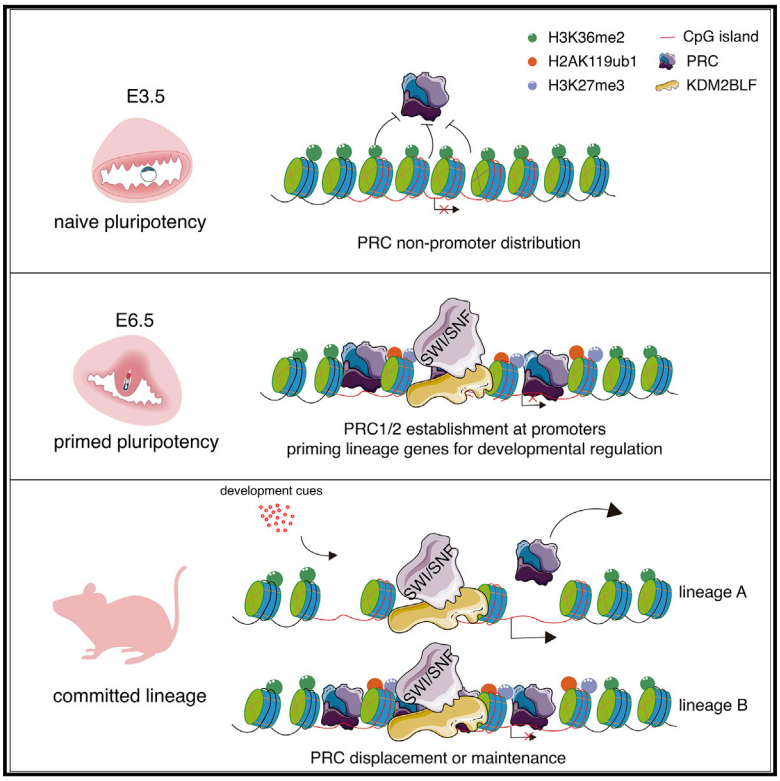
Molecular Cell

Article

CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency

Graphical abstract



Highlights

- KDM2LF expression is spatiotemporally induced during peri-implantation
- KDM2BLF is required for peri-implantation PcG establishment at CGIs
- KDM2BLF orchestrates H3K36me2 demethylation and BAF occupancy for PcG establishment
- KDM2BLF inactivation results in delayed post-implantation development

Authors

Dawei Huo, Zhaowei Yu, Rui Li, ..., Guohong Li, Yong Zhang, Xudong Wu

Correspondence

yzhang@tongji.edu.cn (Y.Z.), wuxudong@tmu.edu.cn (X.W.)

In brief

Chromatin is reconfigured in pluripotent stem cells to prepare for subsequent lineage commitment. Huo et al. demonstrate that spatiotemporally activated KDM2BLF at peri-implantation breaks down the barrier formed by H3K36me2 and nucleosome occupancy, leading to PcG enrichment at non-transcribed CGI-containing promoters.

Huo et al., 2022, Molecular Cell *82*, 1–17
March 17, 2022
https://doi.org/10.1016/j.molcel.2022.01.027

CellPress

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

Molecular Cell

CellPress

Article

CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency

Dawei Huo,[1,2,12] Zhaowei Yu,[3,12] Rui Li,[1] Meihan Gong,[1] Simone Sidoli,[4,5] Xukun Lu,[6] Yuying Hou,[1] Zhongye Dai,[1] Yu Kong,[1] Guifen Liu,[3] Ole Jensen,[4] Wei Xie,[6] Kristian Helin,[7,8] Chaoyang Xiong,[9] Guohong Li,[9,10] Yong Zhang,[3,*] and Xudong Wu[1,2,11,13,*]

[1]State Key Laboratory of Experimental Hematology, The Province and Ministry Co-sponsored Collaborative Innovation Center for Medical Epigenetics, Key Laboratory of Immune Microenvironment and Disease (Ministry of Education), Department of Cell Biology, Tianjin Medical University, Qixiangtai Road 22, Tianjin 300070, China

[2]Department of Neurosurgery, Tianjin Medical University General Hospital, Tianjin 300052, China

[3]Translational Medical Center for Stem Cell Therapy, Institute for Regenerative Medicine, Shanghai East Hospital, School of Life Science and Technology, Shanghai Key Laboratory of Signaling and Disease Research, Tongji University, Shanghai 200092, China

[4]Department of Biochemistry and Molecular Biology, VILLUM Center for Bioanalytical Sciences, University of Southern Denmark, 5230 Odense M, Denmark

[5]Department of Biochemistry, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[6]Center for Stem Cell Biology and Regenerative Medicine, MOE Key Laboratory of Bioinformatics, THU-PKU Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China

[7]Biotech Research and Innovation Centre, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[8]Cell Biology Program and Center for Epigenetics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[9]National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China

[10]University of the Chinese Academy of Sciences, Beijing 100101, China

[11]Tianjin Key Laboratory of Epigenetics for Organ Development of Premature Infants, Tianjin 300450, China

[12]These authors contributed equally

[13]Lead contact

*Correspondence: yzhang@tongji.edu.cn (Y.Z.), wuxudong@tmu.edu.cn (X.W.)

https://doi.org/10.1016/j.molcel.2022.01.027

SUMMARY

Polycomb group (PcG) proteins are essential for post-implantation development by depositing repressive histone modifications at promoters, mainly CpG islands (CGIs), of developmental regulator genes. However, promoter PcG marks are erased after fertilization and *de novo* established in peri-implantation embryos, coinciding with the transition from naive to primed pluripotency. Nevertheless, the molecular basis for this establishment remains unknown. In this study, we show that the expression of the long KDM2B isoform (KDM2BLF), which contains the demethylase domain, is specifically induced at peri-implantation and that its H3K36me2 demethylase activity is required for PcG enrichment at CGIs. Moreover, KDM2BLF interacts with BRG1/BRM-associated factor (BAF) and stabilizes BAF occupancy at CGIs for subsequent gain of accessibility, which precedes PcG enrichment. Consistently, KDM2BLF inactivation results in significantly delayed post-implantation development. In summary, our data unveil dynamic chromatin configuration of CGIs during exit from naive pluripotency and provide a conceptual framework for the spatiotemporal establishment of PcG functions.

INTRODUCTION

How pluripotent stem cells are prepared for lineage commitment is a fundamental question in developmental biology and stem cell biology. Polycomb group (PcG) proteins are conserved transcriptional repressors that form multiprotein complexes, which play pivotal roles in the silencing of non-lineage genes. Biochemical analyses have categorized the polycomb-repressive complexes (PRCs) to at least two subgroups, namely polycomb-repressive complex 2 (PRC2) and PRC1. The two complexes catalyze H3K27 methylation (H3K27me) and H2AK119 monoubiquitylation (H2AK119ub1), respectively (Di Croce and Helin, 2013; Schuettengruber et al., 2017). The core members of each complex are crucial for post-implantation development while generally dispensable for preimplantation development (O'Carroll et al., 2001; Pasini et al., 2004; Voncken et al., 2003). Accordingly, mouse embryonic stem cells (mESCs) derived from the inner cell mass (ICM) of blastocyst have been frequently used as a

model to study how PcG proteins regulate transcription. Consistent with the findings in mouse embryos, PcG proteins are dispensable for mESC self-renewal or naive pluripotency while indispensable for primed pluripotency and subsequent lineage differentiation (Laugesen and Helin, 2014; Weinberger et al., 2016).

In ESCs, the majority of PRCs bind CpG-rich promoters of development regulator genes and typically mark a bivalent chromatin state, i.e., simultaneous histone modifications of H3K4me3 and H3K27me3, which poises these genes for subsequent activation (Laugesen and Helin, 2014; Schuettengruber et al., 2017). In the past two decades, the molecular basis for PRC binding at unmethylated CpG islands (CGIs) has been unveiled. We and others have previously demonstrated that FBXL10/KDM2B, as one of noncanonical PRC1 (ncPRC1)-associated factors, is critical for PRC1 docking at CGIs through its CXXC zinc finger (Farcas et al., 2012; Gearhart et al., 2006; Sánchez et al., 2007; Wu et al., 2013). In PRC2, polycomb-like (PCL) proteins such as PHF1, MTF2, and PHF19 have been found to recognize CGIs through the winged-helix structure (Li et al., 2017; Perino et al., 2018). The identification of these CpG-binding proteins in PRCs provides molecular basis for spontaneous recruitment of PcG proteins to inactive CGI-containing promoters to maintain target gene silencing (Blackledge et al., 2015; Klose et al., 2013; Riising et al., 2014).

