me3 bound regions, DNA isolated from the protein specific ChIP was followed by high-throughput sequencing in *Rnf40*[+/+] and *Rnf40*[-/-] MEFs at passage 3.

The concentration of each DNA sample was first measured using Qubit® 2.0 Fluorometer. 5 ng DNA was diluted in 50 µl EB buffer (10 mM Tris-HCl (pH8) and 0.2% Tween-20) and sonicated for 30 cycles in the biorupter with 30sec on/off setting. Libraries were prepared using the NEBNext Ultra DNA library preparation kit according to the manufacturer's protocol. Each library was quantified by Qubit® 2.0 Fluorometer, and then the fragment size (approximately 300 bp) was determined by using the DNA 1000 chip on the Bioanalyzer 2100.

Finally, 75 bp single-ended tags for H3K4me3 and 51 bp single-ended tags for other samples were sequenced with single indexing using NextSeq or HiSeq 2500 platforms, respectively.

3.3.8 Protein level analysis

3.3.4.1 SDS-PAGE

Protein extracts were prepared by lysing *Rnf40*$^{+/+}$, *Rnf40*$^{-/-}$, and EZH2i MEFs in RIPA buffer supplemented with proteinase inhibitors (1 mM Pefabloc, 1 ng/µl Aprotinin/Leupeptin, 10 mM BGP and 1 mM NEM) followed by sonication (15 sec at 10% power). Before sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE), samples were mixed with 6X laemmli buffer and boiled for 10 min at 95°C. Proteins were separated according to their molecular weight by SDS-PAGE as described before (Laemmli 1970).

3.3.4.2 Western blot analysis

After electrophoresis, the separated proteins were transferred from polyacrylamide gels to PVDF membranes using transfer buffer at 100 V for 1.5 hours. The membranes were blocked in blocking solution for 1 hour to prevention specific binding. Subsequently, the membranes were incubated with protein-specific targeted

primary antibodies at 4°C overnight. After washing out the unspecific bound antibodies with TBS-T, the primary antibody bound membranes were incubated with the corresponding horseradish peroxidase-conjugated anti-rabbit IgG or anti-mouse IgG secondary antibodies for 1 hour at room temperature. After washing thrice with TBS-T, HRP signals were detected by ChemiDoc™ MP Imaging System using enhanced chemiluminescence solution.

3.3.9 Data analysis

3.3.9.1 RT-qPCR and ChIP-qPCR data

Gene expression level in each cDNA sample was calculated using a standard curve as explained previously. Furthermore, the expression of each gene was normalized to *Rplp0* as an internal reference gene. Finally, the relative gene expression levels were calculated, and referred to as "Rel. mRNA level".

The DNA levels in input, IgG, and ChIP samples were quantified using a standard curve made from diluted input samples. DNA levels in IgG and ChIP samples were normalized to input samples, and expressed as "% input".

p-value was calculated using t-test, 'n.s' indicated no significant difference; $*p<0.05$; $**p<0.01$; $***p<0.0001$.

3.3.9.2 RNA-Seq data

RNA-seq data process

Sequencing data were transformed to bcl files by using BaseCaller software, and further transformed to fastq files with CASAVA (version 1.8.2). The quality of fastq data was checked using FastQC (version 0.64) in Galaxy. The raw fastq data from each sample was mapped to mouse reference transcriptome (UCSC, mm9) using

bowtie 2.0 tool under Linux environment, thereby generating sam files (Langmead and Salzberg 2012). Counts of each transcript were aggregated in a CSV table using Ruby Script (https://github.com/judofyr/rubyscript). DESeq analysis (Anders and Huber 2010) was performed and the normalized gene expression counts of each sample were calculated according to the sample count size, and differential gene expression between *Rnf40*[+/+] and *Rnf40*[-/-] conditions were computed. According to the differential gene expression table, *Rnf40*-dependent, -independent, and -suppressed gene clusters were selected as follows: *Rnf40*-dependent genes, baseMean>15, $\log_2$-fold change <-1, and p-value<0.05; *Rnf40*-independent genes, baseMean>15, -0.2<$\log_2$-fold change<0.2, and p-value>0.8; *Rnf40*-suppressed genes, baseMean>15, $\log_2$-fold change>1, and p-value<0.05.

Gene Set Enrichment Analysis (GSEA)

Pathway enrichment scores were calculated using GSEA according to user guide (Subramanian et al. 2005). The normalized gene expression counts, generated from DESeq analysis of RNA-Seq data, were compared with the published C4-curated gene sets database. Genes were sorted from left to right according to fold change in gene expression under *Rnf40*[-/-] vs. *Rnf40*[+/+] conditions. FDR p-value <0.05 was considered as significant enrichment.

Gene ontology analysis (GO)

GO analysis were performed to compare the *Rnf40*-dependent, *Rnf40*-suppressed, or the broadest H3K4me3 bound gene cluster to the published gene sets in "GOTERM_BP_ALL" using DAVID 6.7 (Huang da et al. 2009). The significant enriched GO terms (FDR<0.05) were represented as the Bubble plot generated by using REViGO (Supek et al. 2011).

3.3.9.3 ChIP-Seq data

The raw fastq files of ChIP-Seq data were generated by Dr. Daniela Indenbirken from Heinrich Pette Institute of Hamburg as the similar process as RNA-Seq data. After checking data quality using FastQC, each fastq file was mapped to mouse genome (UCSC, mm9) using bowtie (version 1.0.0) in Galaxy (Langmead et al. 2009). In order to identify the regions where reads were significantly enriched, we further performed Model-based Analysis of ChIP-seq (MACS) (version 1.0.0) for peak calling with the input of each condition as control and "p-value < 0.00001" cutoff for peak detection (Zhang et al. 2008; Liu 2014), thereby generating two important processed data: bed files containing regions significantly bound by the protein of interest and wig files containing ChIP signals on each bound region. Moreover, ChIP signals in each wig file were normalized to the filtered reads per hundred million. The normalized wig files were further transformed to bigwig files using 'Wig/BedGraph-to-bigWig' tool. The ChIP signal profile on each gene could be visualized by loading the bigwig file to Integrative Genomics Viewer (version 2.3.14) (Thorvaldsdottir et al. 2013).

The tables containing mouse genome elements (TSS, gene bodies, etc.) were obtained from UCSC Table Browser (Karolchik et al. 2004). The average signal of H3K4me3, H3K27me3, and H3K27ac near TSS (±1kb) and H2Bub1 on gene bodies were computed from the normalized bigwig files using ComputeMatrix in deepTools (Ramirez et al. 2014). Furthermore, Smooth Script plot analysis investigated the correlation between the given histone modifications and gene expression, and boxplot analysis was used to compare the signals between the given gene clusters, by using those qualified ChIP signals.

The heatmapper in deepTools was used to create heatmaps of each ChIP. CEAS (version 1.0.0) and aggregate profiles analysis were performed in Galaxy/Cistrome (Ji et al. 2006; Shin et al. 2009).

The H3K27me3 targeted distal regions were obtained by considering only the regions further than 5 kb upstream or downstream of gene bodies. Active enhancers were defined as enriched (+) H3K4me1 and H3K27ac but unenriched (-) H3K4me3 regions. Differential binding (DiffBind) analysis of H3K27me3 near TSS (±1kb) and distal regions or H3K27ac on enhancers under $Rnf40^{+/+}$ vs. $Rnf40^{-/-}$ conditions as described before (Ross-Innes et al. 2012).

Enhancer associated coding genes were identified using the **G**enomic **R**egions **E**nrichment of **A**nnotations Tool (GREAT version 3.0.0) with setting as: proximal (5 kb upstream 5 kb downstream) plus distal (up to 300kb) (McLean et al. 2010).

Sequence-based motif analysis for $Rnf40$-suppressed gene associated enhancers in $Rnf40^{-/-}$ MEFs was performed using oPOSSUM (version 3.0) (Kwon et al. 2012). The input file was given the sequences surrounding H3K27ac peak centers (±150 bp) on upregulated genes associated enhancers in $Rnf40^{-/-}$ MEFs. Background file was given all enhancer sequences excluding input sequences.

4. Results

4.1 H2Bub1 and transcription activation

4.1.1 H2Bub1 is enriched in the transcribed regions and gradually decreases toward the 3' end

To investigate the genome-wide H2Bub1 occupancy in the mouse genome, we performed chromatin immunoprecipitation coupled with sequencing (ChIP-seq) for H2Bub1 in mouse embryonic fibroblast cells (MEFs). In agreement with early findings in human cells (Fuchs et al. 2014), H2Bub1 selectively occupies gene bodies and is absent on distal intergenic region (Fig. 5A). In order to further characterize the H2Bub1 distribution within gene bodies, we compared H2Bub1 signals on each element. As shown here (Fig. 5B), H2Bub1 majorly occupies the intron and exon.
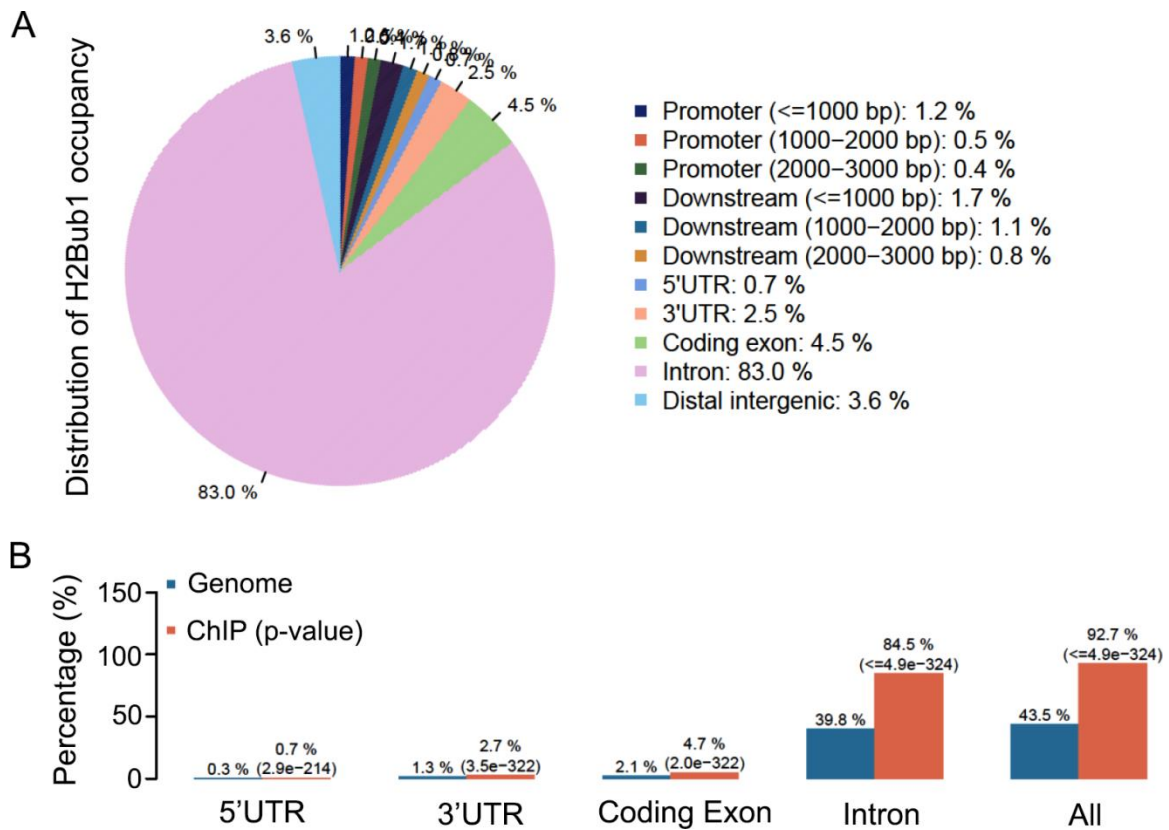
Fig.5 The distribution of H2Bub1 on various genome elements

(**A and B**) Enrichment on chromosome and annotation (CEAS) analysis of H2Bub1 occupancy on various genome elements in MEFS. P-values were calculated by χ2 test.

Furthermore, we investigated the dynamic occupancy pattern of H2Bub1 along gene body. Consistent to the observation in human cells (Jung et al. 2012; Nagarajan et al. 2014), H2Bub1 highly occupied near 5' end of transcribed regions and its occupancy gradually decreased towards the 3' end in MEFs as seen in the case of the *Lrrc1* gene (Fig. 6A and 6B). In addition, the dynamic pattern of H2Bub1 occupancy on exons and introns were observed by CEAS analysis. As shown in average concatenated profiles, H2Bub1 occupancy was remarkably decreased from 5' to 3' ends on exons compared to introns (Fig. 6C and 6D).
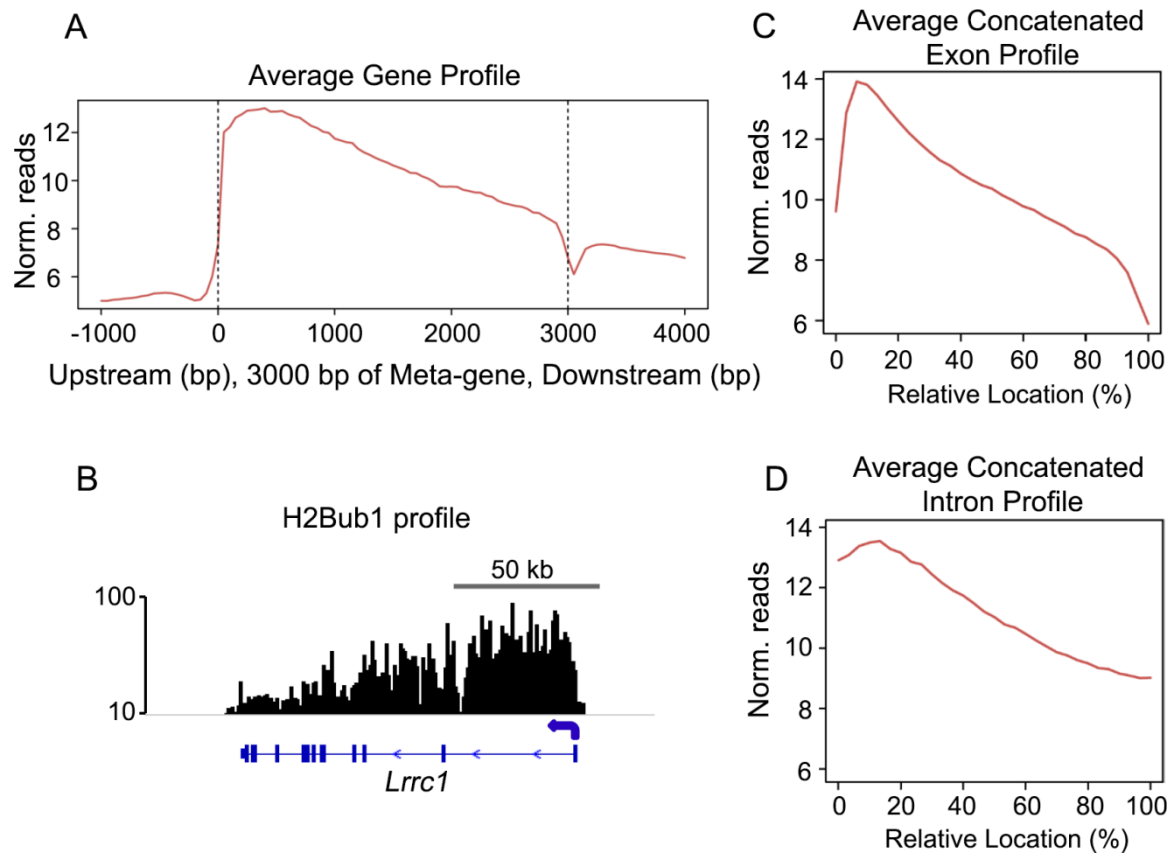
Fig.6 The dynamic pattern of H2Bub1 occupancy on gene body.

(**A**) The average ChIP signals on the meta-gene of 3 kb, which shows that H2Bub1 is enriched on gene bodies and decreases towards the 3' end.

(**B**) The average H2Bub1 signals on the *Lrrc1* gene.

(**C and D**) The average concatenated H2Bub1 signals on exons (C) and introns (D).

4.1.2 H2Bub1 is tightly correlated with gene transcription and active histone modifications

Previous studies have determined that highly transcribed genes are correlated with high H2Bub1 occupancy in different human cell lines (Minsky et al. 2008; Jung et al. 2012; Nagarajan et al. 2014). It was further suggested that H2Bub1 is tightly coupled with RNA polymerase II elongation rate (Johnsen 2012; Fuchs et al. 2014). To investigate the genome-wide correlation between H2Bub1 and gene transcription, we computed the normalized H2Bub1 average signal on each gene body using

deepTools and correlated that to the normalized gene expression level. Consistent with early findings, H2Bub1 is highly correlated with transcription level for most genes (Fig. 7A). Given that the stimulation of transcription begins with the establishment of the transcription preinitiation complex on promoters, and the interplay of various histone modifications on promoters is able to modulate the recruitment of transcription factors, it was interesting to investigate the correlation between H2Bub1 and other active or repressive histone signatures, Chromatin accessibility assessed by DNase-Seq, RNA polymerase II (RNAPII), and nascent RNA transcription detected by GRO-seq on the promoter regions. Indeed, H2Bub1 is highly correlated to transcription initiation associated histone marks H3K4me3 and H3K27ac (Guenther et al. 2007; Karlic et al. 2010), transcription elongation associated H3K36me3 (Guenther et al. 2007), RNAPII, chromatin opened degree (DNase-Seq), and transcription level, while negatively correlated to transcription initiation and elongation repressive marks like H3K27me3 (Schones et al. 2008) and its methyltransferase EZH2 (Cao et al. 2002) (Fig. 7B). Together, these data confirmed H2Bub1 is highly correlated with gene transcription in MEFs.
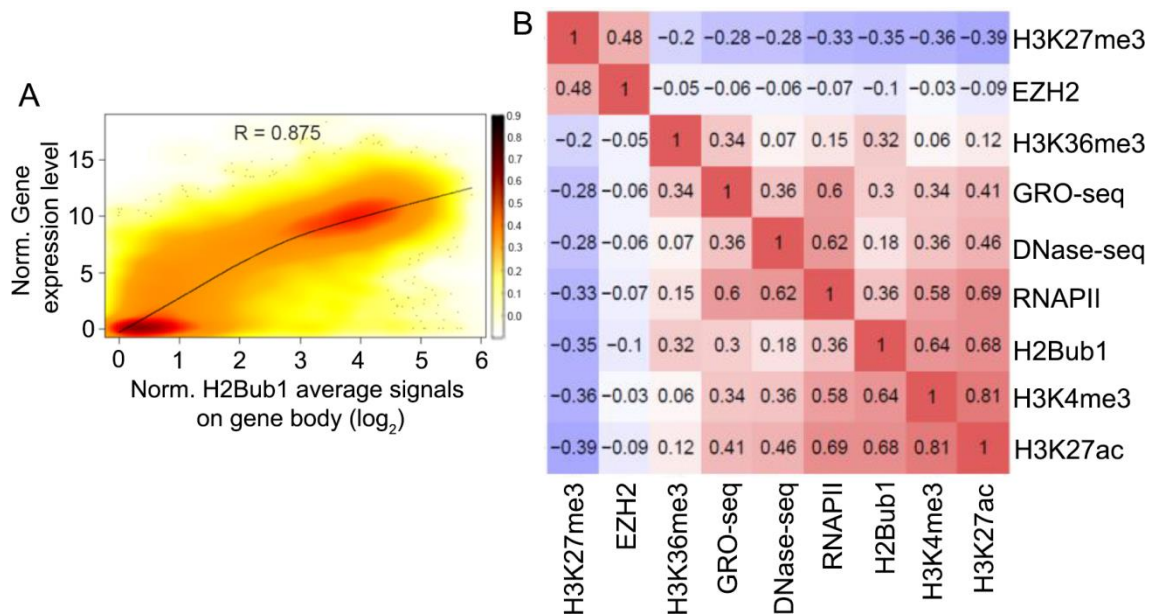
Fig.7 Correlation between H2Bub1, gene expression, active histone modifications, and repressive histone modification

(**A**) SmoothScatter plot analysis displays the relationship between the average H2Bub1 signals on gene body and its own gene expression level. The average H2Bub1 signals on gene body were calculated by ComputeMatrix in deepTools.

(**B**) The heatmap shows the correlation between H3K27me3, EZH2, H3K36me3, GRO-seq, DNase-Seq, mRNA polymerase II (RNAPII), H2Bub1, H3K4me3, and H3K27ac on the region of transcriptional start sits (TSSs) to downstream 1kb. The correlation coefficients were calculated using Pearson's method.

4.1.3 Inducible knockout of RNF40 globally affects active histone modifications

To further study the correlation between H2Bub1 and other active histone marks, we developed a conditional *Rnf40* knockout mouse in which exons 3 and 4 of the mouse Rnf40 gene were flanked by LoxP sites (Fig. 4). This mouse was subsequently crossed to a transgenic line expressing a ubiquitously expressed tamoxifen-inducible Cre recombinase (Rosa26-CreERT2) and mouse embryo fibroblasts were then

obtained from homozygous *Rnf40*[loxP/loxP] embryos containing the CreER[T2] transgene (Hameyer et al. 2007). The deleted RNF40 and H2Bub1 MEFs were effectively generated by treating cells with 4-hydroxytamoxifen (4-OHT) for 5 days and an additional 3 days culture without 4-OHT (Fig. 8A). In agreement with the crosstalk between H2Bub1 and methylation on histone 3 (H3K4me3 and H3K79me3) (Shilatifard 2006; Kim et al. 2009), loss of H2Bub1 resulted in global decrease of H3K4me3 and H3K79me3. In addition, we observed a global decrease of various histone acetylations (H3K9ac, H3K27ac, and H4K12ac) (Fig. 8B). However, there was no significant effect on transcription elongation-associated H3K36me3 (Fig. 8B). Together, we can conclude that H2Bub1 has a critical role in maintaining not only H3K4me3 and H3K79me3 occupancy via histone crosstalk mechanisms but also some other active histone modifications such as histone acetylation.
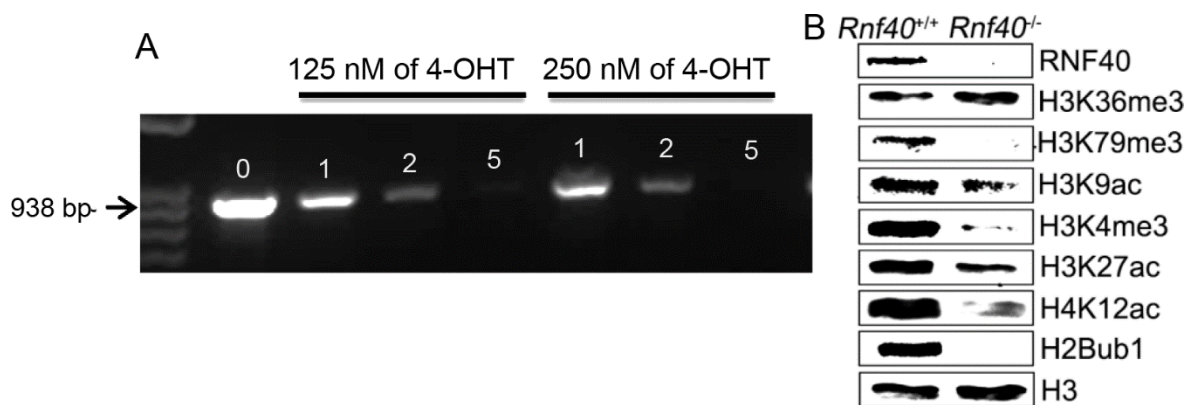

Fig.8 The effects of H2Bub1 deletion on other active histone modifications
(**A**) PCR detection of the sequence containing LoxP site. The *Rnf40* inducible knockout MEFs were treated with 125 nM or 250 nM of 4-OHT for 0, 1, 2, and 5 days. 300 ng of DNA were performed PCR to detect the sequence containing LoxP site (938 bp)

(**B**) Western blot analysis for protein extracts from H2Bub1 presented or deleted MEFs using the antibodies targeting RNF40, H3K36me3, H3K79me3, H3K9ac, H3K4me3, H3K27ac, H4K12ac, H2Bub1, and H3 (the loading control). *Rnf40* knockout was induced by treating with 250 nM of 4-hydroxytamoxifen (4-OHT) for 5 days, and then cultured for 3 days in the absence of 4-OHT.

4.1.4 Loss of RNF40 selectively affects low or moderate H2Bub1 targeted gene transcription

Next, we further categorized genes globally into 4 clusters based on their degree of H2Bub1 occupancy from high to non-enrichment (Fig. 9A), and observed that genes displaying undetectable or abundant levels of H2Bub1 were largely unaffected in their expression level. In contrast, genes displaying low (L) or moderate (M) H2Bub1 occupancy were highly regulated in *Rnf40*-deficient MEFs (Fig. 9B).

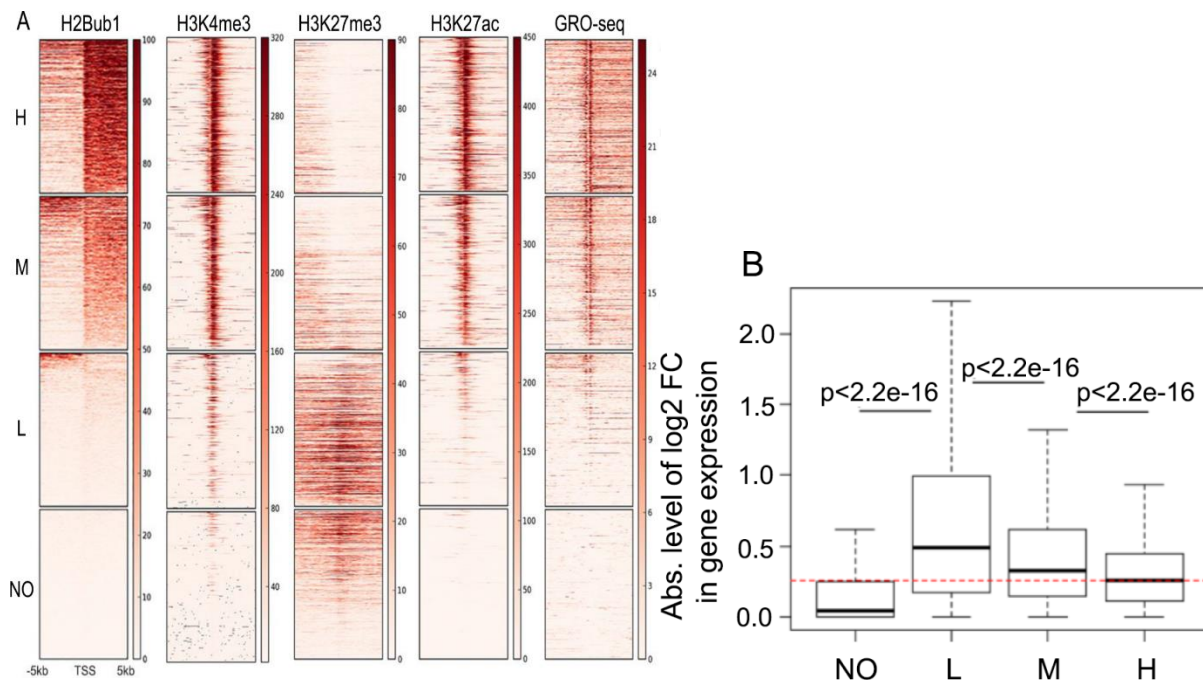

Fig.9 Low or moderate H2Bub1 targeted genes is sensitive to RNF40 deletion

(**A**) Heatmaps show H2Bub1, H3K4me3, H3K27me3, H3K27ac and GRO-seq signal surrounding TSS (±5kb) in wild-type (*Rnf40*$^{+/+}$) MEFs. Genes are sorted according to

H2Bub1 level in descending order. Color key of the heatmaps is shown on their right. According to H2Bub1 occupancy, genes are grouped into high ("H") displaying up to 75% of ranked genes, moderate ("M") displaying 50% ~ 75% of ranked genes, low ("L") displaying 25% ~ 50% of ranked genes, no ("No") displaying ~25% of ranked genes.

(**B**) Boxplot compares the absolute values of log2-fold changes in gene expression (*Rnf40*$^{-/-}$ vs. *Rnf40*$^{+/+}$) in the defined groups. P-value was calculated by unpaired Wilconxon-Mann-Whitney-Test.

