USP16 AND HISTONE H2A DEUBIQUITINATION IN MOUSE EMBRYONIC STEM CELL FUNCTION

by

WEI YANG

HENGBIN WANG, CHAIR
CHENBEI CHANG
LOUISE CHOW
CHRISTOPHER A. KLUG
TIM M. TOWNES

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham, in partial fulfillment of the requirements for the degree of Doctor of Philosophy

BIRMINGHAM, ALABAMA

2014
USP16 AND HISTONE H2A DEUBIQUITINATION IN MOUSE EMBRYONIC STEM CELL FUNCTION

WEI YANG

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR GENETICS

ABSTRACT

In eukaryotic cells, genomic DNA is packaged into a chromatin structure by association with histone and non-histone proteins. Posttranslational modifications of histones play important roles in the regulation of chromatin structure and function. Ubiquitination of histone H2A (ubH2A) represents a predominate modification, occurring on ~10% of total cellular H2A. While H2A ubiquitination is primarily conferred by Polycomb Repressive Complex 1 (PRC1), H2A deubiquitination has been attributed to multiple H2A deubiquitinases. Our laboratory previously reported the purification and functional characterization of a H2A-specific deubiquitinase, USP16 (initially named as Ubp-M) in human cells. However, whether USP16 represents a general and important regulator for H2A deubiquitination and the functional significance of USP16-mediated H2A deubiquitination during development remain largely elusive.

To study the function of USP16 in development, we generated Usp16 complete and conditional mouse models. Previously PRC1 and ubH2A have been shown to contribute to embryonic stem cell (ESC) pluripotency by repressing lineage-specific gene expression. However, whether active deubiquitination co-regulates ubH2A levels in
ESCs and during differentiation is not known. We discovered here that Usp16 regulates H2A deubiquitination and gene expression in ESCs, and importantly, is required for ESC differentiation. Usp16 knockout is embryonic lethal in mice, but does not affect ESC viability or identity. Usp16 binds to the promoter regions of a large number of genes in ESCs, and Usp16 binding is inversely correlated with ubH2A levels, and positively correlated with gene expression levels. Intriguingly, Usp16−/− ESCs fail to differentiate due to ubH2A-mediated repression of lineage-specific genes. Finally, Usp16, but not a catalytically inactive mutant, rescues the differentiation defects of Usp16−/− ESCs. Therefore, this study identifies Usp16 and H2A deubiquitination as critical regulators of ESC gene expression and differentiation.

Our studies on USP16 complete and conditional knockout mouse models reveal USP16 as an important, probably the major, regulator for H2A deubiquitination in vivo. Our studies further demonstrate that USP16 and its mediated H2A deubiquitination play critical roles in mouse ESC lineage commitment.
ACKNOWLEDGMENTS

I would like to thank my thesis mentor Dr. Hengbin Wang for giving me the opportunity to carry out my doctoral study and complete my doctoral degree in his laboratory. I would like to thank all the lab members in Wang lab for their unconditional support and help in completing my thesis research. I would also like to thank my committee members Dr. Chang, Dr. Chow, Dr. Klug and Dr. Townes for their invaluable advices and guidance for my thesis research. I am grateful for other researchers and people from Dr. Townes lab, Dr. Klug lab, Dr. Ryan, Dr. Giles lab, and Dr. Wu lab for their assistance to my thesis project. Finally, I would like to thank my parents and my wife for always being there for me.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Chromatin and Histone Modification</td>
<td>1</td>
</tr>
<tr>
<td>Histone H2B ubiquitination and deubiquitination</td>
<td>3</td>
</tr>
<tr>
<td>Histone H2A ubiquitination and deubiquitination</td>
<td>6</td>
</tr>
<tr>
<td>Questions to be addressed in this thesis</td>
<td>8</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td></td>
</tr>
<tr>
<td>THE HISTONE H2A DEUBIQUITINASE USP16 REGULATES EMBRYONIC STEM CELL GENE EXPRESSION AND LINEAGE COMMITMENT</td>
<td>14</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td></td>
</tr>
<tr>
<td>FUTURE DIRECTIONS</td>
<td>76</td>
</tr>
<tr>
<td>LIST OF REFERENCES</td>
<td>79</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>86</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Tables</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>THE HISTONE H2A DEUBIQUITINASE USP16 REGULATES EMBRYONIC STEM CELL GENE EXPRESSION AND LINEAGE COMMITMENT</td>
<td></td>
</tr>
<tr>
<td>1 Supplementary Table 1</td>
<td>73</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MULTIPLE LEVELS OF CHROMATIN STRUCTURE</td>
</tr>
<tr>
<td>2</td>
<td>POSTTRANSLATIONAL HISTONE MODIFICATIONS</td>
</tr>
<tr>
<td>3</td>
<td>SCHEMATIC REPRESENTATION OF FUNCTIONAL DOMAINS OF USP16</td>
</tr>
</tbody>
</table>

INTRODUCTION

1. MULTIPLE LEVELS OF CHROMATIN STRUCTURE

THE HISTONE H2A DEUBIQUITINASE USP16 REGULATES EMBRYONIC STEM CELL GENE EXPRESSION AND LINEAGE COMMITMENT

1. USP16 IS REQUIRED FOR MOUSE EMBRYONIC DEVELOPMENT

2. USP16 KNOCKOUT DOES NOT AFFECT ESC VIABILITY AND IDENTITY

3. USP16 BINDS TO TRANSCRIPTIOIN START SITE, POSITIVELY CORRELATES WITH GENE EXPRESSION AND INVERSELY CORRELATES WITH H2A DEUBIQUITINATION LEVELS

4. USP16+/− ESCS ARE DEFECTIVE FOR DIFFERENTIATION

5. USP16 IS REQUIRED FOR RESOLVING UBH2A MARKS DURING ESC DIFFERENTIATION

6. THE FAILURE TO ACTIVATE LINEAGE SPECIFIC GENE EXPRESSION IS THE CAUSE OF UNDIFFERENTIATED PHENOTYPE OF USP16+/− ESCS

7. USP16, BUT NOT THE ENZYMATICALLY INACTIVE MUTANT, RESCUED THE UNDIFFERENTIATED PHENOTYPE OF USP16+/− ESCS

S1. GENERATION OF USP16+/− ESCS BY DOUBLE TARGETING

S2. USP16 REGULATES GENE EXPRESSION IN ESCS
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>Usp16 and ubH2A ChIP-sequencing in ESCs</td>
<td>66</td>
</tr>
<tr>
<td>S4</td>
<td>Usp16-regulated gene expression correlates with changes in ubH2A level</td>
<td>68</td>
</tr>
<tr>
<td>S5</td>
<td>Usp16 is required for ESC differentiation</td>
<td>70</td>
</tr>
<tr>
<td>S6</td>
<td>ChIP-qPCR analysis of Usp16 and ubH2A levels in Usp16 target genes Tbx3,</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>HoxD11, and Gsn in Usp16&lt;sup&gt;+/+&lt;/sup&gt; and Usp16&lt;sup&gt;−/−&lt;/sup&gt; EBs</td>
<td></td>
</tr>
</tbody>
</table>


INTRODUCTION

Chromatin and Histone Modification

In eukaryotic cells, genomic DNA is organized into a chromatin structure by association with histone and non-histone proteins. The nucleosome is the basic repeating subunit of chromatin and is composed of 146 base pairs of DNA wrapped around a histone octamer. Each histone octamer contains two copies each of histone H2A, H2B, H3 and H4 (Figure 1). Individual nucleosomes are connected by linker DNA to form the 10 nm “beads on a string” form of chromatin. With the addition of linker histones and other proteins, the “beads on a string” form chromatin is further packaged into higher order chromatin structures. The formation of chromatin not only allows eukaryotic cells to store their genetic information efficiently, but also provides an elegant way to regulate DNA-based nuclear processes including transcription, DNA damage repair, replication, and recombination. Two major mechanisms, ATP-dependent chromatin remodeling and covalent modifications of chromatin, have been shown to modulate the structure of chromatin and facilitate chromatin-based nuclear processes. While ATP-dependent chromatin remodeling activity transiently alters the interaction between histones and DNA, covalent modifications of chromatin components are involved in the stable alteration of chromatin structure. Covalent chromatin modifications include DNA methylation and posttranslational histone modifications such as acetylation, methylation,
phosphorylation, ubiquitination and ADP ribosylation (Figure 2). These mechanisms function coordinately to precisely control the conformation of chromatin structure and facilitate nuclear processes on the chromatin template.

Histones are highly conserved through evolution. Each histone contains a globular domain, which mediates histone-histone interactions and two flexible terminals, which protrude from the chromatin fiber and are subject to a variety of modifications. These histone tails, particularly the N-terminal tails, are enriched for basic amino acids and are the primary sites for many posttranslational modifications. Modifications on histones, individually or in combination, are proposed to form a histone code, which recruits downstream effector proteins and affects specific physiological processes. Specific histone modifications have been correlated with particular chromatin and cellular functions. For example, methylation of histone H3 lysine 4 and lysine 79 correlates with active transcription while methylation of H3 lysine 9 and lysine 27 is associated with transcriptional repression. In addition to transcriptional regulation, histone modifications have also been implicated in other DNA-based nucleosome processes, including DNA damage response. Upon DNA damage, ATM is directed to the damaged loci and phosphorylates histone H2AX, which is proposed to trigger cell cycle arrest. Other histone modifications, including acetylation and methylation, are also implicated in efficient DNA repair processes. Recently, the role of histone ubiquitination, particularly H2A ubiquitination, in DNA damage repair was highlighted. Ubiquitin conjugates are transiently enriched at DNA damage foci. Within these ubiquitin conjugates, RNF168 catalyzed H2AK13/15 ubiquitination is required for the efficient
recruitment of the downstream repair protein, 53BP1\textsuperscript{15-18}. H2AK13/15 ubiquitination recruits 53BP1 to DNA damage foci through the recently identified 53PB1 ubiquitin interaction motif (residues 1604-1631)\textsuperscript{19}. In summary, posttranslational histone modifications play critical roles in virtually all DNA based nucleosome processes.

**Histone H2B ubiquitination and deubiquitination**

Ubiquitin (Ub) is a 76-amino acid globular protein, attached to target proteins by an isopeptide bond between its carboxyl-terminal glycine and the ε-amino group of a lysine residue within the substrate\textsuperscript{20}. Proteins can be mono- or poly-ubiquitinated, generating ubiquitin chain linkages at positions 11, 29, 48 and 63. While polyubiquitination through linkage 48 leads the protein to 26S proteasome-mediated degradation, monoubiquitination as well as polyubiquitination through linkage 63 regulates protein function. Ubiquitination has been implicated in many cellular processes including proteasome-mediated protein degradation, protein trafficking, cellular stress response, cell-cycle progression, and transcriptional regulation\textsuperscript{20}. Ubiquitination occurs through a series of enzyme-mediated processes: first ubiquitin is activated by E1-activating enzyme; the activated ubiquitin is then transferred to E2-conjugating enzymes; and finally E3-ubiquitin ligases mediate the formation of isopeptide between substrates and ubiquitin. Therefore, substrate specificity is largely determined by the E3 ligase\textsuperscript{21,22}. Polyubiquitination occurs through multiple rounds of this process.

Histones are also targets for ubiquitination. Indeed, ubiquitinated H2A was the first protein identified to be ubiquitinated. In mammals, about 5-15\% of total H2A and 1-2\%
of total H2B are ubiquitinated 23, while in *Saccharomyces cerevisiae*, only H2B is ubiquitinated 24. The ubiquitination sites were identified as lysine 119 on H2A and lysine 120 on H2B (in budding yeast H2B lysine 123) 24. Although polyubiquitinated H2A has also been detected in some tissues 25, H2A and H2B are predominately subjected to monoubiquitination.

The ubiquitin conjugating enzyme for H2B was identified as Rad6 26, which functions together with the RING domain containing E3 ligase Bre1 to ubiquitinate H2B in budding yeast 27,28. It was subsequently shown that the enzymatic machineries are conserved 29,30. Although initial studies indicate a role of H2B ubiquitination in gene silencing, subsequent studies revealed a role of H2B ubiquitination in gene transcription 31. Specifically, Rad6/Bre1 complex was recruited to gene promoters by specific transcription factors 28,32 and the machineries then travel with elongating Pol II through a PAF-mediated interaction. Therefore, ubiquitinated H2B is predominately distributed in the gene coding regions. While initial studies indicated a role of H2B ubiquitination in nucleosome disassembly during transcription elongation, recent studies have shown that H2B ubiquitination may function in nucleosome reassembly during transcription elongation 33,34. *In vitro* studies confirmed that H2B ubiquitination helps stabilize nucleosomes 34. Interestingly, recent studies suggest a role of H2B ubiquitination in regulating transcription-associated processes, including pre-mRNA splicing. H2B ubiquitination was found to decorate intron-exon junctions and abolishment of the H2B ubiquitin ligase was found to affect the splicing of pre-mRNAs both in budding yeast and in mammals 35-37.
One intrinsic feature for H2B ubiquitination is that H2B ubiquitination is a prerequisite for subsequent H3K4 and H3K79 methylations\(^{38-41}\). This so-called “trans-histone regulation” was first discovered in budding yeast and recent studies indicate this regulation may be conserved\(^3\). However, a few laboratories have reported conflicting results, in which H2B ubiquitination seems to function independently of histone H3 methylation\(^4\). Mechanistically, H2B ubiquitination correlates with H3K79 methylation, and intriguingly, H2B ubiquitination was found to directly stimulate H3K79 methylation \textit{in vitro}\(^43\). These studies indicated that H2B ubiquitination most likely alters chromatin confirmation and make it a better substrate for the histone methyltransferase. Although H2B ubiquitination was also reported to simulate H3 methylation \textit{in vitro}, the mechanism may be different, as H3K4 methylation and H2B ubiquitination have distinct chromatin localization\(^4\). While H3K4 methylation is predominately localized at gene promoter regions, H2B ubiquitination is at gene coding regions. Therefore, the exact mechanism of H2B ubiquitination in transcription remains to be determined.

Ubiquitination is a reversible process. Accordingly, a number of deubiquitinases have been reported for histone H2B\(^4\). The majority of these deubiquitinases belong to the cysteine protease family, where an active cysteine residue attacks the isopeptide bond through a nucleophilic mechanism. Ubp8 was first discovered as a subunit of the SAGA histone acetyltransferase complex\(^4\). Interestingly, both H2B ubiquitination and deubiquitination are required for optimal transcription. Mechanistically, H2B ubiquitination may function as a barrier for the phosphorylation of Ser2 on the C-terminal
domain (CTD) of RNA polymerase II, which is a marker for elongating Pol II \(^{50}\). It has been hypothesized that Ubp8 and Rad6 might co-regulate transcription and there may be multiple rounds of H2B ubiquitination and deubiquitination during transcription to enable a precise control of gene expression level. However, whether Ubp8 and Rad6 bind and interact simultaneously on the same RNA polymerase molecule is unknown. Interestingly, our laboratory discovered another H2B deubiquitinase, USP49, which only affects the expression of a few genes but is required for efficient co-transcriptional splicing of a large number of exons \(^{35}\). Mechanistically, USP49 regulates the association of the splicing machinery with chromatin \(^{35}\). However, the exact mechanism by which this effect on co-transcriptional splicing occurs remains to be determined.