Nonetheless, it has remained far less known when and how PcG functions are *de novo* established after fertilization. Recent epigenome profiling in pre- and peri-implantation mouse embryos has illustrated that H3K27me3 enrichment at CGIs is infrequent and unstable prior to implantation (Liu et al., 2016b; Zheng et al., 2016). This establishment of H3K27me3 at promoters is also recapitulated to a large extent in the *in vitro* cultured mESCs when switching from naive to primed pluripotent states (Fidalgo et al., 2016; Joshi et al., 2015; Marks et al., 2012; Weinberger et al., 2016). Moreover, recent studies showed that H2AK119ub1 distribution is largely different from H3K27me3 in preimplantation embryos. However, interestingly, it also switches from a pervasive genomic distribution to a promoter-specific pattern at peri-implantation (Chen et al., 2021; Mei et al., 2021; Zhu et al., 2021). Thus, the two modifications are not robustly established at inactive CGIs until after implantation. The established transcription memory contributes to poising development regulator genes either for subsequent silencing maintenance or resolution for transcription activation. However, it remains unclear how PcG binding and functions at CGI-containing promoters are *de novo* established and synchronized at this time window.

In this study, on the basis of unbiased expression profiling, we excitingly observed that the long isoform of KDM2B (KDM2BLF) emerges at peri-implantation in a spatiotemporally specific manner to reconfigure CGIs. For one thing, it erases H3K36me2 at CGIs based on its instinctive histone demethylase activity. For another, KDM2BLF interacts with BRG1/BRM-associated factor (BAF). Meanwhile, KDM2BLF-mediated H3K36me2 demethylation at CGIs is vital for robust BAF occupancy and the accompanied gain of chromatin accessibility. These coordinated chromatin reconfigurations by KDM2BLF generate permissive nucleation sites for PRC2 as well as PRC1 and allow *de novo* establishment of stable PcG domains at inactive chromatin environment, which is critical for post-implantation embryonic development. Therefore, our study highlights KDM2B as a crucial multifaceted chromatin regulator in the selective acquisition of PcG functions at CGIs during exit from naive pluripotency.

RESULTS

KDM2BLF is specifically activated and depletes H3K36me2 at CGIs during exit from naive pluripotency

Since core members of the PRC2 and PRC1 are expressed at the pre-implantation stage, we speculated that certain accessory factors may emerge at peri-implantation to remove unknown barriers at CGIs for the stable and synchronized PRC1/2 deposition on entry into the primed pluripotency. Hence, we took advantage of the published RNA sequencing (RNA-seq) data of early mouse embryos (Liu et al., 2016a; Zhang et al., 2018) and analyzed the expression dynamics of all PcG genes and genes coding for key associated proteins. As shown in Figure S1A, several genes are significantly upregulated at post-implantation stage compared with pre-implantation stage. We were intrigued by the activation of the specific long isoform of *Kdm2b* (*Kdm2b_1*, later as *Kdm2bLF*), which is transcribed from an alternative promoter to the short isoform (*Kdm2b_2*, or *Kdm2bSF*) (Figure 1A), because of its described ability to directly bind CGIs and its role in PRC1 recruitment (Farcas et al., 2012; Gearhart et al., 2006; Sánchez et al., 2007; Wu et al., 2013). KDM2BLF contains a JmjC domain in addition to the CXXC zinc finger, a plant homeodomain (PHD), the SKP1-binding F-box, and leucine-rich repeats (LRRs) shared by the long and short isoforms (Figure 1B). The CXXC domain of KDM2B is responsible for CGI binding (Farcas et al., 2012; He et al., 2013; Wu et al., 2013), and the LRR region is responsible for the interaction with PRC1 members RING1B and NSPc1/PCGF1 (Wu et al., 2013). KDM2BLF was originally identified as a histone demethylase specific for H3K36me2 (He et al., 2008) similar to its paralog FBXL11/KDM2A (Tsukada et al., 2006). A time course expression analysis showed that the *Kdm2bLF* mRNA levels are barely detectable before implantation, while peak at embryonic day 5.5 (E5.5) epiblast (Epi) and later decrease after E6.5. By contrast, the expression levels of *Kdm2bSF* are modestly decreased from ICM to Epi (Figure 1C), whereas the expression levels of *Kdm2a* remain almost unchanged (Figures S1B and S1C). To simplify the mechanistic studies of the implantation process, we took advantage of *in vitro* culture systems of mESCs in 2i (MEK and GSK3β inhibitors plus leukemia inhibitory factor [LIF]), SL (serum plus LIF), and FA conditions (FGF2 plus Activin) to represent the transition from naive to intermediate and primed pluripotent states (Marks et al., 2012; Nichols and Smith, 2009). Similar expression patterns for *Kdm2b* isoforms are observed in mESCs upon culture condition switch from 2i to SL (Figures S2A and S2B). Then, what is the biological significance of spatiotemporally specific activation of KDM2BLF?

Considering that KDM2BLF acts as the demethylase against H3K36me2, we then conducted STAR chromatin immunoprecipitation sequencing (ChIP-seq) (Zhang et al., 2016) for H3K36me2 in early embryos. As shown in Figures 1D and 1E, although H3K36me2 shows low but still substantial level of

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

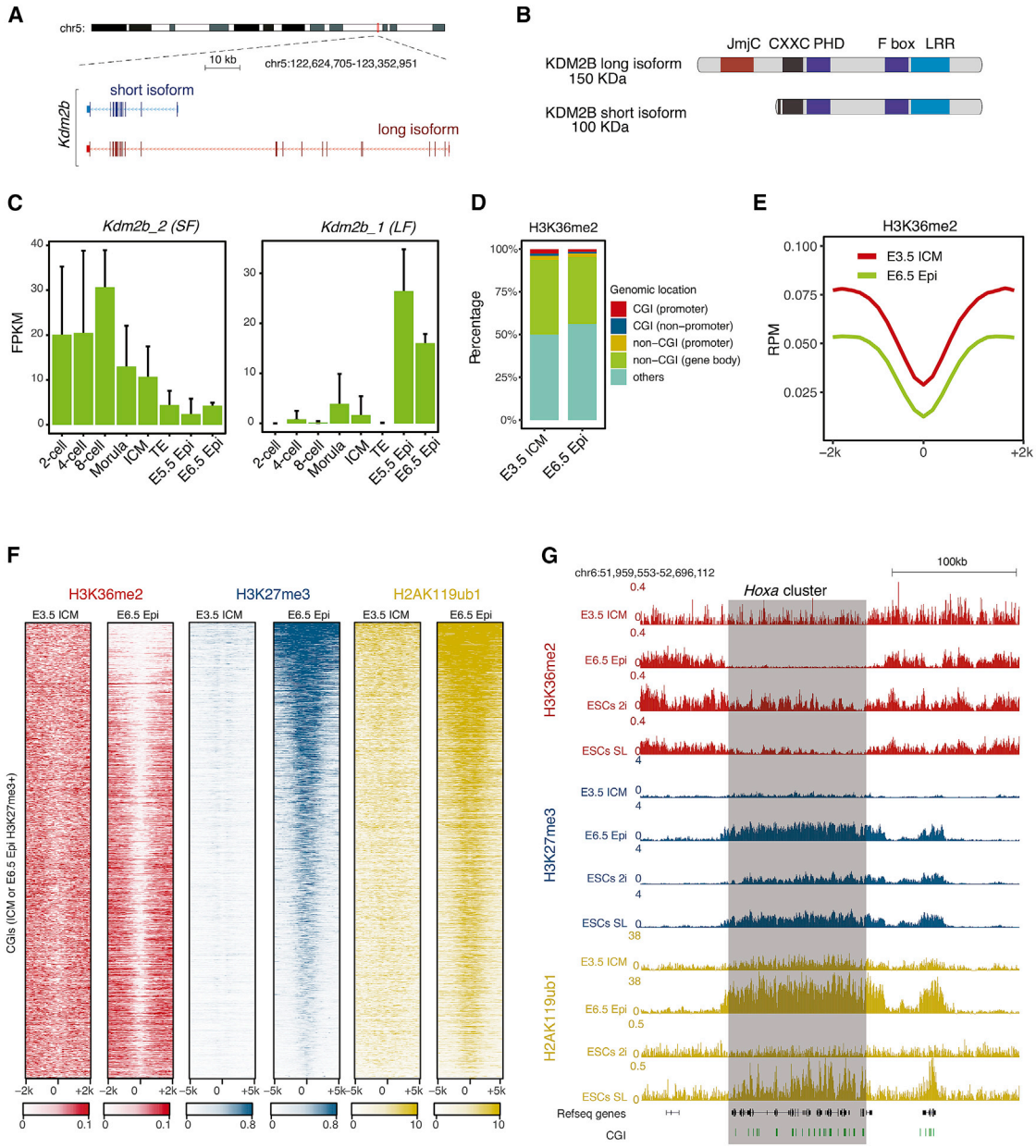


Figure 1. KDM2BLF activation is significantly associated with the establishment of PcG marks at CGIs after implantation

(A and B) Schematic showing UCSC genome reference and domains of KDM2B long and short isoforms.

(C) Bar plot showing the expression levels of Kdm2b isoforms in post-implantation development (E5.5 Epi and E6.5 Epi) and preimplantation development (two-cell, four-cell, eight-cell, morula, ICM, and TE). Signals represent RNA-seq RPM. RNA-seq data were from GSE70605 and GSE76505 with quality control filtration. LF represents long isoform, and SF represents short isoform. Error bars represents standard deviation (SD) for replicates.

(D) The stacked bar plot showing the genomic distribution of H3K36me2 STAR ChIP-seq reads in E3.5 ICM and E6.5 Epi. See more detailed definitions for 5 groups of genomic loci in STAR Methods.