In addition, we studied the effects of loss of H2Bub1 on the occupancy of H3K4me3, H3K27ac and H3K27me3 near the TSS. SmoothScatter analysis showed that active marks (H3K4me3 and H3K27ac) were most strongly decreased on genes displaying high levels of H2Bub1 and slightly increased on non-/low-H2Bub1 marked genes (Fig. 10A and 10B; left panel). Consistent with the dynamic pattern of gene regulation (Fig. 9B), the active and repressive histone marks surrounding TSS in "L" and "M" gene clusters were significantly altered in *Rnf40*-deficient MEFs (Fig. 10A-C; right panel). Notably, majority of genes in the highly regulated clusters (L and M) were at bivalent state with decorated by both active and repressive marks (Fig. 9A). Thus, we hypothesized that the dependency of differential regulation on H2Bub1 is highly associated to histone modification context.
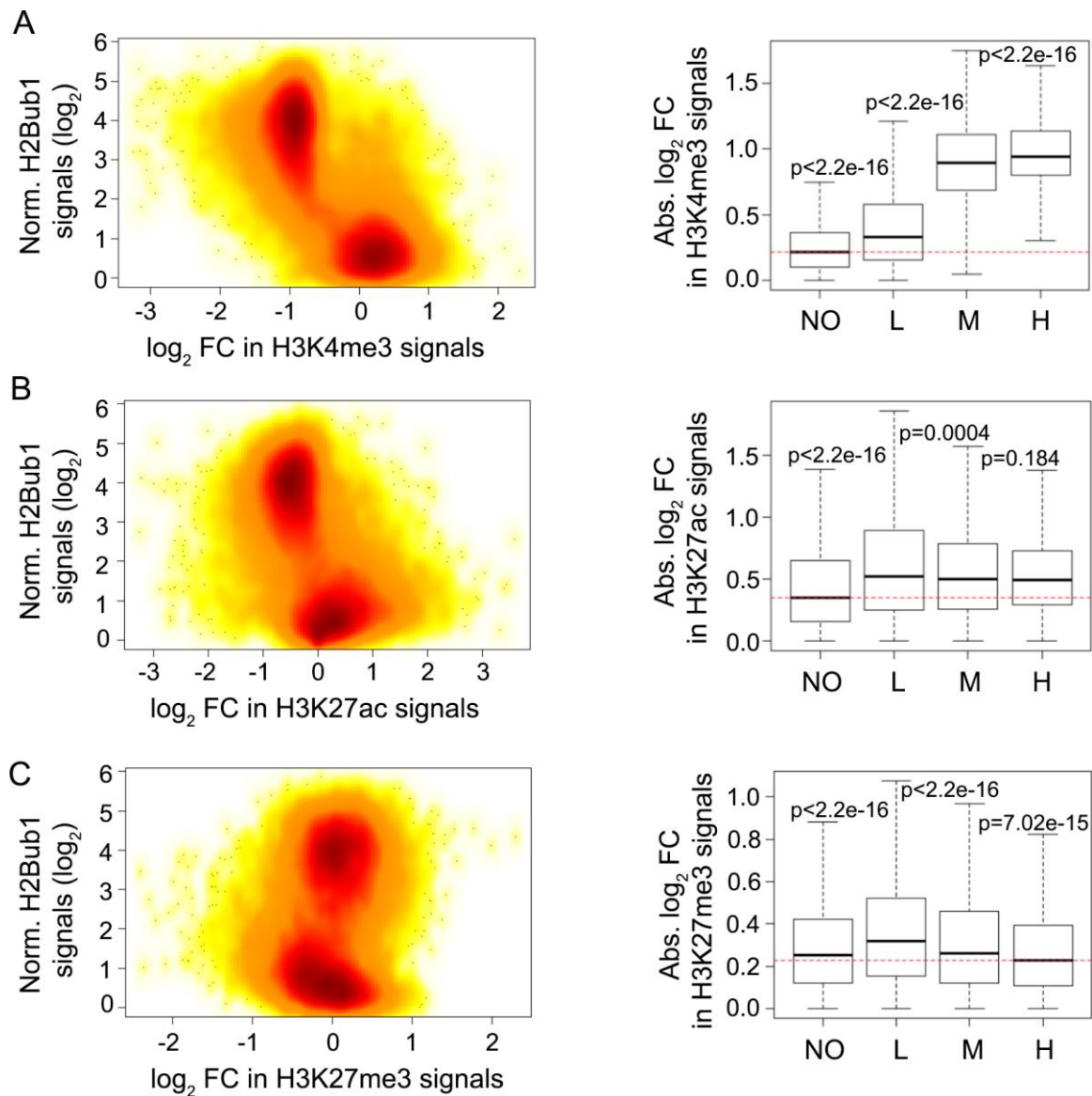
Fig.10 Genome-wide analysis of the alteration of H3K4me3, H3K27me3, and H3K27ac surrounding TSS (±1kb) in the absence of H2Bub1

(**A-C; left panel**) The smoothScatter plots compare average H2Bub1 signals on the gene body to the changing levels of H3K4me3 (A), H3K27ac (B), and H3K27me3 (C) occupancy near the TSS. The correlation coefficient above each smoothScatter plot was calculated using the Pearson method.

(**A-C; right panel**) Boxplots show the absolute value of log2-fold changes in H3K4me3 (A), H3K27ac (B), and H3K27me3 (B) occupancy for genes grouped

according to H2Bub1 occupancy as "No", "L", "M", and "H". P-values were calculated by unpaired Wilcoxon-Mann-Whitney-Test.

4.1.5 Transcription dependency on H2Bub1 is not associated with gene length in RNF40 deleted cells

It is suggested that transcription level of a gene is correlated with its own morphological parameters (such as distance to neighbors, numbers of exons and introns, and gene length). Moreover, early data showed that transcription regulation in response to RNF20 knockdown majorly occurs in longer genes (Fuchs et al. 2012). In contrast to the observation in RNF20 knockdown cells, there is no significant difference in gene length between RNF40 regulated and independent genes (Fig. 11A). Furthermore, we investigated the alteration of gene expression both in the shortest and longest genes. There is no difference in H2Bub1 occupancy between the shortest and longest genes (Fig. 11B). In general, we didn't observe too many significant changes in gene expression in those two gene clusters (Fig. 11C). We conclude that transcription dependency on H2Bub1 is not associated with gene length.
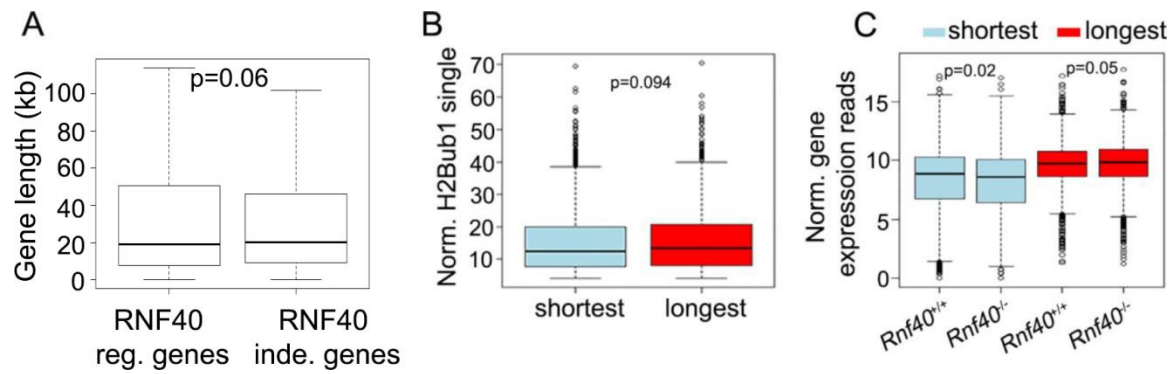
Fig.11 Transcription regulation induced by RNF40 deletion is not associated with gene length

(**A**) Boxplot analysis compares gene length between RNF40-regulated (reg.) and -independent (inde.) gene clusters. RNF40 independent genes: adjustment q-value greater than 0.8, log2-fold change between -0.2 and 0.2. RNF40 regulated genes: adjustment q-value less than 0.05, log2-fold change less than -1 or greater than 1. p-value was calculated by unpaired Wilcoxon-Mann-Whitney-Test.

(**B**) Boxplot analysis compares normalized H2Bub1 average signals on gene body between the shortest genes and longest genes. p-value was calculated by unpaired Wilcoxon-Mann-Whitney-Test.

(**C**) Boxplot analysis compares the alteration in gene expression between *Rnf40*[+/+] and *Rnf40*[-/-] conditions in the shortest and longest gene clusters. P-value was calculated by paired Wilcoxon-Mann-Whitney-Test. The longest gene shows up to 75% ranked gene length; the shortest gene cluster involves genes showing less to 25% ranked gene length.

4.1.6 Three classified states of promoters

Given that H2Bub1-regulated genes are comprehensively occupied by active and repressive histone modifications, such as H3K4me3, H3K27me3, and H3K27ac (Fig. 9A), in order to better investigate the role of H2Bub1 in transcription, it is necessary

to classify global genes according to the epigenetic context. Based on the active H3K4me3 and repressive H3K27me3 occupancy near TSS regions, we grouped global genes into only H3K4me3 enriched (H3K4me3$^+$), both H3K4me3 and H3K27me3 occupied (H3K4me3$^+$ & H3K27me3$^+$), and only H3K27me3 enriched (H3K27me3$^+$) clusters (Fig. 12C)., we identified that out of 4727 genes were marked by H3K27me3 containing 2686 of H3K27me3$^+$ genes (Fig. 12A and 12C); and out of 13605 of H3K4me3 marked genes containing 11564 of H3K4me3$^+$ genes (Fig. 12B and 12C). Through overlapping the H3K4me3 and H3K27me3 marked genes, we further identified 2041 genes marked both by H3K4me3 and H3K27me3 (Fig. 12C).
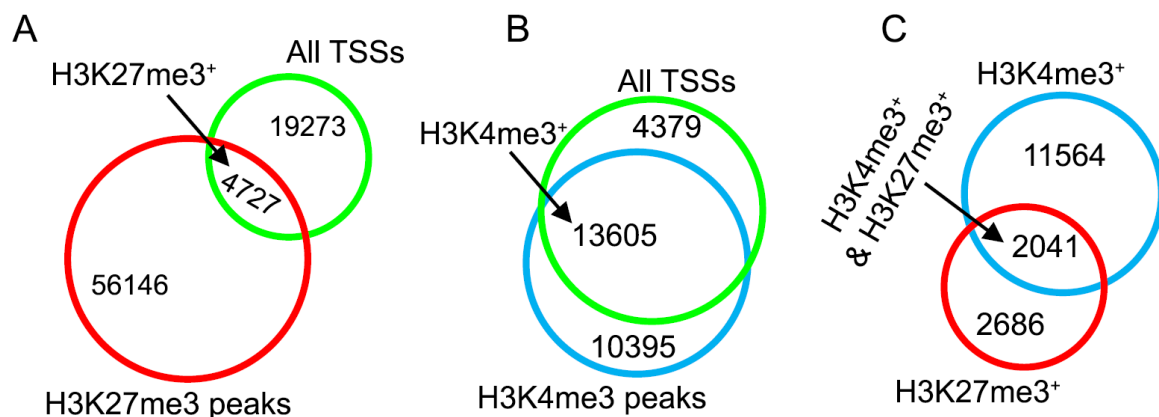
Fig.12 Three chromatin states in MEFs

(**A-C**) Venn diagram analysis identified genes with H3K27me3 (A), H3K4me3 (B), and both H3K4me3 & H3K27me3 enriched (C) near TSS regions. The significant H3K4me3 or H3K27me3 peaks were called using MACS (version 1.0.1) with p-value less than 0.00001.

In addition, we utilized heatmaps which displayed histone modification occupancy surrounding the TSS of the classified gene clusters. Combined with the published data in mouse fibroblasts (Liu et al. 2014; Wei et al. 2015) H3K4me3$^+$ genes show a high degree of H2Bub1 occupancy and transcription detected by nascent RNA

sequencing (GRO-Seq); H3K4me3$^+$ & H3K27me3$^+$ genes display low or no signals of H2Bub1 and transcription; H3K27me3$^+$ genes show no H2Bub1 occupancy and transcriptional silencing (Fig. 13).

Fig.13 Heatmaps displaying the occupancy of H2Bub1, H3K4me3, H3K27me3, and nascent RNA levels (GRO-Seq) surrounding TSSs in the three chromatin states

The color key of each heatmap is at the right. The 'H3K4me3$^+$' and 'H3K4me3$^+$ & H3K27me3$^+$' genes were sorted according to H2Bub1 occupancy from high to low, and the 'H3K27me3$^+$' genes were sorted according to H3K27me3 from low to high.

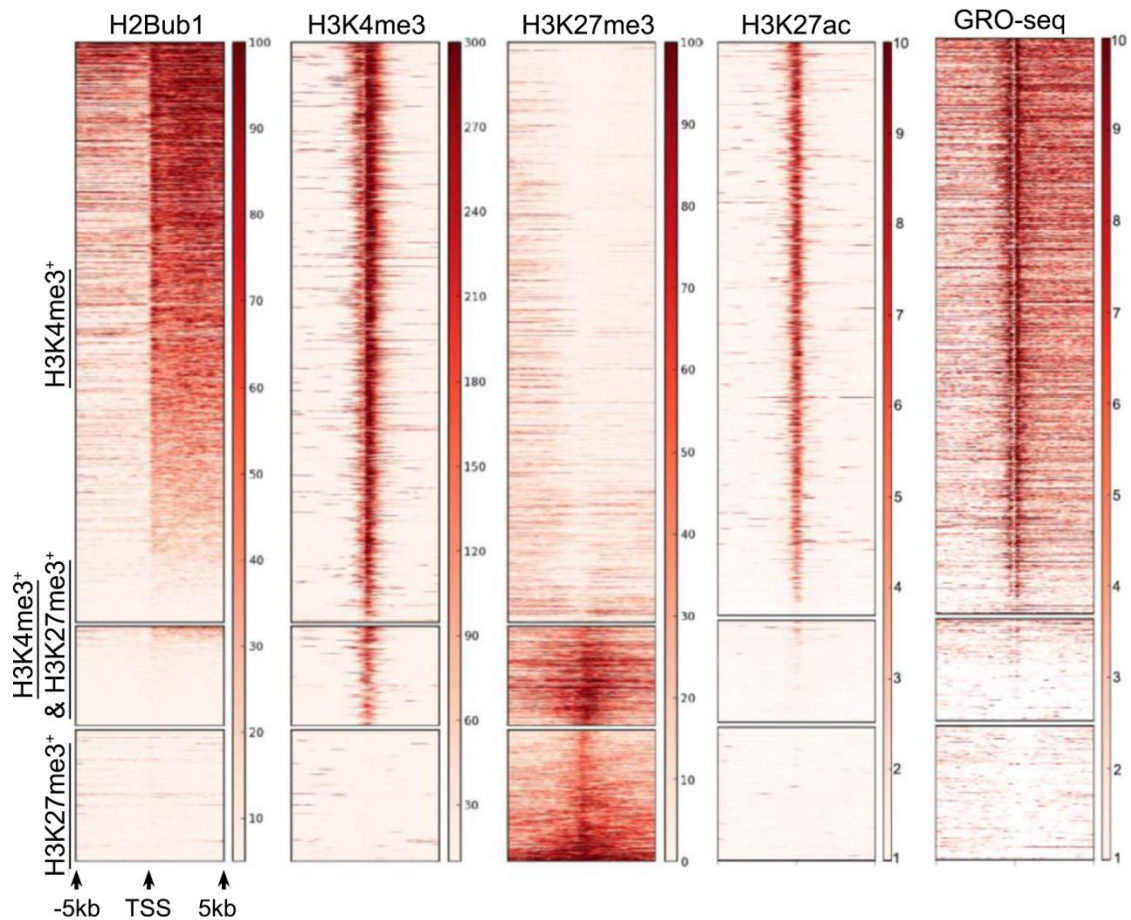Next, we investigated the presence of other histone modifications on TSS regions in response to H2Bub1 deletion. In agreement with crosstalk between H2Bub1 and H3K4me3 (Racine et al. 2012), H3K4me3 was significantly decreased on 'H3K4me3$^+$' gene promoters (Fig. 14A); while slightly but not significantly decreased on 'H3K4me3$^+$ & H3K27me3$^+$' gene promoters (Fig. 14B), possibly because of the low occupancy of H2Bub1 at this chromatin locus (Fig. 13). Surprisingly, loss of H2Bub1 led to a significant decreasing of H3K27me3 on 'H3K4me3$^+$ & H3K27me3$^+$' and 'H3K27me3$^+$' gene promoters (Fig. 15A and 15B). In addition, the absence of H2Bub1 led to a differential alteration of H3K27ac on 'H3K4me3$^+$' and 'H3K4me3$^+$ & H3K27me3$^+$' gene promoters. However, H3K27ac occupancy was differentially altered in those two gene clusters in response to *Rnf40* deletion, which displayed significantly decreased 'H3K4me3$^+$' genes while increased 'H3K4me3$^+$ & H3K27me3$^+$' genes (Fig. 16A and 16B). There was no remarkable signal of H3K27ac on 'H3K27me3$^+$' gene promoters (Fig. 16C). Given that H3K4me3 and H3K27ac serve as transcription coactivator on promoters, while H3K27me3 has the ability of repressing transcription initiation and elongation (Wang et al. 2009; Pasini et al. 2010), we hypothesized that gene regulation at these three clusters might respond differentially to H2Bub1 depletion.
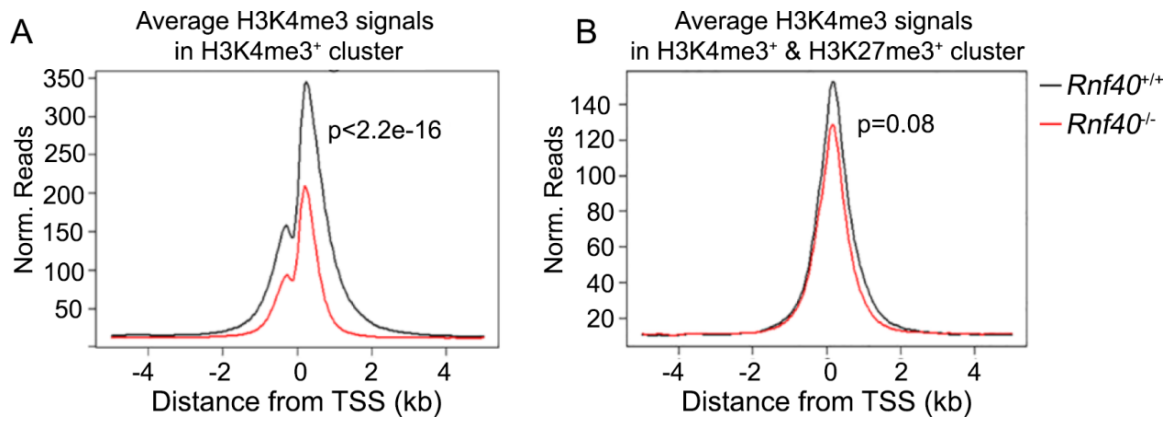
Fig.14 The effects of H2Bub1 deletion on H3K4me3 occupancy at the given gene clusters

Aggregate profile analysis shows the average H3K4me3 signals near TSS regions in 'H3K4me3+' and 'H3K4me3+ & H3K27me3+' gene clusters following RNF40 deletion. P-values were calculated by Wilcoxon-Mann-Whitney-Test.
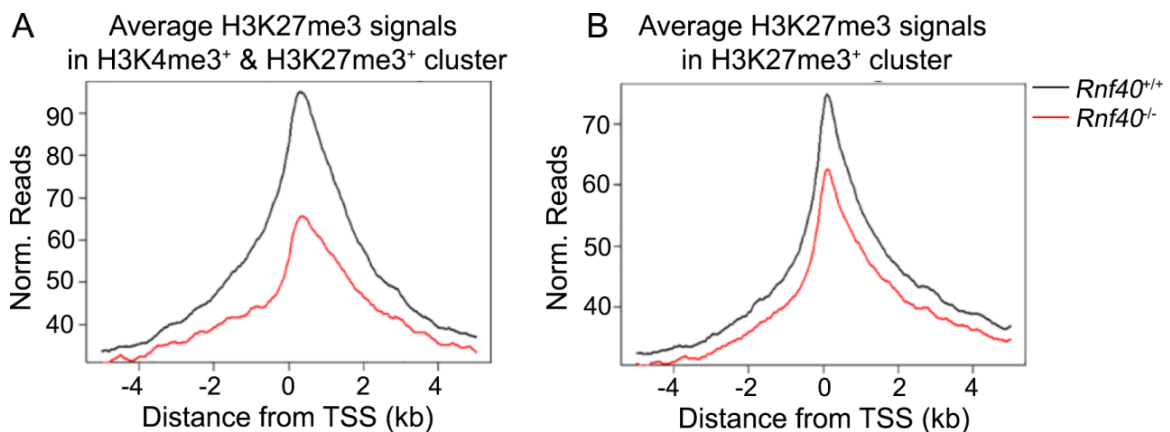
Fig.15 The alteration of H3K27me3 occupancy near TSS regions of given gene clusters following H2Bub1 deletion

Aggregate profile analysis shows the average H3K4me3 signals near TSS regions in 'H3K4me3+' and 'H3K4me3+ & H3K27me3+' gene clusters following RNF40 deletion. P-values were calculated by Wilcoxon-Mann-Whitney-Test.
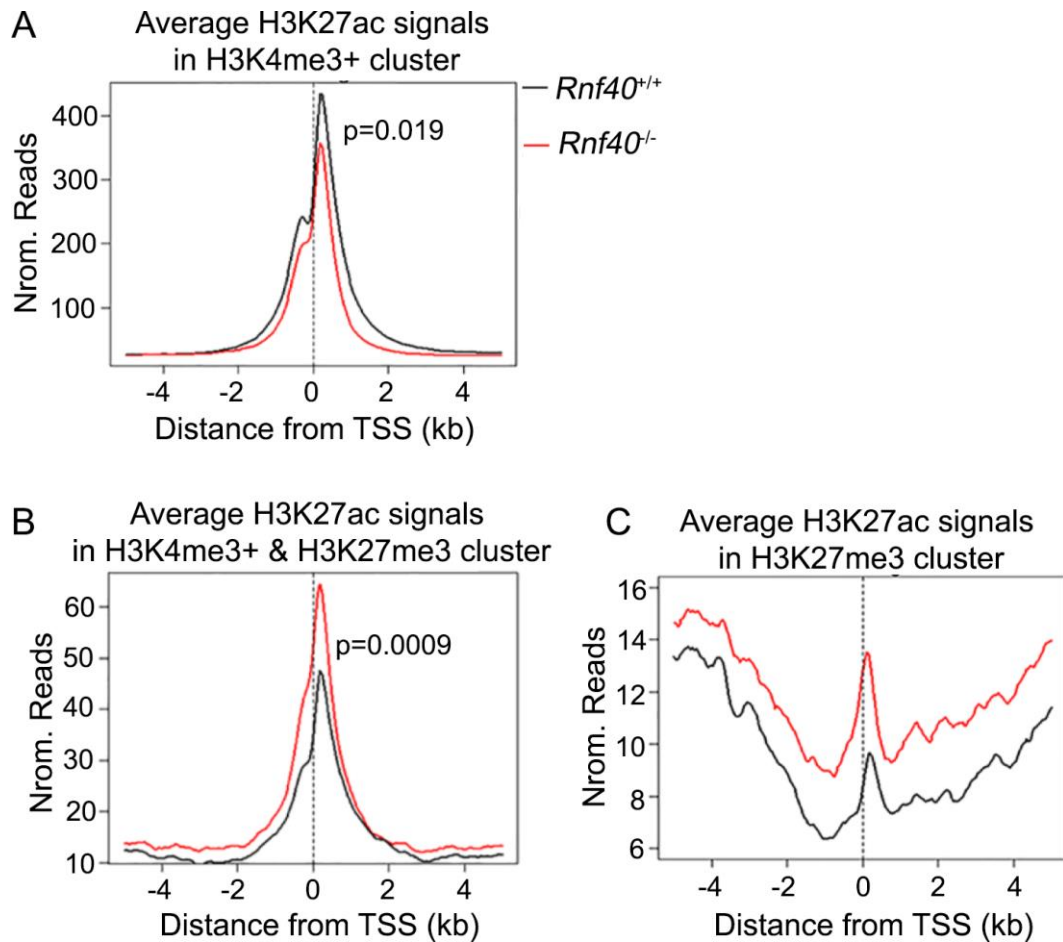
Fig.16 The effects of H2Bub1 deletion on H3K27ac occupancy near TSS regions of the given genes

Aggregate profile analysis shows the average H3K27ac signals near TSS regions in 'H3K4me3+', 'H3K4me3+ & H3K27me3+', and 'H3K27me3+' gene clusters following RNF40 deletion. P-values were calculated by Wilcoxon-Mann-Whitney-Test.

4.1.7 RNF40-dependent genes were enriched with H3K4me3, while RNF40-suppressed genes were enriched with H3K27me3

Next, we investigated the differential alteration in gene expression in the three groups in response to H2Bub1 loss. Here, we showed that RNF40-dependent genes were significantly only marked by H3K4me3 (672/802), which supported the active function of RNF40 in transcription; while RNF40-suppressed genes were remarkably

marked by H3K27me3 (257/672), which indicated another unknown transcriptional function of RNF40 ([Fig. 17](#)).

To investigate the biological function of the differentially expressed genes in $Rnf40^{-/-}$ MEFs, we utilized gene-annotation enrichment analysis. In agreement with the reports in human cells (Chen et al. 2012; Fuchs et al. 2012; Jung et al. 2012), Gene Ontology (GO) analysis showed that RNF40-dependent genes significantly enriched for cell cycle- and development-related gene sets ([Fig. 18A](#)), which supported the selective role of RNF40 in cell cycle associated tumor suppression and stem cell differentiation (Shema et al. 2008; Chen et al. 2012). Interestingly, RNF40-suppressed genes were also significantly enriched for development-related GO terms ([Fig. 18B](#)), which indicated an unknown role of H2Bub1 in tissue development.
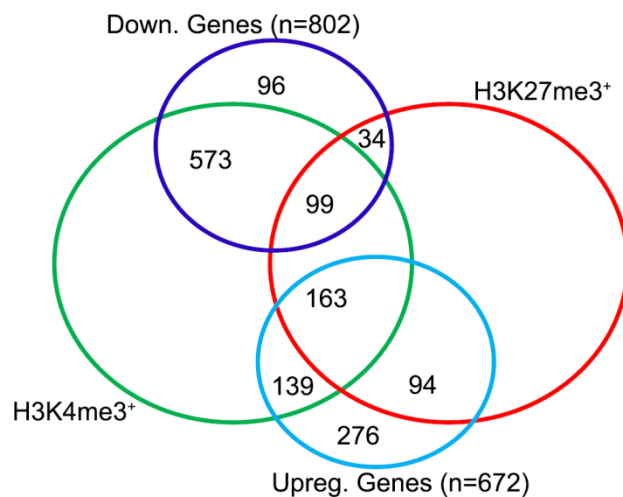
Fig. 17 Venn diagram analysis reveals RNF40-dependent and RNF40-independent gene enrichment

Venn diagram analysis shows the shared number of upregulated ('upreg.'), downregulated ('down.'), H3K4me3 targeted ('H3K4me3⁺'), and H3K27me3 targeted ('H3K27me3⁺') genes. Significantly upregulated genes were defined as those with p-value <0.05, log2-fold change > 1, and average counts ($Rnf40^{+/+}$ and $Rnf40^{-/-}$) > 15;

significantly downregulated genes were defined as those with p-value <0.05, log2-fold change < -1, and average counts ($Rnf40^{+/+}$ and $Rnf40^{-/-}$) > 15.
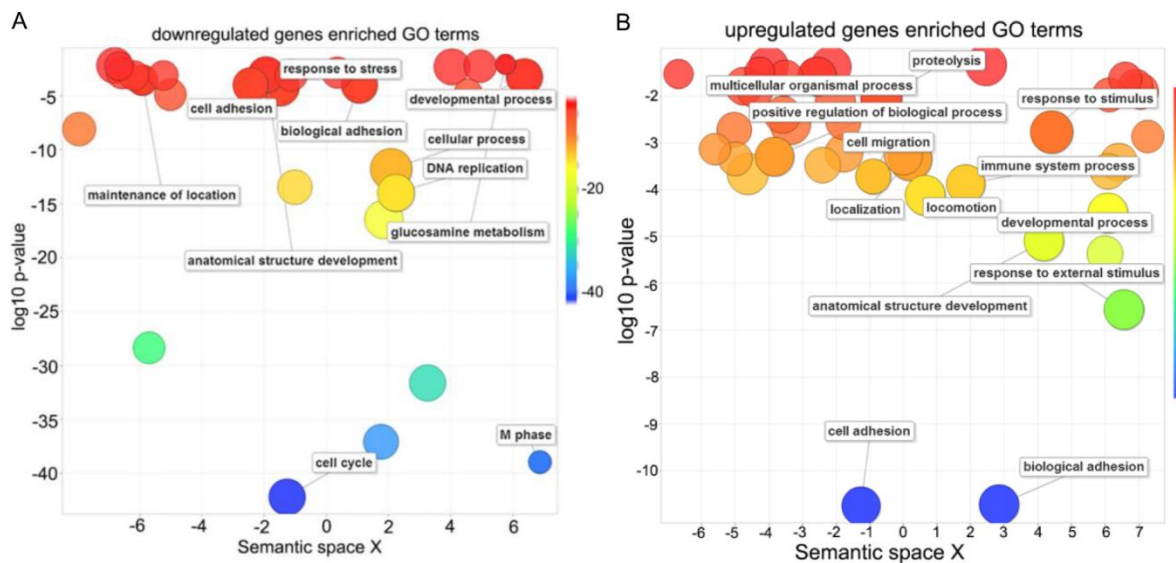
Fig.18 Differential expression induced by H2Bub1 deletion

(**A and B**) The bubble plot shows GO analysis of RNF40-dependent genes (downregulated genes) and RNF40-suppressed genes (upregulated genes). The pathway enrichment was analyzed using DAVID (version 6.7), then the GO terms with FDR value less than 0.05 were plotted by REVIGO. Bubble size indicates the frequency of the GO term in the underlying GO database. Color key of p-value is shown at the right.