**Histone H2A ubiquitination and deubiquitination**

Unlike ubH2B, which is generally distributed in gene coding regions, ubH2A is enriched at transcriptionally inactive regions types \(^ {51,52}\). Recent studies also show that ubH2A is Denriched at gene promoters in various cell types \(^ {51-54}\) and associated with gene repression. The ubiquitin ligase for H2A ubiquitination was identified as Polycomb repressive complex 1, with Ring1/Ring2 as the catalytic subunits \(^ {55,56}\). In addition to Ring2, other components such as Bmi1, were also found to be critical for H2A ubiquitination regulation \(^ {57,58}\). It is generally believed that the critical role of Ring2 in gene silencing depends on H2A ubiquitination. However, direct evidence for this hypothesis is lacking \(^ {59,60}\). PRC1 is recruited to chromatin by the chromodomain in PRC1 with trimethylated lysine 27 on histone H3 (H3K27me3), which is catalyzed by the methytransferase activity of the PRC2 complex \(^ {61-64}\). In addition, the PRC1-RYBP
complex, a variant of PRC1 containing Ring1B, Mel-18 and RYBP, is recruited to target loci to exert repressive function in a PRC2-independent manner. Other than these two mechanisms, specific transcription factors and KDM2B (FBXL10), which specifically recognizes non-methylated DNA in CpG islands, were also proven to function in the recruitment of PRC1. Therefore, PRC1 is probably recruited to chromatin in a context-dependent manner.

Multiple ubH2A deubiquitinases have been identified, including USP16 (Ubp-M), 2A-DUB (MYSM1), USP21, USP7, USP3, and Drosophila calypso. H2A deubiquitination by these enzymes has been shown to be involved in many cellular processes. 2A-DUB specifically deubiquitinates hyperacetylated nucleosomes in vitro, binds to the promoter of genes, and is associated with transcription initiation and gene activation. Usp21 has been reported to be involved in relieving transcriptional repression mediated by ubH2A and to activate transcription in cooperation with H3K4 di- and trimethylation. USP16 is a member of the USP (ubiquitin-specific protease) family. It has three major domains: ZnF-UBP domain, coiled-coil domain, and Peptidase C19 domain. In vitro studies reveal that the ZnF-UBP domain functions as an ubiquitin-binding domain that preferentially binds to the C-terminus of the protein. The two coiled-coil domains, which are located at residues 124-197 and 402-623, might contribute to the protein-protein interactions. The peptidase C19 domain located at the C-terminal and is a ubiquitin hydrolase that can hydrolyze the isopeptide bond between the glycine on ubiquitin and the lysine on acceptor protein. Three key residues form the catalytic center of the Peptidase C19 domain—C205, H758 and D798 (Figure 3). It is reported
that the C205S mutation abolishes the activity of the deubiquitinase. Our group first purified USP16 as a histone H2A specific deubiquitinase in vitro and subsequently demonstrated that USP16 specifically deubiquitinates ubH2A but not ubH2B in vitro and in vivo. Our studies showed that USP16 regulates cell cycle progression and USP16-mediated H2A deubiquitination is a prerequisite for phosphorylation of Ser10 on Histone H3 and chromosome segregation during mitosis. USP16 also regulates HoxD10 gene expression through H2A deubiquitination in mammalian cells and Usp16 knock-down resulted in defective posterior development in Xenopus laevis embryos. Recent studies have also revealed that in Down’s syndrome mouse models, where USP16 gene is triplicated, the self-renewal of hematopoietic stem cells and the expansion of mammary epithelial cells, neural progenitors, and fibroblasts were reduced. Usp16 has also been shown to coordinate with Aurora B-mediated H3S28 phosphorylation to facilitate transcription in quiescent lymphocytes. These results imply the potential rule of USP16 to antagonize PRC1 function in the self-renewal and/or senescence pathways.

**Questions to be addressed in this thesis**

When we purified Usp16 as histone H2A-specific deubiquitinase, multiple H2A deubiquitinases had already been reported. Although these deubiquitinases may function in distinct cell types, physiological conditions, or chromatin loci, it remains to be determined which enzyme functions in which tissue or at which functional stages in vivo. To address these questions, we set out to determine the function of Usp16 during mouse development. Our studies revealed that Usp16 regulates H2A deubiquitination in the mouse embryonic stem cell (ESC), during ESC differentiation. Importantly, Usp16 is
critically required for ESC lineage commitment. Therefore, our studies established USP16 as an important, probably the major, regulator \textit{in vivo} and demonstrated that Usp16 plays an indispensable role in ESC differentiation.
Figure 1. Multiple levels of chromatin structure. Left, schematic of various levels of chromatin compaction, from extended nucleosome arrays to folding of secondary chromatin structures exemplified by the 30-nm chromatin fiber to poorly characterized higher-order structures. There are ten histone tails protruding from each nucleosome core. Right, detail of nucleosome surface showing histones H2A (yellow), H2B (light red), H3 (blue) and H4 (green). Residues comprising the charged pocket are shown in dark red. Model courtesy of J. Chodaparambil. Note: From “Chromatin structure depends on what’s in the nucleosome’s pocket” by Tamara L. Caterino & Jeffrey J. Hayes, *Nature Structural & Molecular Biology*, 14(11), p. 1056. Copyright 2007 by Nature Publishing Group. Adapted with permission.
Figure 2. Posttranslational histone modifications. Modifications include acetylation (ac), methylation (me), phosphorylation (ph) and ubiquitination (ub1). Most of the known histone modifications occur on the N-terminal tails of histones, with some exceptions including ubiquitination of the C-terminal tails of H2A and H2B and acetylation and methylation of the globular domain of H3 at K56 and K79, respectively. Globular domains of each core histone are represented as colored ovals. Note: From “Covalent modifications of histones during development and disease pathogenesis” by Sukesh R Bhaumik, Edwin Smith, and Ali Shilatifard, 2007, *Nature Structural & Molecular Biology*, 14(11), p. 10085. Copyright 2007 by Nature Publishing Group. Adapted with permission.
Figure 3. Schematic representation of functional domains of USP16. Three major domains: ZnF-UBP domain, coiled-coil domain and Peptidase C19 domain are indicated. Three key residues C205, H758 and D798 are indicated as active sites to form the catalytic center of the peptidase C19 domain.
USP16 AND HISTONE H2A DEUBIQUITINATION IN MOUSE EMBRYONIC STEM CELL FUNCTION

by

WEI YANG, YUN-HWA LEE, AMANDA E. JONES, JESSICA L. WOOLNOUGH, DEWANG ZHOU, QIAN DAI, KEITH E. GILES, QIANG WU, TIM M. TOWNES, AND HENGBIN WANG

Copyright
2014
by
Nature Publishing Group
Used by permission
Format adapted and errata corrected for dissertation
CHAPTER 1

THE HISTONE H2A DEUBIQUITINASE USP16 REGULATES EMBRYONIC STEM CELL GENE EXPRESSION AND LINEAGE COMMITMENT

Abstract

Polycomb Repressive Complex 1 and histone H2A ubiquitination (ubH2A) contribute to embryonic stem cell (ESC) pluripotency by repressing lineage-specific gene expression. However, whether active deubiquitination co-regulates ubH2A level in ESCs and during differentiation is not known. Here, we report that Usp16, a histone H2A deubiquitinase, regulates H2A deubiquitination and gene expression in ESCs, and importantly, is required for ESC differentiation. Usp16 knockout is embryonic lethal in mice, but does not affect ESC viability or identity. Usp16 binds to the promoter regions of a large number of genes in ESCs, and Usp16 binding is inversely correlated with ubH2A levels, and positively correlated with gene expression levels. Intriguingly, Usp16−/− ESCs fail to differentiate due to ubH2A-mediated repression of lineage-specific genes. Finally, Usp16, but not a catalytically inactive mutant, rescues the differentiation defects of Usp16−/− ESCs. Therefore, this study identifies Usp16 and H2A deubiquitination as critical regulators of ESC gene expression and differentiation.
INTRODUCTION

Embryonic stem cells (ESCs) derived from blastocyst is an unique cell type which has the ability to differentiate and contribute to all somatic cell types\textsuperscript{79}. Two hallmarks of ESCs are self-renewal identity and pluripotent developmental plasticity\textsuperscript{79}. The pluripotent gene expression program is controlled by core the transcriptional regulatory network centered with Oct4, Sox2 and Nanog\textsuperscript{80,81}. Oct4-Sox2 was first reported to co-regulate Fgf4 expression in ESCs by through the Oct4-Sox2 motif on its promoter\textsuperscript{82}. They also directly regulate Nanog expression\textsuperscript{83}. More interestingly, Oct4 and Sox2 bind to their own promoters and regulate their own expression by a positive feedback loop\textsuperscript{84,85}. Genome wide mapping analysis reveals that these core regulators bind to a large number of actively transcribed gene promoters with substantial overlapping\textsuperscript{81,86,87}. However, recent studies show that Oct4 also binds to repressed genes and lineage specific genes promoters in ESCs\textsuperscript{88}, which implies a more comprehensive regulatory role in regulating ESC gene expression of these core regulators remaining to be elucidated\textsuperscript{89}.

In cooperation with the core transcriptional regulatory network, epigenetic regulation in ESCs is critical and recent studies reveal that the pluripotent state and development potential of ESCs is maintained through a combination of ESC specific transcription factors and recently characterized epigenetic regulators\textsuperscript{80,90,91}. Polycomb group (PcG) proteins are important epigenetic regulators that help maintain cell type-specific gene expression by repressing alternative gene expression programs\textsuperscript{91-93}. PcG proteins effectuate their functions through formation of multi-subunit protein complexes. Two major Polycomb Reppressive Complexes (PRCs) have been described: PRC1 and
PRC2. PRC2 mediates di- and tri-methylation of histone H3K27\(^{61-64}\). Methylated H3K27 serves as a docking site for PRC1 binding, although PRC2 independent PRC1 recruitment mechanisms have also been observed\(^{61,65}\). PRC1, which contains Bmi1, Cbx, Phc, and the catalytic subunit Ring1/2 (mouse Ring1A/B), ubiquitinates histone H2A K119\(^{55,56,94}\). Both H3K27 methylation and H2AK119 ubiquitination (ubH2A) are reported to be associated with transcriptional repression and critical for gene repression in many cell types\(^{55,61,95}\). Recently, it is reported that in ESCs, Polycomb group (PcG) proteins complex PRC1 function to repress the expression of key developmental regulators, thus stabilizing the pluripotent gene expression program\(^{90,95,96}\). PRC1 and PRC2 bind to a large number of lineage specific genes promoters, coincided with ubH2A and H3K27me histone marks enrichment at these loci, respectively\(^{51,95}\). Some of these gene promoters are co-occupied by PRC1 and PRC2, while some are exclusively bound by PRC2. Simultaneous depletion of Ring1A and Ring1B in ESCs causes depletion of ubH2A marks, de-repression of bivalent genes and loss of ESC identity\(^{54,97}\). In addition, knockout of PRC subunits Suz12, Ezh2, Eed, and Ring1B alone impairs ESCs self-renewal and pluripotency and is embryonic lethal in mice\(^{98-101}\). H2A ubiquitination is shown to directly block RNA polymerase II elongation and is one of the determining factors for productive transcription\(^{54}\). In addition to gene repression related studies, Brookes, et al.\(^{53}\) recently reported that in ESCs ubH2A is enriched at promoters of genes involved in metabolism and other processes, suggesting additional roles for Ring1B and ubH2A in ESCs\(^{53}\). Moreover, genes bound by PRC1 and enriched for ubH2A in ESCs are associated with RNAPII-S5P, which does not produce mature mRNA, strengthening the links between Ring1B, ubH2A, and transcriptional silencing. This is consistent with a
role for ubH2A in transcriptional repression\textsuperscript{52,54,97}. Despite an unambiguously demonstrating the functions of ubH2A in PRC1-mediated gene repression \textit{in vivo} remains a challenge in higher eukaryotes.

One signature epigenetic pattern of ESCs is bivalency. Specifically, some genes enriched for repressive marks H3K27me2/3 by PRC2 are also enriched for H3K4me3, a mark generally found at actively transcribed genes\textsuperscript{102}. These bivalent modifications are found principally at key developmental regulators and may contribute to repression of these genes in ESCs while enabling their rapid induction in response to developmental signals\textsuperscript{103}. Recent investigations reveal that PRC1-mediated ubH2A also marks bivalent genes and regulates their expression in ESCs\textsuperscript{51,54,104}. Therefore, PRC1 binding and PRC1-mediated H2A ubiquitination, may also be required for the efficient repression of key developmental related genes to maintain ESC identity.

During ESC differentiation, bivalent histone modifications are resolved to monovalent modifications and pluripotency-specific genes are epigenetically silenced\textsuperscript{103,105}. It is known that two JmjC-domain-containing proteins, UTX and JMJD3, can actively remove H3K27me and, therefore, may counteract PRC2-mediated gene silencing during ESC differentiation\textsuperscript{6,106}. Like H3K27me3, ubH2A is a reversible mark that can be removed through the activity of deubiquitinating enzymes. Although a number of ubH2A deubiquitinases have been reported and shown to regulate a variety of cellular processes in different organism, none of them have yet been linked to ESC differentiation\textsuperscript{68-72} and whether ubH2A deubiquitination is important during ESC
differentiation and which enzyme mediates ubH2A deubiquitination during cell fate transitions is not known. Recent studies on one of these deubiquitinases Usp16 in some somatic progenitors and stem cells implies the potential rule of Usp16 to antagonize PRC1 function in the self-renewal and/or senescence pathways\textsuperscript{77,78} and it is highly possible that Usp16 also regulates ESC pluripotency. This prompts us to explore the rule of Usp16 in mouse ESC lineage commitment.
METHODS

Cell culture
Mouse V6.5 and all derived ESCs were cultured in ESC media, containing DMEM medium (High glucose, Gibco MT-10-013) supplemented with 50 unit/ml penicillin and streptomycin (Life technologies, 15070-063), 18% murine ESC defined FBS (Thermo Scientific, SH30070.03E), 2mM L-glutamine (Cellgro, 25-005-CI), 1mM sodium pyruvate (Gibco, 11360-070), 1X nonessential amino acids (Cellgro, 25-025-CI, 100X stock), 1X nucleoside (Millipore, ES-008-D, 100X stock), 0.007% β-mercaptoethanol (Fisher, O3446I), and 1000 unit/ml mLIF (Millipore, ESGRO) on irradiated mouse embryonic fibroblasts (Millipore, PMEF-NL) and passaged every other days as described previously. For induction of Cre recombinase, Usp162lox:2lox:CAGG Cre ESCs were cultured in medium containing 1µM 4-hydroxytamoxifen (Sigma, H7904) for 3-5 days and then on ESC culture medium.