(E) Line charts showing average H3K36me2 signals at CGIs in E3.5 ICM and E6.5 Epi. The x axis represents the distance to CGI center.

(F) Heatmaps showing H3K36me2 signals in E3.5 ICM and E6.5 Epi, H3K27me3, and H2AK119ub1 signals in ICM and E6.5 Epi at H3K27me3+ CGIs (CGIs overlapping with merged H3K27me3 peaks in ICM and E6.5 Epi). Colors represent ChIP-seq RPM, and rows were ranked by ChIP-seq signals in E6.5 Epi. H3K27me3 ChIP-seq data in ICM and E6.5 Epi were from GSE73952 and GSE98149, and H2AK119ub1 CUT&RUN data were from GSE153531.

(G) The UCSC genome browser view of H3K27me3, HA2K119ub1, and H3K36me2 enrichment in E3.5 ICM and E6.5 Epi and ESCs (2i and SL) at *Hoxa* cluster loci. CpG islands loci are shaded in gray, and signals represent ChIP-seq RPM.

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

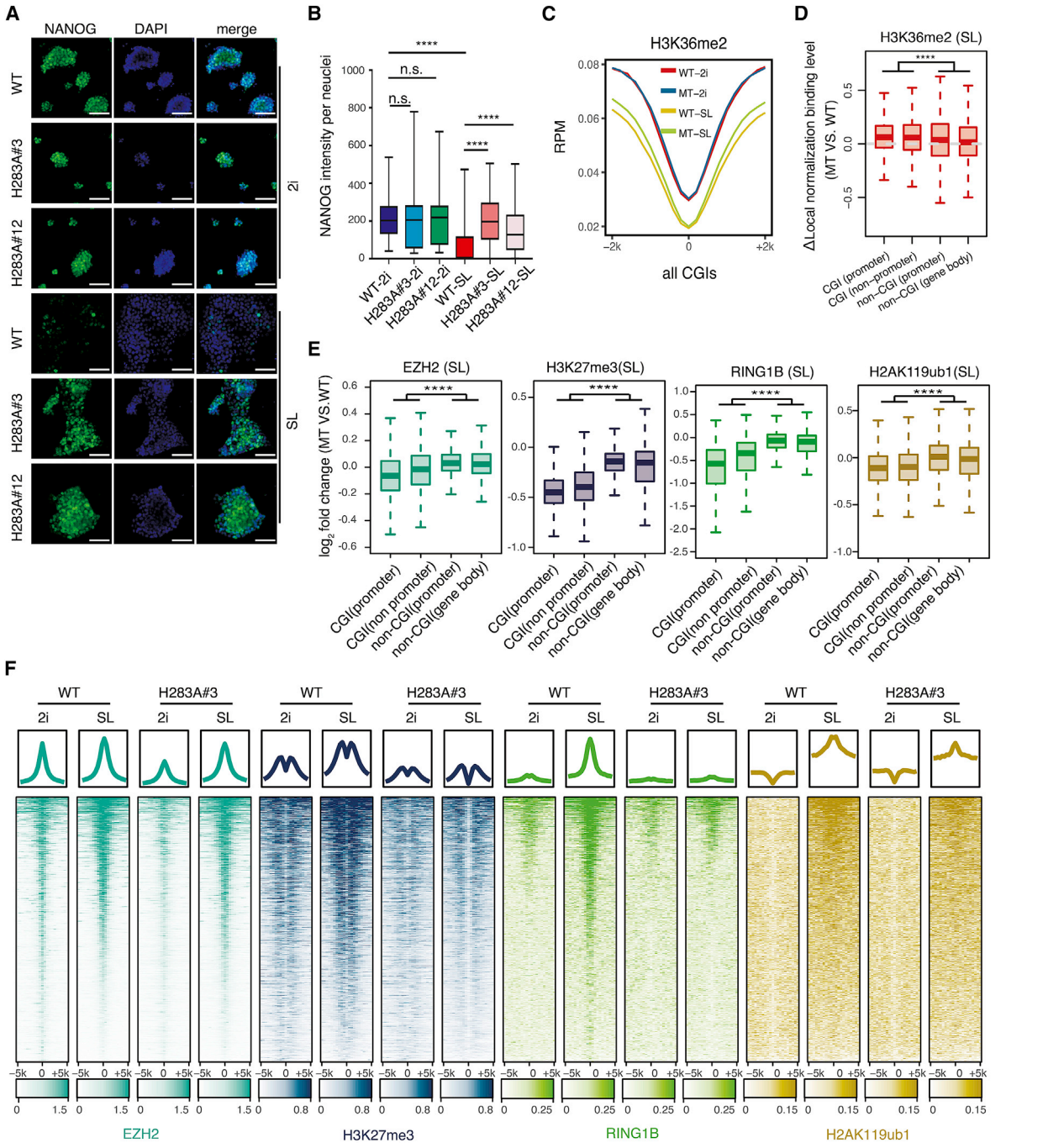


Figure 2. KDM2BLF demethylase activity is required for exit of naive pluripotency and the establishment of PcG functions at CGIs

(A) Immunofluorescence of NANOG in WT and *Kdm2b*-H283A mutant mESCs cultured either in 2i or SL condition. Scale bars, 100 μm.

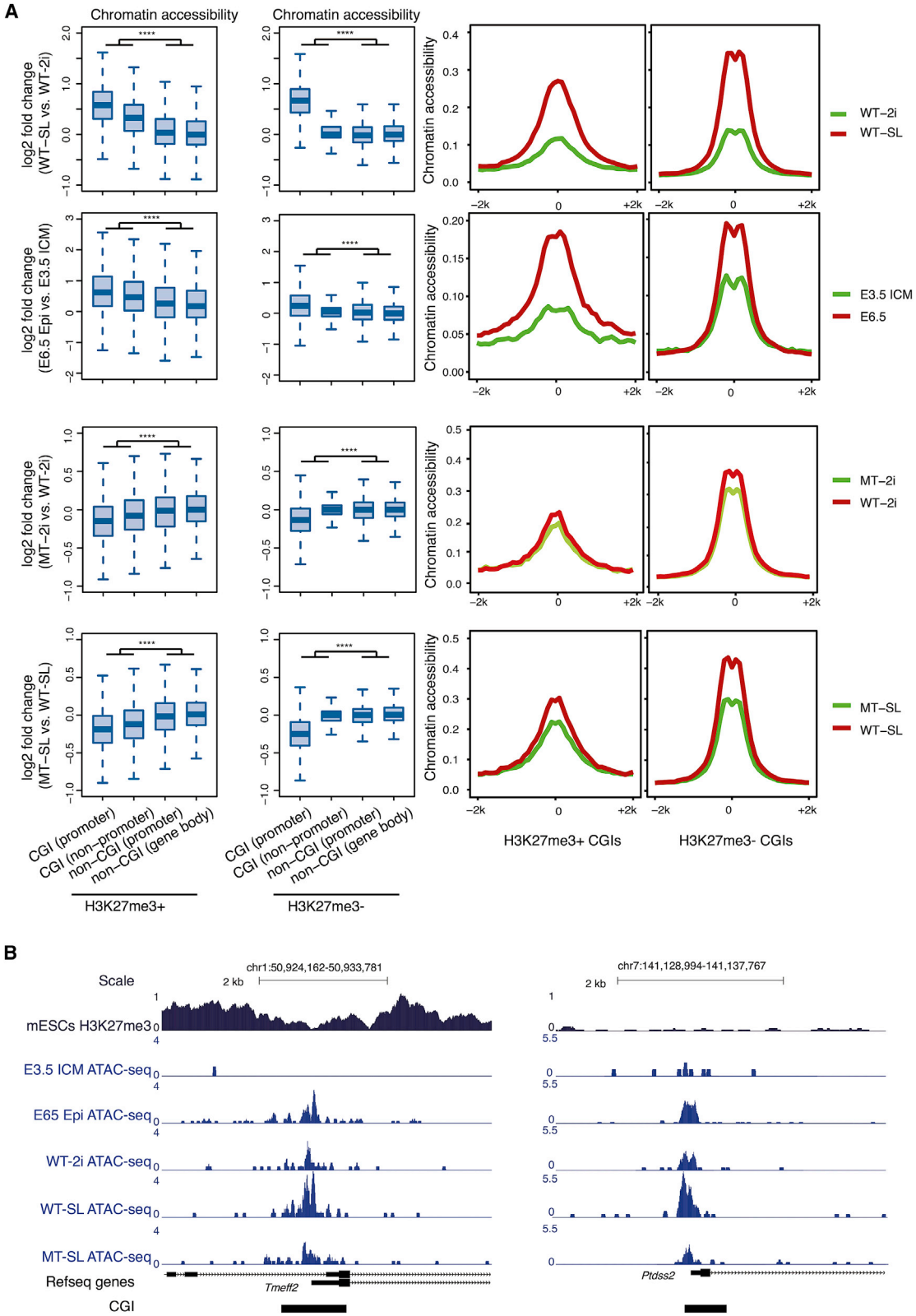
(B) Quantification of NANOG intensity per neuclei of the immunofluorescence assay by the CellProfiler software. Statistical tests were performed using two-tailed unpaired t test with Welch's correction. ****p value < 0.0001, n.s. nonsignificant.