4.1.8 The downregulation of RNF40-dependent genes in $Rnf40^{-/-}$ MEFs is associated to the widespread narrowing of H3K4me3 peaks

Consistent with an intimate crosstalk between H2Bub1 and H3K4me3, RNF40 deficiency resulted in a global decrease in H3K4me3 levels (Fig. 8B). In addition, aggregate analysis of H3K4me3 average signal showed that H3K4me3 (red line) occupancy decreased largely at the 3' side of the TSS-associated H3K4me3 peak coinciding with H2Bub1 occupancy (Fig. 19B). Recent studies reported the breadth

of H3K4me3 shows more important on gene transcription compared to its height (Benayoun et al. 2014; Chen et al. 2015b). Consistently, we observed that the breadth of the H3K4me3 peak is highly correlated with gene expression (Fig. 19A). Moreover, the summits of H3K4me3 peaks on RNF40-dependent (downregulated) genes were shifted by ca. 50 bp towards the 5' end of the gene after *Rnf40* knockout (Fig. 20A). Correlated to the narrowing of H3K4me3 peaks (Fig. 20A), the width of H3K27ac peaks were remarkably narrowed in RNF40-dependent genes following RNF40 deletion (Fig. 20B). However, the width of H3K4me3 and H3K27ac peaks near TSS regions were not altered too much in respond to RNF40 deletion (Fig. 20C and D).
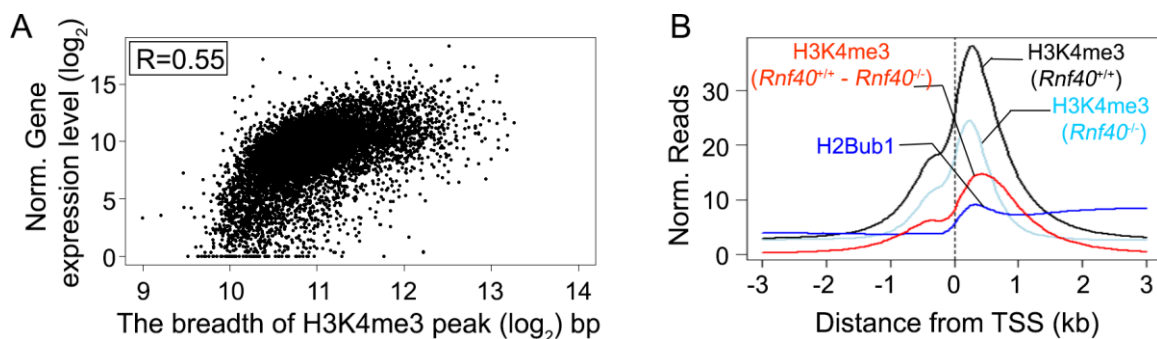
Fig.19 The spread of H3K4me3 is dependent on H2Bub1

(**A**) Plot profile shows the correlation between the width of H3K4me3 domain and their associated gene transcription in *Rnf40* wide type MEFs. The correlation coefficient (R) was calculated by the Pearson method.

(**B**) Aggregate profiles show the average signal of H3K4me3 (black line) and H2Bub1 (blue line), as well as the dynamic changes in H3K4me3 (red line) average occupancy with the loss of H2Bub1 surrounding TSS of genes globally. The dynamic changing of H3K4me3 was calculated as the average occupancy of H3K4me3 in *Rnf40*[+/+] excluding that in *Rnf40*[-/-] MEFs.
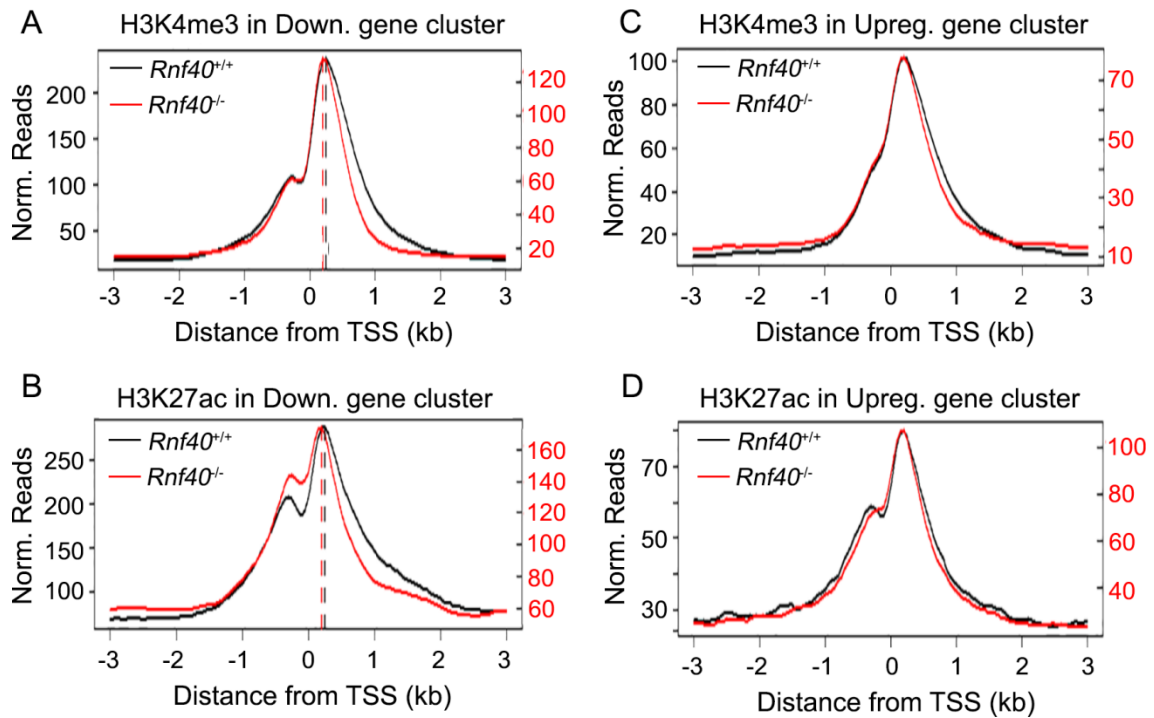
Fig.20 Transcription dependency on H2Bub1 is associated to widespread of H3K4me3 domain

(**A - D**) Aggregate analyses compare H3K4me3 and H3K27ac average occupancy surrounding TSS between *Rnf40*[+/+] (black) and *Rnf40*[-/-] conditions in RNF40-dependent (C) and RNF40-suppressed (E) gene clusters. The dotted lines in Fig. 20C point out the peak center of H3K4me3 in *Rnf40*[+/+] (black) and *Rnf40*[-/-] (red) conditions.

To confirm our bioinformatic analyses, we randomly picked out RNF40-dependent genes, and the single gene expression analysis confirmed the significant downregulation of *Myl9*, *Loxl3*, and *Psrc1* (Fig. 21A). Moreover, ChIP profiles confirmed a significant narrowing of H3K4me3 and H3K27ac peaks on those genes, while the height of H3K4me3 on the promoter of *Psrc1* was not changed following *Rnf40* deletion (Fig. 21B). Together, we propose that the H2Bub1-H3K4me3 trans-

histone pathway is particularly important for broadening of the TSS-associated H3K4me3 domain on a select subgroup of genes.
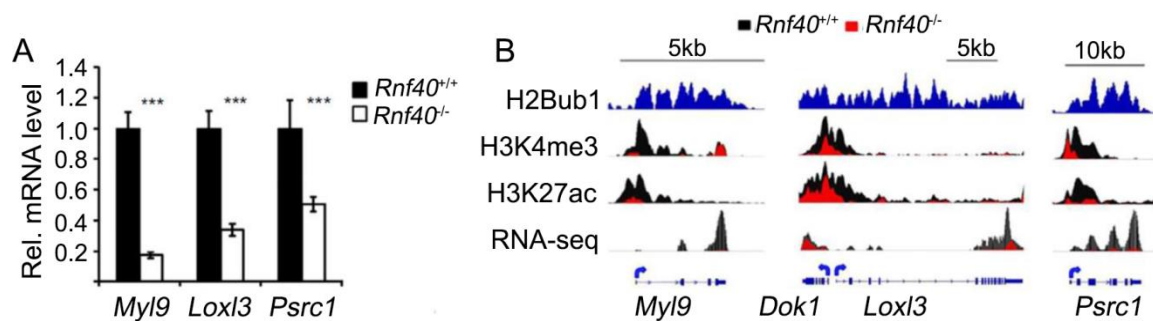
Fig.21 Examples for RNF40-dependent genes are associated to widespread narrowing of H3K4me3 following *Rnf40* deletion

(**A**) qRT-PCR analysis of Rnf40-dependent genes (*Myl9*, *Loxl3*, and *Psrc1*) in *Rnf40*[+/+] and *Rnf40*[-/-] MEF cells. Data are normalized by *Rplp0* gene expression level and shown as "relative mRNA levels", mean ± SD from three independent experiments. *p<0.05, **p<0.01, ***p<0.001, n.s: p>0.05, calculated with two-tailed unpaired t-test.

(**B**) The ChIP profiles show the occupancy of H2Bub1, H3K4me3, and H3K27ac, as well as the normalized mRNA reads on the randomly selected RNF40-dependent genes in *Rnf40*[+/+] (black) and *Rnf40*[-/-] (red) MEFs.

4.1.9 Broadest H3K4me3 is associated with cell development

Given that the broad H3K4me3 domain is tightly linked to the transcription of tumor-suppressed genes and cell lineage-specific genes (Benayoun et al. 2014; Chen et al. 2015b), and the width of H3K4me3 domain is highly dependent on H2Bub1, we next decided to investigate the correlation between H2Bub1 and the broadest H3K4me3 marked genes. Indeed, the width of the broadest H3K4me3 domain (top 3% broadest H3K4me3 peaks) was significantly narrowed in response to *Rnf40* deletion

([Fig. 22A](#)). In addition, we utilized gene ontology analysis of the broadest H3K4me3 occupied genes. Interestingly, consistent with the RNF40-dependent genes enriched GO terms ([Fig. 18A](#)) the broadest H3K4me3 marked genes were significantly enriched for cell cycle and development related genes ([Fig. 22B](#)). For example, the *Wnt5a* ligand, which is suggested to mediate axis induction (Smolich et al. 1993), and *Klf4*, one of Yamanaka factors which takes a function in differentiation and tumor suppression (El-Karim et al. 2013), showed the broadest H3K4me3 occupancy ([Fig. 22C](#)), and its expression was significantly decreased in *Rnf40*$^{-/-}$ MEFs. In agreement with that the broadest H3K4me3 domain is associated to cell identity (Benayoun et al. 2014), *Thy1* gene, a specific marker of MEFs (Tanaka et al. 2002), was significantly downregulated, and the broad H3K4me3 peaks were remarkably narrowed following RNF40 deletion ([Fig. 22D and E](#)). However, the H2Bub1-independent genes such as the housekeeping gene *Actb,* which was also marked by the broadH3K4me3, loss of H2Bub1 didn't significantly shorten the widespread of H3K4me3 instead of the height, thereby didn't affect its transcription much more ([Fig. 22E](#)). Thus, we proposed that the transcription of tissue-specific and cell cycle-related genes requires widening of H3K4me3 domains which are facilitated by H2Bub1.

A

H2Bub1 (*Rnf40*$^{+/+}$)   H3K4me3 (*Rnf40*$^{+/+}$)   H3K4me3 (*Rnf40*$^{-/-}$)

-1kb TSS   10kb

B

cell proliferation

ossification

localization of cell

multicellular organismal process

nitrogen compound metabolism

actin cytoskeleton organization

biosynthesis

macromolecule biosynthesis

biological regulation

cellular process

negative regulation of cellular process

anatomical structure development

developmental process

log10 color key

-10

-30

log10 p-value

semantic space X

C

H3K27ac
H2Bub1
H3K4me3

5kb

1kb

Wnt5a

Klf4

D

Rnf40$^{+/+}$
Rnf40$^{-/-}$

***

Thy1

E

H2Bub1
H3K4me3
H3K27ac
RNA

5kb

5kb

Thy1

Actb

pg. 81

Fig.22 Broadest H3K4me3 marked genes enriched for development in MEFs

(**A**) Heatmaps show H2Bub1 and H3K4me3 occupancy near the broadest H3K4me3 enriched TSS (from TSS upstream 1kb to downstream 10kb) in $Rnf40^{+/+}$ and $Rnf40^{-/-}$ MEFs. Genes were sorted according to the length of H3K4me3 peaks. The top 3% broadest H3K4me3 enriched genes were selected as the broadest H3K4me3 genes. The color key is at the right of each heatmap.

(**B**) The Bubble plot shows the broadest H3k4me3 targeted genes enriched GO terms. The GO enrichment was analyzed using DAVID (version 6.7), then the FDR value less than 0.05 GO terms were plotted by REVIGO. Bubble size indicates the frequency of the GO term in the underlying GO database. Color key of p-value is shown at the right.

(**C**) The profiles show the occupancy of H2Bub1, H3K27ac, and H3K4me3 on the development related genes (Wnt5a and Klf4) in MEFs.

(**D**) qRT-PCR analysis of MEF-specific mark (*Thy1*) in $Rnf40^{+/+}$ and $Rnf40^{-/-}$ MEF cells. Data are normalized by *Rplp0* gene expression level and shown as "relative mRNA levels", mean ± SD from three independent experiments. *p<0.05, **p<0.01, ***p<0.001, n.s: p>0.05, calculated with two-tailed unpaired t-test.

(**E**) The ChIP profiles show the occupancy of H2Bub1, H3K4me3, and H3K27ac, as well as the normalized mRNA reads on the MEF-specific gene (*Thy1*) and housekeeping gene (*Actb*) in $Rnf40^{+/+}$ (black) and $Rnf40^{-/-}$ (red) MEFs.
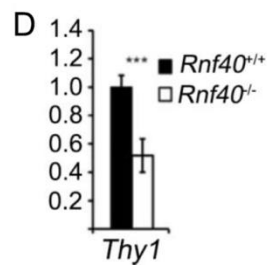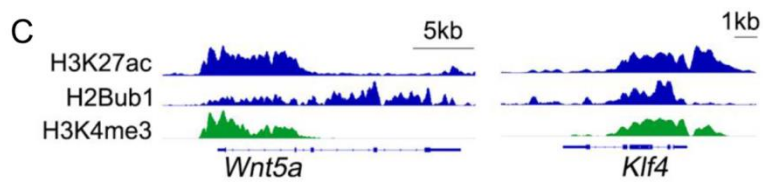
4.1.10 H2Bub1-dependent differentiation genes show wide spreading of H3K4me3 domain during adipocyte differentiation

It was suggested that the transcription of tissue-specific genes requires H2Bub1. Therefore, we further confirmed the relationship between H2Bub1, the broadest H3K4me3, and cell differentiation. After additional analysis of the behavior of

H3K4me3 in hMSC differentiated adipocytes for 5 days from our published data (Karpiuk et al. 2012), we observed that the breadth of H3K4me3 on the RNF40-dependent adipocyte specific genes (which significantly downregulated in response to RNF40 depletion following adipocyte differentiation for 5 days) were remarkably broadened. Moreover, the peak center of H3K4me3 shifted 100 bp downstream (Fig. 23A). For example, the key adipocyte-regulated genes (*PDK4*, *RASD1*, and *PPARG*), whose expression was dependent on RNF40 (Karpiuk et al. 2012), showed a significant widespread of H3K4me3 peak following adipocyte differentiation (ADI.) for 5 days (Fig. 23B). Together, our data provided evidences that H2Bub1-induced widespread of H3K4me3 is essential for cell lineage-specific gene transcription.
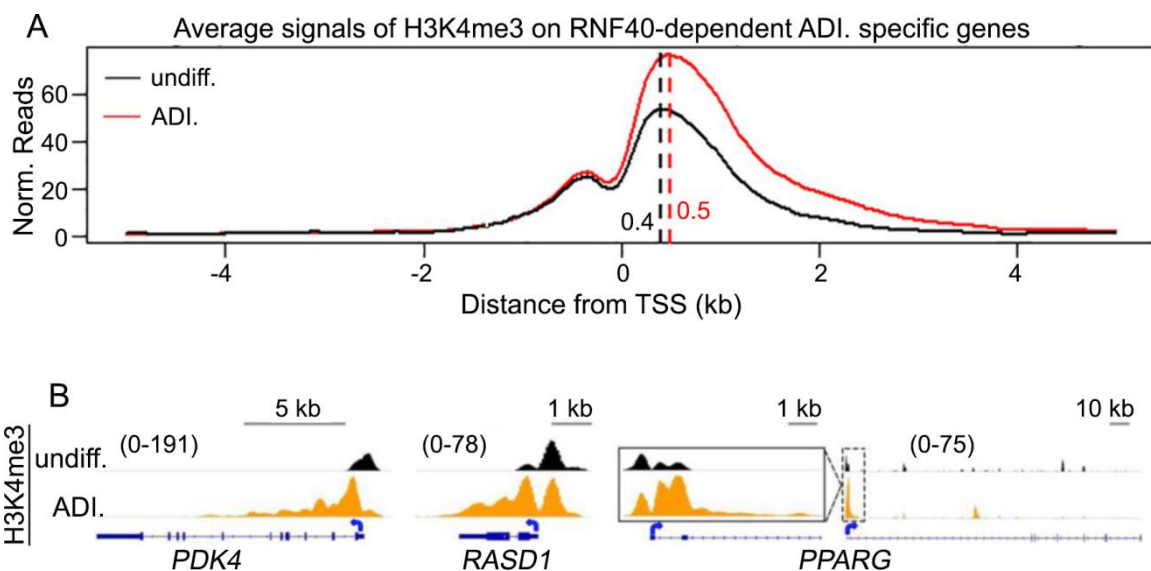
Fig.23 The behavior of H3K4me3 on RNF40-dependent adipocytes specific genes during differentiation

(**A**) Aggregate profiles show the normalized average signal of H3K4me3 surrounding the TSS of RNF40-dependentgenes (± 5kb) in undifferentiated hMSC (undiff., black line) and differentiated adipocytes (ADI., red line). The dotted lines point out the peak center of H3K4me3.

(**B**) ChIP profiles show H3K4me3 occupancy on the key adipocytes differentiation genes (*PDK4*, *RASD1*, and *PPARG*) in un- or differentiated conditions.

4.2   Polycomb complex 2 shows H2Bub1 dependency

Although H2Bub1 is typically correlated to gene transcription, a subset of genes was upregulated in respond to RNF20 depletion (Shema et al. 2008). One early finding suggested that RNF20 suppresses pro-oncogenic gene transcription via blocking the recruitment of the transcription elongation factor TFIIS (Shema et al. 2011). It was further found that the upregulated genes in RNF20-depleted cells show off-targeting of H2Bub1 (Jung et al. 2012). Thus, we speculated that "repressive" functions of H2Bub1 more likely occur via indirect mechanisms.

4.2.1 Moderate H2Bub1 regulates the H3K27 methyltransferase EZH2 expression

To examine whether loss of H2Bub1 impacts specific subsets of genes, we performed gene set enrichment analyses (GSEA) of mRNA-Seq data and identified PRC2- (EZH2, SUZ12, and EED) and PRC1-supressed genes (Fig. 24) as being significantly enriched in *Rnf40*-null cells.
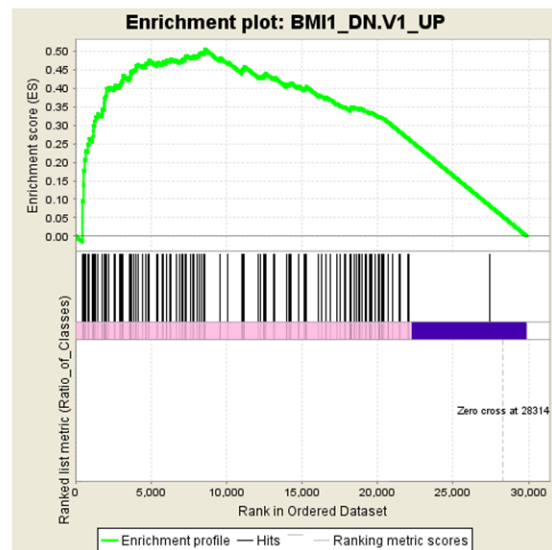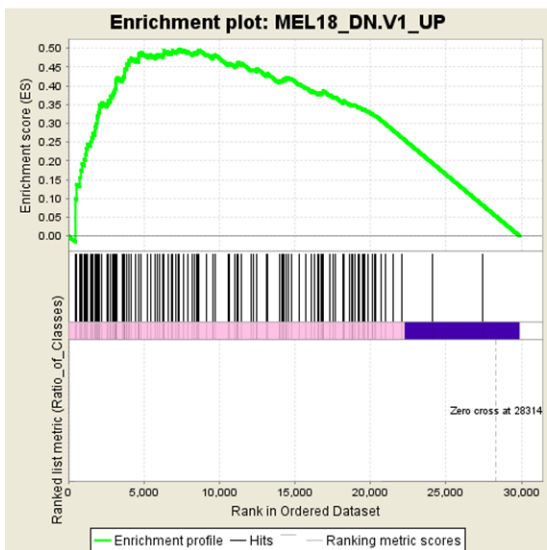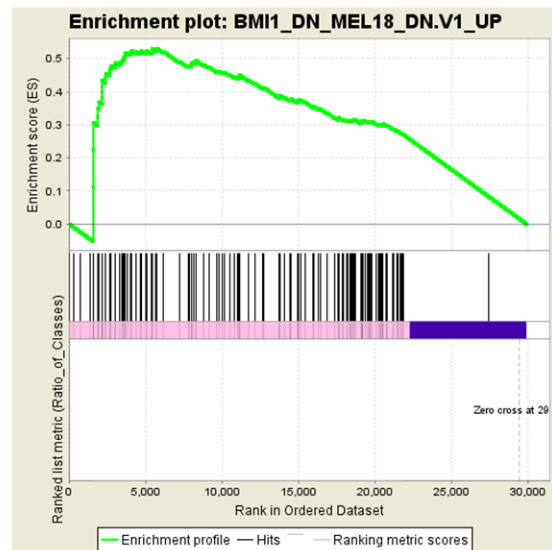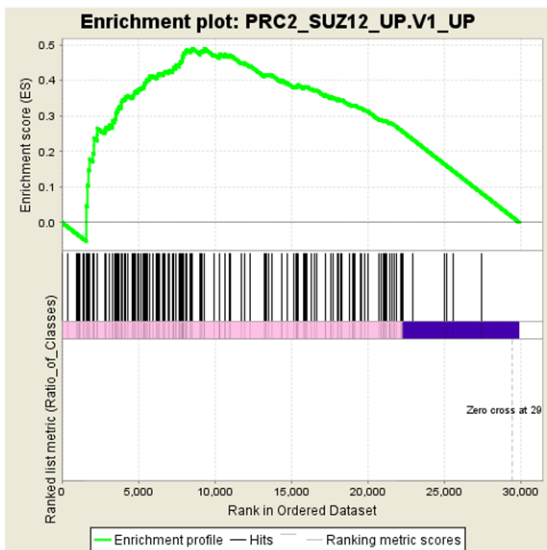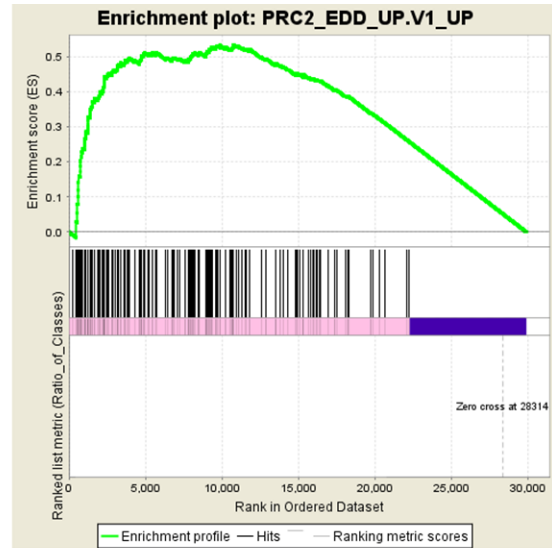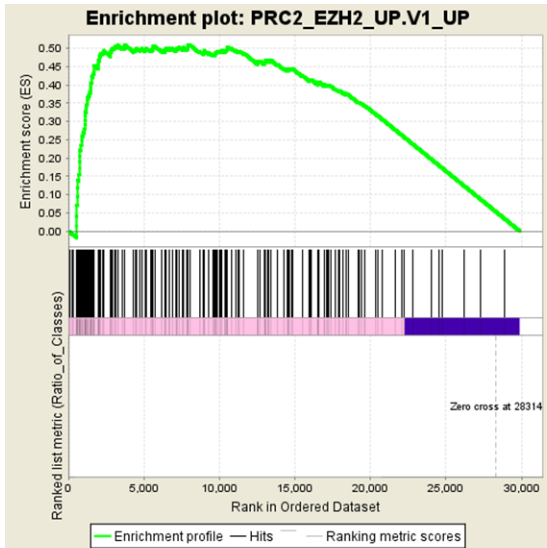
Fig.24 GSEA of RNA-Seq data show enrichment of PcG suppressed genes for upregulation in *Rnf40*[-/-] MEFs

Gene sets of PRC2 (EZH2, SUZ12, and EED) suppressed genes in TIG3 cells (fibroblast). Gene sets of PRC1 (BML1 and MEL18) suppressed genes in DAOY cells (medulloblastoma). GSEA was performed using 'c6.all.v4.0.symbols (Oncogenic signature)' data base. Genes in the graphs were sorted by the gene expression ratio of *Rnf40*[-/-] vs. *Rnf40*[+/+] from left to right.

Strikingly, analysis of RNA- and H2Bub1 ChIP-seq data revealed H2Bub1 occupancy on and a selective decrease in the expression of the Ezh2 gene, while the expression of the remaining members of the PRC2 complex, which catalyzes H3K27 methylation, including Suz12, Eed, and Ezh1 were unaffected (Fig. 25A). These findings could be confirmed by qRT-PCR (Fig. 25B). Consistent with an interdependence in their protein expression levels (Wei et al. 2011; Kim et al. 2015), western blot analysis of other PRC2 subunits revealed decreased protein levels not only of EZH2, but also for SUZ12 and EZH1 in *Rnf40*[-/-] MEFs (Fig. 25C). Consistently, deficiency of H2Bub1 on the body of the *Ezh2* gene resulted in a significant decrease in the transcription initiation associated H3K4me3, H3K27ac, and the normalized mRNA counts on the exon (Fig. 25D). In addition, we analyzed the transcriptional activity of EZH2 in response to *BRE1A* (*RNF20*) and *PAF1* depletion in HCT116 cells (Chen et al. 2015a). Indeed, depletion of BRE1A significantly reduced RNA polymerase II occupancy at the TSS of EZH2. Moreover, loss of H2Bub1 in PAF1-depleted cells resulted in a significant decrease in *Ezh2* transcription (Fig. 25E).
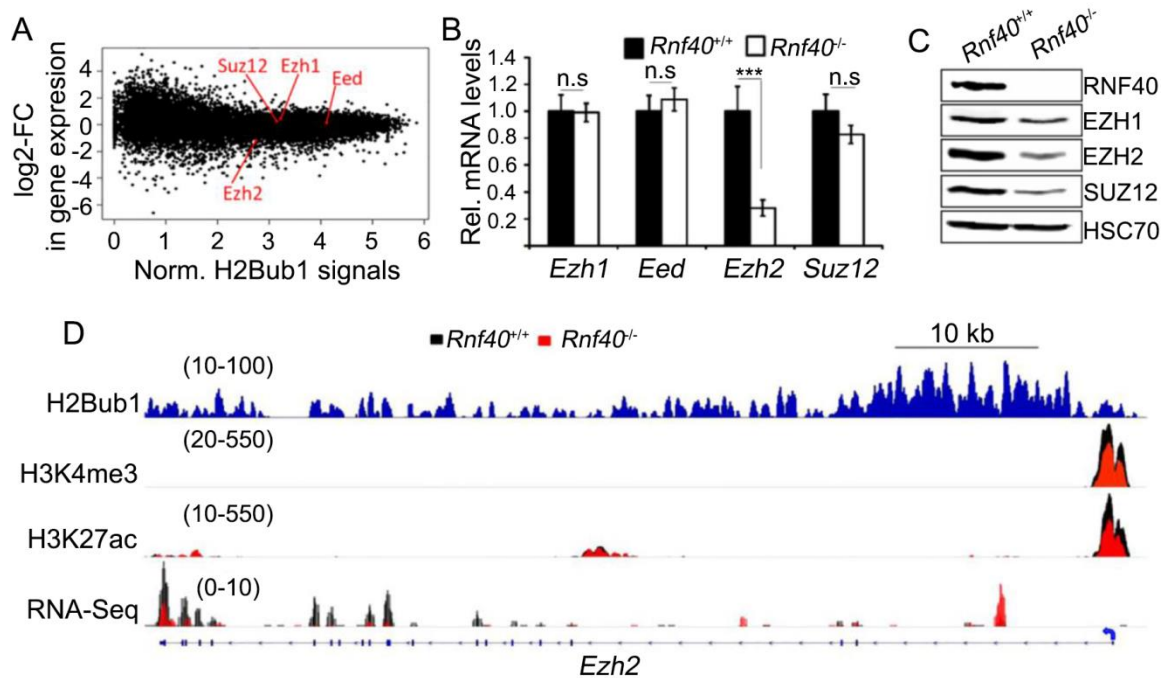
Fig.25 PRC2 complex depends on H2Bub1

(**A**) The scatter plot shows that moderate levels of H2Bub1 on the Ezh2 gene are associated with Rnf40-dependent changes in its expression. The red points denote transcripts encoding the PRC2 subunits EZH1, EZH2, EED, and Suz12.