Generation of Usp16 knockout and conditional knockout mice
To generate Usp16 knockout and conditional knockout mice, targeting vectors were constructed as previously described. Targeting vectors were then electroporated into V6.5 ESCs. After genotyping and sequencing to confirm the correctly recombined knockout allele, Usp16+/− or Usp162lox/− ESCs were injected into C57BL/6 blastocysts to clone mice. The chimeric mice were backcrossed with C57BL/6 and Usp16+/− mice were intercrossed for ESC derivation or embryo characterization. Usp162lox/+ mice were crossed with CAGG-Cre transgenic mice to obtain Usp162lox/+: CAGG-Cre mice, which
was intercrossed or crossed with Usp16^{2lox:2lox} mice for ESC derivation. Primers used for genotypes are labeled in Fig. 1a and Supplementary Fig. 1a. Primer sequences are included in Supplemental Table 1. All animal experiments were carried out according to the guidelines for the care and use of laboratory animals of the University of Alabama at Birmingham.

**Derivation and characterization of ESCs**

Blastocysts were flushed out from the uterus of Usp16^{-/-} or Usp16^{2lox/+}:CAGG-Cre female mice at E3.5 as described^{108} and cultured on feeder cells with 2i medium^{109} or regular ESC medium for 5-8 days. After 5-8 days of culture, ESC-like colonies formed from ICM outgrowths were handpicked using sterile pipette tips and dissociated by 0.25% trypsin and expanded in ESC medium. For proliferation rate assay, 1 x 10^4 ESCs were seeded in 12 well plates and cell numbers were counted and the same seeding procedure was repeated every other day. Western blot assay and qPCR were performed as previously described^{110}. Antibodies include anti-Usp16^{68}, anti-ubH2A (Millipore, 05-678), anti-ubH2B (Millipore, 05-1312), anti-H3 (Abcam, ab100938), anti-GAPDH (Sigma, G9545), Anti-Nanog (R&D Systems, AF2729), Anti-Oct4 (Santa Cruz Biotechnology, SC9081), and anti-Sox2 (Cell signaling, 3579). Alkaline phosphatase substrate kit was from Vector laboratory (SK-5100).

**Cell cycle analysis**

2 x 10^6 control wild type or Usp16^{-/-} ESCs were trypsinized and washed with PBS once. Cell pellets were re-suspended in 50µl PBS and 500µl 0.9% NaCl solution was added to
each sample with gentle vortex. 1.5ml 95% ethanol was then added to the cell suspension with vortex. Cells were fixed at 4°C for overnight. After fixation, cells were re-suspended in 300µl PBS with 1mg/ml RNase (Sigma, R-4875) and incubated at 37°C for 20 minutes. 300µl 40µg/ml propidium iodide (Sigma, P1470) was then added to cell suspension for at least 20 minutes before proceeding to FACS analysis.

**Embroid body formation and analysis**

Embryoid bodies (EBs) were formed in EB medium (ESC medium without mLIF) in hanging drops (800 cells per 20µl drop) on petri dish lids for 48 hours. EBs were then collected and cultured on a horizontal rotator at 40 rpm in petri dishes and collected at indicated time points. Medium was changed every other day. For each time point, the diameter of 50 EBs was counted. For H&E staining, EBs were fixed in 4% paraformaldehyde in PBS for 4 hours and then in 70% ethanol (overnight), 80% ethanol (1 hour), 90% ethanol (1 hour), 95% ethanol (overnight), 100% ethanol (1 hour, twice), followed by standard H&E staining protocol as described previously.

**RNA purification and real time PCR assay**

Total RNA was extracted using TRIzol (Life technologies, 15596-026) following the manufacturer’s instruction. Isolated RNA was reverse transcribed using M-MLV (Promega, M1701) following the manufacturer’s instruction. For quantitative PCR, 0.5µl cDNA was used as template in a 25µl reaction containing 1µg oligo dT 15 primers, 0.5 mM dNTPs, RNase inhibitor, M-MLV, and M-MLV buffer. All quantitative PCR reactions were carried out on a Roche LightCycler 480 thermocycler (Roche). Fold
change of gene expression levels were calculated according to the $2^{\Delta\Delta \text{Ct}}$ method as described previously and normalized to GAPDH. Primers for real time PCR assay are listed in Supplementary Table 1.

**RNA-sequencing and analysis**

RNA-sequencing was performed on Illumina HiSeq2000 in UAB Heflin Genomic Core Facility. cDNA libraries were constructed with TruSeq library generation kits (Illumina). The cDNA ends were repaired, A-tailed, and adaptor ligated for indexing. The cDNA library was then quantitated by real time PCR in a Roche LightCycler 480 with the Kapa Biosystems kit for library quantitation (Kapa Biosystems) prior to cluster generation. Clusters were then generated to yield approximately 725-825 K clusters/mm$^2$ and the density and quality were determined after the first base addition. Paired end 2 x 50bp sequencing runs were performed for all samples. After sequencing, the data was converted to FASTQ Sanger format and then aligned to mouse reference genome mm9 by TopHat. Cufflinks was used to assemble transcripts, estimate the abundances, and test for differential expression. Cuffmerge was used to compare the assembled transcripts to the reference annotation across multiple experiments. Cuffdiff was then used to determine the change in transcript levels. Cufflinks, Cuffmerge and Cuffdiff analyses were performed as described previously. Gene ontology analysis was performed using DAVID (http://david.abcc.ncifcrf.gov/).

**Chromatin immunoprecipitation (ChIP) and ChIP-qPCR assay**
ESCs were cultured on feeder cell free plates and fixed with 1% formaldehyde at room temperature for 10 minutes, followed by quenching with 125mM glycine. Cells were collected by scraping and washed with PBS once. EBs were dissociated with 0.25% trypsin (Cellgro, MT-25-053-CI) at 37°C for 10 minutes prior to fixation. About 5 x 10^7 cells of ESCs and EBs were resuspended in 10ml PBS containing 0.5% Triton X-100 and incubated on ice for 10 minutes. After centrifugation, cell pellets were resuspended in 5ml Tris-EDTA (TE8.0, pH 8.0) containing 0.1% SDS, and protease inhibitors (0.5mM pepstatin A, 1µM leupeptin, 0.5µM aprotinin, 0.5 mM PMSF). Cell suspension was sonicated using the Fisher sonic dismembrator model 500 at 20% output power 20 times for 30 seconds followed by a 30 second rest on ice. After sonication, cell suspension was adjusted with 1x RIPA buffer (10mM Tris-HCl 8.0, 1mM EDTA, 0.1% SDS, 0.1% Na-Deoxycholate and 1% Triton X-100) as described previously before immunoprecipitation. About 4-8µg of affinity purified Usp16 antibody and an equal amount of IgG were incubated with 150 µl Protein A agarose beads (50% slurry in PBS) at room temperature for 1.5 hours and then washed with 1 x RIPA buffer twice to remove unbound antibodies. Antibody-protein A complex was then incubated with 5ml sonicated chromatin from 5 x 10^7 cells in a 15ml canonical tube at 4°C overnight. After incubation, beads were washed in 1.5ml Eppendorf tubes at 4°C with 1ml RIPA buffer (10 minutes, twice), 1ml RIPA buffer with 0.3M NaCl (10 minutes, twice), 1ml LiCl buffer (0.25M LiCl, 0.5% NP40, 0.5% Na-Deoxycholate, 10 minutes, twice), 1ml TE8.0 buffer with 0.2% Triton X-100 (10 minutes, once) and 1ml TE 8.0 buffer (10 minutes, once). Beads were then re-suspended in 100µl TE 8.0 buffer and cross-linking was reversed by adding 0.5M NaCl and 1µg/ml RNase A at 65°C overnight. Additional incubation after
adding 3-5µl of 10% SDS and 5µl of 20 mg/ml protease K was performed at 55°C for 2 hours to fully dissociate protein-antibody complex. DNA template from ChIP was prepared by standard phenol/chloroform DNA extraction and ethanol precipitation. Eluted DNA was dissolved in 50µl TE 8.0 buffer. For ChIP-qPCR assay, 0.5µl of 1:5 dilution of elution was used as templates in a 20µl reactions using GoTag (Promega, M3001) DNA polymerase as described previously. Signals were normalized to input DNA. Primers for ChIP-qPCR are included in Supplementary Table 1.

Native chromatin immunoprecipitati

For native ChIP, ESCs and EB (pre-trypsinized) were collected and washed with pre-chilled PBS once and incubated in 0.5% NP40 in TBS buffer (10mM Tris-HCl pH 7.5, 3mM CaCl2, 2mM MgCl2) on ice for 1hour with gentle vortex every 15 minutes. After incubation, cells were homogenized by 10ml Dounce homogenizer and re-suspended in 25% sucrose in TBS. Cells were then transferred to a 15ml tube prefilled with half volume of 50% sucrose in TBS and spun down at 600g for 20minutes. Pellets were washed in 25% sucrose TBS once and re-suspended in nuclear digestion buffer (50mM Tris-HCl pH 7.5, 1mM CaCl2, 2mM MgCl2, 0.32M sucrose) to reach a rough DNA concentration of 0.5 mg/ml as measured by Nanodrop. 2U/ml micrococcal nuclease (MNase) was added to the chromatin and digestion was carried out at 37°C for 15 minutes and repeated once for the pellet after spinning down the first digestion mixture. Digestion was terminated by addition of 5mM EDTA. Digested chromatin solution was collected and pellet from second digestion was further incubated with nuclear lysis buffer (1mM Tris-HCl pH 7.5, 0.2 EDTA) on ice for 1hour with occasional tapping and then
overnight at 4°C with rotation. The supernatant was collected and combined with the digestion supernatants and stored at -80°C. All buffers used were supplemented with protease inhibitors, 5mM PMSF, 5mM NaF, and 2mM Na₃VO₄.

For each native ubH2A ChIP, 50µl beads were incubated with anti-IgM bridging antibody (Millipore, 12-488) for 1 hour and washed with 1ml RIPA buffer containing 0.3M NaCl. For H2A ChIP, protein-G agarose bead was washed with 1ml RIPA buffer with 0.3M NaCl. Amount of each antibody used to achieve optimal ChIP performance were determined previously. Briefly, 50µg ubH2A antibody (Millipore, 05-678) and 15µg H2A antibody (Millipore, 07-146) were added to treated protein G agrose beads with 400µg chromatin and incubated at 4°C for overnight. After immunoprecipitation, beads were washed with 1ml RIPA buffer (twice, 10 minutes each), 1ml LiCl buffer (10 minutes, once), and 1ml TE8.0 (10 minutes, once). Washed complex were then eluted from beads with 50mM Tris-HCl pH 8.0, 1mM EDTA, and 1% SDS by incubation at 65°C for 5 minutes and room temperature for 20 minutes with shaking. Elution process was repeated once and eluates were combined and then incubated at 55°C for two hours after addition of 5mM EDTA and 200 µg/ml proteinase K. DNA was isolated by phenol-chloroform extraction and ethanol precipitation and dissolved in TE8.0.

**ChIP-Sequencing (ChIP-Seq) and analysis**

ChIP-Sequencing using cross-linked or native chromatin templates was performed at Hudson Alpha Institute Genomic Service Laboratory (GSL) and the National University of Singapore. NEB Next DNA sample prep reagents (New England Biolabs, E6000S/L)
were used for end repair, dA-tailing, and adaptor ligation (custom adaptors from Hudson Alpha GSL). The entire volume ligated product was PCR amplified for 15 cycles using Invitrogen's Pfx DNA polymerase using the custom adaptors designed to allow for 96 unique indexes. Size selection was performed on 2% agarose gel. The final library quality was examined by Qubit and Bioanalysis, and real time PCR was performed using KAPA's library quantification kit for Illumina. Each library was diluted to 12.5pM and clustered using a cBot. Each library was sequenced over 1/6th of a HiSeq 2000 lane following standard Illumina protocols for paired end 50bp sequencing.

For data analysis, the converted fastq files obtained from Hudson Alpha GSL were aligned to mouse reference genome mm9 by BOWTIE. The SAM files after alignment were filtered to remove PCR duplicate and unmapped reads by SAMtools on Galaxy before converting to Bed files. After Bed files were generated, these files were subjected to peak calling with MACS on Galaxy using an FDR (False discovery rate) threshold set to ≤ 0.05. Proximal regions around the transcription start site (-2kb to +1kb) that were occupied by at least one peak were defined as enriched promoters/genes. Coordinates of enriched genes with different binding profiles were then intersected to generate Venn diagrams. To count tag density for all the bed files in the coordinates as indicated in all figures, a 15-step workflow on Galaxy (http://main.g2.bx.psu.edu/) was created and available for sharing upon request. To generate heatmaps, individual mm9 gene (RefSeq annotations) was divided into 100 bins using a custom script. Coverage across binned genes plus 10kb upstream of the TSS and 10kb downstream of the TES was generated using BEDTools ‘coverageBed’ with the -d option. Outputs were converted to data
matrices by a custom script and then visualized as heatmaps in Microsoft Excel 2010 using conditional formatting options. Gene ontology analysis was performed using DAVID (http://david.abcc.ncifcrf.gov/).

**Usp16 rescue experiment**

Usp16−/− ESCs were re-targeted using the same targeting vector or a C205S point mutant vector to generate Usp16 conditional knockout ESCs using procedures similar to those described in the mouse generation method section. Usp16 protein level was measured by western blotting to confirm that wild type or C205S Usp16 was expressed at control levels. EB formation, H&E staining, and RT-qPCR assay were performed as described.

**Oct4 shRNA and HoxB4 overexpression lentivirus infection**

pLKO-Oc4 shRNA and control vectors were kindly provided by Dr. Rui Zhao (UAB Stem Cell Institute). pDL171-HoxB4 and control vectors were kindly provided by Dr. Chia-Wei Chang (UAB Stem Cell Institute). Lentiviral vectors together with packaging vectors were transfected in to 293T cell using calcium phosphate and virus were collected at 48 hour and 72 hour post-infection and pooled. After titration of individual virus, 5x10^5 Usp16−/− ESCs were infected by Oct4 shRNA lentivirus or HoxB4 expression vectors on gelatin coated 6 well plates by spinoculation at room temperature for 2 hours and then incubated at 37°C for overnight before switching to ESC medium. After infection, ESCs were further cultured for 5-7 days before collecting for RNA extraction and RT-qPCR assay. Usp16+/+ and Usp16−/− ECS without infection were used as controls.
**Usp16 and Ring1B overexpression retrovirus infection**

pMigR1-Usp16 and pMigR1-Ring1B retroviral vectors together with packaging vector were transfected in to 293T cell using calcium phosphate and virus were collected at 48 hour and 72 hour post-infection and pooled. After titration of individual virus, $5 \times 10^5$ Usp16$^{+/+}$ ESCs were infected by Usp16 or Ring1B retrovirus on gelatin coated 6 well plates by spinoculation at room temperature for 2 hours and then incubated at 37°C for overnight before switching to ESC medium. After infection, ESCs were further cultured for 2-3 days before collecting for ChIP-qPCR assay.
RESULT

**Usp16 knockout is early embryonic lethal in mice**

USP16 was first identified as a histone H2A-specific deubiquitinase which regulates cell cycle progression and gene expression in human cells. To determine whether Usp16 is essential for mammalian development, we knocked out the gene in murine ESCs and produced mice. The targeting vector was designed to replace exons 5 and 6 of mouse Usp16 with a PGK-Neo/Kan cassette (Fig. 1a, targeting vector). Cysteine 205 in exon 6 is required for Usp16 deubiquitinase activity (Fig. 1a, Usp16 locus). The targeting vector was electroporated into the F1 hybrid (B6/129) ESC line V6.5 and G418 resistant colonies were selected (Fig. 1a). After removing the PGK-Neo/Kan cassette by transient expression of Cre-recombinase, correctly targeted ESCs (Fig. 1B, top panel, right image) were identified by PCR-mediated genotyping (Fig. 1b, bottom panel) and injected into blastocysts to produce mice. After germline transmission, mice heterozygous for Usp16 deletion (Usp16+/−) were inbred to obtain Usp16 knockout mice (Usp16−/−). However, no homozygous Usp16 knockout mice were obtained (Fig. 1c, 89 mice, 12 litters), indicating that Usp16 deletion is embryonic lethal.