(C) Line charts showing average H3K36me2 signals at CGIs in WT and MT ESCs cultured either in 2i or SL condition. The x axis represents the distance to CGI center.

(D and E) Box plot showing log_2-transformed fold change of H3K36me2 (D), EZH2, H3K27me3, RING1B, and H2AK119ub1 (E) signals in mutation and wild-type mESCs. The fold change of ChIP-seq signals at 4 groups of H3K27me3+ genome loci (overlapping with H3K27me3 peaks) was showed, including two groups of CGIs and two groups of non-CGIs (see more details in STAR Methods). Statistical tests to compare log_2-transformed fold change at CGIs and non-CGIs were performed using one-sided Wilcoxon test, ****p value < 0.0001.

(F) Heatmaps showing EZH2, H3K27me3, RING1B, and H3K27me3 signals in wild-type and mutant mESCs at H3K27me3+ CGIs. Colors represent ChIP-seq RPM, and rows were ranked by ChIP-seq signals in WT SL. Line charts on the top of each heatmap showing average signals.

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027



(legend on next page)

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

enrichment at CGIs in E3.5 ICM, it is nearly eradicated at CGIs at E6.5 Epi stage. Similar dynamics of H3K36me2 was observed upon switch of pluripotent states of *in vitro* cultured mESCs, despite the fact that the bulk levels and H3K36me2 density at intergenic regions were modestly increased (Figures S2B–S2D). The increase of global H3K36me2 levels in ESC from 2i to SL was also previously observed in mass spectrometry (MS) analysis (van Mierlo et al., 2019), suggesting a shift from CGI to widespread non-CGI regions. Together, these findings indicate that increased KDM2BLF expression is associated with H3K36me2 depletion at CGIs and genomic redistribution during exit from naive pluripotency.

Since H3K27me3 and H2AK119ub1 are specifically enriched at promoters at this time window (Chen et al., 2021; Fidalgo et al., 2016; Joshi et al., 2015; Marks et al., 2012; Mei et al., 2021; Weinberger et al., 2016), we wondered whether their enrichment is correlated with H3K36me2 depletion. Indeed, at H3K27me3+ regions in E6.5 Epi, the increase of H3K27me3 and H2AK119ub1 densities from E3.5 ICM to E6.5 Epi stage was specifically accompanied with H3K36me2 depletion at CGIs, but not at non-CGIs (Figures 1F, 1G, and S2E). These dynamics of histone modifications at CGIs were also recapitulated in the 2i/SL-ESCs (Figures 1G and S2F). Furthermore, when we ranked H3K36me2 depletion at H3K27me3+ CGIs, we found that the gain of H3K27me3 and H2AK119ub1 at CGIs is significantly correlated with the loss of H3K36me2 in ESCs as well as in early embryos (Figures S2G and S2H). These genome-wide changes were validated by independent ChIP assays followed by locus-specific qPCR (ChIP-qPCR) analysis in ESCs (Figures S4D, S4F, and S4H), although we could not do similar analysis in early embryos due to the limited amount of material. Together, these data illustrate that PRC1/2 activities are established and synchronized at CGIs in concert with H3K36me2 redistribution from CGIs to non-CGI genomic regions during exit from naive pluripotency.

KMD2BLF demethylase activity is required for PcG establishment at CGIs

The specific upregulation of KDM2BLF expression accompanied with H3K36me2 depletion at CGIs during exit from naive pluripotency prompted us to find out whether the demethylase activity of KDM2BLF is required for the *de novo* establishment of PcG binding at this transition. To do this, we first generated ESC lines expressing the catalytically inactive mutant of KDM2B with a point mutation of His283 to Ala (H283A) by targeting the endogenous *Kdm2b* gene (Figure S3A). The expression levels of H283A-mutant (MT) were comparable with WT, and the mutant cells adapted to the culture condition switch from 2i to SL to FA (Figure S3B). Two positive subclones (#3 and #12) were used for functional characterization. The H283A-MT ESCs keep a similar proliferation rate to WT (Figure S3C). However, unlike WT, the H283A-MT ESCs cannot properly exit from naive pluripotency as manifested by the round and packed colonies (Figures S3D and S3E), significantly more homogenous expression of NANOG than WT in SL condition (Figures 2A and 2B). RNA-seq analysis demonstrated that in H283A-MT (referred to as MT) SL-ESCs, the primed pluripotent genes are less sufficiently activated than WT SL-ESCs, whereas the naive pluripotent genes remain as high as in WT 2i-ESCs (Figure S3F). Even when switching to FA condition, we found that the cell morphology and NANOG expression in H283A-MT cells still remain naive-like states (Figures S3G and S3H). These data suggest that the activity of spatiotemporally activated KDM2BLF is required for the exit from naive pluripotency.

To understand how the catalytic activity of KDM2BLF affects PcG establishment at CGIs, we first confirmed that the mutation does not affect its interaction with NSPc1 or its specific localization at CGIs (Figures S4A and S4B). Subsequently, we profiled H3K36me2, EZH2, RING1B, H3K27me3, and H2AK119ub1 in WT and MT ESCs by ChIP-seq, respectively, in 2i and SL conditions. As expected, H3K36me2 levels were modestly but significantly increased at CGI promoters or non-promoters in MT compared with WT cells in SL condition, although it remained constant or slightly decreased at non-CGI regions (Figures 2C, 2D, and S4C). In contrast, EZH2, RING1B, H3K27me3, and H2AK119ub1 levels were all significantly decreased at CGIs in MT ESCs, especially when grown in SL condition, although they remained unaffected at non-CGI regions (Figures 2D–2F and S4C). All these changes at CGIs were validated by independent ChIP-qPCR analyses (Figures S4D–S4H). Notably, these changes are not due to the deregulated expression of EZH2 or RING1B (Figure S4I). Moreover, in contrast to the local changes at CGIs, the global levels of H3K36me2 are even modestly decreased, whereas H3K27me3 levels increased in MT compared with WT, irrespective of culture conditions (Figure S4I), indicating a redistribution of H3K36me2 and PcG marks upon KDM2BLF inactivation. These global changes are consistent with previous proteomic profiling of histone modifications at the two pluripotent states (van Mierlo et al., 2019) and consistent with the hypernaive phenotype in KDM2BLF MT mESCs (Figures 2A, 2B, and S3). Collectively, these data strongly suggest that KDM2BLF-mediated H3K36me2 demethylation at CGIs is required for the corresponding establishment of PcG domain. As a result, KDM2BLF inactivation leads to insufficient PcG enrichment at CGIs and delayed exit from naive pluripotency.

It is noteworthy that PcG binding and DNA methylation is generally mutual exclusive (Blackledge et al., 2015; Li et al., 2017; Riising et al., 2014). Moreover, a previous study has shown that KDM2B is selectively bound at nonmethylated

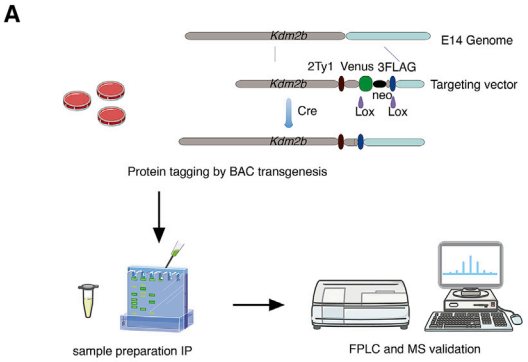
Figure 3. KDM2B-LF demethylates required for establishment of accessibility at CGIs

(A) Boxplots showing $\log_2$-transformed fold change of chromatin accessibility at 4 groups of H3K27me3+ (left) or H3K27me3− (right) genomic loci from WT 2i to SL, E3.5 ICM to E6.5 Epi, and WT 2i/SL to mutant 2i/SL (bottom). Line charts showing the average ATAC-seq signals at H3K27me3+/− CGIs. Public wild-type 2i, SL ATAC-seq data used in the top panel were from GSE92411. Public E3.5 ICM and E6.5 Epi ATAC-seq data were from GSE66390 and GSE125318. Statistical tests to compare $\log_2$-transformed fold change at CGIs and non-CGIs were performed using one-sided Wilcoxon test, ****p value < 0.0001.

(B) The UCSC genome browser view of H3K27me3 enrichment in mESCs and ATAC-seq signal from WT mESCs 2i to SL, ICM to E6.5 Epi, and *Kdm2b* mutant mESCs in SL at H3K27me3-positive-CGI-containing promoter (left) and H3K27me3-negative-CGI-containing promoter (right). Signals represent ChIP-seq RPM.