(**B**) qRT-PCR analysis of *Ezh1*, *Eed*, *Ezh2*, and *Suz12* in *Rnf40*[+/+] and *Rnf40*[-/-] MEF cells. Data are normalized by *Rplp0* gene expression level and shown as "relative mRNA levels", mean ± SD from three independent experiments. *$p<0.05$, **$p<0.01$, ***$p<0.001$, n.s: $p>0.05$, calculated by two-tailed unpaired t-test.

(**C**) Western blots analysis of protein extract from *Rnf40*[+/+] and *Rnf40*[-/-] MEFs using antibodies for RNF40, EZH1, EZH2, SUZ12, and HSC70.

(**D**) The profiles show the occupancy of H3K4me3, H2Bub1, H3K27ac , and the normalized expression reads on *Ezh2* gene in *Rnf40*[+/+] (black) and *Rnf40*[-/-] (red) MEFs.
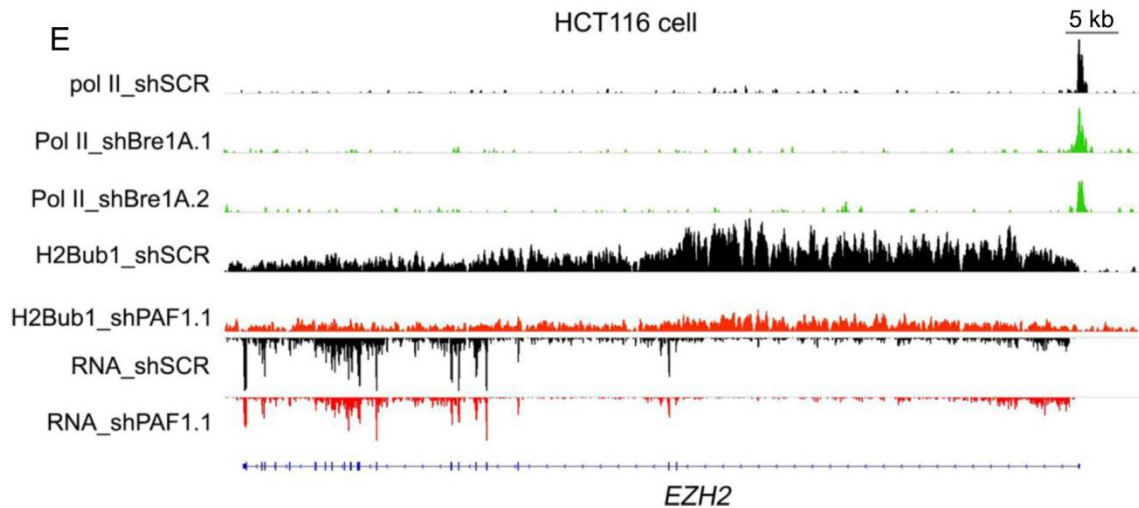
(**E**) The profiles show the occupancy of RNA polymerase II (Pol II), H2Bub1, and nascent RNA (RNA) in control, *BRE1A* knockdown, and *PAF1* knockdown HCT116 cells.

Given that the gene body specific enriched H2Bub1 is involved in the regulation of transcription elongation (Johnsen 2012; Fuchs et al. 2014), we performed ChIP for RNA polymerase II (RNAP II) and H3K4me3. qPCR analyses confirmed that H3K4me3 occupancy near TSS (Fig. 26A) as well as RNA polymerase II (RNAP II) occupancy near the TSS (Fig. 26B) and on the gene body (Fig. 26C) of the *Ezh2* gene were significantly decreased in *Rnf40*-null MEFs, consistent with the observed decrease in *Ezh2* mRNA levels. Thus, we suggested that H2Bub1 directly controls the transcription of *Ezh2* gene both at transcription initiation and elongation phase.
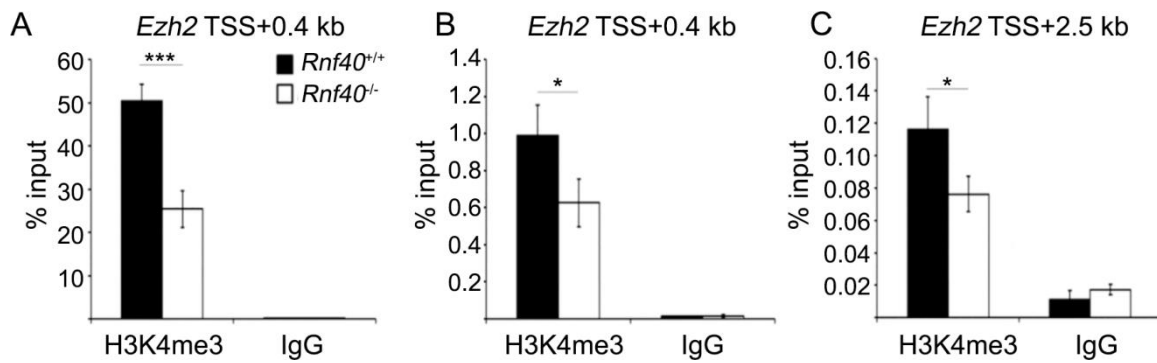
Fig.26 Loss of H2Bub1 affects the recruitment of RNA polymerase II on *Ezh2* gene

(**A – C**) ChIP-qPCR analysis of the occupancy of H3K4me3, RNA polymerase II (RNAP II) near TSS (at downstream 400 bp) and gene body (at downstream 2.5 kb) of *Ezh2* gene. ChIP signal was normalized by input signal, and represented as '%input' ± SD (n=3). Non-specific IgG signal indicated the background binding level. *p<0.05, **p<0.01, ***p<0.001, calculated with two-tailed unpaired t-test.

4.2.2 A significant reduction of H3K27me3 occupancy near TSS regions in *Rnf40*[-/-] MEFs

Given that PRC2 complex majorly carried out tri-methylation of H3 on lysine 27 (H3K27me3), we further investigated the effects of decreased Ezh2 expression on the distribution of H3K27me3 in *Rnf40*-null MEFs by performing Cis-regulatory Element Annotation System (CEAS) (Shin et al. 2009) analysis of H3K27me3 signals in *Rnf40*[+/+] and *Rnf40*[-/-] MEFs. Consistent to the observation before (Tie et al. 2014; Kim et al. 2015), H3K27me3 broadly occupies various genome elements including promoters, gene body, and distal intergenic regions; and shows high enrichment near TSS region (Fig. 27A and B). In relation to EZH2 occupancy on the mouse genome (Boulard et al. 2015), loss of H2Bub1 led to the redistribution of H3K27me3

occupancy on the genome, which showed a decreasing level on the promoter and gene body while a slight increase on distal intergenic (Fig. 27A).
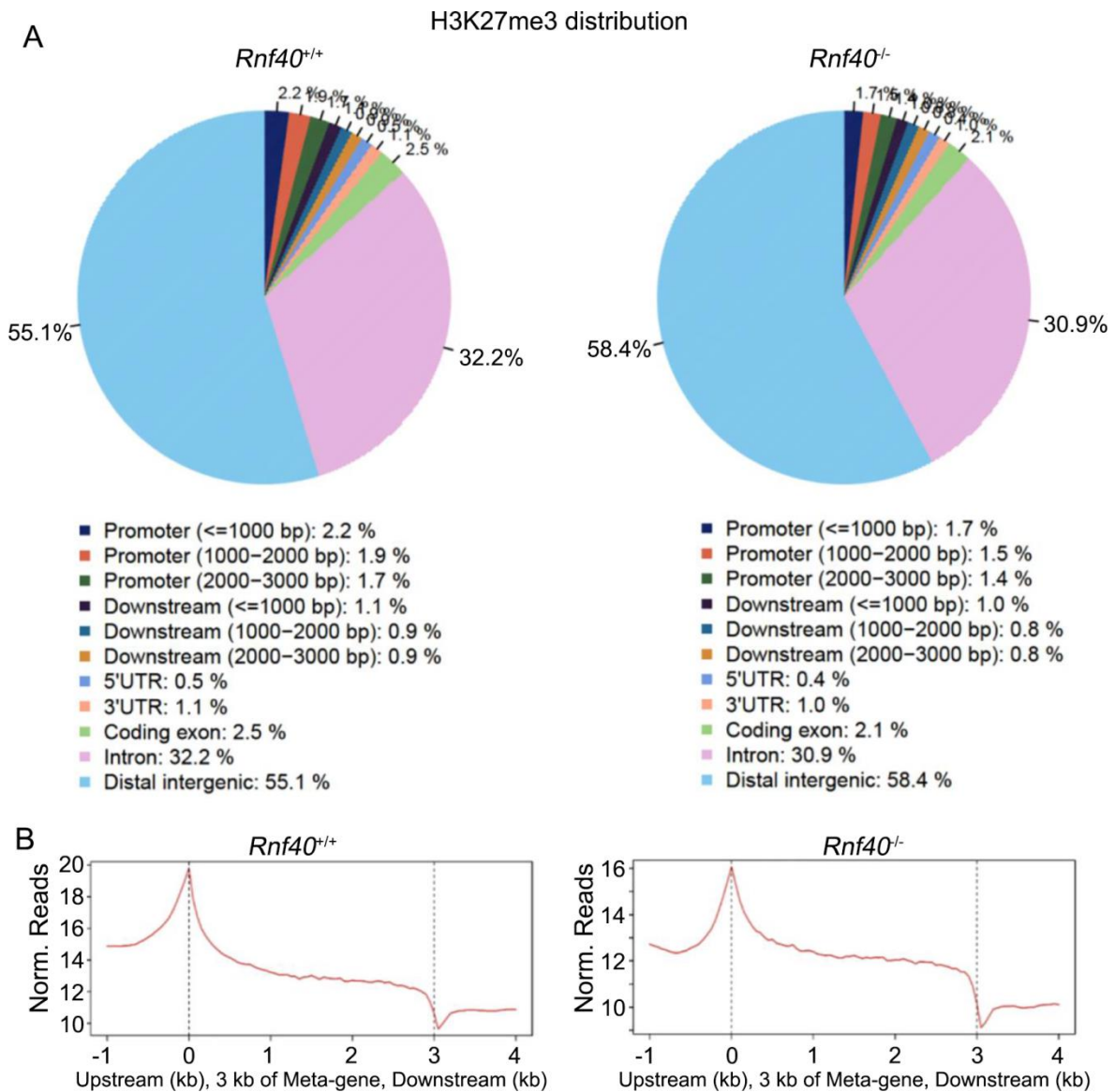
Fig.27 Distribution of H3K27me3 signals on various genome elements in *Rnf40^{+/+}* and *Rnf40^{-/-}* MEFs

(**A**) Enrichment on chromosome and annotation (CEAS) analysis of H3K27me3 distribution various elements of the mouse genome in *Rnf40^{+/+}* and *Rnf40^{-/-}* MEFs.

(**B**) Aggregate analysis of the average ChIP signals of H3K27me3 on the meta-gene of 3 kb, which shows that H3K27me3 enriches near TSS region and decreases towards the 3' end, and has decreased occupancy in *Rnf40*$^{-/-}$ condition (right panel).

Given that H3K27me3 majorly occupies near TSS regions and distal intergenic regions (Fig. 27A), we further performed differential binding (DiffBind) analysis of H3K27me3 behavior near TSS regions (±1kb) and on distal intergenic regions in response to RNF40 deletion. We observed 4727 genes occupied by H3K27me3 near TSS (±1 kb), in which 97% (4241/4727) of those genes displayed a significant reduction in H3K27me3 occupancy following *Rnf40* loss (Fig. 28A), while H3K27me3-enriched distal intergenic regions were differentially affected, displayed a significant increasing on 6.4% (2528/39481) regions and a significant decreasing on 8.7% (3419/39481) distal regions (Fig. 28B). Aggregate analysis of H3K27me3 confirmed significant decrease near TSS regions (±5kb) in *Rnf40*-null MEFs (Fig. 28C). Thus, we suggested that the decreased *Ezh2* expression in *Rnf40*-deleted MEFs majorly reduced H3K27me3 occupancy on promoters.
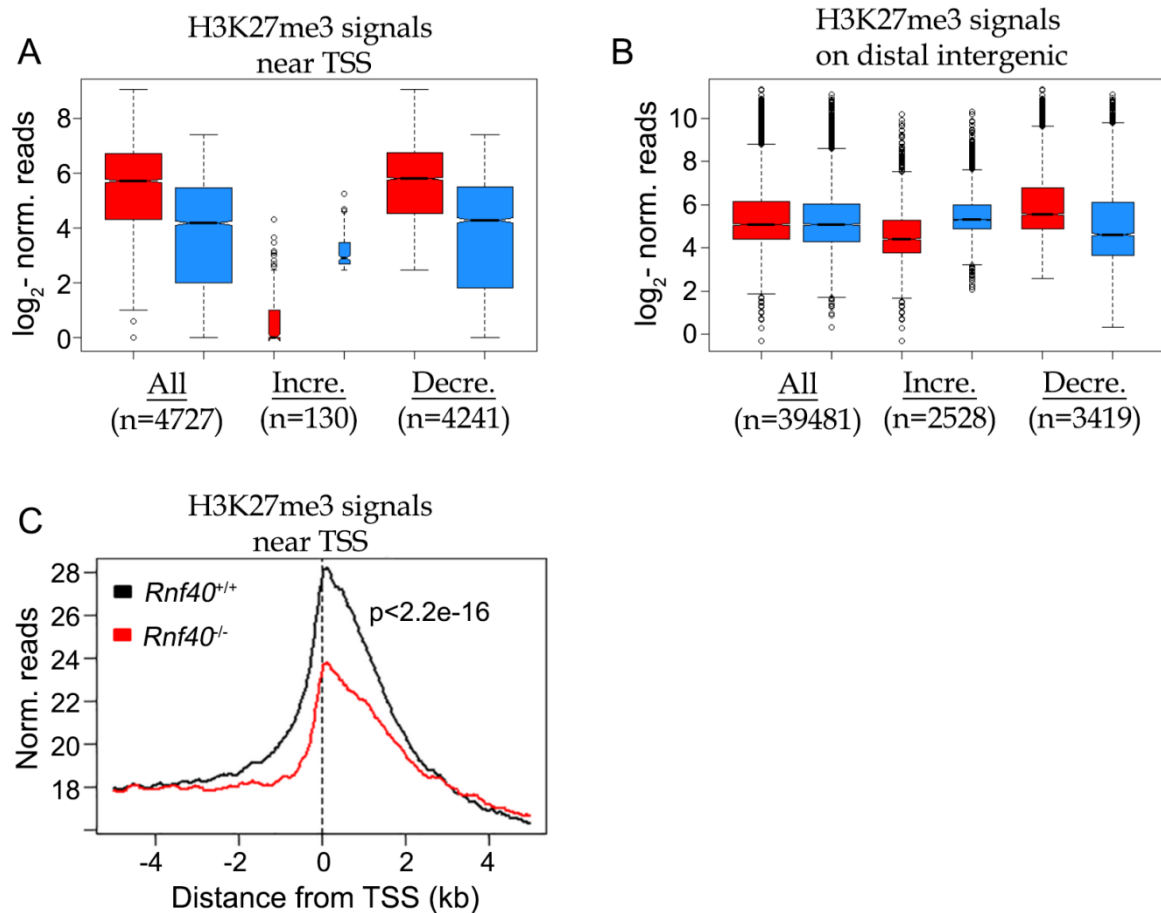
Fig.28 Reduction of *Ezh2* majorly decreases H3K27me3 occupancy near TSS regions

(**A and B**) Differential Binding Analysis (DiffBind) of H3K27me3 occupancy on promoter (A) and distal intergenic region (B) in *Rnf40^+/+* and *Rnf40^-/-* MEFs. Promoters were considered near TSS regions from upstream 1kb to downstream 1kb of the TSS. The region more than 5 kb from gene body was considered as distal intergenic region. The size of the boxplot indicates the number of genes. The regions near TSS or on distal intergenic, displaying a significant increase (incre.) or decrease (decre.) in H3K27me3 signals, were selected based on FDR less than 0.05.

(**C**) Aggregate plot analysis of average H3K27me3 profiles surrounding TSS (±5kb) in wildtype and *Rnf40^-/-* MEFs. p-value was calculated using unpaired Wilcoxon-Mann-Whitney-Test.

4.2.3 EZH2 and H3K27me3 targeted genes show a significant increasing in gene expression in response to *Rnf40* deletion

We next sought to characterize the relationship between EZH2 or H3K27me3 occupancy near the TSS and the induction of gene expression following *Rnf40* deletion. After additional analysis of the published EZH2 ChIP-Seq data in MEFs (Pinter et al. 2012), we identified 861 EZH2 target genes and could demonstrate that H3K27me3 occupancy near the TSS (±5 kb) of these genes was significantly decreased in *Rnf40*-null MEFs (Fig. 29A). Furthermore, GSEA analysis of mRNA-seq data confirmed a significant enrichment for EZH2-targeted genes that were up-regulated in *Rnf40*-null MEFs (Fig. 29B). Additionally, we identified a gene set enriched for H3K27me3 which displayed a greater than 2-fold decrease in H3K27me3 levels surrounding the TSS, and could observe that a large fraction of these genes was upregulated in *Rnf40*-deficient MEFs (Fig. 29C). Moreover, GSEA analysis using genes displaying decreased H3K27me3 occupancy in *Rnf40*-deficient MEFs further confirmed a significant enrichment for genes that were upregulated following *Rnf40* deletion (Fig. 29D). Thus, we proposed that the increased gene expression in *Rnf40*$^{-/-}$ MEFs is related to loss of H3K27me3 on promoters.
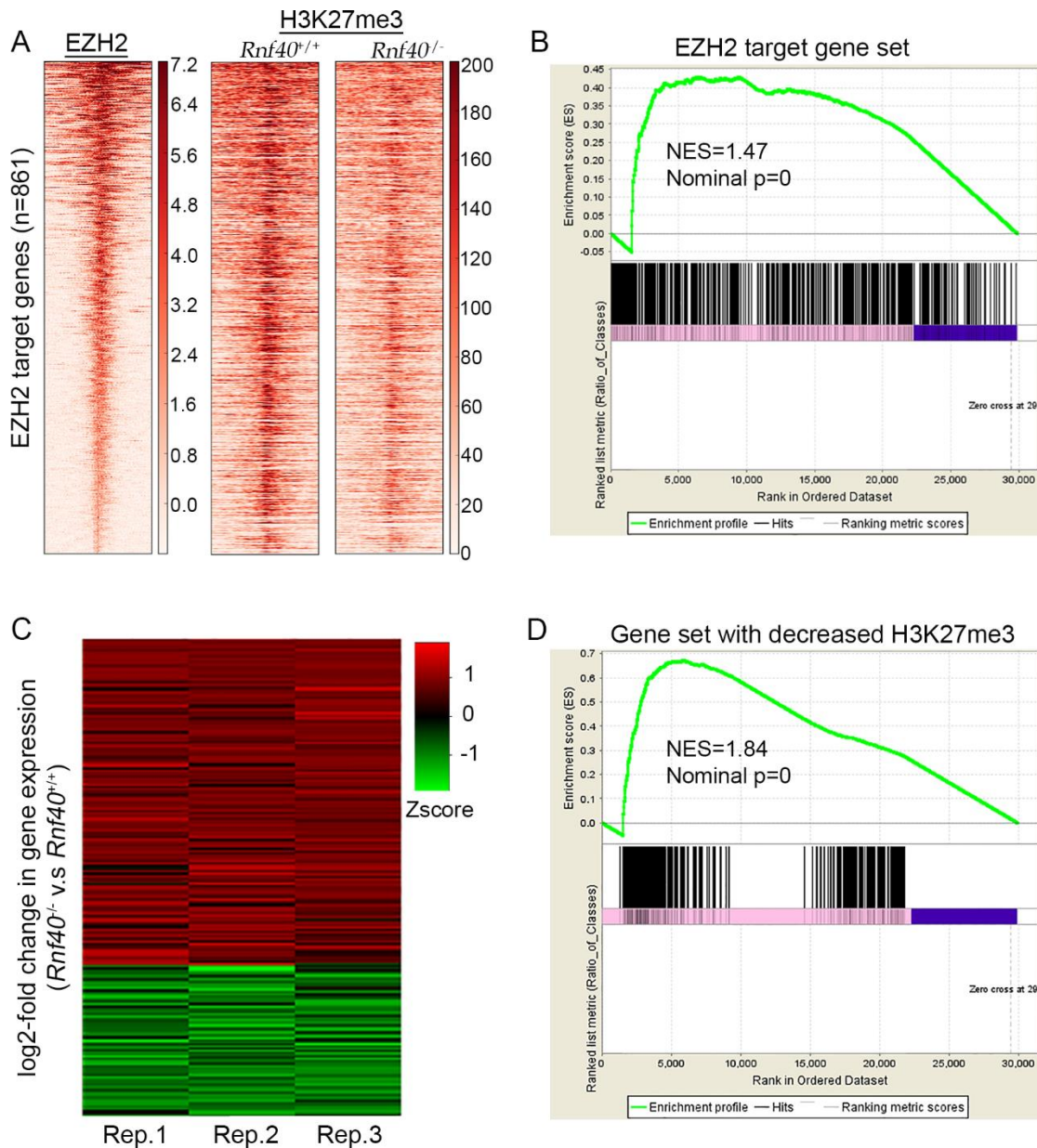
Fig.29 Regulation on EZH2 or H3K27me3 enriched genes

(**A**) Heatmaps show H3K27me3 occupancy of EZH2 targeted genes surrounding TSS in *Rnf40*$^{+/+}$ and *Rnf40*$^{-/-}$ MEFs. EZH2 target genes were identified using MACS for peak calling with "p-value < 0.001" cutoff for peak detection.

(**B**) GSEA analysis shows EZH2 targeted genes are significantly enriched for upregulation following *Rnf40* deletion. NES: normalized enrichment score.

(**C**) Heatmaps show the differential expression of gene displaying significantly decrease in H3K27me3 occupancy following *Rnf40* deletion. The Z score value of

log2-fold change in gene expression (*Rnf40$^{-/-}$* vs. *Rnf40$^{+/+}$*) is given for heatmaps. Red indicates upregulated genes, green indicates downregulated genes.

(**D**) GSEA analysis indicates the decreased H3K27me3 genes significantly enriched for upregulation in *Rnf40$^{-/-}$* MEFs. Genes from left to right are sorted by the ratio of gene expression (*Rnf40$^{-/-}$* / *Rnf40$^{+/+}$*).

4.2.4 Increased H3K27ac signals on PRC2-taget promoters are related to gene upregulation in *Rnf40$^{-/-}$* MEFs

Decrease of repressive signals is not enough to stimulate gene transcription, which requires additional active transcription signals. Recent findings suggested that p300 (one of histone acetyltransferases) is able to pre-occupy H3K27me3 enriched promoters, while its enzymatic activity is blocked by H3K27me3 (Rada-Iglesias et al. 2011; Zentner et al. 2011). Therefore, we hypothesized that loss of H3K27me3 might elevate H3K27ac occupancy. Indeed, Aggregate analysis of H3K27ac signals on EZH2 or decreased H3K27me3 target genes confirmed its significant increase near TSS regions (±5kb) in *Rnf40*-null MEFs (Fig. 30A and B).
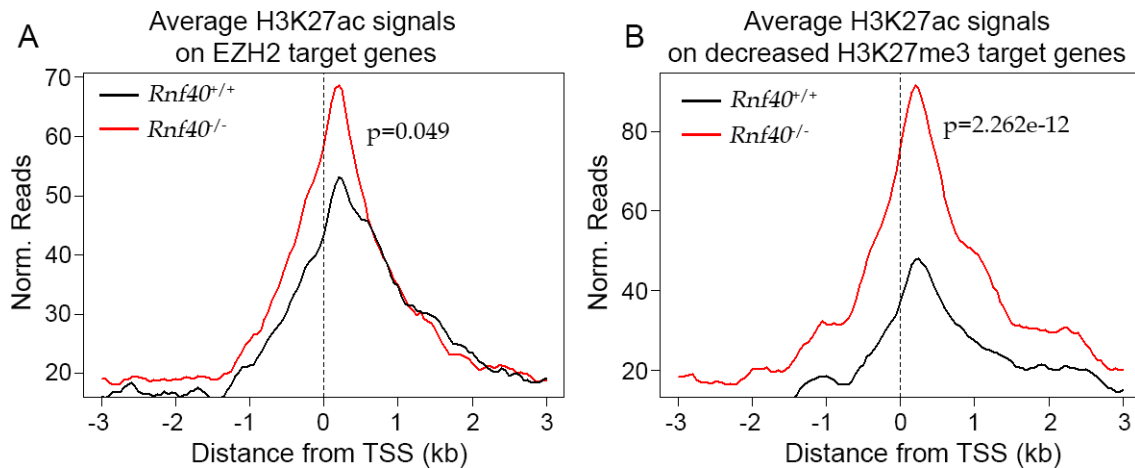
Fig.30 Increased H3K27ac signals on EZH2 or decreased H3K27me3 target promoters

Aggregate plot analysis of average H3K27ac profiles surrounding TSS (±3kb) in wildtype and *Rnf40*[-/-] MEFs. P-value was calculated using unpaired Wilcoxon-Mann-Whitney-Test.

4.2.5 Inhibiting EZH2 leads H3K27me3 targeted RNF40-suppressed genes upregulated in *Rnf40*[+/+] MEFs

To further confirm the influence of the loss of *Ezh2* induced by *Rnf40* deletion on gene transcription in our system, we compared the effects of treating *Rnf40*[+/+] MEFs with the EZH2 selective inhibitor EPZ-6438 to *Rnf40*-deleted MEFs. In agreement with EPZ-6438 specifically blocking EZH2 methyltransferase domain (Knutson et al. 2013), the treatment of EPZ-6438 for two days (EZH2i) significantly decreased global H3K27me3 level and did not affect H2Bub1 and H3K4me3, which was consistent to our findings that deletion of H2Bub1 also significantly decreased global H3K27me3 and H3K4me3 levels (Fig. 31A). Interestingly, we observed inhibiting EZH2 significantly elevated global H3K27ac levels (Fig. 31A). Similar to the effects of *Rnf40* deletion, the inhibition of EZH2 significantly upregulated several randomly

picked RNF40-supressed genes (*Foxl2, Foxl2os, Nat8l, Tgfa, Kcnc3, and Chd5*) that displaying significant H3K27me3 occupancy; while didn't affect H3K27me3 unenriched gene (*Psrc1*) ([Fig. 31B](#)). In addition, ChIP profiles confirmed significant decrease of H3K27me3 occupancy and increase of H3K27ac near TSS of those PRC2 targeted genes which displayed low- or unoccupied H2Bub1 ([Fig. 31C and D](#)). This indicated the upregulation of those genes was a result of the loss or inhibition of EZH2 instead of a direct effect of H2Bub1.
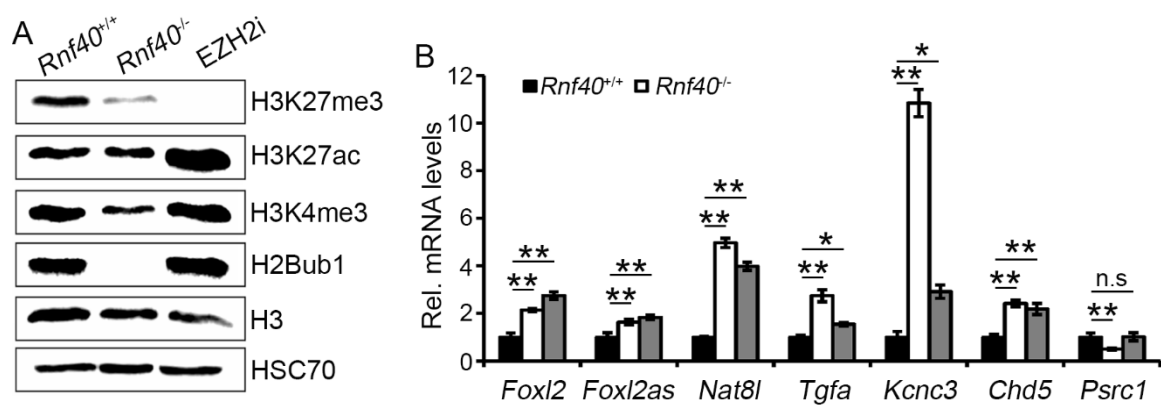
Fig.31 The effects of EZH2 small molecular inhibitor on RNF40-suppressed genes

(**A**) Western blot analysis of protein extracts from *Rnf40$^{+/+}$*, *Rnf40$^{-/-}$*, and EZH2 inhibitor treated *Rnf40$^{+/+}$* (EZH2i) MEFs with specific antibodies for targeting H3K27me3, H3K27ac, H3K4me3, H2Bub1, H3, and HSC70. The detections of H3 and HSC70 were considered as loading controls.