To determine the stage at which Usp16 deletion is lethal, Usp16+/− male and female mice were mated and embryos were isolated. No Usp16−/− embryos were detected at embryonic day (E) 13.5 or E10.5 (data not shown). However, Usp16−/− embryos were detected at E7.5. The majority of E7.5 embryos displayed normal morphology, with closed neural tubes (Fig. 1d, top panel, left image). In contrast, a small number of
embryos appeared to be partially re-absorbed (Fig. 1d, top panel, right image).

Morphologically normal embryos were either Usp16\textsuperscript{+/+} or Usp16\textsuperscript{+/-} while the partially re-absorbed embryos were Usp16\textsuperscript{-/-} (Fig. 1d, bottom panel). These results indicate that defects in Usp16\textsuperscript{-/-} embryos occurred prior to E7.5. To determine whether Usp16 deletion affects embryo viability before or after implantation, we collected blastocysts at E3.5. Genotyping analysis of 33 blastocysts recovered from 5 female mice identified 6 Usp16\textsuperscript{-/-} embryos (Fig. 1e). This result indicates that Usp16 knockout causes lethal developmental defects after implantation but before the E7.5 developmental stage.

**Usp16 knockout does not affect ESC viability and identity**

Similar to Usp16 deletion, knockout of any of the PRC subunits in mice, such as Suz12, Ezh2, Eed, or Ring1B, results in early embryonic lethality\textsuperscript{98-101}. ESCs deficient for these PRC subunits are viable, but are prone to spontaneously differentiate during culture. Therefore, we reasoned that Usp16 knockout ESCs should also be viable. To study the role of Usp16 in early mouse embryonic development, we first derived Usp16\textsuperscript{+/-} ESCs by culturing blastocysts in 2i medium. Surprisingly, of 45 ESC lines derived by this method, 17 were Usp16\textsuperscript{+/-}, 28 were Usp16\textsuperscript{+/-}, and none were Usp16\textsuperscript{-/-}. The successful generation of multiple Usp16\textsuperscript{+/-} and Usp16\textsuperscript{+/-} ESC lines suggests that our inability to generate Usp16\textsuperscript{-/-} ESC lines was not due to technical reasons. One possible explanation is that Usp16 is required for gene expression reprogramming during the blastocyst to ESCs transition. To test this hypothesis, we derived Usp16\textsuperscript{-/-} ESCs by targeting the remaining wild-type allele in the Usp16\textsuperscript{+/-} ESCs, which were used to generate Usp16 knockout mice. For this purpose, we constructed a conditional targeting vector by inserting a LoxP site upstream
of exon 5 and a PGK-Neo/Kan cassette (which was flanked with two FRT sites and one downstream LoxP site) downstream of Usp16 exon 6 (Supplementary Fig. 1a, targeting vector). This vector was then electroporated into the Usp16+/− ESCs to replace the remaining wild-type Usp16 allele. Since the homologous regions of the conditional targeting vector have been deleted in the Usp16 knockout allele, the wild-type allele should be specifically replaced. ESCs in which the remaining wild-type allele was replaced were selected by G418 resistance. The PGK-Neo/Kan cassette was then removed by transient expression of Flippase and used for the generation of Usp16 conditional knockout mice (Supplementary Fig. 1a, Usp16lox, see below). Usp16 was deleted from these ESCs by transient expression of Cre recombinase (Supplementary Fig. 1a, Usp16lox). Usp16−/− ESCs were identified by PCR-mediated genotyping (Supplementary Fig. 1c) and confirmed by western blot and real-time PCR analysis (Supplementary Fig. 1d; Fig. 1e, first panel). Usp16−/− ESCs exhibited normal morphology (Supplementary Fig. 1b) and expressed pluripotent genes at levels similar to control V6.5 ESCs (Supplementary Fig. 1e, second to fourth panels, compare lanes 2, 3 with 1). Therefore, even though direct derivation of Usp16−/− ESCs from blastocysts could not be achieved, Usp16 does not appear to be required for ESC viability.

Since Usp16−/− ESCs generated by the approach described above have been subjected to multiple rounds of electroporation, which might alter cellular functions, we also derived Usp16−/− ESCs by an alternative approach. We took advantage of Usp16 conditional knockout mice, which were generated by injecting Usp162Lox ESCs (Supplementary Fig. 1a) into blastocysts (data not published). These conditional Usp16
knockout (Usp16\(^{2\text{lox}/2\text{lox}}\)) mice were crossed with the inducible Cre expression mouse line, CAGG-Cre, to obtain Usp16\(^{2\text{lox}+/+}\): CAGG-Cre mice. These mice were then interbred or bred with Usp16\(^{2\text{lox}/2\text{lox}}\) mice. ESCs were derived from the offspring of these crossings and Usp16\(^{2\text{lox}/2\text{lox}}\): CAGG-Cre ESCs were identified by PCR mediated genotyping and confirmed by sequencing (Fig. 2a, bottom image). Usp16 was deleted by culturing Usp16\(^{2\text{lox}/2\text{lox}}\): CAGG-Cre ESCs with 1\(\mu\)M 4-hydroxytamoxifen (4-OHT), which induces nuclear localization of the Cre recombinase, for 3-5 days (Fig. 2a) before switching to 4-OHT free medium. The efficiency of Usp16 deletion was confirmed by western blot assay (Fig. 2c, first two panels, compare lanes 2, 4 with 1, 3). Consistent with our previous studies, Usp16 knockout caused a specific increase of ubH2A levels, but had no effect on ubH2B levels (Fig. 2c, third and sixth panels, compare lanes 2, 4 with 1, 3, the intensity of ubH2A signal were quantified and labeled). Usp16\(^{-/-}\) ESCs exhibited normal morphology (Fig. 2b, top panel) and have growth rates (Fig. 2e) and cell cycle profiles (Fig. 2f) similar to Usp16\(^{+/+}\) ESCs. This contrasts with studies in HeLa cells, where USP16 deletion causes a slow growth phenotype, partially due to the defects of H2A deubiquitination during cell cycle G2/M phase progression\(^{68,118}\). Usp16\(^{-/-}\) ESCs stained positively for alkaline phosphatase (Fig. 2b, bottom panels). Pluripotent genes Oct4, Sox2, and Nanog (Fig. 2d, top three panels, compare lanes 2, 4 with 1, 3) and PRC subunits Ring1B, Suz12, and Ezh2 (Fig. 2d, fourth to sixth panels, compare lanes 2, 4 with 1, 3) were not affected by Usp16 deletion. Together, these studies demonstrate that Usp16 regulates ubH2A levels in ESCs, and is not required for ESC viability and identity.
Usp16 binds to gene promoters, positively correlates with gene expression and inversely correlates with H2A ubiquitination levels

Since Usp16 regulates ubH2A levels in ESCs, and PRC1 and H2A ubiquitination are enriched at developmental and metabolic genes in ESCs\textsuperscript{51,53}, we investigated whether Usp16 regulates gene expression in ESCs by RNA-sequencing (RNA-seq). Usp16 knockout was associated with an apparent decrease of transcript abundance (Supplementary Fig. 2a), consistent with the role of Usp16 as a histone H2A deubiquitinase functioning in gene activation. 1,094 genes were significantly differentially expressed (Log\textsubscript{2} [Fold change] >2 cutoff), with 330 genes up-regulated and 764 genes down-regulated in Usp16\textsuperscript{−/−} ESCs (Fig. 3a). Up-regulated genes are particularly enriched for myofibril assembly, macromolecular complex assembly, chromatin assembly, and nucleosome organization, while down-regulated genes are enriched for transcription regulation, RNA, nitrogen, macromolecule metabolism (Supplementary Fig. 2b), indicating that further increasing ubH2A levels by inactivation of Usp16 affects many basic cellular processes. However, removing ubH2A by knocking out Ring1A/1B and further increasing ubH2A by inactivation of Usp16 primarily affect distinct groups of genes, only 12.6\% of Usp16 knockout affected genes (1,094, including both up- and down- regulated, Fig. 3a) overlaps with genes affected by Ring1A/B inactivation (1,156, including both up- and down- regulated)\textsuperscript{97} (Supplementary Fig. 2d). Since Usp16\textsuperscript{−/−} ESCs maintain ESC identity, the deregulation of these genes does not appear to impede ESC self-renewal. However, Usp16\textsuperscript{−/−} ESCs do have a higher tendency for differentiation under certain culture conditions compared to Usp16\textsuperscript{+/+} ES cells (Supplementary Fig. 2e), indicating that Usp16-regualted gene expression also contributes to the stabilization of
the pluripotent gene expression program. We further validated the RNA-seq results by real-time PCR of ten randomly selected genes (Supplementary Fig. 2c). These results confirmed our RNA-seq data and revealed that Usp16 regulates gene expression in ESCs.

To determine whether Usp16 regulates H2A ubiquitination level at PRC1 target genes in ESCs, we first determined the Usp16 binding profile in ESCs by ChIP sequencing (ChIP-seq). ChIP-sequencing results revealed that Usp16 binds to all chromosomes with no particular bias for any individual chromosomes (Supplementary Fig. 3a). Tag counts normalized to the coverage of different gene structures revealed that Usp16 binding is particularly enriched in the proximal regions (2kb upstream) of transcription start site (TSS) and 5’UTR with relatively low strength in 5’distal regions (50 kb upstream of TSS), gene coding regions (both exons and introns), 3’UTR, 2 kb downstream of TTS, 3’distal region (50 kb), and non-genic regions compared to IgG (Fig. 3b). To gain insight into the role of Usp16 in gene regulation, we performed a meta-gene analysis combining all genes in the mouse genome, including up and downstream regions. This analysis revealed that Usp16 specifically binds to gene promoter regions, defined as -2 kb to +1 kb of TSS sites (Fig. 3b). Parallel ChIP-seq experiments with Usp16 antibody in Usp16−/− ESCs did not reveal any significant binding (Supplementary Fig. 3c), indicating the specificity of ChIP-seq signals. To better understand the significance of Usp16 binding, we employed peak calling analysis (see methods) and identified ~7,000 Usp16 binding sites, which occupy 4,814 promoters. Usp16 bound genes are enriched for genes encoding proteins involved in RNA processing, RNA binding, protein localization, macromolecule catabolic process, protein transport, DNA binding, cell cycle,
chromosome organization, protein biosynthesis, DNA repair, RNA polymerase activity, and proteasome complex, indicating a potential and wide-spread function of Usp16 in regulating a variety of fundamental cellular processes, although only limited amounts of genes show significant changes when Usp16 is inactivated (Supplementary Fig. 3d).

To determine the site-specific effects of Usp16 on H2A ubiquitination, we measured ubH2A levels in Usp16+/+ and Usp16−/− ESCs by native ChIP-seq. Consistent with recent studies, ubH2A was enriched at gene promoters (Fig. 3c)53,104. By using peak calling analysis, we identified 2,433 gene promoters significantly enriched for ubH2A. These genes strongly overlap with ubH2A enriched genes reported in a recent study (75% overlapping, similar binding profiles, and similar GO categories, Supplementary Fig. 3e)53, indicating the reliability of our native ChIP-seq experiments. Genes bound by ubH2A are mainly developmental and transcription regulators (Supplementary Fig. 4a). 2,022 genes become newly enriched for ubH2A in Usp16−/− ESCs, indicating a significant role for Usp16 in regulating ubH2A levels at these genes (Supplementary Fig. 4a). These genes include some developmental regulators but are mainly enriched for general biological processes, consistent with recent report53 (Supplementary Fig. 4a). Supporting our biochemical assays, Usp16 knockout caused a significant increase of ubH2A levels without any apparent effects on H2A levels (Fig. 3c). To better define the relationships between Usp16, ubH2A level, PRC1 and gene expression on a gene-by-gene basis we generated binding profile heat maps. Genes were sorted based on expression level (Fig. 3d, the 1,000 highest and least expressed genes are shown). Since the bottom genes are largely unexpressed (FPKM<1), we re-sorted these genes using Ring1B binding extracted
from a recent publication 119. The heatmaps of Usp16 binding reinforces the metagene analysis, illustrating localization specific to the promoter regions of virtually all genes (Fig. 3d, first column). Furthermore, consistent with a role in promoting transcriptional activity, Usp16 was more highly enriched at active promoters than silent genes (Fig. 3d, first column, top vs bottom). Ring1B localization was also specific to the promoter regions of virtually all genes, but its binding was more prominent among the silent genes than in the most highly expressed (Fig. 3d, second column). Consistent with each protein’s enzymatic activity, ubH2A was virtually undetectable at active genes (where Usp16 was most abundant), and accumulated among silent genes (where Ring1B was most abundant) (Fig. 3d, third and fourth columns). Furthermore, the heatmaps clearly illustrate the strong correlation between the presence of Ring1B and that of ubH2A, as previously reported51,53,104. Both the heatmap and metagene analysis illustrate that ubH2A levels are markedly increased in Usp16−/− ESCs (Fig. 3d,e, f). Visualization of H2A signal did not reveal changes in H2A signal in Usp16+/− or Usp16−/− ESCs (Supplementary Fig. 4b, right two columns), indicating the specificity of the detected ubH2A changes. Further evidence that Usp16 and Ring1B competitively regulate ubH2A levels was provided by conventional ChIP-qPCR assay. As shown in Supplementary Fig. 4d, overexpression of Ring1B and Usp16 increased the levels of these proteins at gene promoters (compare Usp16 and Ring1B panel). Importantly, overexpression of Ring1B correlates with an increase of ubH2A and overexpression of Usp16 correlates with a decrease of ubH2A at gene promoters (Supplementary Fig. 4d, compare ubH2A and IgG panel). Consistent with a role of ubH2A in gene repression, ubH2A enriched genes have decreased expression in response to Usp16 knockout (Fig. 3g, compare Usp16+/− and
Usp16 knockouts). In summary, these data reveal that Usp16 binds to gene promoter, positively correlates with gene expression, and inversely correlates with Ring1B binding and H2A ubiquitination levels.