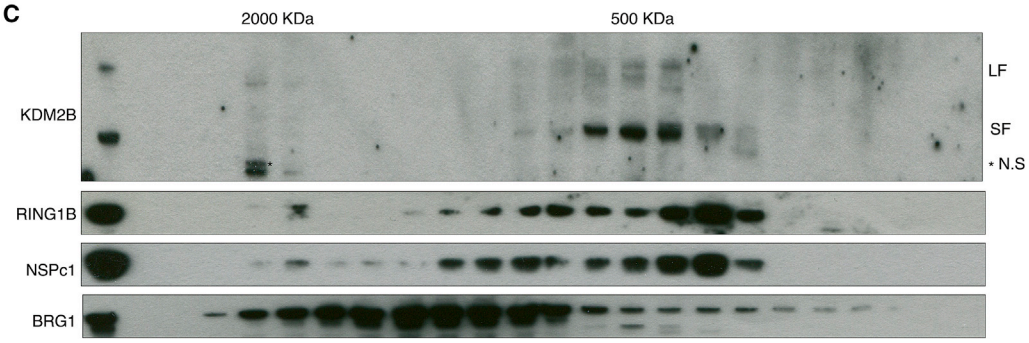
Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

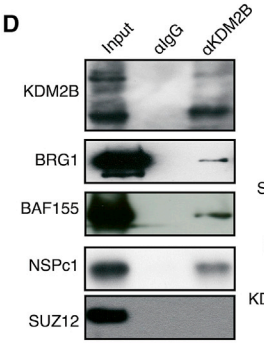
CellPress

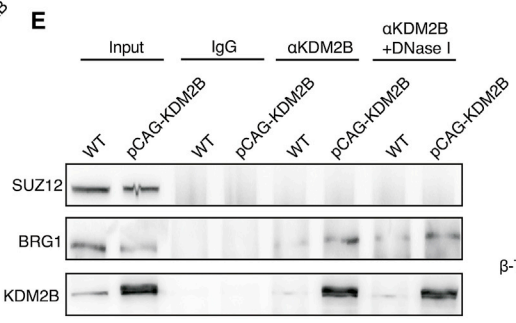


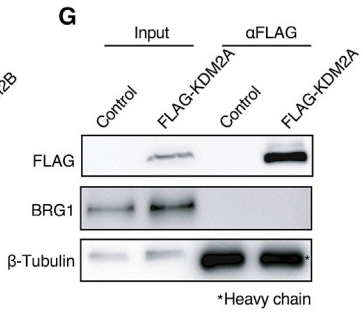
B

Protein	# Peptides	Coverage	MW [kDa]
KDM2B	103	57.60	149.6
PcG complex			
REST	38	35.12	117.7
BCOR	80	56.79	192.5
BCOR1	29	19.71	190.4
RING1B	11	37.5	37.6
NSPc1	11	25.87	30.3
SKP1	4	25.15	18.7
RYBP	1	14.47	24.8
SWI/SNF complex			
SMARCA4/BRG1	5	3.72	180.3
SMARCC1/BAF155	3	5.71	155









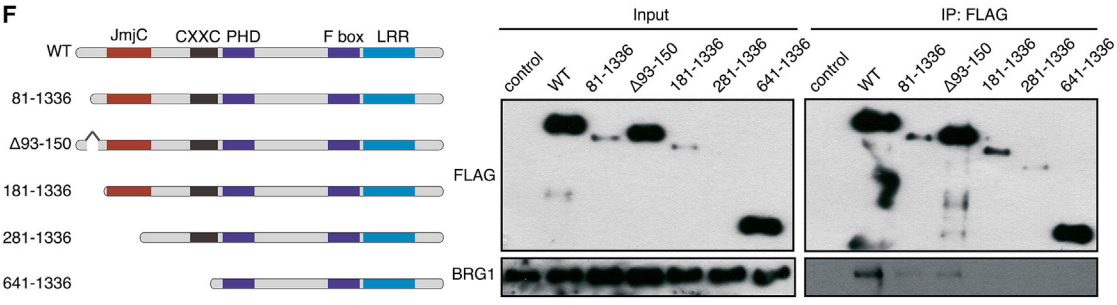


Figure 4. KDM2BLF specifically interacts with BAF

(A) Scheme describes the main steps of the tag transgenesis through BAC engineering techniques and FPLC-MS.

(B) The list of proteins specifically identified in FLAG affinity purifications followed by MS analysis. The numbers of tryptic peptides, coverage of the whole protein, and the molecular weight (MW, KDa) are shown.

(C) Cofraction of KDM2B, RING1B, NSPc1, and BRG1 by gel filtration analysis. The FLAG pull-down fraction from BAC-KDM2B mESCs was eluted by a gel filtration column and then analyzed by WB assay with KDM2B, RING1B, NSPc1, and BRG1 antibodies. Molecular wight are shown at top (KDa).

(legend continued on next page)

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

CGIs, and loss of *Kdm2b* results in gain of CGI DNA methylation (Boulard et al., 2015). Then, we wondered whether KDM2BLF inactivation affects PcG establishment due to gain of DNA methylation at CGIs. To examine this, we performed the genome-wide bisulfite sequencing analysis. The correlation analysis showed that the profiles of DNA methylation at both H3K27me3+ and H3K27me3− CGIs were almost identical in WT and MT cells (R = 0.978) (Figures S5A and S5B). This apparent discrepancy from previous findings (Boulard et al., 2015) may be explained by the fact that our cells are not deleted for *Kdm2b* but express mutant KDM2B that could be sufficient to protect CGIs from *de novo* methylation. Even when comparing the genome-wide changes of H3K36me2 and DNA methylation in MT versus WT cells, we failed to observe any correlation of their dynamics (Figure S5C). These data suggest that the negative effects on PcG establishment at CGIs in KDM2B-inactive ESCs are not caused by the gain of DNA methylation.

KDM2BLF inactivation affects gain of CGI accessibility during exit from pluripotency

Previously, we have shown that PcG proteins are preferentially bound to nucleosome-depleted CGIs of untranscribed genes, rather than inaccessible regions (Riising et al., 2014). Since PcG marks are not established at CGIs until peri-implantation (Chen et al., 2021; Mei et al., 2021; Zheng et al., 2016), we wondered whether there exists a local increase of chromatin accessibility during the transition. To test this, we analyzed assay for transposase accessible chromatin followed by sequencing (ATAC-seq) data of ESCs grown in 2i or SL. Genome-wide analysis showed an increase of chromatin accessibility at CGIs in SL-ESCs compared with 2i-ESCs. In contrast, no significant changes were observed at non-CGI regions. When analyzing published ATAC-seq data in early mouse embryos (Wu et al., 2016; Xiang et al., 2020), we observed similar CGI accessibility increases from ICM to E6.5 Epi (Figures 3A and 3B). These data are also in line with the finding that H3K36me2/3 prevent spurious promoter activation in mouse and yeast (Li et al., 2009; Neri et al., 2017; Venkatesh et al., 2012; Xu et al., 2019).

Then, we assessed whether and how KDM2BLF inactivation would affect the increase in chromatin accessibility. Interestingly, we showed that the chromatin accessibility in MT ESCs is significantly lower when compared with WT cells, especially when cultured in SL condition (Figures 3A and 3B). These data suggest that the KDM2BLF demethylase activity accounts for the increased accessibility at CGIs at peri-implantation, which is associated with timely PcG enrichment and synchronization.