(**B**) qRT-PCR analysis of RNF40-suppressed genes in *Rnf40$^{+/+}$*, *Rnf40$^{-/-}$*, and EZH2i conditions. Data are normalized by *Rplp0* gene expression level and shown as "relative mRNA levels", mean ± SD from three independent experiments. *$p<0.05$, **$p<0.01$, ***$p<0.001$, calculated by two-tailed unpaired t-test.
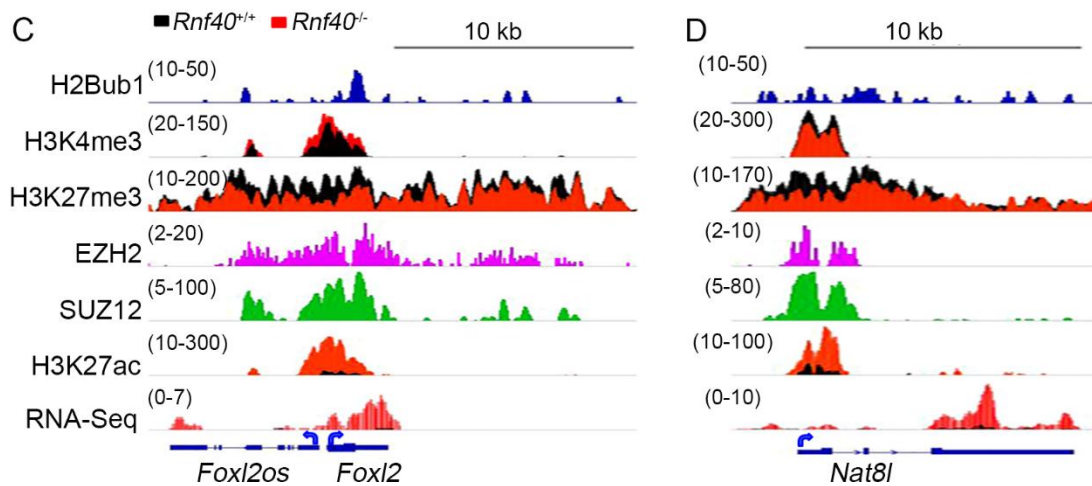
(**C**) qRT-PCR analysis of RNF40-dependent gene (*Psrc1*) in EZH2i MEFs.

(**D**) The profiles show the occupancy of H3K4me3, H2Bub1, H3K27ac , EZH2, SUZ12, and the normalized expression reads on the given RNF40-supressed genes in *Rnf40⁺/⁺* (black) and *Rnf40⁻/⁻* (red) MEFs.

4.3 H2Bub1 and bivalent domain

A genome-wide understanding of the chromatin landscape has revealed that a large number of genes are simultaneously decorated by both active H3K4me3 and repressive H3K27me3, which is referred to as bivalency (Bernstein et al. 2006; Voigt et al. 2013; Vlaming et al. 2014). It is believed that developmental genes are frequently enriched for bivalent chromatin marks (Sachs et al. 2013), which allows this subset of genes to be rapidly activated following developmental stimuli. Changes in either H3K4me3 or H3K27me3 can affect the expression of developmental genes resulting in either their full activation or repression (Agger et al. 2007; Wang et al. 2009; Agarwal and Jothi 2012). Interestingly, our above findings suggested that H2Bub1 genome-wide supervised H3K4me3 as well as H3K27me3 occupancy on promoters via different mechanisms. In addition, it suggested that H2Bub1 activated some important tissue-specific bivalent gene transcription via inducing the

demethylation of H3K27me3 (Karpiuk et al. 2012). However, Venn diagram overlapping of H3K4me3 enriched (+), H3K27me3 enriched (+), *Rnf40*-dependent (down.), and *Rnf40*-independent (upreg.) genes revealed that bivalent genes were differentially regulated in response to *Rnf40* deletion (Fig. 17). Thus, it is necessary to confirm the genome-wide role of H2Bub1 in bivalent activity.

4.3.1 H2Bub1 is required for the transcription of bivalent genes

We first identified 2041 bivalent genes in *Rnf40*$^{+/+}$ MEFs (Fig. 12C). Interestingly, H2Bub1 was more highly correlated to gene transcription of bivalent genes detected by GRO-Seq compared to that in global genes (0.42 v.s 0.28) as well as to RNAP II (0.41 v.s 0.36), while H3K4me3 was less correlated to bivalent gene transcription (0.24 v.s 0.34) as well as to RNAP II (0.3 v.s 0.58) (Fig. 32A and Fig. 7B). For example, *Zfp275* and *Bmp3* show similar occupancy of H3K4me3, the difference being the *Zfp275* is enriched by H2Bub1, RNAPII, and H3K27ac whileBmp3 is targeted by EZH2 and occupied by H3K27me3 to a greater degree. Therefore *Zfp275* is transcribed, while *Bmp3* is silenced (Fig. 32B). This finding suggested that H2Bub1 may play a more crucial role in controlling the transcriptional activity of bivalent genes compared to H3K4me3.
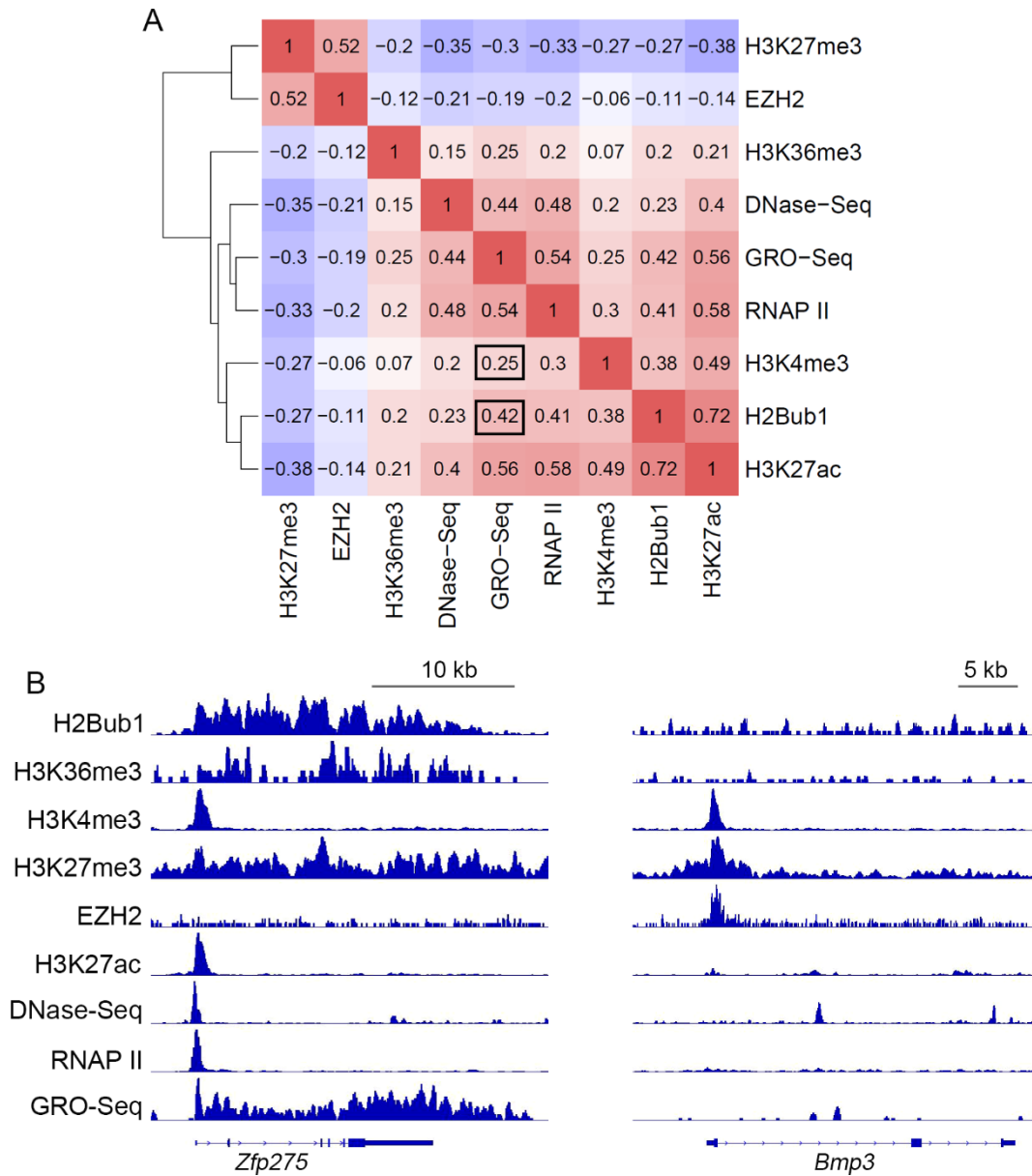
Fig.32 Correlation of H3K27me3, EZH2, H3K36me3, DNase-Seq, GRO-Seq, RNAP II, H3K4me3, H2Bub1, and H3K27ac near bivalent gene TSSs

(**A**) Heatmap coupled with correlation coefficients displaying the correlations on the region from TSS to downstream 1kb.

(**B**) ChIP profiles show H3K27me3, EZH2, H3K36me3, DNase-Seq, GRO-Seq, RNAP II, H3K4me3, H2Bub1, and H3K27ac occupancy on *Zfp275* and *Bmp3* genes.

4.3.2 H2Bub1 modulates bivalent gene transcription in an epigenetic-context manner

It was suggested that the bivalency could be considered as a metastable equilibrium between active and repressive state, which was supervised by various chromatin-modifying complexes (Voigt et al. 2013). However, less is known about their resolution. In agreement with our previous reports (Karpiuk et al. 2012), H2Bub1 is required to activate bivalent genes. To understand in detail the role of H2Bub1 in the shifting of this metastable equilibrium, we classified the 2041 bivalent genes into H2Bub1 enriched and unenriched groups. In addition to confirm the correlation between H2Bub1 and other factors, heatmaps and boxplots analyses found out that loss of H2Bub1 significantly decreased H3K27me3 (Fig. 33D) and increased H3K27ac (Fig. 33E) on all bivalent genes. However, the behavior of H3K4me3 was dependent on H2Bub1. In H2Bub1+ genes this mark significantly decreased while H2Bub1- gene it was unaffected. Notably H3K4me3 on H2Bub1+ genes showed significantly higher occupancy than on H2Bub1- genes (Fig. 33C). It can be speculated that the additional H3K4me3 occupancy on H2Bub1+ genes compared to H2Bub1- genes is highly dependent on H2Bub1. Moreover, H2Bub1+ bivalent genes were averagely downregulated and H2Bub1- bivalent genes were significantly upregulated following H2Bub1 deletion (Fig. 33B). This data confirmed H2Bub1 coordinates bivalent gene transcription by controlling the shifting of the metastable equilibrium between H3K4me3 and H3K27me3.
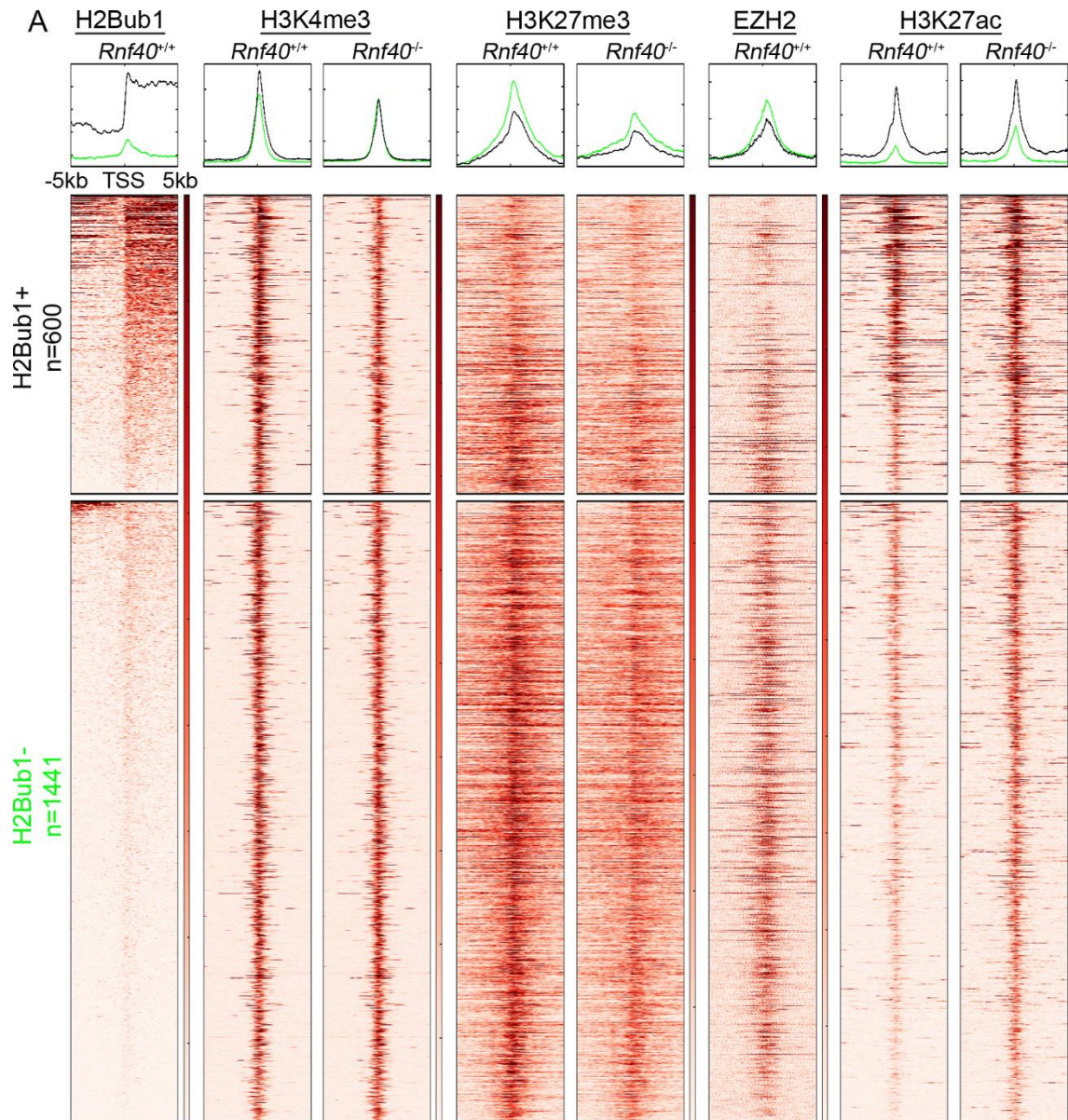
Fig.33 The behavior of multiple histone modifications on H2Bub1 enriched (+) or H2Bub1 unenriched (-) bivalent gene TSS following *Rnf40* deletion

(**A**) Heatmaps showing H2Bub1, H3K4me3, H3K27me3, SUZ12, EZH2, and H3K27ac occupancy on the promoter (the region within ± 5kb from TSS) of classified bivalent genes. Genes were grouped according to H2Bub1 occupancy. H2Bub1 enriched (+) bivalent genes were identified according to H2Bub1 occupancy on gene body ($\log_2$-normalized value greater than 2). Genes were sorted according to H2Bub1 occupancy from high to low.
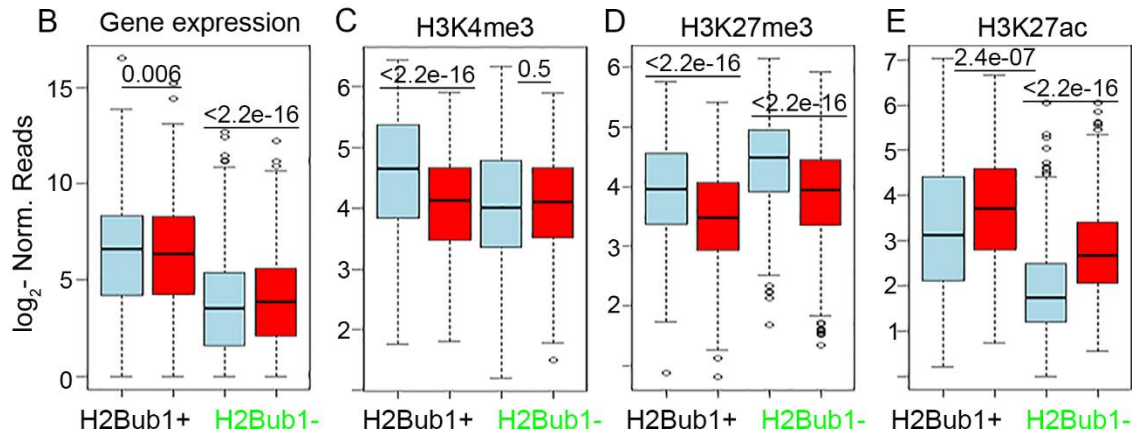
(**B-E**) Boxplots compared the alteration of H3K4me3 (B), H3K27me3 (C), H3K27ac (D) occupancy near TSS (from TSS to downstream 1 kb), as well as gene expression (E) in the grouped genes. P-value was calculated by paired Wilcoxon-Mann-Whitney-Test.

To investigate the biological function of the differentially regulated bivalent genes, we conducted GO analysis of the 163 upregulated or 99 downregulated bivalent genes (Fig. 17). Interestingly, similar to global upregulated genes enriched GO terms, upregulated bivalent genes were significantly enriched for development associated genes while downregulated bivalent genes did not display significant enrichment (Table 1).

Downregulated bivalent gene enriched GO terms				
GO Terms	Count	%	p value	FDR
homeostatic process	9	9.09	0.009	13
chemical homeostasis	7	7.07	0.010	14
system development	19	19.19	0.010	15
cellular calcium ion homeostasis	4	4.04	0.011	15
calcium ion homeostasis	4	4.04	0.012	17
cellular metal ion homeostasis	4	4.04	0.014	19
extracellular matrix organization	4	4.04	0.014	20

metal ion homeostasis	4	4.04	0.016	22
Upregulated bivalent gene enriched GO terms				
tissue development	20	12.27	6.9E-07	0.001
multicellular organismal development	40	24.54	1.7E-05	0.026
system development	34	20.86	3.6E-05	0.056
anatomical structure development	35	21.47	5.7E-05	0.090
localization	40	24.54	5.8E-05	0.091
cell communication	14	8.59	7.1E-05	0.111
regulation of cell adhesion	7	4.29	1.1E-04	0.167
developmental process	40	24.54	1.2E-04	0.184

Table1 Gene Oncology analysis

GO analysis shows the enrichment pathway of differentially regulated bivalent genes in response to *Rnf40* deletion. GO analysis was performed based on 'GOTERM_BP_ALL ' database. FDR < 0.05 was considered as significant enrichment.

4.3.3 H2Bub1 coordinates homeobox genes activation and repression

The *Hox* gene clusters represent a prototypical evolutionarily conserved example of coordinated transcriptional and epigenetic regulation during development. Importantly, H2Bub1 was previously reported to be required for the activation of some Hox genes (Zhu et al. 2005). Thus, in order to examine the regulatory mechanisms by which H2Bub1 functions and coordinates the epigenomic landscape on *Hox* genes we analyzed the effects of *Rnf40*-depletion on histone modification profiles on *Hox* gene clusters. Interestingly, these profiles showed that *Hox* clusters were decorated by both active and repressive histone modifications (Fig. 34). For example, the *Hoxc* cluster displays increasing levels of the active histone marks H3K4me3 and H3K27ac in a 5' to 3' manner while the repressive mark H3K27me3

displayed an inverse pattern ([Fig. 34](#)). Interestingly, following *Rnf40* deletion and the loss of H2Bub1, the levels of H3K4me3 on the active 3' portion of the cluster containing *Hoxc4-10* decreased significantly compared to other less active genes. Notably, consistent with our other findings, H3K27me3 levels at the 5' end of the cluster (e.g. *Hoxc9-13*) decreased. Finally, the levels of H3K27ac decreased at the 3' end *Hoxc* genes (*Hoxc4-8*), and increased at the 5' end of the cluster (*Hoxc13-10*) ([Fig. 34](#)).
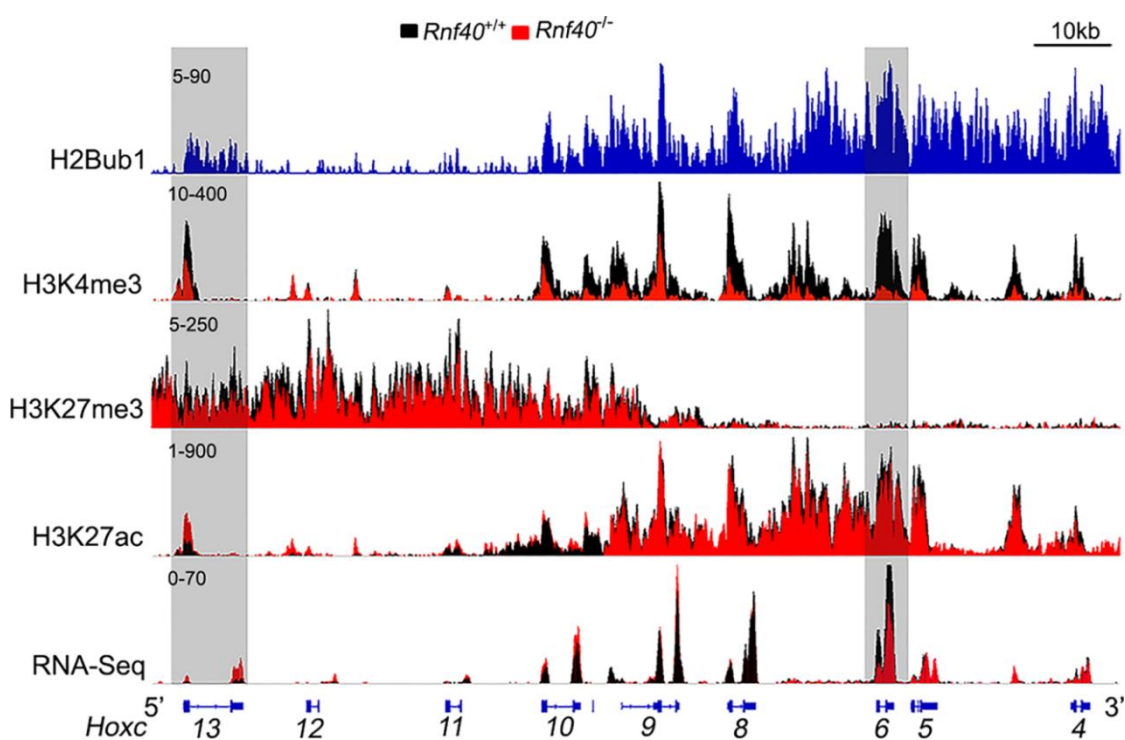
Fig.34 ChIP and RNA-seq profiles on Hoxc genes

Moreover, qPCR analysis confirmed that the H3K27me3 enriched 5' Hoxc gene expression (*Hoxc13* and *Hoxc10*) significantly increased while H3K4me3 and H2Bub1 enriched 3' Hoxc gene expression (*Hoxc6* and *Hoxc10*) significantly decreased in response to H2Bub1 deletion; However, EZH2 inhibitor treatment consistently upregulated *Hoxc13* and *Hoxc10* while didn't affect *Hoxc6* and *Hoxc8*

([Fig. 35A](#)). In addition, ChIP-qPCR analysis confirmed that both depletion of RNF40 or inhibition of EZH2 methyltransferase activity significantly decreased H3K27me3 occupancy on promoter of *Hoxc13* gene while H3K4me3 levels were unaffected and RNAPII occupancy increased. In contrast, RNF40 loss resulted in decreased H3K4me3 and RNAP II occupancy on promoter of the *Hoxc6* gene, while EZH2 inhibition had no effect on their occupancy ([Fig. 35B](#)). Together, these data suggest that H2Bub1 differentially regulates *Hox* genes in a context-dependent manner by coordinating the equilibrium between active (H3K4me3 and H3K27ac) and repressive (H3K27me3) histone modifications.
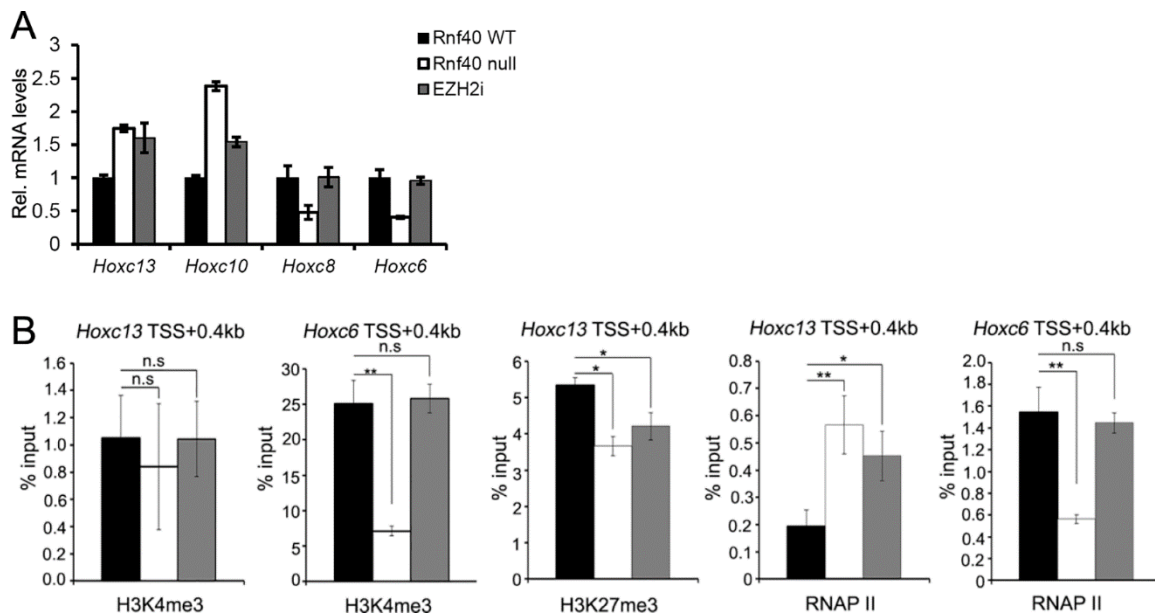
Fig.35 Single gene analysis confirming gene regulation and the alteration of histone modification on Hoxc genes

(**A**) qPCR analysis of *Hoxc13, Hoxc10, Hoxc8,* and *Hoxc6* gene expression in *Rnf40* wide type (Rnf40 WT), *Rnf40* knock out (Rnf40 null), and EPZ-6438 treated *Rnf40* wide type (EZH2i) MEFs. Data are normalized by *Rplp0* gene expression level and shown as "relative mRNA levels", mean ± SD from three independent experiments.

(**B**) ChIP-qPCR analysis of the occupancy of H3K4me3, H3K27me3, and RNA polymerase II (RNAP II) near TSS (at downstream 400 bp) of *Hoxc13* and *Hoxc6* genes. ChIP signal was normalized by input signal, and represented as '%input' ± SD (n=3). *$p<0.05$, **$p<0.01$, n.s represents $p>0.05$, calculated with two-tailed unpaired t-test.

4.4 H2Bub1 and Enhancers

In agreement to the results in siRNA-mediated RNF20 or RNF40 knockdown human cells (Shema et al. 2008), a number of H2Bub1 target genes were upregulated after *Rnf40* deletion. One report proposed that RNF20, another E3-specific ligase for H2Bub1, suppresses pro-oncogenic genes by blocking the recruitment of the transcription elongation factor S-II (TFIIS) (Shema et al. 2011). Additionally, we confirmed that a subset of increased genes in *Rnf40*[-/-] MEFs is related to the demethylation of H3K27me3 at promoters ([Fig.17](#)). However, we further observed more than 41% (276/672) of RNF40-suppressed genes displayed no significant occupancy of H3K4me3, H3K27me3, and H2Bub1 at promoters ([Fig. 17](#)). In the past few years, it has been confirmed that enhancers have the ability of modulating gene transcription from a great distance, by recruiting multiple transcription factors to loop out the intervening sequences facilitating interactions with promoters. Thus, we hypothesized that a subset of H2Bub1-suppressed genes is associated with enhancer activity.

4.4.1 H2Bub1 modulates enhancer activity

It is believed that H3K4me1 initially occupies enhancers, and H3K27ac typically represents enhancer activity, while H3K4me3 poorly enriches enhancers. We identified 30893 active enhancers decorated with H3K4me1 and H3K27ac and

unoccupied by H3K4me3 (H3K4me1$^+$H3K27ac$^+$H3K4me3$^-$) in MEFs ([Fig. 36A](#)). In an attempt to understand enhancer activity following *Rnf40* deletion, DiffBind analysis was performed to identify H3K27ac-enriched enhancer regions which change between *Rnf40* wt and null MEFs. In general, H3K27ac signals on enhancers were increased in *Rnf40*$^{-/-}$ MEFs ([Fig. 36B](#)). Differentially, 7852 of enhancers displayed increased H3K27ac signals, while 6913 of enhancers exhibited decreased H3K27ac signals (FDR<0.05) ([Fig 36B](#)). These results indicate that gene body enriched H2Bub1 could affect H3K27ac occupancy on enhancers.
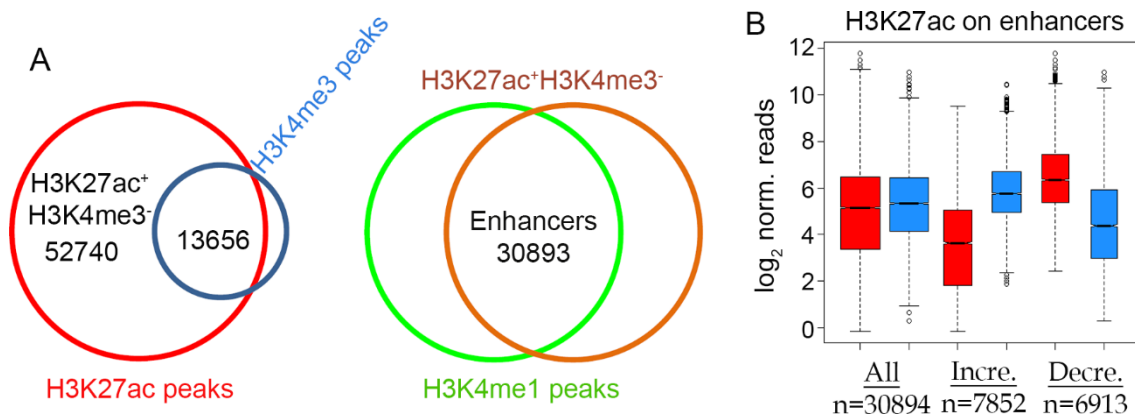
Fig.36 Loss of H2Bub1 affects H3K27ac occupancy on enhancers

(**A**) Identification of the activated enhancers in *Rnf40*$^{+/+}$ or *Rnf40*$^{-/-}$ MEFs. Venn diagram show the shared binding region of H3K27ac and H3K4me3 (left panel) as well as H3K4me1 and H3K27ac (right panel). Enhancer was defined as H3K4me1 enriched (+), H3K27ac enriched (+), but H3K4me3 unenriched (-) regions.