**Usp16−/− ESCs are defective for differentiation**

Although Usp16 is not required for ESC identity, Usp16 knockout is lethal during early embryogenesis in mice, suggesting that Usp16 knockout may affect ESC differentiation. To determine whether Usp16 regulates ESC differentiation, we differentiated Usp16−/− and Usp16+/+ ESCs by forming embryoid bodies (EBs) and compared their phenotypes. The number of EBs formed from Usp16−/− ESCs (Usp16−/− EBs) is similar to the number of Usp16+/+ ESC derived EBs (Usp16+/+ EBs), indicating that Usp16 knockout does not affect EB formation (data not shown). The morphology of Usp16+/+ EBs was heterogeneous and contained multiple condensed regions refractory to light (Fig. 4a, top panels). In contrast, Usp16−/− EBs was more homogeneous, containing fewer light refractory regions and a smaller condensed central region (Fig. 4a, bottom panels). The transparency difference indicates that, even though both ESCs can form EBs, the internal structure of Usp16−/− and Usp16+/− EBs differ. In addition, Usp16−/− EBs exhibited a distinct growth curve (Fig. 4b). Usp16+/+ EBs grew rapidly, showing a dramatic increase in diameter between day 3 to day 7 before reaching a plateau between day 7-8. In contrast, Usp16−/− EBs grew more slowly, did not exhibit staged growth, and had not reached maximal growth by the end of the experiment (day 9-12). To better identify the differences between these EBs, we sectioned day 12 EBs and analyzed them by hematoxylin and eosin staining. As shown in Fig. 4c, Usp16+/+ EBs exhibited...
heterogeneous staining and showed different germ layer morphology, such as small cavity structures (top and middle panels). In contrast, Usp16−/− EBs stained more homogenously and contained undifferentiated cell clumps, with central regions of dead and apoptotic cells (Fig. 4c, bottom panels). These studies reveal that Usp16−/− ESCs are unable to undergo normal differentiation.

To determine whether the defects in Usp16−/− ESC differentiation are linked to H2A ubiquitination, we analyzed the levels of ubH2A, Usp16, and pluripotent gene markers during EB formation. Usp16 expression exhibits a dynamic change during EB formation, increasing at early stages of ESC differentiation, peaking at day 3-6, and then decreasing to ESC levels after day 6 (Fig. 4d, top panel). ubH2A levels during EB formation were largely inversely correlated with Usp16 levels and decreased during the initial stage of EB formation before gradually returning to the levels in ESCs (Fig. 4d, sixth panel, the intensity of ubH2A was measured and labeled). In contrast, Usp16−/− EBs did not exhibit the same dynamic change of ubH2A levels (Fig. 4d, first and sixth panels). Moreover, while pluripotent genes Oct4, Nanog and Sox2 were quickly repressed in Usp16+/− EBs, becoming undetectable after day 6 (Fig. 4d, second to fourth panels), these genes remained at high levels in day 12 Usp16−/− EBs (Fig. 4d, second to fourth panels). To further confirm the western blot results, we measured the RNA levels of pluripotent genes as well as markers for the three primary germ layers in day 6 and day 12 EBs by RT-qPCR assay. Consistent with the western blot results, the expression of Usp16 in Usp16+/− EBs at day 6 is higher than both day 0 and day 12. There is no detectable expression of pluripotent genes Oct4, Nanog, and Sox2 at day 12 (Fig. 4e), but
lineage-specific markers for ectoderm (Vimentin, Nestin, Igf2), mesoderm (Mef2c, Pax3, MyoD), and endoderm (Gata4, Gata6 and FoxA2) are expressed at high levels (Fig. 4f). In contrast, Usp16−/− EBs did not express these lineage-specific markers and maintained high expression levels of pluripotent genes even at day 12 (Fig. 4e and 4f). Taken together, these data reveal that Usp16 is required to activate cell type-specific gene expression programs and repress the expression of pluripotent markers during ESC differentiation.

**Usp16 is required for removal of the ubH2A mark during ESC differentiation**

PcG proteins regulate gene repression, at least in part, through covalent modification of histones55,61. During cell fate transitions, the repressive chromatin marks H3K27me and ubH2A are selectively resolved at bivalent genes to specify new, lineage specific gene expression profiles. Our studies on EBs formed from Usp16−/− ESCs suggest an intriguing working model wherein Usp16-mediated H2A deubiquitination relieves ubH2A-mediated repression to activate genes required for proper cell lineage commitment. Our studies revealed that Usp16−/− ESCs fail to activate germ layer markers when stimulated to differentiate (Fig. 4f). To determine the effects of Usp16 deletion on gene expression during differentiation, we performed RNA-seq on day 12 EBs. In contrast to the relatively small numbers of differentially expressed genes in Usp16−/− ESCs, 2,371 genes exhibited significant changes (Log₂ [Fold change] >2), with 785 genes up-regulated and 1,586 genes down-regulated (Fig. 5a). Gene ontology analysis revealed that the majority of down-regulated genes are involved in development (Supplementary Fig. 5a). Genes up-regulated are mainly related to metabolism, biosynthesis or early stage
developmental processes such as mesoderm development, or blastocyst development (Supplementary Fig. 5b). This indicates that Usp16 is required for the normal ESC differentiation process. By comparing the gene expression profiles of Usp16+/+ ESCs and EBs, we identified approximately 1,200 genes up-regulated (Fig. 5b, top panel, fold change>2, black line) and 700 genes down-regulated during normal ESC differentiation (Fig. 5b, bottom panel, fold change>2, black line). These differentially expressed genes reflect the normal reprogramming of the gene expression profile during ESC differentiation. Interestingly, the majority of these genes are misregulated during in vitro differentiation of Usp16−/− ESCs into EBs (Fig. 5b, top and bottom panels, red dots). Thus, Usp16 is required for the proper regulation of gene expression during ESC differentiation.

Since severe defects in differentiation were observed in Usp16−/− EBs, we investigated how Usp16 regulates ESC differentiation by measuring Usp16 binding in Usp16+/− EBs and comparing it to the Usp16 binding profile in Usp16+/+ ESCs. Usp16 binding in both EBs and ESCs is significantly enriched at promoter regions (Fig. 5c, compare to Fig. 3b). Parallel experiments in Usp16−/− EBs did not detect any significant Usp16 binding (Supplementary Fig. 5c), confirming the specificity of the ChIP-seq results. About 40% of Usp16 bound promoters in EBs overlap with promoters bound by Usp16 in ESCs (Supplementary Fig. 5e). Interestingly, 4,481 genes, which were not bound by Usp16 in ESCs were bound by Usp16 in EBs (Supplementary Fig. 5e). These newly bound genes are enriched for transcription regulation, embryonic development, various developmental processes, regulation of biosynthetic process, and regulation of
macromolecule biosynthetic process (Supplementary Fig. 5f). These data indicate that Usp16 might be dynamically recruited to novel targets to regulate ubH2A levels and gene expression during ESC differentiation. Importantly, many of these genes are activated during ESC differentiation. To determine the importance of ubH2A in ESC differentiation, we performed ChIP-seq analysis of ubH2A in Usp16\(^{+/+}\) and Usp16\(^{-/-}\) EBs. Consistent with the western blot results, we found that ubH2A levels at transcription start sites are significantly increased in Usp16\(^{-/-}\) EBs compared to Usp16\(^{+/+}\) EBs (Fig. 5d). Interestingly, only 57 transcription start sites are significantly enriched for ubH2A in Usp16\(^{+/+}\) EBs, while 3,100 transcription start sites are enriched for ubH2A in Usp16\(^{-/-}\) EBs (Fig. 5d). This demonstrates that Usp16 is required for global H2A deubiquitination during ESC differentiation. The low levels of ubH2A in Usp16\(^{+/+}\) EBs may indicate that cells in EBs are actively undergoing cell fate transition and, therefore, not many genes are repressed. Parallel experiments with H2A antibody reveal the specificity of the ubH2A ChIP-seq (Supplementary Fig. 5d, middle two columns). During ESCs differentiation, ubH2A marks were largely removed (Fig. 5e, compare first two columns, genes were sorted based on ubH2A level in Usp16\(^{+/+}\) ESCs). Deubiquitination of ubH2A is strongly linked to Usp16 binding (Fig. 5e, fourth column). Importantly, Usp16 bound genes remain enriched for ubH2A in Usp16\(^{-/-}\) EBs (Fig. 5e, compare third column with the first two columns). Direct comparison of Usp16 binding revealed that there is a clear increase of Usp16 binding in EBs as compared to ESCs within these genes (Fig. 5e, compare last two columns; qualified in Fig. 5f). The aberrant enrichment of ubH2A in Usp16\(^{-/-}\) EBs correlated with decreased expression of these genes as compared to Usp16\(^{+/+}\) EBs (Fig. 5g). These data suggest that Usp16 is required to remove the ubH2A repressive marks at
developmental genes during ESC differentiation. To confirm our ChIP-seq results, we performed conventional ChIP. As shown in Supplementary Fig. 6, Usp16 binding to promoter regions of Tbx3, HoxD11, and Gsn genes, which are activated during Usp16+/+ ESC differentiation, were significantly reduced in Usp16−/− EBs as compared to Usp16+/+ EBs (compared red lines with dashed red lines). Importantly, the absence of Usp16 binding in Usp16−/− EBs correlated with increased ubH2A levels at promoter regions (compared blue line with dashed blue line). These results suggest that Usp16 is required to resolve the ubH2A repressive marks at developmental genes during ESC differentiation.

The failure to activate lineage-specific gene expression causes differentiation defects in Usp16−/− ESCs

Our studies reveal that Usp16−/− ESCs do not activate lineage-specific genes and repress pluripotency genes when induced to differentiate (Figs. 4 and 5). To determine the primary cause for the undifferentiated phenotype of Usp16−/− EBs, we tested whether knockdown of pluripotent genes or overexpression of lineage-specific genes could rescue this phenotype. Since EB formation requires prolonged culture and selection of infected cells caused substantial cell death during EB formation (data not shown), we tested whether the undifferentiated phenotype of Usp16−/− ESCs could be recapitulated in other culturing systems. For this purpose, we cultured control and Usp16−/− ESCs in gelatin-coated plates without LIF to induce ESC differentiation. As shown in Fig. 6a and 6b, when cultured for 6 to 8 days, control Usp16+/+ ESCs successfully differentiated, as evidenced by the repression of pluripotent genes Oct4 and Nanog, and activation of
lineage-specific genes including Gata4, Gata6, Hoxc6 (control Usp16+/+). In this culture system, Usp16−/− ESCs also failed to differentiate, as evidenced by the high expression of Oct4 and Nanog, and low levels of expression of differentiation-associated genes including Gata4, Gata6, Hoxc6, etc. (Fig. 6a and 6b, Control Usp16+/−). We next tested whether knockdown of Oct4 could rescue the undifferentiated phenotypes. As shown in Fig. 6a, when Oct4 was significantly knocked down in Usp16−/− ESCs, Nanog remains highly expressed and lineage-associated genes Gata4, Gata6, Igfbf5 and Hoxc6 remains silenced in these cells (compare Usp16+/+ and Usp16−/− with Oct4 KD). These data indicate that the failure to repress pluripotent gene expression is unlikely the primary cause for the undifferentiated phenotype of Usp16−/− ESCs.

To determine whether the failure to activate lineage-specific gene expression is the primary cause for the undifferentiated phenotype of Usp16−/− ESCs, we tested whether the ectopic over-expression of Usp16-target genes, which overcome the repressive ubH2A chromatin marks in Usp16−/− ESCs, could rescue the undifferentiated phenotype. For this purpose, we expressed Usp16 target gene, HoxB4, in control wild type Usp16+/+ and Usp16−/− ESCs by lentiviral infection. As shown in Fig. 6b, HoxB4 is highly expressed in differentiated Usp16+/+ and Usp16+/− ESCs when compared to the uninfected ESCs (top panel). Although the expression level of HoxB4 in Usp16−/− ESCs is slightly lower than in Usp16+/+ ESCs, pluripotent gene Oct4 was significantly repressed (Fig. 6b, second panel) and differentiation-associated genes Gata4, Gata6, Runx1, and Gata2 were significantly activated in HoxB4-infected Usp16−/− ESCs (Fig. 6B, third to sixth panels). These data suggest that the inability to activate lineage-specific gene expression, but not
the failure to repress pluripotent genes, is the primary cause for the undifferentiated phenotype of Usp16\(^{-/-}\) EBs.

**Usp16, but not the catalytic inactive mutant, rescues the undifferentiated phenotypes of Usp16\(^{-/-}\) ESCs**

To determine whether defects in Usp16\(^{-/-}\) EB differentiation are due to Usp16 deubiquitinase activity, we electroporated wild type and catalytically inactive Usp16 targeting vectors into Usp16\(^{-/-}\) ESCs. As shown in Fig. 7a, wild-type and mutant Usp16 were expressed at levels similar to that in Usp16\(^{+/+}\) ESCs (compare lane 3, 4 with 1). Expression of wild-type but not enzymatically inactive mutant Usp16 decreased the ubH2A level in Usp16\(^{-/-}\) ESCs (compare lane 3, 4 with 1, 2). When these ESCs were used for EB formation, we found that a significant number of EBs formed from Usp16\(^{-/-}\) ESCs rescued with wild-type Usp16 displayed a differentiated phenotype (Fig. 7b, see insertion for relative ratio of differentiated EBs and top panels for examples) while virtually all EBs formed from ESCs rescued by catalytically inactive Usp16 exhibited an undifferentiated phenotype (Fig. 7b, bottom panel). The differentiated phenotype was further confirmed by RT-qPCR analysis of lineage-specific genes in these EBs. As shown in Fig. 7c, expression of the Igf2, Vimentin, Gata4, Gata6, and Foxa2 genes was significantly increased in EBs formed from Usp16\(^{-/-}\) ESCs rescued by wild-type Usp16. Surprisingly, although there was a modest reduction in the expression of pluripotent genes Oct4 and Nanog, we did not observe a difference between EBs rescued by wild-type or catalytically inactive mutant Usp16 (Fig. 7c). This is likely due to the presence of
both undifferentiated and differentiated EB populations derived from Usp16\textsuperscript{+/−} ESCs rescued by wild-type Usp16 (Fig. 7b, see insertion for relative ratio of differentiated EB).