KDM2BLF interacts with BAF

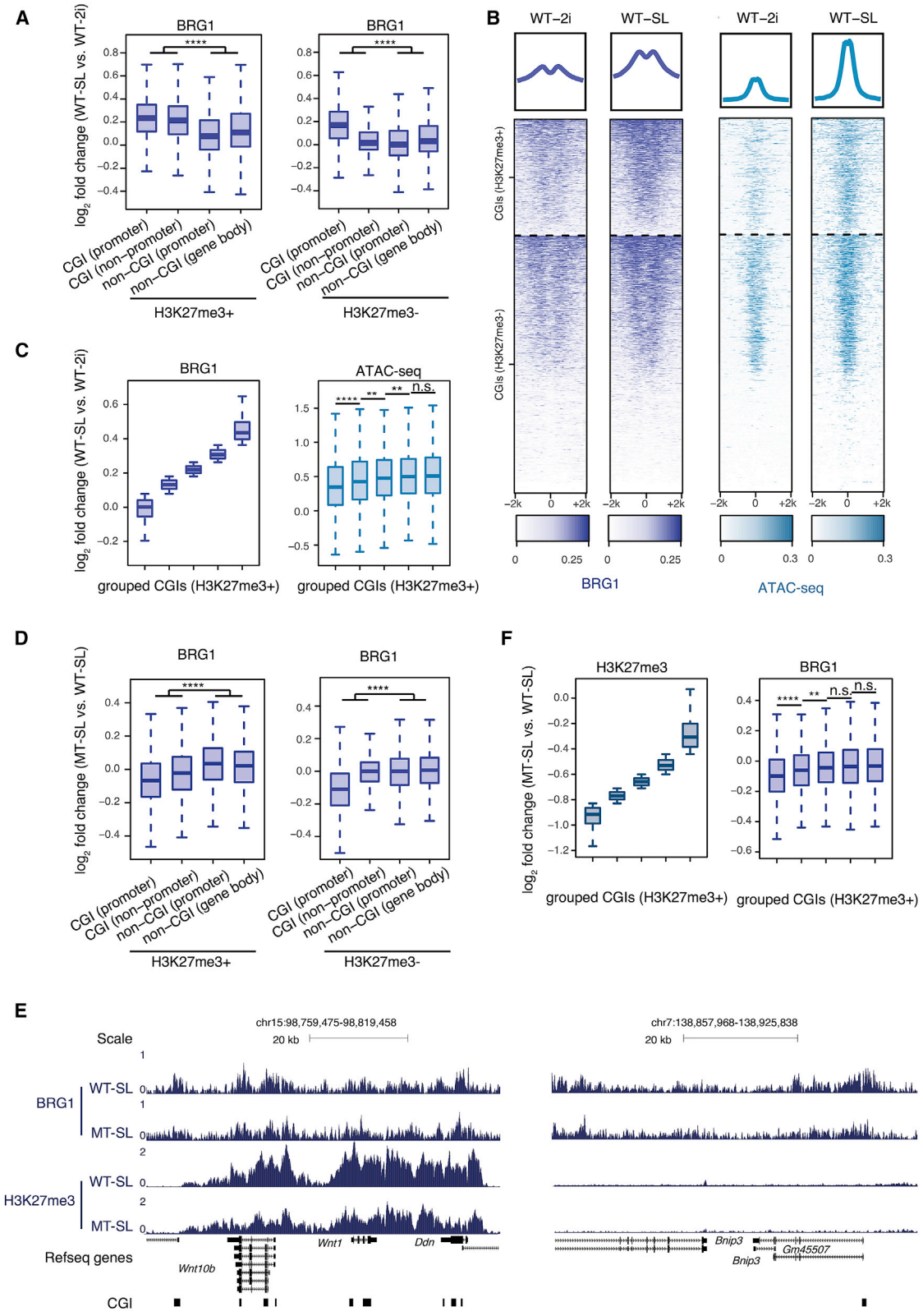
Next, we addressed the mechanism by which the catalytic activity of KMD2BLF could regulate CGI accessibility during exit from naive pluripotency. To do this, we managed to identify potential KDM2B-interacting proteins (Figure 4A). First, we generated an ESC line expressing KDM2B fused to two tags (2 × Ty1 + 3 × FLAG) at the C terminus (Figure S6A; detailed in STAR Methods). Using a FLAG antibody for immunoprecipitation (IP), we showed that RING1B and NSPc1, but not SUZ12, interacted with KDM2B, confirming the specificity of the affinity purification (Figure S6B). Subsequently, MS analysis showed that SMARCA4/BRG1 and SMARCC1/BAF155, the core components of mammalian SWI/SNF (also referred to as Brg/Brahma-associated factors [BAF]) ATP-dependent chromatin remodeling complexes (Ho et al., 2009), were enriched, in addition to the known ncPRC1 components (Figure 4B). Through gel filtration analysis, we found that KDM2BLF specifically co-eluted in a bigger molecular weight (MW) fraction with BRG1 than the other two isoforms and PRC1 components (Figure 4C). Moreover, we validated the interaction between KDM2B and BAF as well as NSPc1 at the endogenous levels (Figure 4D). When we performed IP assay with overexpressed KDM2BLF with or without DNase I in ESCs, we observed a much stronger interaction with BRG1 than the endogenous one. In contrast, neither endogenous nor ectopic KDM2BLF could pull down SUZ12, although they are colocalized at inactive CGIs (Figure 4E). The seemingly weak interaction between KDM2B and BAF is probably due to the transient and low expression levels of KDM2BLF at pluripotent states. Moreover, by expression of different truncated versions of KDM2BLF, we found that the interaction between KDMB2B and BRG1 requires an intact N terminus of KDM2B (Figure 4F). In contrast, we did not detect any interaction between KDM2A and BRG1 (Figure 4G). Taken together, these data confirm that KDM2BLF interacts with BAF, which may play important roles at pluripotent states when KDM2BLF expression is specifically activated.

Kdm2bLF inactivation leads to a decrease of BAF occupancy at CGIs and changes in chromatin accessibility

Given that BAF alters nucleosome positions and generates accessible DNA (Clapier et al., 2017), we examined whether BAF may contribute to the gain of CGI accessibility during exit from naive pluripotency. Globally, BRG1 binds 54.3% of H3K27me3+ CGIs and also 31.2% of H3K27me3− CGIs. Although the role of BAF at active promoters is well studied, we mainly focused on its binding at polycomb targets. Interestingly, BRG1 occupancy at CGIs was significantly increased in SL conditions as compared with 2i conditions, but not at non-CGI regions (Figures 5A, 5B, and S6C). The gain of BRG1 occupancy was significantly correlated with the increased accessibility at CGIs (Figures 5B and 5C). Then, we compared BRG1 occupancy in KDM2BLF-MT and WT ESCs in SL condition. As shown in Figures 5D, 5E, and S6C, BRG1 enrichment is significantly compromised at CGIs in MT ESCs but is unaffected at the non-CGI regions. This is true for both polycomb targets (H3K27me3+) and others (H3K27me3−). More attractively, the decreased H3K27me3 enrichment at CGIs is significantly correlated with reduced BRG1 binding upon

(D) Coimmunoprecipitation of KDM2B with PRC1 and BAF in mESCs. IgG served as a negative control.
(E) KDM2B immunoprecipitation of endogenous (WT) or overexpressed KDM2BLF (pCAG- KDM2B) showing that overexpressed KDM2B pulls down more BRG1 than the endogenous one, but not SUZ12. IgG was served as negative control.
(F) Coimmunoprecipitation showing the interaction between KDM2B truncates and BRG1 in 293T cells.
(G) Immunoprecipitation assay of ectopic KDM2A and BRG1 in 293T cells.

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027



(legend on next page)

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

KDM2B inactivation (Figures 5E and 5F). Taken together, KDM2BLF demethylates H3K36me2 at CGIs meanwhile facilitating BRG1 positioning to gain chromatin accessibility, which is vital for stable PcG enrichment at untranscribed genes. In another word, KDM2BLF couples with BAF to prevent nucleosome retention at CGIs for robust PcG deposition. Further investigations are warranted to clarify how KDM2B coordinates the two multiprotein complexes and how dynamic control of H3K36me2 affects BAF deposition and chromatin remodeling activity *in vivo*.

BRG1 is indispensable for KDM2BLF-induced chromatin remodeling and PcG deposition

We then followed to test if BRG1 is indeed required for gain of accessibility and subsequent PcG enrichment at CGIs. Considering that majority of CGIs are already accessible in ESCs, we sought to determine the locus-specific requirement of BRG1. To do this, we generated a doxycline (Dox) inducible FLAG-dCas9-KDM2BLF fusion system (details in STAR Methods) and tethered the fusion proteins to an H3K36me2+ region by specific sgRNA (Figures 6A and 6B). KDM2B ChIP-qPCR analysis showed that sgRNA successfully directed KDM2BLF-WT and KDM2BLF-MT to the specific target site (Figure 6C). As expected, KDM2BLF-WT but not MT removed H3K36me2 mark (Figure 6D), recruited BRG1 (Figure 6E), increased the local accessibility (Figure 6F), and facilitated NSPc1 binding and PcG enrichment (Figures 6G–6K). Subsequently, we depleted BRG1 in the inducible FLAG-dCas9-KDM2BLF-WT cells (Figure 6B) and performed a similar ChIP-qPCR analysis. Interestingly, even though the H3K36me2 levels were still decreased with less BRG1 (Figures 6D and 6E), the tethering of KDM2BLF (Figure 6C) failed to induce accessibility as in cells expressing BRG1 (Figure 6F). Strikingly, similar to KDM2BLF-MT cells, the deposition of PcG was significantly affected by BRG1 depletion (Figures 6G–6K). These data indicate that BRG1-induced gain of local chromatin accessibility is an essential prerequisite for KDM2BLF-mediated PcG establishment.

KDM2BLF activity is required for PcG establishment and post-implantation embryonic development

To further get insight into the biological significance of *Kdm2bLF* inactivation *in vivo*, we generated mice bearing a homozygous point mutation (H283A) at mouse *Kdm2b* locus by CRISPR-Cas9-mediated genome engineering. We found that homozygous MT mice were not born at the expected Mendelian ratio, although surviving mice are fertile (Figure S7A; only 9.9% of the homozygotes survive post-natally). Through intercross mating of homozygotes, we found that significantly fewer implanted embryos could be observed in MT than WT (Figures 7A and 7B). Around a quarter of MT embryos was significantly smaller than WT at E6.5 (Figures 7C and S7B). Hematoxylin-eosin staining showed that the E6.5 MT embryos were deformed and relatively deficient in adopting the typical post-implantation Epi egg cylinder shape (Figure S7C), indicating early post-implantation development defects. Another quarter of embryos died with improper organogenesis after gastrulation (Figure S7D). The pleotropic developmental defects of KDM2BLF inactive mutants are likely due to the different extent of compensation by KDM2A. Meanwhile, we performed *in vitro* culturing of E2.5 embryos for 28 h. We did not observe significant difference between WT and MT embryos, suggesting that the H283A mutant does not affect pre-implantation development (Figure S7E). Therefore, *Kdm2bLF* inactivation significantly affects post-implantation embryonic development.