(**B**) DiffBind analysis of H3K27ac behaviors on enhancers following *Rnf40* deletion. Increased (Incre.) or decreased (Decre.) activity of enhancer were separately selected as follows: log$_2$-fold change >0, FDR<0.05; log$_2$-fold change <0, FDR<0.05.

4.4.2 The activation of RNF40-suppressed genes in *Rnf40*$^{-/-}$ MEFs is highly related to increased enhancer activity

In order to characterize the relationship between increased or decreased enhancer activity and the induction of gene expression following *Rnf40* deletion, we separately predicted the significant increased and decreased enhancer annotated genes by conducing GREAT analysis (McLean et al. 2010). Interestingly, we identified 7641 of significant increased enhancers which annotated 4725 genes, involving more than 46% (308/672) of upregulated genes while only 19% (156/802) of downregulated genes (Fig. 37A). In addition, 3222 of significant decreased enhancers annotated 3433 genes, only containing 14.2% (114/802) of downregulated genes and 17% (117/672) of upregulated genes (Fig. 37B). Hence, we hypothesized that the upregulation of a subgroup of RNF40-suppressed genes is independent of a direct effect on PRC2 and may be instead be associated with the de novo activation of distal enhancers.
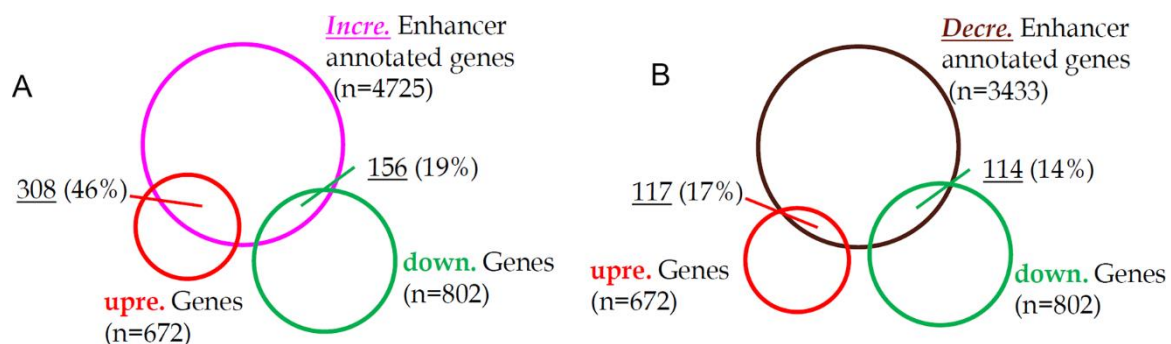
Fig.37 Distal enhancer activation affects RNF40-suppressed gene transcription
(**A**) Venn diagram analysis shows the shared gene numbers during RNF40-suppresed genes (upre. genes), RNF40-dependent genes (down. genes), and RNF40-suppressed enhancer associated genes.

(**B**) Venn diagram shows the shared gene numbers between RNF40-dependent genes, RNF40-independent genes and RNF40-dependent enhancer associated genes. Significantly increased enhancers show 2-fold increased H3K27ac, while significantly decreased enhancers show 2-fold decreased H3K27ac (FDR<0.05).

4.4.3 The increased *Foxl2* expression is responsible for the activation of enhancers associated with RNF40-suppressed genes

To uncover potential transcription factors which may contribute to enhancer activation and upregulation of genes in *Rnf40*$^{-/-}$ MEFs, we performed sequence-based motif analysis of the 308 gene-associated enhancers and identified a significant enrichment of Forkhead box protein binding motifs ([Fig. 38A](#)). Given that the expression of *Foxl2* was significantly increased in *Rnf40*$^{-/-}$ MEFs ([Fig. 31C and B](#)), we further identified 3223 enhancers in our study which were found to be occupied by FOXL2 in a published ChIP-seq dataset (Georges et al. 2014) ([Fig. 38B](#)). In addition, GREAT analysis of those regions identified the FOXL2-enriched enhancer-associated genes, which contained more than 27% (184/672) of the upregulated genes while 16% (136/802) of downregulated genes ([Fig. 38C](#)), and 100 genes (more than 60%) which were upregulated and displayed enhancer activation following *Rnf40* deletion ([Fig. 38D](#)). Consistent with increased enhancer activation, the H3K27ac occupancy surrounding these FOXL2 enriched distal regions was significantly increased in *Rnf40*$^{-/-}$ MEFs ([Fig. 38E](#)).
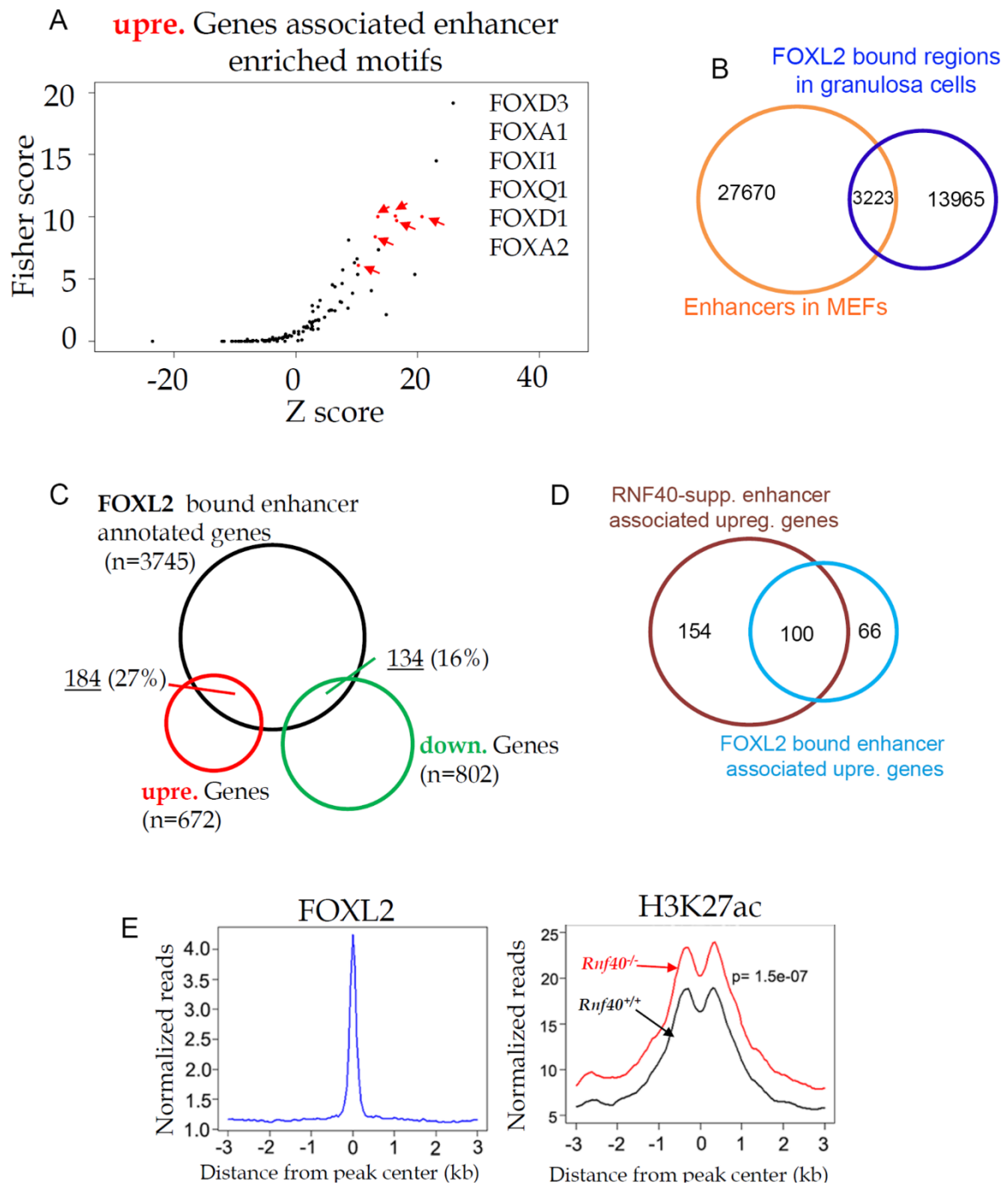
Fig.38 FOXL2 is distributed to activate RNF40-suppressed gene associated enhancers

(**A**) Plot showing Fisher Score and Z Score in motif analysis of RNF40-suppressed gene associated activating enhancers. The Fisher Score >5 & Z Score > 10 indicates significantly enriched motif pointed out red.

(**B**) Venn diagram shows the shared regions between enhancer in MEFs and FOXL2 binding sites in mouse granulosa cells.

(**C**) Venn diagrams showing the numbers shared by downregulated genes, upregulated genes, and FOXL2 enriched enhancer associated genes. FOXL2 enriched enhancer associated genes were obtained from GREAT analysis of FOXL2 enriched enhancers.

(**D**) Venn diagram shows the shared genes between RNF40-suppressed enhancer associated and FOXL2 enriched enhancers associated upregulated genes.

(**E**) Aggregate profiles show FOXL2 and H3K27ac average signal surrounding the FOXL2 binding center (±3kb) in $Rnf40^{+/+}$ or $Rnf40^{-/-}$ MEFs. The regions were given from 184 (136 +48) genes associated FOXL2 enriched enhancers according to Fig. 38C. P-value was calculated by unpaired Wilcoxon-Mann-Whitney-Test.

In order to confirm the role of FOXL2 in the upregulation of this subset of genes in $Rnf40^{-/-}$ MEFs, we examined the effects of siRNA-mediated FOXL2 depletion in MEFs following $Rnf40$ deletion. Consistent with a previous study demonstrating the importance of FOXL2 for their expression (Georges et al. 2014), we observed that both the $Esr2$ and $Efna5$ genes were significantly upregulated following $Rnf40$ deletion as well as following EZH2 inhibitor treatment. Importantly and consistent with an indirect effect mediated by FOXL2, these effects were blocked by FOXL2 depletion (Fig. 39A). Moreover, ChIP-seq profiles confirmed that H3K27ac occupancy on each of these genes was increased at FOXL2-bound enhancers following $Rnf40$ deletion (Fig. 39B and 39C). Together these data support a central

role for FOXL2 in mediating enhancer activation and increased gene expression of a subset of genes whose expression increases following *Rnf40* deletion.
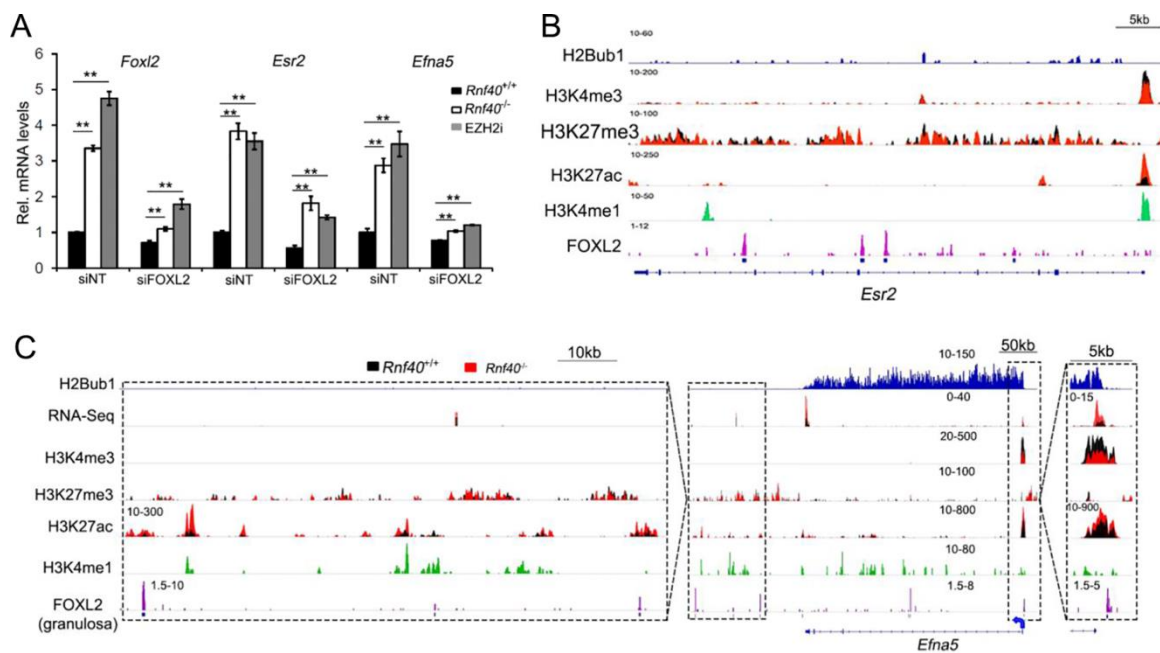
Fig.39 SiRNA-mediated *Foxl2* knockdown blocks a fragment of RNF40-suppressed genes

(**A**) qPCR analysis of *Foxl2, Esr2,* and *Efna5* gene expression in siRNA-mediated *Foxl2* knockdown *Rnf40*[+/+], *Rnf40*[-/-], and EZH2i MEFs. Data are normalized by *Rplp0* gene expression level and shown as "relative mRNA levels", mean ± SD from three independent experiments. **p<0.01, calculated with two-tailed unpaired t-test.

(**B and C**) The ChIP-seq profiles show occupancy of H2Bub1, H3K4me3, H3K27me3, H3K27ac, H3K4me1 and FOXL2 on *Esr2* and *Efna5* genes and their associated distal regions.

5. Discussion

The complex regulatory network of post-translational histone modifications has long been hypothesized to play a significant role in controlling the timely activation or repression of gene transcription (Strahl and Allis 2000), which is associated to cell fate determination during embryogenesis and tumorigenesis. H2Bub1 is one such histone modification, studied extensively in the past decade. Several mechanisms have been proposed to explain the role of H2Bub1 in facilitating gene transcription, such as the CDK9-WAC-RNF20/40 axis (Johnsen 2012; Karpiuk et al. 2012), FACT and H2A:H2Bub1 cooperative interaction (Pavri et al. 2006), H2Bub1-H3 methylation trans-histone tail crosstalk mechanism (Kim et al. 2009; Kim et al. 2012), etc. In contrast, transcription repressive mechanisms were also established for H2Bub1, such as RNF20 inhibiting TFIIS recruitment (Shema et al. 2011) and H2Bub1 stabilizing nucleosome model (Fleming et al. 2008). Here we investigated the genome-wide occupancy of H2Bub1 and examined the effects of its loss on promoter and enhancer activity at diverse chromatin states and how this is related to alterations in gene transcription. Furthermore, we propose context-dependent mechanisms of direct or indirect regulation of gene expression by H2Bub1 via comprehensive mapping of active histone modifications (H3K4me3 and H3K27ac) and repressive H3K27me3 (Fig. 40). In addition to validating the previously recognized H2Bub1 correlation with promoter activity via trans-histone tail crosstalk (Dover et al. 2002; Sun and Allis 2002), excitingly, our data provide the first mechanistic explanations of how loss of RNF40 can modulate enhancer activity and suggest that the role of H2Bub1 in transcription are far beyond facilitating transcription elongation.
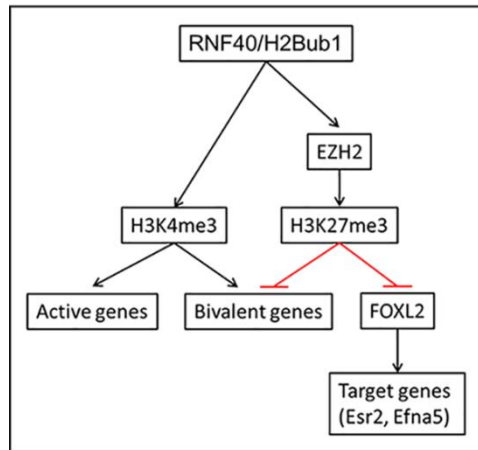
Fig. 40 Model describing RNF40 mediated H2Bub1 regulatory network.

RNF40-mediated H2Bub1 facilitates genes transcription involving bivalent genes via H2Bub1-H3K4me3 crosstalk. On the other hand, H2Bub1 indirectly maintains an abundance of genes in a silenced state by controlling EZH2 level which is a subunit of PRC2 complex responsible for repressive H3K27me3 marks.

5.1 H2Bub1 crosstalk with H3K4me3 to modulate gene transcription

In addition to the agreement with the H2Bub1-H3K4me3 crosstalk model in which H2Bub1 facilitates the trimethylation of H3K4 by the SET/COMPASS complex (Kim et al. 2013), we first provide genome-wide data that H2Bub1 is tightly linked to H3K4me3. However, the absence of H2Bub1 results in decreased (but not a total loss) of H3K4me3 occupancy on each gene. Notably, the decrease in H3K4me3 occupancy was most apparent at regions downstream of the TSS, which was also co-occupied by H2Bub1. Following loss of H2Bub1, these H3K4me3 peaks displayed a significant narrowing of their TSS-associated peaks. We speculate that the bulk of H3K4me3 near the TSS may be catalyzed by SET/COMPASS or other H3K4 methyltransferases in an RNF40/H2Bub1-independent manner, but that transcriptional elongation-associated spreading of H3K4me3 into genes is highly dependent upon RNF20/40-mediated H2B monoubiquitination. This effect can also

be observed on the *Hoxc* gene cluster where H3K4me3 on each of the *Hoxc* genes decreases, but some degree of H3K4me3 remains and becomes more focused around the TSS. These effects closely resemble those observed in *Mll1*-deficient MEFs (Wang et al., 2009), suggesting that H2Bub1 may selectively direct MLL-dependent H3K4 methylation near TSS.

Recent studies provided a previously unrecognized association of broad H3K4me3 peaks with the expression of tumor suppressor and cell identity genes (Benayoun et al. 2014; Chen et al. 2015b). We consistently observe broadest H3K4me3 domains are significantly enriched in cell cycle-related tumor suppressors and development-associated genes in MEFs. However, the knowledge about the factors promoting H3K4me3 spreading is limited. In mammals, multiple methyltransferases are responsible for H3K4me3, SET1A/B complex carries out the bulk of H3K4me3 (Wu et al. 2008), while the MLL complexes more likely function at specific genes (Shilatifard 2012). CFP1 containing CxxC domain, the unique component of SET1A/B complex, specifically directs SET1A/B complex into non-methylated CpG-islands enriched promoters to carry out H3K4me3 (Lee and Skalnik 2005). Excitingly, deletion of *CFP1* and depletion of *WDR82* (another unique subunit of SET1A/B complex) narrow the width of H3K4me3 peaks genome-wide (Clouaire et al. 2012; Clouaire et al. 2014; Austenaa et al. 2015), suggesting that SET1A/B complex is the major player responsible for widening H3K4me3 peaks ([Fig. 41](#)). Given that the H2Bub1-H3K4me3 crosstalk mechanism was established early between H2Bub1 and SET1A/B complex in which H2Bub1 facilitates H3K4me3 via providing a binding site for WDR82 (Wu et al. 2008), it is fascinating to speculate that H2Bub1 is associated to the wide spread of H3K4me3 domain. Indeed, in addition to agreement with potential tumor suppressor functions of RNF20/40 and H2Bub1 and their

requirement for stem cell differentiation (Johnsen 2012; Karpiuk et al. 2012), we could observe a widening of H3K4me3 peaks on RNF40-dependent lineage-specific genes during differentiation (Fig. 41). In turn, loss of H2Bub1 significantly narrowed the broadest H3K4me3 domain. Together, we suggest that the presence of broad H3K4me3 domains on tumor suppressor and lineage-specific genes is highly dependent on H2Bub1.
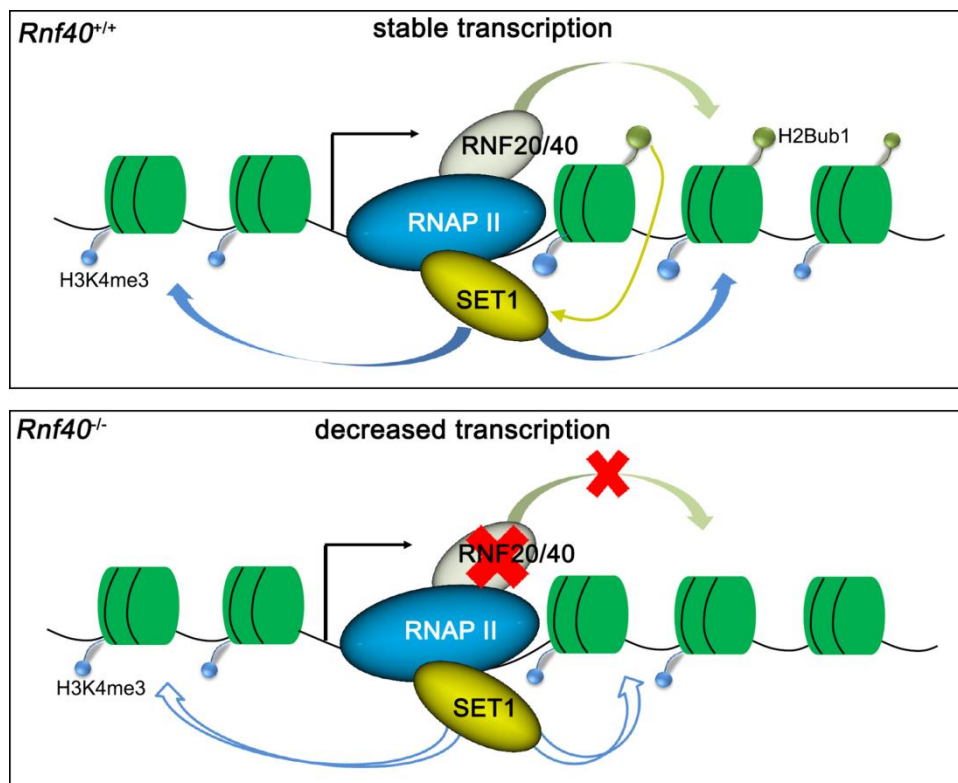
Fig. 41 Transcriptional dependency on H2Bub1 is associated to H3K4me3 width

The SET1A/B complex is directed into the non-methylated CpG island-enriched gene via the CxxC domain-containing CFP1, and then catalyzes trimethylation on H3 lysine 4. The widening of H3K4me3 into the gene is facilitated by H2Bub1 via interacting with WDR82, which promotes gene transcription. Loss of H2Bub1 leads to significant narrowing of H3K4me3 peaks, resulting in reduced transcription level.

5.2 Transcriptional dependency on H2Bub1 is not associated with gene length in deletion system

The coactive function of H2Bub1 in transcription has been confirmed in various independent studies. However, loss of H2Bub1 decreases the expression of only a subset of genes. One early study suggests gene expression level is tightly associated to its own morphological parameters, such as distance to neighbors, gene or 3'-UTR length, number of exons, etc. (Chiaromonte et al. 2003). Indeed, depletion of *RNF20* in human cells selectively leads to significant decreasing of longer gene expression (Fuchs et al. 2012). However, we don't observe that transcriptional dependency on H2Bub1 is associated to gene length in our *Rnf40* inducible knockout system. A possible reason is that depletion of RNF20 cannot lead to complete absence of H2Bub1 compared to our knockout system. There are no significant differences of H2Bub1 level on gene body between the shortest and longest transcribed genes. Given that H2Bub1 occupies the gene body in a gradually decreasing manner from 5' to 3' region, the average level of H2Bub1 at 3' regions in the longest genes is significantly lower than that in shortest genes. Even though H2Bub1 occupancy at 3' regions is low, it plays a crucial role in mRNA 3'-end processing (Pirngruber et al. 2009). Therefore, depletion of *RNF20* would induce absence of H2Bub1 at 3' regions, resulting in a significant narrowing of H2Bub1 peaks in longest genes compared to that in shortest genes. The transcriptional product at 3' regions can only be observed at the presence of H2Bub1 (Pavri et al. 2006). Thus, the long genes show much more dependency on H2Bub1 in the RNF20 depleted cells. However, H2Bub1 is completely lost both in short and long genes in the deletion systems. Thus, we observed that transcriptional dependency on H2Bub1 is not associated to gene length.

pg. 118

5.3 RNF40-regulated genes display low and moderate H2Bub1 occupancy

In order to obtain efficient activation of gene transcription, the signals enabling transcriptional activity, including active histone modifications, need to reach a certain threshold after which gene expression can occur (Voigt et al. 2013). According to our data, genes, which display the highest occupancy of H2Bub1 and other active histone modifications, appear to be more robustly expressed and less sensitive to changes in the presence of individual histone modification. Even in the absence of H2Bub1, these genes retain sufficient additional active signals to maintain high levels of transcription. In another case, some highly transcribed genes such as housekeeping genes are controlled by simpler promoters and enhancers showing lower sequence conservation, and do not require additional particular transcription complexes to activate transcription (Farre et al. 2007). The transcription of these genes shows low degree of dependency on histone modifications. In contrast, fully repressed genes such as those marked by only repressive histone modifications may require extremely high levels of additional activation signals to switch from a repressed to an active state. Furthermore, there seems to be a complex regulatory mechanism at genes marked by varying degrees of both active and repressive histone modifications as we observed for low to moderate H2Bub1 occupied genes, whose transcription require histone modifications to cooperate with the recruitment of tissue-specific transcription factors. Thus, it is particularly vulnerable to changes in expression elicited by the loss of either active or repressive marks.

5.4 H2Bub1 indirectly maintains cell identity via controlling EZH2 transcription

Although H2Bub1 occupancy has been shown to correlate with transcriptional elongation and gene activity (Minsky et al. 2008; Shema et al. 2008), a number of genes have also been shown to be upregulated following loss of H2Bub1 in mammalian cells (Shema et al. 2008; Shema et al. 2011; Bonnet et al. 2014). Consistently, in our studies we observed similar numbers of genes up- and down-regulated in *Rnf40*$^{-/-}$ MEFs. One previous explanation for a repressive role of H2Bub1 was that it obstructs the recruitment of transcription elongation factor TFIIS to genes, thereby suppressing transcription (Shema et al. 2011). However, consistent with findings following *RNF20* or *RNF40* knockdown (Jung et al. 2012), we find that the vast majority of RNF40-suppressed genes do not display significant levels of H2Bub1, thereby suggesting that their regulation may occur through more indirect mechanisms. Consistently, we find that the *Ezh2* gene, encoding the catalytic component of the PRC2 complex, which mediates H3K27 methylation, displays a significant level of H2Bub1 occupancy and requires RNF40 for its full expression. Moreover, the dependency of EZH2 transcription on H2Bub1 is conserved, and can be observed in other cell lines such as HCT116 (Chen et al. 2015a).

Notably, in agreement with the finding that EZH2 primarily occupies regions surrounding TSS in mouse and human cells (Xu et al. 2012; He et al. 2013), decreased *Ezh2* expression following *Rnf40* deletion resulted in decreased H3K27me3 occupancy near the TSS. Furthermore, consistent with a crucial role for EZH2 as a central mediator of H2Bub1-dependent gene repression, small molecule inhibition of EZH2 enzymatic activity resulted in a similar de-repression of

H3K27me3 targeted genes which were up-regulated in *Rnf40*-deficient MEFs. However, the upregulation of H3K27me3 targeted genes induced by RNF40 deletion could be blocked by overexpressing EZH2 but not in EZH2 SET domain mutated cells. Together, we observed a previously unrecognized regulatory pathway by which gene is silenced via the RNF40/H2Bub1-EZH2-H3K27me3 axis ([Fig. 40](#)).