Based on these studies, we propose a model for Usp16 in regulating ESC gene expression and differentiation (Fig. 7d). In ESCs, Usp16 binds to a large number of genes involved in RNA-processing, metabolic process, chromosome organization. The binding of Usp16 is positively correlated with gene expression levels and inversely correlated with ubH2A levels and Ring1B binding. When ESCs are differentiated, Usp16 is required to reverse ubH2A mediated gene repression and enables gene activation and subsequent ESC differentiation. When Usp16 is absent, the uH2A repressive marks persist at developmental regulators and ESCs do not differentiate.
DISCUSSION

ESCs are valuable resources for regenerative medicine and powerful tools for studying mammalian development. Understanding the molecular mechanisms controlling ESC pluripotency and lineage specification is a prerequisite for its medical application. Epigenetic mechanisms play important roles in establishing and maintaining cell type-specific gene expression programs. Previous studies have revealed critical roles for PRC2 and P?RC1 in repression of lineage-specific gene expression programs and stabilizing pluripotent gene expression programs in ESCs. Previous studies from our laboratory and others reveal that USP16 functions as a histone H2A deubiquitinase to regulate gene expression at specific genes and loci. However, the genome binding profile of Usp16 and the genes regulated by Usp16 genome wide were not assayed in these studies. Recent studies also reveal that in Down’s syndrome, where Usp16 is triplicated, Usp16 may contribute to somatic stem cell defects. Usp16 has also been shown to coordinate with Aurora B-mediated H3S28 phosphorylation to ensure gene transcription in quiescent lymphocytic cells. Understanding the molecular mechanism of Usp16 is necessary for understanding the function of Usp16 in these processes.

In this study, we demonstrated that Usp16 regulates ubH2A levels and gene expression in ESCs. Our data reveal that Usp16 binds to a large number of genes in ESCs, and Usp16 binding is inversely correlated with ubH2A levels and Ring1B binding, and positively correlated with gene expression levels. Usp16 knockout results in a significant increase of ubH2A levels, correlating with a decrease of gene expression.
Although Usp16 is not required for ESC viability and self-renewal, Usp16 knockout ESCs have altered gene expression patterns and are not stable at certain culture conditions. Therefore, this study identifies Usp16 and H2A deubiquitination as critical regulators for gene expression to stabilize pluripotent gene expression programs in ESCs. This study adds Usp16 and H2A deubiquitination to the long list of factors regulating ESC function.

More significantly, our studies revealed that Usp16 and H2A deubiquitination are required for ESC lineage commitment or differentiation. Usp16 knockout ESCs failed to differentiate. Intriguingly, the failure to activate developmental regulators was the primary cause for the undifferentiated phenotype of Usp16$^{-/-}$ ESCs. The failure to activate lineage specific gene expression could be attributed to the inability of Usp16$^{-/-}$ ESCs to remove repressive ubH2A marks from developmental regulators (Figs. 5 and 6). This result highlights the opposing regulatory role of Usp16-mediated ubH2A deubiquitination to PRC1-mediated ubH2A. Therefore, these studies identify Usp16 and H2A deubiquitination as previously uncharacterized epigenetic mechanisms, which function independently of H3K27me and H3K4me, to control ESC lineage commitment. This study not only provides further evidence for ubH2A in gene silencing$^{52,104}$ but also a more comprehensive understanding of the functions of H2A ubiquitination in the regulation of chromatin structures and cell differentiation programs.
Figure 1. Usp16 is required for mouse embryonic development.

a. Schematic representation of the strategies used to generate Usp16 knockout mice.

Partial regions of Usp16 locus (from exon 2 to exon 7) are shown. Exons are shown as filled boxes and LoxP sites as filled triangles. PCR primers used for
genotyping are shown as arrows. Position of cysteine 205, an amino acid essential for Usp16 deubiquitination activity, is indicated.

b. Identification of correctly targeted mouse embryonic stem cells. Top panels, phase contrast of the morphology of correctly targeted Usp16 ESCs as compared to wild type V6.5 ESCs. Bottom panel, images of genotyping results of ESCs as shown in the top panel. PCR reactions were performed using primers as indicated in panel A. Scale car, 50µm.

c. Usp16 is required for mouse viability. Images of PCR genotyping of offspring from Usp16+/− mice intercrossing. The number and ratio of wild type and heterozygote adult mice genotyped were shown (The number in the parenthesis indicates litters examined).

d. Usp16 knockout results in early embryonic lethal. Image and PCR genotyping of E7.5 embryos from Usp16+/− mice intercrossing showing partially re-absorption of Usp16−/− embryos. The number and ratio of embryos genotyped were shown (The number in the parenthesis indicates litters examined).

e. Usp16−/− blastocysts are viable. PCR genotyping of blastocysts from Usp16+/− intercrossing. The number and ratio of blastocysts genotyped were shown under the PCR image (The number in the parenthesis indicates female mice sacrificed)
Figure 2. Usp16 knockout does not affect ESC viability and identity.

a. Schematic representation of the strategy used to delete Usp16 (top panels) and an image of genotyping result of Usp16 deleted ESCs (bottom panel). PCR primers used for genotyping are indicated in the top panel.
b. Phase contrast images (top panels) and alkaline phosphatase staining images (bottom panels) of Usp16+/+ and Usp16−/− ESCs. Scale bar, 50µm.

c. Western blot analysis of Usp16 (top panel), H2A ubiquitination (third panel), and H2B ubiquitination (fifth panel) levels in two independent Usp16+/+ and Usp16−/− ESC lines. The quantitation of ubH2A signals was labeled. Signals in control Usp16+/+ ESCs were arbitrarily set as 1. GAPDH, histone H2A, H2B, and H3 were used as loading controls.

d. Western blot analysis of the levels of pluripotent genes and PRC subunits in two independent control Usp16+/+ and Usp16−/− ESC lines. β-tubulin was used as a loading control.

e. Growth curve of two independent control Usp16+/+ and Usp16−/− ESC lines. 1 x 10^4 ESCs were seeded in 12 well plates and cell numbers were counted every other day. The same seeding procedure was repeated every other day.

f. Fluorescence activated cell sorting (FACS) analysis of control Usp16+/+ and Usp16−/− ESC lines. The percentages of cell populations at each cell cycle phase were labeled.
Figure 3. Usp16 binds to transcription start site, positively correlates with gene expression and inversely correlates with H2A deubiquitination levels.

a. Scatter plot of gene expression levels in Usp16\textsuperscript{+/+} and Usp16\textsuperscript{-/-} ESCs as determined by RNA-seq. The fragments per kilobase of exon model per million mapped reads (FPKM), was calculated for each transcript using Cufflinks and transformed to Log2 (FPKM) as relative RNA level. Genes that are significantly deregulated (Log\textsubscript{2} [Fold change] > 2) in Usp16\textsuperscript{-/-} ESCs were shown in the bottom table.
b. Binding of Usp16 (red line) and IgG (black line) to gene coding regions plus 10kb upstream of the transcription start site (TSS) and downstream of transcription ending site (TES) in Usp16\(^{+/+}\) ESCs. Parallel experiments in Usp16\(^{-/-}\) ESCs were shown in supplementary Fig. 3c.

c. ubH2A (orange line, Usp16\(^{-/-}\); blue line, Usp16\(^{+/+}\)) and H2A (brown line, Usp16\(^{-/-}\); green line, Usp16\(^{+/+}\)) signals across gene transcribing regions plus 10kb upstream of TSS and downstream of TES.

d. Usp16 (first column), Ring1B (second column), and ubH2A (third and fourth columns) distribution at individual genes in Usp16\(^{+/+}\) and Usp16\(^{-/-}\) ESCs. Top row, top 1,000 expressed genes sorted by relative expression (FPKM). Bottom row, bottom 1,000 expressed genes in gene list re-sorted based on Ring1B signal. Analyses span 10kb upstream of the TSS, the gene body, and 10kb downstream of the TES.

e. Binding profiles of ubH2A in bottom 1,000 expressed genes re-sorted by Ring1B binding in Usp16\(^{+/+}\) and Usp16\(^{-/-}\) ESCs.

f. ubH2A distribution at individual genes in Usp16\(^{+/+}\) and Usp16\(^{-/-}\) ESCs. Genes was sorted based on ubH2A signals in Usp16\(^{+/+}\) ESCs.

g. Expression levels of genes in Fig.3f in Usp16\(^{+/+}\) and Usp16\(^{-/-}\) ESCs.
Figure 4. Usp16<sup>−/−</sup> ESCs are defective for differentiation.

**a.** Phase-contrast images of EB morphology during EB formation from Usp16<sup>+/+</sup> and Usp16<sup>−/−</sup> ESCs.
b. Growth curve of EBs formed from Usp16+/+ and Usp16-/- ESCs. The diameter of EBs at each day was measured and compared. For each time point, 50 EB diameters were measured. Means and standard deviations (error bars) are shown.

c. Hematoxylin and eosin staining of day12 EBs formed from Usp16+/+ and Usp16-/- ESCs. Day 12 EBs from Usp16-/- ESCs lack organized structure and failed to differentiate. Scale car, 100µm.

d. Western blot analysis of the levels of Usp16 (top panel), ubH2A (sixth panel), and pluripotent genes (second to fourth panels) during EB formation. ubH2A signals were quantified as in Fig. 2c. Signals in Usp16+/+ EBs at day 0 were arbitrarily set as 1. GAPDH and histone H3 were used as loading controls.

e. Real-time PCR analysis of Usp16 and pluripotent genes expression at day 0, day 6 and day 12 during EB formation of two independent Usp16+/+ and Usp16-/- ESC lines. Gene expression levels at day 0 were arbitrary set as 1.

f. Real-time PCR analysis of lineage-specific genes expression at day 0, day 6 and day 12 during EB formation from two independent Usp16+/+ and Usp16-/- ESC lines. Gene expression levels at day 0 were arbitrary set as 1.
Figure 5. Usp16 is required for resolving ubH2A marks during ESC differentiation.

a. Scatter plot of gene expression levels in EBs formed by Usp16+/+ (black line) and Usp16−/− ESCs (red dots). mRNA expression level was calculated as log₂ (FPKM).

Summary of genes up- or down-regulated in Usp16−/− ESCs was shown in the table.
b. Scatter plot of up- (top panel) and down- (bottom panel) regulated genes during Usp16^{+/+} ESC differentiation (black line) as compared to their expression during Usp16^{-/-} ESC differentiation (red dots). Genes were sorted according to log2 fold change in expression.

c. Binding of Usp16 (red line) and IgG (black line) to gene transcribing regions plus 10kb upstream of TSS and downstream of TES in Usp16^{+/+} EBs. Parallel experiments with Usp16^{-/-} EBs were shown in supplementary Fig. 5d.

d. ubH2A signal (orange line, Usp16^{-/-} EBs; blue line, Usp16^{+/+} EBs) and H2A signal (brown line, Usp16^{-/-} EBs; green line, Usp16^{+/+} EBs) across gene transcribing regions plus 10kb upstream of TSS and downstream of TES.

e. ubH2A (first three columns) and Usp16 (fourth and fifth columns) distribution at individual genes in Usp16^{+/+} ESCs, Usp16^{+/+} EBs, and Usp16^{-/-} EBs. Genes were sorted based on ubH2A binding in Usp16^{+/+} ESCs. Top 1,000 expressed genes were shown.

f. Usp16 signal in Usp16^{+/+} ESCs (red line) and EBs (blue line) of genes as shown in Fig. 5e.

g. A box and whisker plot, as described in Fig. 3a, of the relative expression of ubH2A enriched genes in Usp16^{+/+} and Usp16^{-/-} EBs. P-value is calculated by the Student’s t-test.
Fig. 6. The failure to activate lineage specific gene expression is the cause of undifferentiated phenotype of Usp16<sup>+/−</sup> ESCs.

a. Knockdown of Oct4 has no effect on Usp16<sup>+/−</sup> ESC differentiation. RT-qPCR analysis of genes in differentiated Usp16<sup>+/+</sup> and Usp16<sup>+/−</sup> ESCs with or without infection of shRNA against Oct4.
b. Expression of HoxB4 triggers Usp16−/− ESC differentiation. RT-qPCR analysis of genes in differentiated control and Usp16−/− ESCs with or without infection of HoxB4 lentivirus.
Fig. 7. Usp16, but not the enzymatically inactive mutant, rescued the undifferentiated phenotype of Usp16<sup>−/−</sup> ESCs.

a. Western blot analysis of Usp16 (top panel) and H2A ubiquitination levels (third panel) in Usp16<sup>−/−</sup> ESCs, and Usp16<sup>−/−</sup> ESCs rescued with wild type or enzymatically inactive C205S mutant Usp16. ubH2A signals were quantified as in
Fig. 2c. Signals in Usp16\textsuperscript{+/+} ESCs were arbitrarily set as 1. GAPDH and histone H3 were used as loading controls.

b. Hematoxylin and eosin staining of day12 EBs formed by Usp16\textsuperscript{-/-} ESCs rescued with wild type or enzymatically inactive C205S mutant Usp16. A bar chart summary of EBs exhibiting differentiated phenotype is shown. Scale car, 100\(\mu\)m.

c. RT-qPCR analysis of genes in EBs formed by control, Usp16\textsuperscript{+/+} ESCs, and Usp16\textsuperscript{-/-} ESCs rescued with wild type or enzymatic inactive C205S mutant Usp16.

d. A proposed model for Usp16 and H2A deubiquitination in ESC gene expression and lineage commitment. In ESCs, Usp16 binds to a large number of genes and Usp16 binding inversely correlates with ubH2A levels and positively correlates with gene expression. During ESC differentiation, Usp16 is responsible for reversing H2A ubiquitination at developmental genes, enabling ESC differentiation.
Figure S1 (Yang et al.)

a. Schematic representation of the strategies used to generate Usp16^{−/−} ESCs. Partial regions of Usp16 locus (from exon 2 to exon 7) are shown. Exons are shown as filled boxes, Loxp sites as filled triangles, and FRT sites as empty triangles. PCR primers used
for genotyping are shown as arrows. Position of cysteine 205, an amino acid essential for Usp16 deubiquitination activity, is indicated. b. Phase contrast images of two independent Usp16−/− ESCs. These ESC lines exhibit normal ESC morphology. Scale bar, 50µm. c. Images of PCR genotyping results of control and two independent Usp16−/− ESC lines. Primers used for genotyping are indicated in panel A (CKO-F/R) and Fig. 1a (KO-F/R). d. Real-time PCR analysis of Usp16 expression levels in control and two independent Usp16−/− ESC lines. The expression level of Usp16 in control ESCs was set as 1. e. Western blot assay of Usp16 and pluripotent genes in control and two independent Usp16−/− ESC lines. GAPDH was used as loading control.
Figure S2 (Yang et al.)