Then, we tested if the observed defects in the KDM2BLF mutant mice were related to the compromised CGI accessibility and PcG establishment. E6.5 epiblasts from both WT and MT embryos that showed developmental defects were harvested for profiling of H3K36me2 (STAR ChIP-seq), H3K27me3 (CUT&Tag), and chromatin accessibility (ATAC-seq). Focusing on the top CGIs with H3K27me3 enrichment in WT E6.5 Epi embryos, we found that accompanied with significantly higher residual levels of H3K36me2, the accessibility and H3K27me3 levels were significantly lower in MT embryos (Figures 7D–7F). Furthermore, these chromatin deregulations occur specifically at CGIs, rather than at non-CGI regions (Figure 7G). Taken together, the demethylase activity of KDM2B is indispensable for the peri-implantational establishment of PcG functions at CGIs, which is essential for post-implantation development.

DISCUSSION

PRCs are well recognized as critical forces at CGIs that poise development regulator genes at primed pluripotent state for subsequent transcriptional regulation until lineage specification and are therefore indispensable for post-implantation development (Bernstein et al., 2006; Laugesen and Helin, 2014; Marks et al., 2012; O'Carroll et al., 2001; Pasini et al., 2004; Voncken et al., 2003). In this study, we have taken both cell and animal models

Figure 5. KDM2B-LF inactivation affects BRG1 occupancy at CGIs

(A and D) Boxplot comparing log_2-transformed fold change of Brg1 signals at different genomic regions in WT-SL and 2i ESCs (A), or in WT and MT-SL ESCs (D). H3K27me3+ (with H3K27me3 peaks, left) and H3K27me3− (without H3K27me3 peaks, right) genomic loci were showed. Statistical tests to compare log_2-transformed fold change at CGIs, and non-CGIs were performed using one-sided Wilcoxon test, ****p value < 0.0001.

(B) Heatmaps showing BRG1 (left) and ATAC-seq (right) signals at H3K27me3+ and H3K27me3− CGIs in WT-2i and SL. Colors represent ChIP-seq, or ATAC-seq RPM and rows were ranked by BRG1 signal in SL. Line charts showing average signals.

(C) The association between gained BRG1 and chromatin accessibility from 2i to SL at H3K27me3+ CGIs. Left panel, boxplot showing H3K27me3+ CGIs were classified into 5 groups with increased log_2-transformed fold change of BRG1 signals from 2i to SL. Right panel, boxplots showing log_2-transformed fold change of ATAC-seq signals at 5 groups of loci. Statistical tests were performed using one-sided Wilcoxon test, **p value < 0.01, ****p value < 0.0001, and n.s., nonsignificant.

(E) The UCSC genome browser view of H3K27me3 and BRG1 enrichment from WT and MT mESCs in SL at H3K27me3-positive-CGI containing promoter (left) and H3K27me3-negative-CGI containing promoter (right). Signals represent ChIP-seq RPM.

(F) The association between decreased H3K27me3 and decreased BRG1 from WT to MT at H3K27me3+ CGIs. Grouping and statistical tests were performed as (C), **p value < 0.01, ****p value < 0.0001, and n.s., nonsignificant.

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

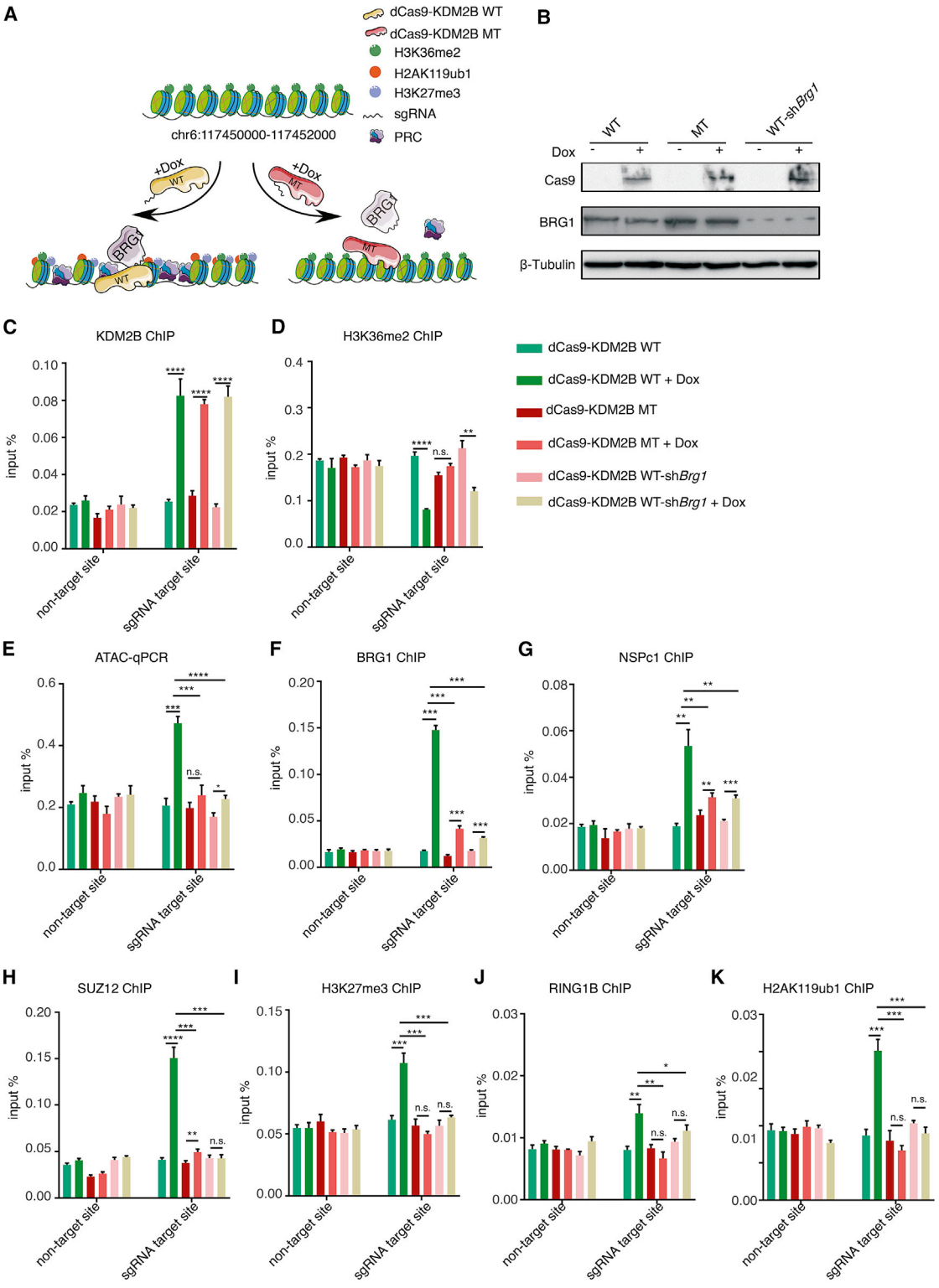


Figure 6. KDM2BLF requires BRG1 to initiate chromatin accessibility for PcG establishment in an inducible H3K36me2 depletion system

(A) Schematic describes the inducible H3K36me2 depletion system, and the doxycycline induced dCas9-KDM2B WT or dCas9-KDM2B mutant fusion protein could be recruited to the H3K36me2+ region by a specific sgRNA. Chromatin accessibility and enrichment of polycomb marks at this locus by dCas9-KDM2B mutant and dCas9-KDM2B with Brg1 depletion were compared to dCas9-KDM2B WT (details in STAR Methods).

(legend continued on next page)

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

to track gene expression changes and chromatin kinetics during exit from naive pluripotency to understand how PcG functions are *de novo* established at CGIs.

CGIs, as core components of most vertebrate gene promoters, play a central role in modulating gene expression by providing a gene regulatory platform that is capable of contributing to a bistable chromatin environment. Being characterized in ESCs or somatic tissues, CGIs have been regarded as chromatin regions with nucleosome depletion and free of DNA methylation and H3K36me2 in physiological conditions (Deaton and Bird, 2011; Fenouil et al., 2012; Rose and Klose, 2014; Stadler et al., 2011). In this study, we unexpectedly provide direct evidence that these chromatin features are not established until peri-implantation. Our data and recent studies showed that the loss of DNA methylation does not necessarily lead to accumulation of PcG marks (Shirane et al., 2020), although DNA methylation is well known to counteract PcG deposition (Bartke et al., 2010). Instead, we raise that inaccessible chromatin with H3K36me2 mark at CGIs is a pre-implantation barrier for PcG nucleation. Although H3K36me2/3 has been known to counteract PRC2 activity (Dorighi and Tamkun, 2013; Schmitges et al., 2011; Streubel et al., 2018), in this study, we confirm that H3K36me2 depletion at CGIs are required for the establishment of PcG repression at inactive loci. Besides, a future genome-wide profiling of H3K36me2 at preimplantation stages will be helpful to dissect how these chromatin reconfigurations are gradually achieved. As a multidomain chromatin regulator, KDM2BLF erases H3K36me2 and couples with BAF to gain local chromatin accessibility and thereby breaks down barriers for PcG establishment. This orchestrated chromatin configuration at CGIs for PcG establishment may be of great significance for robust poising of lineage regulator genes before exposure to complexed development cues (Figure S7F). As a consequence, KDM2BLF activity is indispensable for post-implantation embryonic development.