Only decreasing repressive H3K27me3 is not enough to switch on gene transcription. For this, additional coactivators to facilitate transcription factor recruitment are required. H3K27ac functions as a binding platform for transcription factors such as BRD4. This modification is mediated by HATs such as P300/CBP. Several independent studies have observed that P300/CBP pre-occupies the PRC2-targeted promoters, while its enzymatic activity is blocked by H3K27me3 (Pasini et al. 2010; Tie et al. 2014; Lee et al. 2015b). Therefore, loss of H3K27me3 by EZH2 inhibition resulted in increased global H3K27ac levels. In addition, we confirmed that the upregulation of RNF40-suppressed genes appears to be related to a shift in the balance between H3K27me3 and H3K27ac whereby decreased H3K27 methylation enables the acetylation of the same residue at these loci ([Fig. 42](#)).
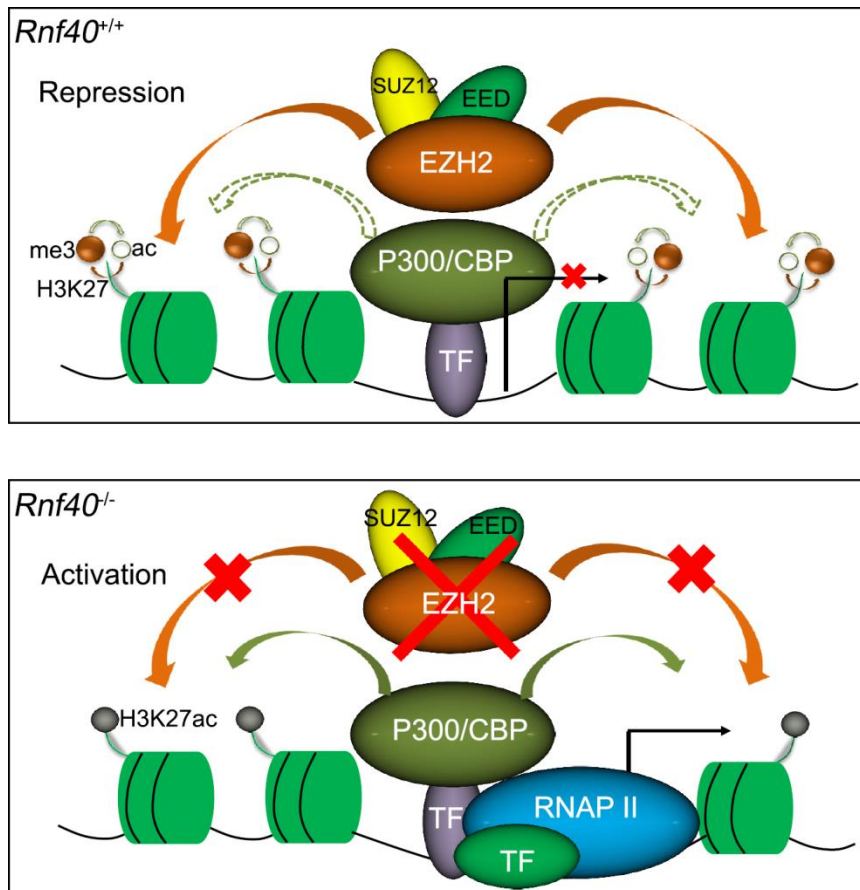
Fig. 42 Model depicting how RNF40 indirectly monitors the dynamic antagonism between H3K27me3 and H3K27ac

PRC2 complex is required to maintain the unique poised state. The enzymatic activity of P300 is restricted by PRC2-mediated H3K27me3. RNF40-mediated H2Bub1 selectively controls EZH2 transcription. Loss of H3K27me3 resulted in increased P300 activity and H3K27ac level, thereby promoting gene transcription in *Rnf40*[-/-] MEFs.

Given the essential role of EZH2 in cell fate determination and proper development (Aloia et al. 2013), we further determined that RNF40-supressed genes were enriched for developmental regulators, further supporting a critical function of RNF40 in directing cell fate decision. Consistent with a context-dependent function of H2Bub1 in regulating different groups of genes, we previously demonstrated a

central role for RNF20/40-dependent H2B monoubiquitination in differentiation to the osteoblast- and adipocyte lineages (Karpiuk et al. 2012), while another group reported that H2Bub1 levels decrease during myoblast differentiation (Vethantham et al. 2012). Thus, it is possible that the RNF20/40-H2Bub1 pathway may promote cell differentiation to one lineage and suppress that of another lineage in a given epigenetic context while promoting differentiation to other lineages in a different context.

5.5 H2Bub1 and Bivalent domain

Early studies in mouse ESC cells identified a unique chromatin state, which is decorated both by active H3K4me3 and repressive H3K27me3 and referred as to "bivalency". According to "The transition-state stabilization" theory (Pauling 1946), there is a transition-state between substrates and products in enzyme-promoted biology reactions. Transcriptional activation and repression is comprehensively controlled by multiple enzymes and epigenetic factors. In a wider meaning, bivalency could be considered as a metastable-state between transcriptional activation and repression. Changes in either H3K4me3 or H3K27me3 can remarkably affect bivalent gene transcription. We observed some bivalent genes transcribed in varying degrees, while others were fully silenced. We further confirmed that the activation of bivalency is associated with the shifting of the equilibrium of repression towards activation. In addition, genome-wide mapping of H2Bub1 on the bivalent genes revealed that H2Bub1 is an important factor in shifting the metastable equilibrium and correlated to a higher degree to bivalent gene transcription than H3K4me3.

H2Bub1 plays a dual function to switch on bivalent gene transcription. H2Bub1-H3K4me3 crosstalk can further be confirmed on bivalent genes. Loss of H2Bub1

leads H3K4me3 occupancy on H2Bub1 enriched (+) bivalent genes decrease to the similar degree of H2Bub1 unenriched (-) bivalent genes. It was suggested that the SWI/SNF chromatin remodeling complex binds preferentially to H2Bub1-rich chromatin (Shema-Yaacoby et al. 2013) and functions as an epigenetic antagonist of Polycomb complex (Wilson et al. 2010)., We observed lower signals of H3K27me3 on H2Bub1$^+$ bivalent genes compared to H2Bub1$^-$ bivalent genes. Moreover, additional analysis of BRG1, H3K4me3, and H3K27me3 ChIP-seq data in mouse ESC cells revealed a subset of bivalent genes occupied by the SWI/SNF complex (Min et al. 2011; de Dieuleveult et al. 2016). In addition, the decreased level of H3K27me3 on H2Bub1$^+$ bivalent genes could be related to the recruitment of H3K27 demethylase UTX. UTX has been identified in multiple H3K4 methyltransferase complexes such as MLL1/MLL2 and MLL3/MLL4 (Cho et al. 2007; Issaeva et al. 2007; Patel et al. 2007). Moreover, UTX was also observed on a subset of bivalent genes (Voigt et al. 2013). Depletion of UTX led to a significant decrease of bivalent gene transcription and increased H3K27me3 occupancy (Agger et al. 2007). Thus, H2Bub1 is also capable of inducing the demethylation of H3K27me3 on bivalency via facilitating the recruitment of UTX-contained H3K4 methyltransferase complexes.

In agreement with the crosstalk between H3K4me3 and H3 acetylation (Zhang 2006; Karmodiya et al. 2012), we observed a significant decrease of H3K27ac on only H3K4me3 occupied genes in respond to H2Bub1 deletion. In contrast, H3K27ac acted in an inverse manner on bivalent genes (either H2Bub1$^+$ or H2Bub1$^-$ bivalent genes). Notably, the metastable equilibrium between H3K27ac and H3K27me3 is another important effect on bivalent activity as discussed above. In fact, increased H3K27ac is directly associated to the decreasing of H3K27me3. Thus, we suggest that H3K27me3 is the key limitation on histone acetylation on bivalent domain.

To summarize, we mapped a potential regulatory network on controlling bivalent activity, in which H2Bub1 plays multiple roles in modulating bivalent activity. The activation of bivalent genes is a cooperative process ([Fig. 43](#)).
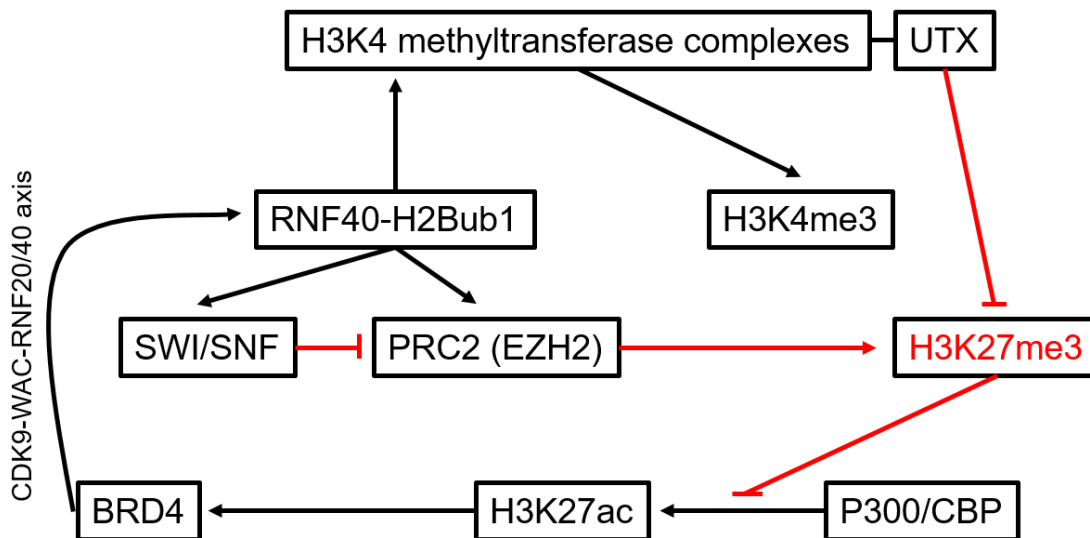
Fig. 43 The regulatory network of H2Bub1 on bivalent domain

RNF40-directed H2Bub1 modulates bivalent activity through multiple functions: i) Facilitates the recruitment of the SWI/SNF chromatin remodeling complex, ii) Promotes methylation on H3K4 via activating multiple H3K4 methyltransferase complexes containing UTX, iii) Monitors H3K27me3 occupancy by controlling *Ezh2* transcription (the methyltransferase in PRC2 complex). The silencing of bivalent genes is carried out by H3K27me3, which has abilities of restricting the enzymatic activity of p300, blocking recruitment of transcriptional co-activators, and limiting histone exchange. The abolishment of H3K27me3 is the key limitation of activating bivalent genes, which could be directly or indirectly achieved by H2Bub1. The recruitment of H3K4 methyltransferase complexes which constitute UTX directly mediates demethylation of H3K27me3. On the other hand, the SWI/SNF complex recruited by H2Bub1 antagonizes PRC2 activity. Therefore, the abolishment of H3K27me3 further liberates the HAT activity of p300/CBP complex to elevate histone

acetylation. Furthermore, readers of histone acetylation such as BRD4 promote H2B monoubiquitination via CDK9-WAC-RNF20/RNF40 axis.

5.6 Does the gene body-specific H2Bub1 affect enhancer activity?

H3K4me3 has been recently identified as an additional active chromatin feature at enhancers, which is essential for cell fate determination and tumour stimulation (Clouaire et al. 2012; Austenaa et al. 2015; Shen et al. 2016). Release of SET1 complex from CpG enriched promoters by depleting the CxxC domain-containing CFP1 and WDR82 results in overactivation of enhancers characterized by the deposition of H3K4me3 and histone acetylation (Clouaire et al. 2014; Austenaa et al. 2015). Thus, the H2Bub1-H3K4me3 crosstalk via interaction with WDR82 led us to believe that gene body-specific H2Bub1 could function at enhancers as WDR82 does. Indeed, loss of RNF40-mediated H2Bub1 significantly elevates the bulk of H3K27ac at enhancers. However, we didn't observe any appearance of H3K4me3 at those enhancers in $Rnf40^{-/-}$ MEFs. Considering the different cell type, the deposition of H3K4me3 at those enhancers would possibly require potential machinery which is undetectable in MEFs.

Excitingly, we observed the overactivation of enhancers displaying increased H3K27ac level was highly correlated to the upregulation of RNF40-suppressed genes. Activation of enhancers is required for the recruitment of multiple transcription factors such as the pioneer proteins. We further determined FOXL2, whose expression was increased following $Rnf40$ deletion, was responsible for overactivating those enhancers. Thus, these findings indicate upregulation of genes in response to $Rnf40$ deletion depends on the activity of other epigenetic regulatory proteins.

In conclusion, we provide evidence and insight into the apparent discrepancy between the association of H2Bub1 with active gene transcription and the unexpected finding that a nearly equal fraction of genes become up- or downregulated following its loss. Our results support a model in which the direct function of RNF40 and H2Bub1 lies primarily in transcriptional activation. However, given the finding that the *Ezh2* gene is a major target of RNF40 and H2Bub1, and the demonstration that the effects of *Rnf40* deletion on these "H2Bub1-suppressed" genes can be mimicked by inhibition of EZH2 catalytic activity, support a model in which "suppression" of gene transcription by H2Bub1 is mediated via indirect effects through PRC2. These findings, together with our results supporting a role for H2Bub1 in controlling H3K4me3 on RNF40-dependent genes, provide an important step in the elucidation of the enigmatic role of H2Bub1 in transcription. Further studies examining the effects of Rnf40 deletion in additional cell types and tissues, in conjunction with in vivo disease models, will shed further light into the biological and mechanistic functions of H2Bub1 and further elucidate its context-dependent function.

6. Reference

Agarwal SK, Jothi R. 2012. Genome-wide characterization of menin-dependent H3K4me3 reveals a specific role for menin in the regulation of genes implicated in MEN1-like tumors. PloS one 7(5): e37952.

Agger K, Cloos PA, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini AE, Helin K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. Nature 449(7163): 731-734.

Akhtar-Zaidi B, Cowper-Sal-lari R, Corradin O, Saiakhova A, Bartels CF, Balasubramanian D, Myeroff L, Lutterbaugh J, Jarrar A, Kalady MF et al. 2012. Epigenomic enhancer profiling defines a signature of colon cancer. Science 336(6082): 736-739.

Aloia L, Di Stefano B, Di Croce L. 2013. Polycomb complexes in stem cells and embryonic development. Development 140(12): 2525-2534.

Anders S, Huber W. 2010. Differential expression analysis for sequence count data. Genome biology 11(10): R106.

Ardehali MB, Mei A, Zobeck KL, Caron M, Lis JT, Kusch T. 2011. Drosophila Set1 is the major histone H3 lysine 4 trimethyltransferase with role in transcription. The EMBO journal 30(14): 2817-2828.

Austenaa LM, Barozzi I, Simonatto M, Masella S, Della Chiara G, Ghisletti S, Curina A, de Wit E, Bouwman BA, de Pretis S et al. 2015. Transcription of Mammalian cis-Regulatory Elements Is Restrained by Actively Enforced Early Termination. Molecular cell 60(3): 460-474.

Bapat SA, Jin V, Berry N, Balch C, Sharma N, Kurrey N, Zhang S, Fang F, Lan X, Li M et al. 2010. Multivalent epigenetic marks confer microenvironment-responsive epigenetic plasticity to ovarian cancer cells. Epigenetics 5(8): 716-729.

Bataille AR, Jeronimo C, Jacques PE, Laramee L, Fortin ME, Forest A, Bergeron M, Hanes SD, Robert F. 2012. A universal RNA polymerase II CTD cycle is orchestrated by complex interplays between kinase, phosphatase, and isomerase enzymes along genes. Molecular cell 45(2): 158-170.

Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D. 2003. FACT facilitates transcription-dependent nucleosome alteration. Science 301(5636): 1090-1093.

Benayoun BA, Pollina EA, Ucar D, Mahmoudi S, Karra K, Wong ED, Devarajan K, Daugherty AC, Kundaje AB, Mancini E et al. 2014. H3K4me3 breadth is linked to cell identity and transcriptional consistency. Cell 158(3): 673-688.

Bentley DL. 2014. Coupling mRNA processing with transcription in time and space. Nature reviews Genetics 15(3): 163-175.

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR et al. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. Cell 120(2): 169-181.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125(2): 315-326.

Bisgrove DA, Mahmoudi T, Henklein P, Verdin E. 2007. Conserved P-TEFb-interacting domain of BRD4 inhibits HIV transcription. Proceedings of the National Academy of Sciences of the United States of America 104(34): 13690-13695.

Blazek D, Kohoutek J, Bartholomeeusen K, Johansen E, Hulinkova P, Luo Z, Cimermancic P, Ule J, Peterlin BM. 2011. The Cyclin K/Cdk12 complex maintains genomic stability via regulation of expression of DNA damage response genes. Genes & development 25(20): 2158-2172.

Bogdanovic O, Fernandez-Minan A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, van Kruysbergen I, van Heeringen SJ, Veenstra GJ, Gomez-Skarmeta JL. 2012. Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis. Genome research 22(10): 2043-2053.

Bonnet J, Wang CY, Baptista T, Vincent SD, Hsiao WC, Stierle M, Kao CF, Tora L, Devys D. 2014. The SAGA coactivator complex acts on the whole transcribed genome and is required for RNA polymerase II transcription. Genes & development 28(18): 1999-2012.

Boulard M, Edwards JR, Bestor TH. 2015. FBXL10 protects Polycomb-bound genes from hypermethylation. Nature genetics 47(5): 479-485.

Buro LJ, Chipumuro E, Henriksen MA. 2010. Menin and RNF20 recruitment is associated with dynamic histone modifications that regulate signal transducer and activator of transcription 1 (STAT1)-activated transcription of the interferon regulatory factor 1 gene (IRF1). Epigenetics & chromatin 3(1): 16.

Calo E, Wysocka J. 2013. Modification of enhancer chromatin: what, how, and why? Molecular cell 49(5): 825-837.

Cao R, Tsukada Y, Zhang Y. 2005. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Molecular cell 20(6): 845-854.

Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298(5595): 1039-1043.

Caterino TL, Hayes JJ. 2007. Chromatin structure depends on what's in the nucleosome's pocket. Nature structural & molecular biology 14(11): 1056-1058.

Chandrasekharan MB, Huang F, Sun ZW. 2009. Ubiquitination of histone H2B regulates chromatin dynamics by enhancing nucleosome stability. Proceedings of the National Academy of Sciences of the United States of America 106(39): 16686-16691.

Chapman RD, Heidemann M, Albert TK, Mailhammer R, Flatley A, Meisterernst M, Kremmer E, Eick D. 2007. Transcribing RNA polymerase II is phosphorylated at CTD residue serine-7. Science 318(5857): 1780-1782.

Chen FX, Woodfin AR, Gardini A, Rickels RA, Marshall SA, Smith ER, Shiekhattar R, Shilatifard A. 2015a. PAF1, a Molecular Regulator of Promoter-Proximal Pausing by RNA Polymerase II. Cell 162(5): 1003-1015.

Chen K, Chen Z, Wu D, Zhang L, Lin X, Su J, Rodriguez B, Xi Y, Xia Z, Chen X et al. 2015b. Broad H3K4me3 is associated with increased transcription elongation and enhancer activity at tumor-suppressor genes. Nature genetics 47(10): 1149-1157.

Chen S, Li J, Wang DL, Sun FL. 2012. Histone H2B lysine 120 monoubiquitination is required for embryonic stem cell differentiation. Cell research 22(9): 1402-1405.

Chiaromonte F, Miller W, Bouhassira EE. 2003. Gene length and proximity to neighbors affect genome-wide expression levels. Genome research 13(12): 2602-2608.

Chipumuro E, Henriksen MA. 2012. The ubiquitin hydrolase USP22 contributes to 3'-end processing of JAK-STAT-inducible genes. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 26(2): 842-854.

Cho YW, Hong T, Hong S, Guo H, Yu H, Kim D, Guszczynski T, Dressler GR, Copeland TD, Kalkum M et al. 2007. PTIP associates with MLL3- and MLL4-containing histone H3 lysine 4 methyltransferase complex. The Journal of biological chemistry 282(28): 20395-20406.

Ciccia A, Elledge SJ. 2010. The DNA damage response: making it safe to play with knives. Molecular cell 40(2): 179-204.

Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. Molecular cell 9(2): 279-289.

Clouaire T, Webb S, Bird A. 2014. Cfp1 is required for gene expression-dependent H3K4 trimethylation and H3K9 acetylation in embryonic stem cells. Genome biology 15(9): 451.

Clouaire T, Webb S, Skene P, Illingworth R, Kerr A, Andrews R, Lee JH, Skalnik D, Bird A. 2012. Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes & development 26(15): 1714-1728.

Cole AJ, Clifton-Bligh R, Marsh DJ. 2015. Histone H2B monoubiquitination: roles to play in human malignancy. Endocrine-related cancer 22(1): T19-33.

Coudreuse D, Nurse P. 2010. Driving the cell cycle with a minimal CDK control network. Nature 468(7327): 1074-1079.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA et al. 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America 107(50): 21931-21936.

Daniel JA, Grant PA. 2007. Multi-tasking on chromatin with the SAGA coactivator complexes. Mutation research 618(1-2): 135-148.

de Almeida SF, Grosso AR, Koch F, Fenouil R, Carvalho S, Andrade J, Levezinho H, Gut M, Eick D, Gut I et al. 2011. Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. Nature structural & molecular biology 18(9): 977-983.

de Dieuleveult M, Yen K, Hmitou I, Depaux A, Boussouar F, Bou Dargham D, Jounier S, Humbertclaude H, Ribierre F, Baulard C et al. 2016. Genome-wide nucleosome specificity and function of chromatin remodellers in ES cells. Nature 530(7588): 113-116.

Devaiah BN, Lewis BA, Cherman N, Hewitt MC, Albrecht BK, Robey PG, Ozato K, Sims RJ, 3rd, Singer DS. 2012. BRD4 is an atypical kinase that phosphorylates serine2 of the RNA polymerase II carboxy-terminal domain. Proceedings of the National Academy of Sciences of the United States of America 109(18): 6927-6932.

Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. 1999. Structure and ligand of a histone acetyltransferase bromodomain. Nature 399(6735): 491-496.

Di Croce L, Helin K. 2013. Transcriptional regulation by Polycomb group proteins. Nature structural & molecular biology 20(10): 1147-1155.

Dincer A, Gavin DP, Xu K, Zhang B, Dudley JT, Schadt EE, Akbarian S. 2015. Deciphering H3K4me3 broad domains associated with gene-regulatory networks and conserved epigenomic landscapes in the human brain. Translational psychiatry 5: e679.

Egloff S, Murphy S. 2008. Cracking the RNA polymerase II CTD code. Trends in genetics : TIG 24(6): 280-288.

Eick D, Geyer M. 2013. The RNA polymerase II carboxy-terminal domain (CTD) code. Chemical reviews 113(11): 8456-8490.

El-Karim EA, Hagos EG, Ghaleb AM, Yu B, Yang VW. 2013. Kruppel-like factor 4 regulates genetic stability in mouse embryonic fibroblasts. Molecular cancer 12: 89.

Emre NC, Ingvarsdottir K, Wyce A, Wood A, Krogan NJ, Henry KW, Li K, Marmorstein R, Greenblatt JF, Shilatifard A et al. 2005. Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing. Molecular cell 17(4): 585-594.

Endoh M, Zhu W, Hasegawa J, Watanabe H, Kim DK, Aida M, Inukai N, Narita T, Yamada T, Furuya A et al. 2004. Human Spt6 stimulates transcription

elongation by RNA polymerase II in vitro. Molecular and cellular biology 24(8): 3324-3336.

Espinosa JM. 2008. Histone H2B ubiquitination: the cancer connection. Genes & development 22(20): 2743-2749.

Farley FW, Soriano P, Steffen LS, Dymecki SM. 2000. Widespread recombinase expression using FLPeR (flipper) mice. Genesis 28(3-4): 106-110.

Farre D, Bellora N, Mularoni L, Messeguer X, Alba MM. 2007. Housekeeping genes tend to show reduced upstream sequence conservation. Genome biology 8(7): R140.

Feaver WJ, Gileadi O, Li Y, Kornberg RD. 1991. CTD kinase associated with yeast RNA polymerase II initiation factor b. Cell 67(6): 1223-1230.

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. Nature 438(7071): 1116-1122.

Fleming AB, Kao CF, Hillyer C, Pikaart M, Osley MA. 2008. H2B ubiquitylation plays a role in nucleosome dynamics during transcription elongation. Molecular cell 31(1): 57-66.

Fragkos M, Ganier O, Coulombe P, Mechali M. 2015. DNA replication origin activation in space and time. Nature reviews Molecular cell biology 16(6): 360-374.

Fuchs G, Hollander D, Voichek Y, Ast G, Oren M. 2014. Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. Genome research 24(10): 1572-1583.

Fuchs G, Shema E, Vesterman R, Kotler E, Wolchinsky Z, Wilder S, Golomb L, Pribluda A, Zhang F, Haj-Yahya M et al. 2012. RNF20 and USP44 regulate stem cell differentiation by modulating H2B monoubiquitylation. Molecular cell 46(5): 662-673.

Fujinaga K, Irwin D, Huang Y, Taube R, Kurosu T, Peterlin BM. 2004. Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Molecular and cellular biology 24(2): 787-795.

Gardner RG, Nelson ZW, Gottschling DE. 2005. Ubp10/Dot4p regulates the persistence of ubiquitinated histone H2B: distinct roles in telomeric silencing and general chromatin. Molecular and cellular biology 25(14): 6123-6139.

Georges A, L'Hote D, Todeschini AL, Auguste A, Legois B, Zider A, Veitia RA. 2014. The transcription factor FOXL2 mobilizes estrogen signaling to maintain the identity of ovarian granulosa cells. eLife 3.

Glover-Cutter K, Larochelle S, Erickson B, Zhang C, Shokat K, Fisher RP, Bentley DL. 2009. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. Molecular and cellular biology 29(20): 5455-5464.

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA. 2007. A chromatin landmark and transcription initiation at most promoters in human cells. Cell 130(1): 77-88.

Hahn MA, Dickson KA, Jackson S, Clarkson A, Gill AJ, Marsh DJ. 2012. The tumor suppressor CDC73 interacts with the ring finger proteins RNF20 and RNF40 and is required for the maintenance of histone 2B monoubiquitination. Human molecular genetics 21(3): 559-568.

Hameyer D, Loonstra A, Eshkind L, Schmitt S, Antunes C, Groen A, Bindels E, Jonkers J, Krimpenfort P, Meuwissen R et al. 2007. Toxicity of ligand-dependent Cre recombinases and generation of a conditional Cre deleter mouse allowing mosaic recombination in peripheral tissues. Physiological genomics 31(1): 32-41.

Han Y, Luo J, Ranish J, Hahn S. 2014. Architecture of the Saccharomyces cerevisiae SAGA transcription coactivator complex. The EMBO journal 33(21): 2534-2546.

Hanks S, Perdeaux ER, Seal S, Ruark E, Mahamdallie SS, Murray A, Ramsay E, Del Vecchio Duarte S, Zachariou A, de Souza B et al. 2014. Germline mutations in the PAF1 complex gene CTR9 predispose to Wilms tumour. Nature communications 5: 4398.

Hawes MC, O'Brien JP. 2008. Scoliosis and the human genome project. Studies in health technology and informatics 135: 97-111.

He J, Shen L, Wan M, Taranova O, Wu H, Zhang Y. 2013. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. Nature cell biology 15(4): 373-384.

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature genetics 39(3): 311-318.

Henry KW, Wyce A, Lo WS, Duggan LJ, Emre NC, Kao CF, Pillus L, Shilatifard A, Osley MA, Berger SL. 2003. Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. Genes & development 17(21): 2648-2663.

Herz HM, Mohan M, Garruss AS, Liang K, Takahashi YH, Mickey K, Voets O, Verrijzer CP, Shilatifard A. 2012. Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. Genes & development 26(23): 2604-2620.

Hintermair C, Heidemann M, Koch F, Descostes N, Gut M, Gut I, Fenouil R, Ferrier P, Flatley A, Kremmer E et al. 2012. Threonine-4 of mammalian RNA polymerase II CTD is targeted by Polo-like kinase 3 and required for transcriptional elongation. The EMBO journal 31(12): 2784-2797.

Holmqvist PH, Mannervik M. 2013. Genomic occupancy of the transcriptional co-activators p300 and CBP. Transcription 4(1): 18-23.

Hondele M, Stuwe T, Hassler M, Halbach F, Bowman A, Zhang ET, Nijmeijer B, Kotthoff C, Rybin V, Amlacher S et al. 2013. Structural basis of histone H2A-H2B recognition by the essential chaperone FACT. Nature 499(7456): 111-114.

Hsin JP, Sheth A, Manley JL. 2011. RNAP II CTD phosphorylated on threonine-4 is required for histone mRNA 3' end processing. Science 334(6056): 683-686.

Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature protocols 4(1): 44-57.

Hwang WW, Venkatasubrahmanyam S, Ianculescu AG, Tong A, Boone C, Madhani HD. 2003. A conserved RING finger protein required for histone H2B monoubiquitination and cell size control. Molecular cell 11(1): 261-266.

Issaeva I, Zonis Y, Rozovskaia T, Orlovsky K, Croce CM, Nakamura T, Mazo A, Eisenbach L, Canaani E. 2007. Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Molecular and cellular biology 27(5): 1889-1903.

Iwasaki W, Miya Y, Horikoshi N, Osakabe A, Taguchi H, Tachiwana H, Shibata T, Kagawa W, Kurumizaka H. 2013. Contribution of histone N-terminal tails to the structure and stability of nucleosomes. FEBS open bio 3: 363-369.

Ji X, Li W, Song J, Wei L, Liu XS. 2006. CEAS: cis-regulatory element annotation system. Nucleic acids research 34(Web Server issue): W551-554.

Jiang T, Zhou X, Taghizadeh K, Dong M, Dedon PC. 2007. N-formylation of lysine in histone proteins as a secondary modification arising from oxidative DNA damage. Proceedings of the National Academy of Sciences of the United States of America 104(1): 60-65.

Johnsen SA. 2012. The enigmatic role of H2Bub1 in cancer. FEBS letters 586(11): 1592-1601.

Johnson DP, Spitz GS, Tharkar S, Quayle SN, Shearstone JR, Jones S, McDowell ME, Wellman H, Tyler JK, Cairns BR et al. 2015. HDAC1,2 inhibition impairs EZH2- and BBAP-mediated DNA repair to overcome chemoresistance in EZH2 gain-of-function mutant diffuse large B-cell lymphoma. Oncotarget 6(7): 4863-4887.

Joo HY, Jones A, Yang C, Zhai L, Smith ADt, Zhang Z, Chandrasekharan MB, Sun ZW, Renfrow MB, Wang Y et al. 2011. Regulation of histone H2A and H2B deubiquitination and Xenopus development by USP12 and USP46. The Journal of biological chemistry 286(9): 7190-7201.

Jung I, Kim SK, Kim M, Han YM, Kim YS, Kim D, Lee D. 2012. H2B monoubiquitylation is a 5'-enriched active transcription mark and correlates with exon-intron structure in human cells. Genome research 22(6): 1026-1035.