**a**

![Box plot showing relative mRNA level with P=3.84x10^-7](image)

**b**

**Up regulated genes (330)**
- myofibril assembly
- macromolecular complex assembly
- chromatin assembly
- nucleosome organization
- DNA packaging
- macromolecular complex assembly

**Down regulated genes (754)**
- regulation of transcription
- regulation of RNA metabolism process
- regulation of nitrogen compound metabolism process
- positive regulation of macromolecule metabolism process
- positive regulation of biosynthetic process
- phosphate metabolism process
- regulation of cell proliferation
- positive regulation of developmental process
- embryonic morphogenesis

**d**

![Venn diagram with 956, 138, 1017](image)

USP16 KO  Ring1A/B KO

**e**

![mLIF Day3 images](image)

USP16^+/+  USP16^-/

- Flattened rough edge colony
- Round smooth edge colony

**c**

![Bar chart showing fold change](image)
Supplementary Figure 2. Usp16 regulates gene expression in ESCs. a. A box and whisker plot of the estimated gene expression levels of transcripts in Usp16<sup>+/+</sup> and Usp16<sup>−/−</sup> ESCs as determined by RNA-seq. The lower end of the box represents the 25<sup>th</sup> percentile, the upper end of the box represents the 75<sup>th</sup> percentile, the bar represents the median, and the upper and lower whiskers represent the 95<sup>th</sup> and 5th percentiles, respectively. The fragments per kilobase of exon model per million mapped reads (FPKM), was calculated for each transcript using Cufflinks. Unexpressed genes (RPKM < 1) were excluded from the comparison. The Student’s t-test was performed after log transformation of the FPKM values. b. Gene ontology analysis of genes up- (top panel) and down- (bottom panel) regulated in Usp16<sup>−/−</sup> ESCs. c. RT-qPCR analysis of selected genes up- (top panel) and down- (bottom panel) regulated in Usp16<sup>−/−</sup> ESCs as compared to Usp16<sup>+/+</sup> ESCs. Levels in Usp16<sup>+/+</sup> ESCs were set as 1 (dished lines). d. Venn diagram showing genes that exhibit expression changes in Usp16 and Ring1A/B knockout ESCs. Gene numbers are labeled inside of the diagrams. e. Usp16<sup>−/−</sup> ESCs tend to differentiate on feeder cell-free plates. Up panel, phase contrast images of Usp16<sup>+/+</sup> and Usp16<sup>−/−</sup> ESCs cultured on feeder-free plates; bottom, numbers and percentage of ESC colonies with distinct morphologies.
Supplementary Figure 3. Usp16 and ubH2A ChIP-sequencing in ESCs. a. Reads count normalized to genomic coverage of control IgG and Usp16 ChIP-Seq on different chromosomes. b. Tag count for control IgG and Usp16 ChIP-Seq, as normalized to genomic coverage, in different genomic structures. Usp16 specifically binds to regions around the transcription start site (from -2kb to +1kb). c. Binding of Usp16 (green line) and IgG (black line) to gene transcribing regions plus 10kb upstream of the transcriptional start site (TSS) and downstream of transcriptional ending site (TES) in Usp16<sup>−/−</sup> ESCs. d. Gene ontology analysis of Usp16 bound genes in ESCs identified by
MACS. e. Venn diagram showing high overlapping between the ubH2A enriched genes identified in this study and in previous publications (Brookes et al, 2012). Genes identified as enriched for ubH2A in Usp16<sup>+/+</sup> and Usp16<sup>−/−</sup> ESCs are shown. Numbers of genes in each category are indicated.
Figure S4 (Yang et al.)

a) **ubH2A enriched genes in Usp16**

<table>
<thead>
<tr>
<th>Category</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>transcription regulator activity</td>
<td></td>
</tr>
<tr>
<td>Homeobox</td>
<td></td>
</tr>
<tr>
<td>transcription regulation pattern</td>
<td></td>
</tr>
<tr>
<td>specification process</td>
<td></td>
</tr>
<tr>
<td>embryonic organ development</td>
<td></td>
</tr>
<tr>
<td>embryonic morphogenesis</td>
<td></td>
</tr>
<tr>
<td>neuron differentiation</td>
<td></td>
</tr>
<tr>
<td>cell fate commitment</td>
<td></td>
</tr>
<tr>
<td>skeletal system development</td>
<td></td>
</tr>
<tr>
<td>chordate embryonic development</td>
<td></td>
</tr>
<tr>
<td>sensory organ development</td>
<td></td>
</tr>
<tr>
<td>tissue morphogenesis</td>
<td></td>
</tr>
<tr>
<td>differentiation</td>
<td></td>
</tr>
<tr>
<td>tube development</td>
<td></td>
</tr>
<tr>
<td>epithelium development</td>
<td></td>
</tr>
<tr>
<td>regulation of cell development</td>
<td></td>
</tr>
<tr>
<td>embryonic appendage morphogenesis</td>
<td></td>
</tr>
<tr>
<td>positive regulation of cell differentiation</td>
<td></td>
</tr>
</tbody>
</table>

b) **Newly ubH2A enriched genes in Usp16**

<table>
<thead>
<tr>
<th>Category</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>glycoprotein</td>
<td></td>
</tr>
<tr>
<td>cell junction</td>
<td></td>
</tr>
<tr>
<td>cell adhesion</td>
<td></td>
</tr>
<tr>
<td>calcium signaling pathway</td>
<td></td>
</tr>
<tr>
<td>pathways in cancer</td>
<td></td>
</tr>
<tr>
<td>ion channel activity</td>
<td></td>
</tr>
<tr>
<td>intracellular signaling cascade</td>
<td></td>
</tr>
<tr>
<td>neuron projection morphogenesis</td>
<td></td>
</tr>
<tr>
<td>axonogenesis</td>
<td></td>
</tr>
<tr>
<td>neurogenesis</td>
<td></td>
</tr>
<tr>
<td>gland development</td>
<td></td>
</tr>
<tr>
<td>cell morphogenesis</td>
<td></td>
</tr>
</tbody>
</table>

c) **Highly expressed genes**

Graph showing gene expression (FPKM) across different conditions.

d) **Relative enrichment**

- **Usp16**
- **ubH2A**
- **Ring1B**
- **IgG**

Graphs illustrating the relative enrichment of different conditions.
Supplementary Figure 4. Usp16-regulated gene expression correlates with changes in ubH2A level. a. Gene ontology analysis of genes identified to be enriched for ubH2A in Usp16<sup>+/+</sup> (left) and newly identified ubH2A enriched genes in Usp16<sup>−/−</sup> ESCs (right). b. Usp16 (first column) and H2A (middle and right columns) distribution at individual genes in Usp16<sup>−/−</sup> ESCs. Top rows, top 1,000 expressed genes sorted by relative expression (FPKM). Bottom, bottom 1,000 genes sorted based on Ring1B signal. Analyses span 10kb upstream of the TSS, the gene body, and 10kb downstream of the TES. c. Binding profiles of ubH2A in top 1,000 expressed genes sorted by gene expression levels in Usp16<sup>+/+</sup> and Usp16<sup>−/−</sup> ESCs. d. ChIP-qPCR analysis of ubH2A enriched genes Pax3, Tbx3, Nkx2.2, HoxA7, Pax9 and HoxD10 after Usp16 overexpression (Red) and Ring1B overexpression (Blue) and control expression vector (Black). IgG was used a negative control. Mean and standard deviations (SD) from three replicates are shown.
Figure S5 (Yang et al.)

(a) Up-regulated genes in Usp16<sup>-/-</sup> EB (1,586)
- negative regulation gene expression
- negative reg. of macromolecule metabolism
- negative reg. of nucleic acid metabolic process
- mesoderm development
- negative reg. of cellular biosynthetic process
- embryonic morphogenesis
- gastrulation
- mesoderm morphogenesis
- endoderm formation
- blastocyst development

(b) Down-regulated genes in Usp16<sup>-/-</sup> EB (785)
- blood vessel development
- skeletal system development
- developmental protein
- lung development
- embryonic morphogenesis
- kidney development
- angio genesis
- gland development
- transcription factor activity
- regulation of cell proliferation
- reproductive developmental process
- differentiation
- epithelium development
- genitalia development
- heart development

(c) Usp16<sup>-/-</sup> EB
- IgG

(d) RPKM
- WT ES
- hH2A

(e) Usp16 bound new target genes in EB
- ESC
- EB

(f) Down-regulated genes in Usp16<sup>-/-</sup> EB (785)
- transcription regulation
- chordate embryonic development
- regulation of biosynthetic process
- regulation of macromolecule biosynthetic process
- cellular macromolecule catabolic process
- regulation of macromolecule metabolic process
- regulation of nitrogen compound metabolic process
- embryonic morphogenesis
- cell cycle
- chromatin modification
- protein transport
- skeletal system development
- heart development
- in utero embryonic development
- chromosome organization
- regulation of RNA metabolic process
- Pathways in cancer
- apoptosis
- embryonic limb morphogenesis
- embryonic organ morphogenesis
- phosphorus metabolic process
- heart development
- tissue morphogenesis
- mesoderm morphogenesis
- kidney development
Supplementary Figure 5. Usp16 is required for ESC differentiation.  

a. Gene ontology analysis of genes up-regulated in Usp16\(^{-/-}\) EBs identified by MACS. 
b. Gene ontology analysis of genes down-regulated in Usp16\(^{-/-}\) EBs identified by MACS. 
c. Binding of Usp16 (green line) and IgG (black line) to gene transcribing regions plus 10 kb upstream of the transcriptional start site (TSS) and downstream of transcriptional ending site (TES) in Usp16\(^{-/-}\) EBs. 
d. H2A (first three columns) and Usp16 (fourth column) distribution at individual genes in Usp16\(^{+/+}\) ESCs, Usp16\(^{+/+}\) EBs, and Usp16\(^{-/-}\) EBs. Genes were sorted as in Figure 5e. Top 1,000 expressed genes were shown. 
e. Venn diagram showing genes enriched for Usp16 binding in Usp16\(^{+/+}\) ESCs and EBs. Numbers of genes in each category are indicated. 
f. Gene ontology analysis of genes newly bound by Usp16 in Usp16\(^{+/+}\) EBs.
Supplementary Figure 6. ChIP-qPCR analysis of Usp16 and ubH2A levels in Usp16 target genes Tbx3, HoxD11, and Gsn in Usp16^{+/+} and Usp16^{-/-} EBs. A diagram of each gene is shown at the bottom.
## Supplementary Table 1

### Primers List

#### Usp16 knock out ES cell and mice

<table>
<thead>
<tr>
<th></th>
<th>Forward primer (5' to 3')</th>
<th>Reverse primer (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usp16 WT allele</td>
<td>GCTGTGAAACACTGAGTCCGT</td>
<td>TTGGTCTGTCCCCCTCTGCAC</td>
</tr>
<tr>
<td>(WT-F, WT-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usp16 KO allele</td>
<td>TCTGCTGCGCTCTGCCCTCCCA</td>
<td>CTCATGAAAGCAGGGCCACGAC</td>
</tr>
<tr>
<td>(KO-F, KO-R)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Usp16 conditional knock out ES cell and mice

<table>
<thead>
<tr>
<th></th>
<th>Forward primer (5' to 3')</th>
<th>Reverse primer (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usp16 WT allele</td>
<td>GACCAAGCAGCTGGTGTAACAGA</td>
<td>CACTTGCTACCACTCCAACCT</td>
</tr>
<tr>
<td>(WT-F, WT-R)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usp16 CKO allele</td>
<td>GTGGGTTCTTGTGTGGTCTG</td>
<td>CACTTGCTACCACTCCAACCT</td>
</tr>
<tr>
<td>(CKO-F, CKO-R)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### RT-qPCR for ESCs and EBs characterization

<table>
<thead>
<tr>
<th>Genes</th>
<th>Forward primer (5' to 3')</th>
<th>Reverse primer (5' to 3')</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>TGGATTTGGACGCATTGGTC</td>
<td>TTTGCACCTGGTGATC</td>
</tr>
<tr>
<td>Usp16</td>
<td>TTGGGCAAATGGTTGATTATTGTAAGA</td>
<td>AAGAAACAAGGTATTCCAGAAATT</td>
</tr>
<tr>
<td>Oct4</td>
<td>GCTACCCCTGGGCGTTCTCT</td>
<td>GGCCGCAGTTTACACATGTTCTC</td>
</tr>
<tr>
<td>Sox2</td>
<td>AAAACCAACACTGGTCTCTCA</td>
<td>CGAAGGCTCACCCTAAGCGCA</td>
</tr>
<tr>
<td>Nanog</td>
<td>AGCCCTCCAGCAGATGCA</td>
<td>GGTGTTAAGGAGGGAGGTCCTAAG</td>
</tr>
<tr>
<td>Vimentin</td>
<td>GGCGTACGGAGAATGTCG</td>
<td>CACTTCTACGAGGATGTTTAGC</td>
</tr>
<tr>
<td>ActinA2</td>
<td>CCAGACACATCAGGGAGTAATG</td>
<td>TCTATCGGATACAATCTGCAAGT</td>
</tr>
<tr>
<td>Igf2</td>
<td>CAGTTTGTGTTGCTGGAGCCA</td>
<td>ACGTTTGCTCCCCGATTTTGT</td>
</tr>
<tr>
<td>Mef2c</td>
<td>GACGTAGCCTGTTGTCCTG</td>
<td>GTGTTGAAATCCGCGAAGGTG</td>
</tr>
<tr>
<td>Pax3</td>
<td>TCCCATGGTCTGGCTCTCTAAAG</td>
<td>CTGCCACGGACGGTCTGCCTAAG</td>
</tr>
<tr>
<td>MyoD</td>
<td>CCCAGACACGACTGCTTTCTCT</td>
<td>CACACCGGCTGCTGCTTCAAG</td>
</tr>
<tr>
<td>Gata4</td>
<td>CACAAGATGAACGGCGATCAAC</td>
<td>CAGGCGTGGTGTTGAGTCTAGT</td>
</tr>
<tr>
<td>Gata6</td>
<td>GGTCCTCATCAGCAGATGAAATG</td>
<td>TGGCAGCAGGACAGTCCAGGGA</td>
</tr>
<tr>
<td>Foxa2</td>
<td>GATGGAAGGCGACGAGGC</td>
<td>GTATGCTGTTCATGCACCTCACCATT</td>
</tr>
<tr>
<td>HoxC6</td>
<td>CACTGAGTACGCGGTGTAATG</td>
<td>CTCGGATTGTAAGTGTTCGTTG</td>
</tr>
<tr>
<td>Igfbp5</td>
<td>TTCCCTCCAGGAGTTCAAAGC</td>
<td>GGTCAATGGTGCGAGGGTACA</td>
</tr>
<tr>
<td>Runx1</td>
<td>GCAGGGAACGATGAAGACTAC</td>
<td>GCAACTTGGGGGATTTGTA</td>
</tr>
<tr>
<td>Gata2</td>
<td>CACCCCGCCGTATTGAATG</td>
<td>CCTGCGAGTCGAGATGGTTG</td>
</tr>
<tr>
<td>--------</td>
<td>---------------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>