Kanno T, Kanno Y, LeRoy G, Campos E, Sun HW, Brooks SR, Vahedi G, Heightman TD, Garcia BA, Reinberg D et al. 2014. BRD4 assists elongation of both coding and enhancer RNAs by interacting with acetylated histones. Nature structural & molecular biology 21(12): 1047-1057.

Kari V, Shchebet A, Neumann H, Johnsen SA. 2011. The H2B ubiquitin ligase RNF40 cooperates with SUPT16H to induce dynamic changes in chromatin structure during DNA double-strand break repair. Cell cycle 10(20): 3495-3504.

Karlic R, Chung HR, Lasserre J, Vlahovicek K, Vingron M. 2010. Histone modification levels are predictive for gene expression. Proceedings of the National Academy of Sciences of the United States of America 107(7): 2926-2931.

Karmodiya K, Krebs AR, Oulad-Abdelghani M, Kimura H, Tora L. 2012. H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. BMC genomics 13: 424.

Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ. 2004. The UCSC Table Browser data retrieval tool. Nucleic acids research 32(Database issue): D493-496.

Karpiuk O, Najafova Z, Kramer F, Hennion M, Galonska C, Konig A, Snaidero N, Vogel T, Shchebet A, Begus-Nahrmann Y et al. 2012. The histone H2B monoubiquitination regulatory pathway is required for differentiation of multipotent stem cells. Molecular cell 46(5): 705-713.

Kim HE, Symanowski JT, Samlowski EE, Gonzales J, Ryu B. 2010a. Quantitative measurement of circulating lymphoid-specific helicase (HELLS) gene transcript: a potential serum biomarker for melanoma metastasis. Pigment cell & melanoma research 23(6): 845-848.

Kim J, Guermah M, McGinty RK, Lee JS, Tang Z, Milne TA, Shilatifard A, Muir TW, Roeder RG. 2009. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. Cell 137(3): 459-471.

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

Kim J, Hake SB, Roeder RG. 2005. The human homolog of yeast BRE1 functions as a transcriptional coactivator through direct activator interactions. Molecular cell 20(5): 759-770.

Kim J, Kim JA, McGinty RK, Nguyen UT, Muir TW, Allis CD, Roeder RG. 2013. The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. Molecular cell 49(6): 1121-1133.

Kim KH, Kim W, Howard TP, Vazquez F, Tsherniak A, Wu JN, Wang W, Haswell JR, Walensky LD, Hahn WC et al. 2015. SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of EZH2. Nature medicine 21(12): 1491-1496.

Kim SK, Jung I, Lee H, Kang K, Kim M, Jeong K, Kwon CS, Han YM, Kim YS, Kim D et al. 2012. Human histone H3K79 methyltransferase DOT1L protein [corrected] binds actively transcribing RNA polymerase II to regulate gene expression. The Journal of biological chemistry 287(47): 39698-39709.

Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. 2002. Nucleosome remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. Molecular cell 9(3): 541-552.

Knutson SK, Warholic NM, Wigle TJ, Klaus CR, Allain CJ, Raimondi A, Porter Scott M, Chesworth R, Moyer MP, Copeland RA et al. 2013. Durable tumor regression in genetically altered malignant rhabdoid tumors by inhibition of methyltransferase EZH2. Proceedings of the National Academy of Sciences of the United States of America 110(19): 7922-7927.

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128(4): 693-705.

Krogan NJ, Kim M, Ahn SH, Zhong G, Kobor MS, Cagney G, Emili A, Shilatifard A, Buratowski S, Greenblatt JF. 2002. RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Molecular and cellular biology 22(20): 6979-6992.

Kwon AT, Arenillas DJ, Worsley Hunt R, Wasserman WW. 2012. oPOSSUM-3: advanced analysis of regulatory motif over-representation across genes or ChIP-Seq datasets. G3 2(9): 987-1002.

Laemmli UK. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227(5259): 680-685.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nature methods 9(4): 357-359.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology 10(3): R25.

Laribee RN, Krogan NJ, Xiao T, Shibata Y, Hughes TR, Greenblatt JF, Strahl BD. 2005. BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. Current biology : CB 15(16): 1487-1493.

Larochelle S, Amat R, Glover-Cutter K, Sanso M, Zhang C, Allen JJ, Shokat KM, Bentley DL, Fisher RP. 2012. Cyclin-dependent kinase control of the initiation-to-elongation switch of RNA polymerase II. Nature structural & molecular biology 19(11): 1108-1115.

Lee HG, Kahn TG, Simcox A, Schwartz YB, Pirrotta V. 2015a. Genome-wide activities of Polycomb complexes control pervasive transcription. Genome research 25(8): 1170-1181.

Lee JH, Skalnik DG. 2005. CpG-binding protein (CXXC finger protein 1) is a component of the mammalian Set1 histone H3-Lys4 methyltransferase complex, the analogue of the yeast Set1/COMPASS complex. The Journal of biological chemistry 280(50): 41725-41731.

Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, Shilatifard A. 2007. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. Cell 131(6): 1084-1096.

Li XS, Trojer P, Matsumura T, Treisman JE, Tanese N. 2010. Mammalian SWI/SNF--a subunit BAF250/ARID1 is an E3 ubiquitin ligase that targets histone H2B. Molecular and cellular biology 30(7): 1673-1688.

Li Z, Gadue P, Chen K, Jiao Y, Tuteja G, Schug J, Li W, Kaestner KH. 2012. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. Cell 151(7): 1608-1616.

Lindeman LC, Andersen IS, Reiner AH, Li N, Aanes H, Ostrup O, Winata C, Mathavan S, Muller F, Alestrom P et al. 2011. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. Developmental cell 21(6): 993-1004.

Liu L, Xu Y, He M, Zhang M, Cui F, Lu L, Yao M, Tian W, Benda C, Zhuang Q et al. 2014. Transcriptional pause release is a rate-limiting step for somatic cell reprogramming. Cell stem cell 15(5): 574-588.

Liu T. 2014. Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells. Methods in molecular biology 1150: 81-95.

Liu Z, Oh SM, Okada M, Liu X, Cheng D, Peng J, Brat DJ, Sun SY, Zhou W, Gu W et al. 2009. Human BRE1 is an E3 ubiquitin ligase for Ebp1 tumor suppressor. Molecular biology of the cell 20(3): 757-768.

Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. 2006. Interchromosomal interactions and olfactory receptor choice. Cell 126(2): 403-413.

Lu H, Zawel L, Fisher L, Egly JM, Reinberg D. 1992. Human general transcription factor IIH phosphorylates the C-terminal domain of RNA polymerase II. Nature 358(6388): 641-645.

Ma MK, Heath C, Hair A, West AG. 2011. Histone crosstalk directed by H2B ubiquitination is required for chromatin boundary integrity. PLoS genetics 7(7): e1002175.

Magklara A, Yen A, Colquitt BM, Clowney EJ, Allen W, Markenscoff-Papadimitriou E, Evans ZA, Kheradpour P, Mountoufaris G, Carey C et al. 2011. An epigenetic signature for monoallelic olfactory receptor expression. Cell 145(4): 555-570.

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

Marshall NF, Price DH. 1992. Control of formation of two distinct classes of RNA polymerase II elongation complexes. Molecular and cellular biology 12(5): 2078-2090.

Mbogning J, Nagy S, Page V, Schwer B, Shuman S, Fisher RP, Tanny JC. 2013. The PAF complex and Prf1/Rtf1 delineate distinct Cdk9-dependent pathways regulating transcription elongation in fission yeast. PLoS genetics 9(12): e1004029.

McGarvey KM, Van Neste L, Cope L, Ohm JE, Herman JG, Van Criekinge W, Schuebel KE, Baylin SB. 2008. Defining a chromatin pattern that characterizes DNA-hypermethylated genes in colon cancer cells. Cancer research 68(14): 5753-5759.

McGinty RK, Kim J, Chatterjee C, Roeder RG, Muir TW. 2008. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature 453(7196): 812-816.

McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. 2010. GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology 28(5): 495-501.

McPherson CE, Shim EY, Friedman DS, Zaret KS. 1993. An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. Cell 75(2): 387-398.

Messner S, Hottiger MO. 2011. Histone ADP-ribosylation in DNA repair, replication and transcription. Trends in cell biology 21(9): 534-542.

Metzger MB, Hristova VA, Weissman AM. 2012. HECT and RING finger families of E3 ubiquitin ligases at a glance. Journal of cell science 125(Pt 3): 531-537.

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP et al. 2007. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448(7153): 553-560.

Min IM, Waterfall JJ, Core LJ, Munroe RJ, Schimenti J, Lis JT. 2011. Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. Genes & development 25(7): 742-754.

Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M. 2008. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. Nature cell biology 10(4): 483-488.

Morey L, Pascual G, Cozzuto L, Roma G, Wutz A, Benitah SA, Di Croce L. 2012. Nonoverlapping functions of the Polycomb group Cbx family of proteins in embryonic stem cells. Cell stem cell 10(1): 47-62.

Morgan MT, Haj-Yahya M, Ringel AE, Bandi P, Brik A, Wolberger C. 2016. Structural basis for histone H2B deubiquitination by the SAGA DUB module. Science 351(6274): 725-728.

Nagarajan S, Hossan T, Alawi M, Najafova Z, Indenbirken D, Bedi U, Taipaleenmaki H, Ben-Batalla I, Scheller M, Loges S et al. 2014. Bromodomain protein BRD4 is required for estrogen receptor-dependent enhancer activation and gene transcription. Cell reports 8(2): 460-469.

Ng HH, Xu RM, Zhang Y, Struhl K. 2002. Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. The Journal of biological chemistry 277(38): 34655-34657.

Nicassio F, Corrado N, Vissers JH, Areces LB, Bergink S, Marteijn JA, Geverts B, Houtsmuller AB, Vermeulen W, Di Fiore PP et al. 2007. Human USP3 is a chromatin modifier required for S phase progression and genome stability. Current biology : CB 17(22): 1972-1977.

Ntziachristos P, Tsirigos A, Welstead GG, Trimarchi T, Bakogianni S, Xu L, Loizou E, Holmfeldt L, Strikoudis A, King B et al. 2014. Contrasting roles of histone 3 lysine 27 demethylases in acute lymphoblastic leukaemia. Nature 514(7523): 513-517.

O'Loghlen A, Munoz-Cabello AM, Gaspar-Maia A, Wu HA, Banito A, Kunowska N, Racek T, Pemberton HN, Beolchi P, Lavial F et al. 2012. MicroRNA regulation of Cbx7 mediates a switch of Polycomb orthologs during ESC differentiation. Cell stem cell 10(1): 33-46.

Ogawa H, Ishiguro K, Gaubatz S, Livingston DM, Nakatani Y. 2002. A complex with chromatin modifiers that occupies E2F- and Myc-responsive genes in G0 cells. Science 296(5570): 1132-1136.

Ong CT, Corces VG. 2011. Enhancer function: new insights into the regulation of tissue-specific gene expression. Nature reviews Genetics 12(4): 283-293.

Pasini D, Malatesta M, Jung HR, Walfridsson J, Willer A, Olsson L, Skotte J, Wutz A, Porse B, Jensen ON et al. 2010. Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. Nucleic acids research 38(15): 4958-4969.

Patel MC, Debrosse M, Smith M, Dey A, Huynh W, Sarai N, Heightman TD, Tamura T, Ozato K. 2013. BRD4 coordinates recruitment of pause release factor P-TEFb and the pausing complex NELF/DSIF to regulate transcription elongation of interferon-stimulated genes. Molecular and cellular biology 33(12): 2497-2507.

Patel SR, Kim D, Levitan I, Dressler GR. 2007. The BRCT-domain containing protein PTIP links PAX2 to a histone H3, lysine 4 methyltransferase complex. Developmental cell 13(4): 580-592.

Pauling L. 1946. Molecular Architecture and Biological Reactions. Chem Eng News 24(24): 1375-1377.

Pavri R, Zhu B, Li G, Trojer P, Mandal S, Shilatifard A, Reinberg D. 2006. Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. Cell 125(4): 703-717.

Pekowska A, Benoukraf T, Ferrier P, Spicuglia S. 2010. A unique H3K4me2 profile marks tissue-specific gene regulation. Genome research 20(11): 1493-1502.

Peng JC, Valouev A, Swigut T, Zhang J, Zhao Y, Sidow A, Wysocka J. 2009. Jarid2/Jumonji coordinates control of PRC2 enzymatic activity and target gene occupancy in pluripotent cells. Cell 139(7): 1290-1302.

Pirngruber J, Shchebet A, Schreiber L, Shema E, Minsky N, Chapman RD, Eick D, Aylon Y, Oren M, Johnsen SA. 2009. CDK9 directs H2B monoubiquitination and controls replication-dependent histone mRNA 3'-end processing. EMBO reports 10(8): 894-900.

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E et al. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122(4): 517-527.

Racine A, Page V, Nagy S, Grabowski D, Tanny JC. 2012. Histone H2B ubiquitylation promotes activity of the intact Set1 histone methyltransferase complex in fission yeast. The Journal of biological chemistry 287(23): 19040-19047.

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, Wysocka J. 2011. A unique chromatin signature uncovers early developmental enhancers in humans. Nature 470(7333): 279-283.

Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. 2014. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic acids research 42(Web Server issue): W187-191.

Robzyk K, Recht J, Osley MA. 2000. Rad6-dependent ubiquitination of histone H2B in yeast. Science 287(5452): 501-504.

Rodriguez-Navarro S. 2009. Insights into SAGA function during gene expression. EMBO reports 10(8): 843-850.

Roh TY, Cuddapah S, Cui K, Zhao K. 2006. The genomic landscape of histone modifications in human T cells. Proceedings of the National Academy of Sciences of the United States of America 103(43): 15782-15787.

Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR et al. 2012. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481(7381): 389-393.

Roth SY, Allis CD. 1992. Chromatin condensation: does histone H1 dephosphorylation play a role? Trends in biochemical sciences 17(3): 93-98.

Sachs M, Onodera C, Blaschke K, Ebata KT, Song JS, Ramalho-Santos M. 2013. Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo. Cell reports 3(6): 1777-1784.

Sadeghi L, Siggens L, Svensson JP, Ekwall K. 2014. Centromeric histone H2B monoubiquitination promotes noncoding transcription and chromatin integrity. Nature structural & molecular biology 21(3): 236-243.

Samara NL, Datta AB, Berndsen CE, Zhang X, Yao T, Cohen RE, Wolberger C. 2010. Structural insights into the assembly and function of the SAGA deubiquitinating module. Science 328(5981): 1025-1029.

Samara NL, Ringel AE, Wolberger C. 2012. A role for intersubunit interactions in maintaining SAGA deubiquitinating module structure and activity. Structure 20(8): 1414-1424.

Sarkari F, Sanchez-Alcaraz T, Wang S, Holowaty MN, Sheng Y, Frappier L. 2009. EBNA1-mediated recruitment of a histone H2B deubiquitylating complex to the Epstein-Barr virus latent origin of DNA replication. PLoS pathogens 5(10): e1000624.

Sarma K, Margueron R, Ivanov A, Pirrotta V, Reinberg D. 2008. Ezh2 requires PHF1 to efficiently catalyze H3 lysine 27 trimethylation in vivo. Molecular and cellular biology 28(8): 2718-2731.

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. Cell 132(5): 887-898.

Schones DE, Zhao K. 2008. Genome-wide approaches to studying chromatin modifications. Nature reviews Genetics 9(3): 179-191.

Schreiber SL, Bernstein BE. 2002. Signaling network model of chromatin. Cell 111(6): 771-778.

Schubeler D, MacAlpine DM, Scalzo D, Wirbelauer C, Kooperberg C, van Leeuwen F, Gottschling DE, O'Neill LP, Turner BM, Delrow J et al. 2004. The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. Genes & development 18(11): 1263-1271.

Schulman BA, Harper JW. 2009. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nature reviews Molecular cell biology 10(5): 319-331.

Schulte A, Czudnochowski N, Barboric M, Schonichen A, Blazek D, Peterlin BM, Geyer M. 2005. Identification of a cyclin T-binding domain in Hexim1 and biochemical analysis of its binding competition with HIV-1 Tat. The Journal of biological chemistry 280(26): 24968-24977.

Schulze JM, Hentrich T, Nakanishi S, Gupta A, Emberly E, Shilatifard A, Kobor MS. 2011. Splitting the task: Ubp8 and Ubp10 deubiquitinate different cellular pools of H2BK123. Genes & development 25(21): 2242-2247.

Schwer B, Shuman S. 2011. Deciphering the RNA polymerase II CTD code in fission yeast. Molecular cell 43(2): 311-318.

Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-dimensional folding and functional organization principles of the Drosophila genome. Cell 148(3): 458-472.

Shchebet A, Karpiuk O, Kremmer E, Eick D, Johnsen SA. 2012. Phosphorylation by cyclin-dependent kinase-9 controls ubiquitin-conjugating enzyme-2A function. Cell cycle 11(11): 2122-2127.

Shema-Yaacoby E, Nikolov M, Haj-Yahya M, Siman P, Allemand E, Yamaguchi Y, Muchardt C, Urlaub H, Brik A, Oren M et al. 2013. Systematic identification of proteins binding to chromatin-embedded ubiquitylated H2B reveals recruitment of SWI/SNF to regulate transcription. Cell reports 4(3): 601-608.

Shema E, Kim J, Roeder RG, Oren M. 2011. RNF20 inhibits TFIIS-facilitated transcriptional elongation to suppress pro-oncogenic gene expression. Molecular cell 42(4): 477-488.

Shema E, Tirosh I, Aylon Y, Huang J, Ye C, Moskovits N, Raver-Shapira N, Minsky N, Pirngruber J, Tarcic G et al. 2008. The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. Genes & development 22(19): 2664-2676.

Shen H, Xu W, Guo R, Rong B, Gu L, Wang Z, He C, Zheng L, Hu X, Hu Z et al. 2016. Suppression of Enhancer Overactivation by a RACK7-Histone Demethylase Complex. Cell 165(2): 331-342.

Shen X, Kim W, Fujiwara Y, Simon MD, Liu Y, Mysliwiec MR, Yuan GC, Lee Y, Orkin SH. 2009. Jumonji modulates polycomb activity and self-renewal versus differentiation of stem cells. Cell 139(7): 1303-1314.

Shilatifard A. 2006. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annual review of biochemistry 75: 243-269.

-. 2012. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual review of biochemistry 81: 65-95.

Shin H, Liu T, Manrai AK, Liu XS. 2009. CEAS: cis-regulatory element annotation system. Bioinformatics 25(19): 2605-2606.

Smith E, Lin C, Shilatifard A. 2011. The super elongation complex (SEC) and MLL in development and disease. Genes & development 25(7): 661-672.

Smolich BD, McMahon JA, McMahon AP, Papkoff J. 1993. Wnt family proteins are secreted and associated with the cell surface. Molecular biology of the cell 4(12): 1267-1275.

Spitz F, Furlong EE. 2012. Transcription factors: from enhancer binding to developmental control. Nature reviews Genetics 13(9): 613-626.

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

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 102(43): 15545-15550.

Sun ZW, Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature 418(6893): 104-108.

Supek F, Bosnjak M, Skunca N, Smuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. PloS one 6(7): e21800.

Svensson JP, Shukla M, Menendez-Benito V, Norman-Axelsson U, Audergon P, Sinha I, Tanny JC, Allshire RC, Ekwall K. 2015. A nucleosome turnover map reveals that the stability of histone H4 Lys20 methylation depends on histone recycling in transcribed chromatin. Genome research 25(6): 872-883.

Takahashi K, Yamanaka S. 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126(4): 663-676.

Tanaka TS, Kunath T, Kimber WL, Jaradat SA, Stagg CA, Usuda M, Yokota T, Niwa H, Rossant J, Ko MS. 2002. Gene expression profiling of embryo-derived stem cells reveals candidate genes associated with pluripotency and lineage specificity. Genome research 12(12): 1921-1928.

Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in bioinformatics 14(2): 178-192.

Tie F, Banerjee R, Conrad PA, Scacheri PC, Harte PJ. 2012. Histone demethylase UTX and chromatin remodeler BRM bind directly to CBP and modulate acetylation of histone H3 lysine 27. Molecular and cellular biology 32(12): 2323-2334.

Tie F, Banerjee R, Saiakhova AR, Howard B, Monteith KE, Scacheri PC, Cosgrove MS, Harte PJ. 2014. Trithorax monomethylates histone H3K4 and interacts directly with CBP to promote H3K27 acetylation and antagonize Polycomb silencing. Development 141(5): 1129-1139.

Tie F, Banerjee R, Stratton CA, Prasad-Sinha J, Stepanik V, Zlobin A, Diaz MO, Scacheri PC, Harte PJ. 2009. CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. Development 136(18): 3131-3141.

Trujillo KM, Osley MA. 2012. A role for H2B ubiquitylation in DNA replication. Molecular cell 48(5): 734-746.

Turner BM. 2000. Histone acetylation and an epigenetic code. BioEssays : news and reviews in molecular, cellular and developmental biology 22(9): 836-845.

van Wijk SJ, Timmers HT. 2010. The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. FASEB journal : official publication of the Federation of American Societies for Experimental Biology 24(4): 981-993.

Vethantham V, Yang Y, Bowman C, Asp P, Lee JH, Skalnik DG, Dynlacht BD. 2012. Dynamic loss of H2B ubiquitylation without corresponding changes in H3K4 trimethylation during myogenic differentiation. Molecular and cellular biology 32(6): 1044-1055.

Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM et al. 2006. The Polycomb group protein EZH2 directly controls DNA methylation. Nature 439(7078): 871-874.

Vitaliano-Prunier A, Menant A, Hobeika M, Geli V, Gwizdek C, Dargemont C. 2008. Ubiquitylation of the COMPASS component Swd2 links H2B ubiquitylation to H3K4 trimethylation. Nature cell biology 10(11): 1365-1371.

Vlaming H, van Welsem T, de Graaf EL, Ontoso D, Altelaar AF, San-Segundo PA, Heck AJ, van Leeuwen F. 2014. Flexibility in crosstalk between H2B ubiquitination and H3 methylation in vivo. EMBO reports 15(10): 1077-1084.

Voigt P, LeRoy G, Drury WJ, 3rd, Zee BM, Son J, Beck DB, Young NL, Garcia BA, Reinberg D. 2012. Asymmetrically modified nucleosomes. Cell 151(1): 181-193.

Voigt P, Tee WW, Reinberg D. 2013. A double take on bivalent promoters. Genes & development 27(12): 1318-1338.

Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. 2004. Role of histone H2A ubiquitination in Polycomb silencing. Nature 431(7010): 873-878.

Wang P, Lin C, Smith ER, Guo H, Sanderson BW, Wu M, Gogol M, Alexander T, Seidel C, Wiedemann LM et al. 2009. 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. Molecular and cellular biology 29(22): 6074-6085.

Weake VM, Workman JL. 2008. Histone ubiquitination: triggering gene activity. Molecular cell 29(6): 653-663.

-. 2010. Inducible gene expression: diverse regulatory mechanisms. Nature reviews Genetics 11(6): 426-437.

Wei C, Qiu J, Zhou Y, Xue Y, Hu J, Ouyang K, Banerjee I, Zhang C, Chen B, Li H et al. 2015. Repression of the Central Splicing Regulator RBFox2 Is Functionally Linked to Pressure Overload-Induced Heart Failure. Cell reports.

Wei Y, Chen YH, Li LY, Lang J, Yeh SP, Shi B, Yang CC, Yang JY, Lin CY, Lai CC et al. 2011. CDK1-dependent phosphorylation of EZH2 suppresses methylation of H3K27 and promotes osteogenic differentiation of human mesenchymal stem cells. Nature cell biology 13(1): 87-94.

Wen H, Ao S. 2000. RBP95, a novel leucine zipper protein, binds to the retinoblastoma protein. Biochemical and biophysical research communications 275(1): 141-148.

Wilson BG, Wang X, Shen X, McKenna ES, Lemieux ME, Cho YJ, Koellhoffer EC, Pomeroy SL, Orkin SH, Roberts CW. 2010. Epigenetic antagonism between polycomb and SWI/SNF complexes during oncogenic transformation. Cancer cell 18(4): 316-328.

Wisniewski JR, Zougman A, Mann M. 2008. Nepsilon-formylation of lysine is a widespread post-translational modification of nuclear proteins occurring at residues involved in regulation of chromatin function. Nucleic acids research 36(2): 570-577.

Wood A, Schneider J, Dover J, Johnston M, Shilatifard A. 2005. The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. Molecular cell 20(4): 589-599.

Workman JL. 2006. Nucleosome displacement in transcription. Genes & development 20(15): 2009-2017.

Wu M, Wang PF, Lee JS, Martin-Brown S, Florens L, Washburn M, Shilatifard A. 2008. Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. Molecular and cellular biology 28(24): 7337-7344.

Xu K, Wu ZJ, Groner AC, He HH, Cai C, Lis RT, Wu X, Stack EC, Loda M, Liu T et al. 2012. EZH2 oncogenic activity in castration-resistant prostate cancer cells is Polycomb-independent. Science 338(6113): 1465-1469.

Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. 2006. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. Molecular cell 21(2): 227-237.

Zaret KS, Carroll JS. 2011. Pioneer transcription factors: establishing competence for gene expression. Genes & development 25(21): 2227-2241.

Zentner GE, Tesar PJ, Scacheri PC. 2011. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome research 21(8): 1273-1283.

Zhang F, Yu X. 2011. WAC, a functional partner of RNF20/40, regulates histone H2B ubiquitination and gene transcription. Molecular cell 41(4): 384-397.

Zhang Y. 2006. It takes a PHD to interpret histone methylation. Nature structural & molecular biology 13(7): 572-574.

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). Genome biology 9(9): R137.

Zhao Y, Lang G, Ito S, Bonnet J, Metzger E, Sawatsubashi S, Suzuki E, Le Guezennec X, Stunnenberg HG, Krasnov A et al. 2008. A TFTC/STAGA module mediates histone H2A and H2B deubiquitination, coactivates nuclear receptors, and counteracts heterochromatin silencing. Molecular cell 29(1): 92-101.

Zhou YB, Gerchman SE, Ramakrishnan V, Travers A, Muyldermans S. 1998. Position and orientation of the globular domain of linker histone H5 on the nucleosome. Nature 395(6700): 402-405.

Zhu B, Zheng Y, Pham AD, Mandal SS, Erdjument-Bromage H, Tempst P, Reinberg D. 2005. Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. Molecular cell 20(4): 601-611.

Zinner R, Albiez H, Walter J, Peters AH, Cremer T, Cremer M. 2006. Histone lysine methylation patterns in human cell types are arranged in distinct three-dimensional nuclear zones. Histochemistry and cell biology 125(1-2): 3-19.

7. Curriculum Vitae

Wanhua Xie

Address: Zimmermann straße 16h, 37075 Göttingen, Germany

Email: wanhuaxie18@gmail.com; Phone: +49 17683429537

Education and Research Experience

10/2012 – Present   **PhD student**

5/2014 – Present   Department of General, Visceral and Pediatric Surgery, University Medical Center Göttingen, Germany.

10/2012 – 5/2014   Institute for Tumor Biology, University Hospital Hamburg-Eppendorf, Hamburg, Germany

Supervisor: Prof. Dr. Steven A. Johnsen

Project: The role of RNF40 mediated H2Bub1 in gene transcription

09/2009 – 07/2012   **Master Student**, Department of Biochemistry and Molecular Biology, Jilin University, Changchun, China

Supervisor: Prof. Dr. Daxin Pang

Project: The role of Wnt signaling in porcine embryo development

09/2005 – 07/2009   **Bachelor Student**, Department of Biotechnology, Jilin University, Changchun, China

Publications

Hossan T, Nagarajan S, Baumgart SJ, **Xie W**, Magallanes RT, Hernandez C, Chiaroni PM, Indenbirken D, Spitzner M, Thomas-Chollier M et al. 2016. Histone Chaperone SSRP1 is Essential for Wnt Signaling Pathway Activity During Osteoblast Differentiation. *Stem cells* 34(5): 1369-1376.

Huang Y, Ouyang H, **Xie W**, Chen X, Yao C, Han Y, Han X, Song Q, Pang D, Tang X. 2013. Moderate expression of Wnt signaling genes is essential for porcine parthenogenetic embryo development. *Cellular signalling* 25(4): 778-785.

Huang Y, Tang X, **Xie W**, Zhou Y, Li D, Zhou Y, Zhu J, Yuan T, Lai L, Pang D et al. 2011. Vitamin C enhances in vitro and in vivo development of porcine somatic cell nuclear transfer embryos. *Biochemical and biophysical research communications* 411(2): 397-401.

Main laboratory techniques

Bioinformatic

- Chromatin immunoprecipitation coupled with sequencing, RNA sequencing, and library perpetration for high throughout sequencing

- Analysis of high throughout sequencing data: in Linux system, using various tools such as Bowtie and MACS; some public servers like Galaxy, Galaxy/deepTools, Galaxy/Cistrome, DAVID, oPOSSUM, and so on; some bioinformatic software based on java like IGV, GSEA, and seqMINER; analyses in R program such as DESeq for RNA-seq data, DiffBind, Epicseg, and so on

Generation of transgenic pigs using somatic cell nuclear transfer

- In vitro maturation of porcine oocytes and embryos, in vitro fertilization, somatic cell nuclear transfer, porcine embryo transfer, single embryo PCR

Reference

Prof. Dr. Steven A Johnsen,
Department of General, Visceral and Pediatric Surgery, University Medical Center Göttingen, 37075 Göttingen, Germany.
steven.johnsen@med.uni-goettingen.de

Prof. Dr. Hans Will,
Institute for Tumor Biology, University Hospital Hamburg-Eppendorf, 20246 Hamburg, Germany.
hanskilianwill@gmail.com