**RT-qPCR for ESCs RNA-Seq confirmation**

<table>
<thead>
<tr>
<th><strong>Up-regulated genes</strong></th>
<th><strong>Forward primer (5’ to 3’)</strong></th>
<th><strong>Reverse primer (5’ to 3’)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bclaf1</td>
<td>CGTCCAGAACATATTGCAGG TC</td>
<td>CCCATAAGGGTCGTCTCATTCCTT</td>
</tr>
<tr>
<td>Rbfox2</td>
<td>TCCGAGGAGACCTATCTGAGG</td>
<td>AATCCGCTCTGGTAAACCACA</td>
</tr>
<tr>
<td>Hectd2</td>
<td>GCAGCTTGTGCTGCAGTCTCA</td>
<td>ACTTCAGATGGTAACGCTTCCA</td>
</tr>
<tr>
<td>Igsf9b</td>
<td>TCAGCACCAGGTCTTGTG</td>
<td>GATTACGTGACATCGCAGAAC</td>
</tr>
<tr>
<td>Nhs</td>
<td>GCTCATGTGACACTGTGCG</td>
<td>GTGAGTAACGTGGATGTGCT</td>
</tr>
<tr>
<td>Dnmt3b</td>
<td>CTGTCGAACCCCGACATAGC</td>
<td>CCGGAAACTCCACAGGGTA</td>
</tr>
<tr>
<td>Dmc1</td>
<td>CCCCTGTGTGACAGTCTAAC</td>
<td>GGTAAGATGTGACCCGAAG</td>
</tr>
<tr>
<td>Clcn3</td>
<td>AGCTTACACAGCCACAGTGCG</td>
<td>GTCCGGCCTCTAACAAATTTGTCAT</td>
</tr>
<tr>
<td>Pou2fl</td>
<td>CTGTCACCTCTGTTCCTGAACC</td>
<td>AAAGTTCTCCAACTCATGAAGCA</td>
</tr>
<tr>
<td>Dnmt1</td>
<td>ACTCCCCTCGGGCAGATGCAT</td>
<td>AGGTTCGAGACGACAGAAC</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Down-regulated genes</strong></th>
<th><strong>Forward primer (5’ to 3’)</strong></th>
<th><strong>Reverse primer (5’ to 3’)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tcea1</td>
<td>GAGAAGAAAGTAGTGTCCCGAGC</td>
<td>GCCCCGAGGAAAAGATGAACAT</td>
</tr>
<tr>
<td>Clcn3</td>
<td>GACTGGGTGCGAGAGAAGTG</td>
<td>CATCCTGACCAGGCGTCTAC</td>
</tr>
<tr>
<td>Tpm1</td>
<td>TTGAAGCCAGGCACACAAAGG</td>
<td>TCATACTTCGGGTACGACATCTT</td>
</tr>
<tr>
<td>Cnot2</td>
<td>CTGTGAAGAAGGACAGCGACT</td>
<td>CCACCCCTCTACAAACTCT</td>
</tr>
<tr>
<td>Ubap2l</td>
<td>TTGAGGGGTCCAATATTCCCGC</td>
<td>TCCATGCACCTGGATGATCAT</td>
</tr>
<tr>
<td>Fam122b</td>
<td>AATGGATCTCTGTGTTTTGGAGGC</td>
<td>CACTGTGAGGACGTCGTCA</td>
</tr>
<tr>
<td>Rhox1</td>
<td>GTGTTCACAGGGGGAAGAAG</td>
<td>CTCCATGTTCCGAAAGCCGT</td>
</tr>
<tr>
<td>Taf11</td>
<td>GCAGAAGGTGGACGAGATG</td>
<td>AAAGCTGACCTGGCTGATACATTS</td>
</tr>
<tr>
<td>Rhbdl2</td>
<td>GGTCCACAGGATCGTCTCACA</td>
<td>GGTCCTCCACACCGCATAATAAAT</td>
</tr>
<tr>
<td>Siva1</td>
<td>CGCCCATCGCTGTTCATC</td>
<td>CTACACATCCTTCGCGCATAGTC</td>
</tr>
</tbody>
</table>

**ChIP-qPCR**

<table>
<thead>
<tr>
<th>HoxD11 Locus</th>
<th><strong>Forward primer (5’ to 3’)</strong></th>
<th><strong>Reverse primer (5’ to 3’)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair1</td>
<td>CAGGAATCTGGGTGTGGTGTG</td>
<td>TCCTCCCTCTTTCTCCTCAAT</td>
</tr>
<tr>
<td>Pair2</td>
<td>TATGCCACAGGAGATGCTCAG</td>
<td>CAAAGCTCAGGAGGCTGGAAG</td>
</tr>
<tr>
<td>Pair3</td>
<td>AGCTTCTGGGAGTCTGTTTTT</td>
<td>CTTCTGTGGTCAAGGGGGCGTA</td>
</tr>
<tr>
<td>Pair4</td>
<td>ATTTCTCTCTATGGGTCGAC</td>
<td>GCACCTGTCAAAGATGTGTTCA</td>
</tr>
<tr>
<td>Pair5</td>
<td>TTCCCCTACTCTCTCAACC</td>
<td>GGCGTAGTAGGGAGCGTAGC</td>
</tr>
<tr>
<td>Pair6</td>
<td>ATAGTGCGAGAAGACAGGC</td>
<td>AGGCTGGCACTGAGATAGGA</td>
</tr>
<tr>
<td>Tbx3 Locus</td>
<td>Forward primer (5’ to 3’)</td>
<td>Reverse primer (5’ to 3’)</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Pair1</td>
<td>GAAGAGTTGGACACCTTCTTG</td>
<td>CGGCCAATTATGACGCTATCG</td>
</tr>
<tr>
<td>Pair2</td>
<td>CCCTACCTGGAGAAGGAAC</td>
<td>TTTCAATTCTCTGCTGAGTA</td>
</tr>
<tr>
<td>Pair3</td>
<td>GGAGAGTTTCTACGTTTTCT</td>
<td>GTGCACAGCGTCTGATCATCT</td>
</tr>
<tr>
<td>Pair4</td>
<td>CGCGAGAGCTAGAGGATCTTG</td>
<td>CGGCCAATAAGCCTAACAAC</td>
</tr>
<tr>
<td>Pair5</td>
<td>TTCCAGTTTCCACACCTTCTC</td>
<td>TGGTATGTCCTCTGGACG</td>
</tr>
<tr>
<td>Pair6</td>
<td>GCTCTTCGCTGACTGAAAC</td>
<td>GGCAGGAGATCTGGGAGACAT</td>
</tr>
<tr>
<td>Pair7</td>
<td>GCCTCTGAAGACATGGAAC</td>
<td>CATCTCTGTACCCCGCTGTT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gsn Locus</th>
<th>Forward primer (5’ to 3’)</th>
<th>Reverse primer (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pair1</td>
<td>TCAGAGCGAGGGCAATAGAACC</td>
<td>TGTAAGGCCCTACCTCTCAA</td>
</tr>
<tr>
<td>Pair2</td>
<td>CATCAAGCCCAACAATCTTT</td>
<td>AGTCTGCTGCTGACACACT</td>
</tr>
<tr>
<td>Pair3</td>
<td>TTCAAGAAACACAGCGACAC</td>
<td>CTCTGACTCTGCTGGGCTCT</td>
</tr>
<tr>
<td>Pair4</td>
<td>GAAGCCAATTTTGCTTAGCAA</td>
<td>GAGCTCCTTGCCCATAAATGC</td>
</tr>
<tr>
<td>Pair5</td>
<td>CAGTGAGCAGCCCTGTATGTT</td>
<td>GAAACCCCTAGCTCCAAGGAC</td>
</tr>
<tr>
<td>Pair6</td>
<td>CTTTCAGGAGCTTTCCACAGT</td>
<td>GGGACTTCAAATGAGGAAAA</td>
</tr>
<tr>
<td>Pair7</td>
<td>CACACACATTGGGCACACAGA</td>
<td>GGGACAGAACCCAAACAGAA</td>
</tr>
</tbody>
</table>

ChIP-qPCR for Usp16 and Ring1B over expression in ESCs

<table>
<thead>
<tr>
<th>Forward primer (5’ to 3’)</th>
<th>Reverse primer (5’ to 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pax3</td>
<td>GGATATGACGCGAGTAGT</td>
</tr>
<tr>
<td>Tbx3</td>
<td>GGCAATAAGCCTAACAAC</td>
</tr>
<tr>
<td>Nkx2.2</td>
<td>GCGCGGCGCTAGTTGAA</td>
</tr>
<tr>
<td>HoxA7</td>
<td>GACAACTCCTATACCTATCCTG</td>
</tr>
<tr>
<td>Pax9</td>
<td>ATACTCTGGCTGAGGACA</td>
</tr>
<tr>
<td>HoxD10</td>
<td>GGCAGGACTTCAAGGATCTAC</td>
</tr>
</tbody>
</table>
CHAPTER 2

FUTURE DIRECTIONS

Our studies revealed that Usp16 and H2A deubiquitination play an indispensable role in mouse embryonic development and ESC lineage commitment. We also, for the first time, showed the binding profile of Usp16 across the entire mouse genome. Our studies reveal that the binding of Usp16 is particularly enriched at gene promoters. Interestingly, Usp16 binding inversely correlates with ubH2A levels and positively correlates with gene expression levels. More importantly, Usp16 deleted ESCs fail to differentiate, primarily due to the inability of these ESCs to remove the repressive ubH2A marks and activate developmental regulator genes and the differentiation gene expression program. Therefore, this study adds Usp16 and H2A deubiquitination to the list of epigenetic regulators controlling ESC gene expression and lineage commitment.

One interesting question raised from our study is whether the requirement for Usp16 in removal of ubH2A and gene activation applies to other cell types during development. A recent study has revealed that Usp16, which is located on human chromosome 21 and triplicated in Down syndrome, contributes to somatic stem cell defects. This contribution has been attributed to the effects of Usp16 in gene expression. Using the Usp16
conditional knockout mouse model generated from this study, we explored the potential function of Usp16 in several developmental processes. Our studies revealed that when Usp16 is deleted in the mouse mammary gland by crossing with MMTV-cre mice, the development of mammary gland was not affected. This result seems to conflict with recent studies showing that Usp16 up-regulation in human Down syndrome affects somatic stem cell function\textsuperscript{77}. Since these experiments were carried out in different experimental settings (up-regulation vs down-regulation), the exact function of Usp16 in mammary gland stem cells and mammary gland development remains to be determined.

When Usp16 was deleted in the heart, the development of heart was not affected. Instead we observed a Non-compaction, hypomyocardial wall phenotype. In contrast to these observations, we identified a critical role of Usp16 in hematopoiesis. Specific deletion of Usp16 in the hematopoietic system results in dramatic decreases of the cell numbers of multiple cell lineages. The cell autonomous effect of Usp16 was demonstrated by in vitro colony-forming assay of bone marrow cells. Non-competitive repopulation assays revealed that the majority of mice transplanted with Usp16 deleted bone marrow died. Furthermore, we showed that mice that survived after transplantation of Usp16 deleted bone marrow had an incomplete deletion of Usp16. The critical role of Usp16 in mouse hematopoiesis was further demonstrated by a competitive repopulation assay, where reconstitution of all hematopoietic cell lineages was abolished. We also observed that hematopoietic stem cells were also not preserved in the presence of competitive cells. RNA-Seq studies revealed that Usp16 regulates expression of many genes involved in hematopoiesis. Therefore, Usp16 does not regulate somatic stem cell function uniformly, but instead exhibits a tissue specificity (affecting HSC/progenitor cell differentiation, not
affecting early development of the mammary gland or the heart). The Usp16 conditional knockout mouse models generated from this study are expected to contribute greatly to our understanding of the role of Usp16 and H2A deubiquitination in many important processes.

Since complete deletion of Usp16 causes embryonic lethality, our studies on Usp16 knockout ESCs are purely mechanistic. However, recent studies have revealed that Usp16 up-regulation in human Down syndrome contributes to somatic stem cell defects. Our unpublished data also revealed that Usp16 is down-regulated in human AML patients and cancer patients. Whether down-regulation of Usp16 has any functional significance remains unknown. Our preliminary studies revealed that the Usp16 heterozygous knockout mice, which express Usp16 at 50%, are more resistant to DNA damage. Consistent with this phenotype, we observed that certain p53 target genes were not properly induced in Usp16 knockdown cells. These studies suggest that the levels of Usp16 may affect DNA damage response and the outcome of cancer treatment.

In summary, this study reveals that Usp16 and H2A deubiquitination play critical roles in ESC gene expression and lineage commitment. The conditional knockout mice generated from this study may help reveal the function of Usp16 in other physiological processes.


APPENDICES
NOTICE OF RENEWAL

DATE: January 11, 2012

TO: HENGBIN WANG, Ph.D.
KAUL-402A 0024
FAX (205) 934-0758

FROM: Judith A. Kapp, Ph.D., Chair
Institutional Animal Care and Use Committee (IACUC)

SUBJECT: Title: Role of Ubp-M and H2A Deubiquitination in Chromatin and Cellular Function
Sponsor: NIH
Animal Project Number: 120208431

As of February 24, 2012, the animal use proposed in the above referenced application is renewed. The University of Alabama at Birmingham Institutional Animal Care and Use Committee (IACUC) approves the use of the following species and numbers of animals:

<table>
<thead>
<tr>
<th>Species</th>
<th>Use Category</th>
<th>Number in Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mice</td>
<td>A</td>
<td>1152</td>
</tr>
</tbody>
</table>

Animal use must be renewed by February 23, 2013. Approval from the IACUC must be obtained before implementing any changes or modifications in the approved animal use.

Please keep this record for your files, and forward the attached letter to the appropriate granting agency.

Refer to Animal Protocol Number (APN) 120208431 when ordering animals or in any correspondence with the IACUC or Animal Resources Program (ARP) offices regarding this study. If you have concerns or questions regarding this notice, please call the IACUC office at (205) 934-7692.
THE UNIVERSITY OF ALABAMA AT BIRMINGHAM
Institutional Animal Care and Use Committee (IACUC)

Notice of Approval for Protocol Modification

DATE: July 1, 2014

TO: HENG BIN WANG, Ph. D.
KAUL-402A
(205) 934-5286

FROM: 
Robert A. Kesterson, Ph.D., Chair
Institutional Animal Care and Use Committee (IACUC)

SUBJECT: Title: Ubp-M and H2A Deubiquitination in Hematopoiesis and Leukemia
Sponsor: Leukemia and Lymphoma Society
Animal Project Number: 140709691

On July 1, 2014, the University of Alabama at Birmingham Institutional Animal Care and Use Committee (IACUC) reviewed the animal use proposed in the above referenced application. It approved the modification as described. Personnel: Wei Yang. The sponsor for this project may require notification of modification(s) approved by the IACUC but not included in the original grant proposal/experimental plan; please inform the sponsor if necessary.

The following species and numbers of animals reflect this modification:

<table>
<thead>
<tr>
<th>Species</th>
<th>Use Category</th>
<th>Number In Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mice</td>
<td>A</td>
<td>0 - Procedural modification only</td>
</tr>
</tbody>
</table>

The IACUC is required to conduct continuing review of approved studies. This study is scheduled for annual review on or before July 17, 2015. Approval from the IACUC must be obtained before implementing any changes or modifications in the approved animal use.

Please keep this record for your files.

Refer to Animal Protocol Number (APN) 140709691 when ordering animals or in any correspondence with the IACUC or Animal Resources Program (ARP) offices regarding this study. If you have concerns or questions regarding this notice, please call the IACUC office at (205) 934-7692.

Mailing Address:
CH19 Suite 403
933 19th Street South
(205) 934-7692
Birmingham, AL 35294-0019

Institutional Animal Care and Use Committee (IACUC)
CH19 Suite 403
933 19th Street South
(205) 934-7692
FAX (205) 934-1188
Birmingham, AL 35294-0019