Clearly, our data demonstrate that KDM2B acts far more than to recruit PRC1. KDM2BSF is already strongly expressed before implantation, and it should be sufficient to recruit PRC1 to CGIs by biochemical instinct. However, upon implantation, the spatiotemporally emerging KDM2BLF provides permissive chromatin environment for PRC2 as well as PRC1, despite the lack of physical interaction with PRC2. Notably, H2AK119ub1 and H3K27me3, the two PcG marks in preimplantation embryos, are not as tightly coupled as in post-implantation ones. H2AK119ub1 rather than H3K27me3 is partially retained at a subset of promoters of developmental genes after fertilization (Chen et al., 2021; Mei et al., 2021; Zhu et al., 2021), which may be due to insufficient H2A deubiquitylation or unstable H2AK119ub1 deposition, which is not strongly affected by H3K36me2 (Lagarou et al., 2008). Given that early embryos are free of lineage differentiation cues, the remnant H2AK119ub1 at development regulator genes may be sufficient for gene silencing at preimplantation stages.

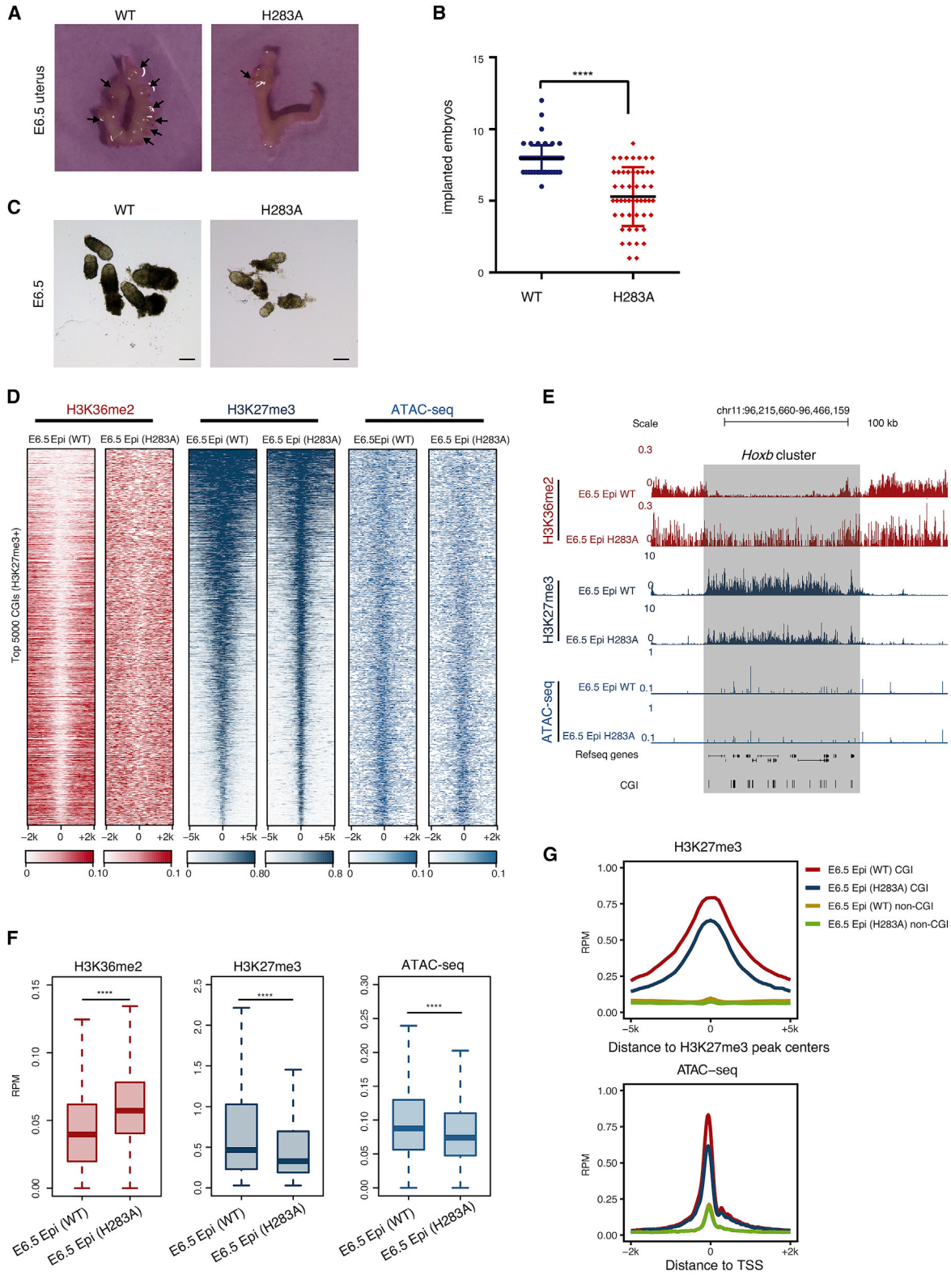
Interestingly, we found that KDM2BLF specifically interacts with BAF. Notably, only BRG1 and BAF155 were identified by MS. It may be accounted by low expression levels of KDM2BLF and insufficient sensitivity of MS detection. Alternatively, KDM2BLF may form a context-specific SWI/SNF complex, which is typically high dynamic (Ho et al., 2009). Another recent study to identify KDM2B partners in LPS-induced macrophages also detected only two components, BRG1 and SMARCC2/BAF170, although they detected far more peptides of BRG1 than our assay (Zhou et al., 2019). As seen from the size exclusion chromatography (Figure 4C), BRG1 exists in a wide range of fractions, whereas KDM2BLF coexists specifically in a unique big-molecular-weight fractions.

In addition to the physical interaction, H3K36me2 depletion by KDM2BLF at CGIs may favor BAF occupancy or activity so as to create chromatin accessibility. Consistent with this hypothesis, pervasive H3K36me2 throughout genome seems to be an indicator of nucleosome retention and chromatin stability (Lee et al., 2013). Moreover, H3K36me2 recruits HDAC complex in yeast (Li et al., 2009) and recruits DNMT3A to catalyze DNA methylation in mammals (Weinberg et al., 2019; Xu et al., 2020). Therefore, H3K36me2 acts as a vital mark to restrict chromatin accessibility and prevent spurious transcription initiation. Accordingly, peri-implantation erasure of H3K36me2 at CGIs is a crucial step to set up regulatory platforms for lineage regulator genes, either transcription poising or activation. Importantly, our data indicate that this chromatin setting requires BAF. In support, the development of *Brg1*-null mouse embryos is arrested at the peri-implantation stage (Bultman et al., 2000). Given that KDM2BLF enhances somatic cell reprogramming (Wang et al., 2011; Zhou et al., 2017), it is worthwhile testing whether it involves BAF-facilitated chromatin remodeling.

Our finding that gain of chromatin accessibility precedes PcG establishment seems counterintuitive. As repressors of transcription, PcG proteins compact chromatin to limit the access of the transcriptional machinery to DNA (Yuan et al., 2012). In agreement with this, CGIs associated with PcG proteins are generally less accessible than CGIs of actively transcribed genes. Nevertheless, PcG proteins do not define chromatin accessibility. First, PRCs sense low accessibility and inactive transcription state for stable binding (Blackledge et al., 2015; Riising et al., 2014). Second, PcG loss in ESCs does not lead to increased accessibility at CGIs (King et al., 2018). Actually, increased nucleosome occupancy caused by either inhibition of ERK1/2 activity (Tee et al., 2014) or by loss of *Brg1* or *Baf250a* in ESCs (Ho et al., 2009; Lei et al., 2015; Weber et al., 2021) has been shown to negatively affect stable PcG deposition at transcriptionally poised CGIs. Thus, it is likely that PRCs prefer locally accessible regions as nucleation sites before spreading for chromatin compaction (Oksuz et al.,

(B) Immunoblot analysis showing expression of induced dCas9 fusion protein and the Brg1 knockdown efficiency. β-tubulin served as loading control.
(C, D, and F) ChIP-qPCR analysis showing the enrichment of dCas9 fusion protein (C), H3K36me2 (D), BRG1 (F), and PcG proteins and marks (G–K) at the designed H3K36me2+ region among the three groups.
(E) ATAC-qPCR analysis comparing chromatin accessibility at the designed H3K36me2+ region among the three groups. For all these qPCR analyses ($n \geq 3$), a non-target site was served as negative control. Statistical tests were performed using two-tailed unpaired t test with Welch's correction. *p value < 0.05, **p value < 0.01, ***p value < 0.001, ****p value < 0.0001, n.s., nonsignificant.

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027



(legend on next page)

2018). Consistently, protein-free linker DNA has been shown to dominate the PRC2-nucleosome interaction *in vitro* (Wang et al., 2017). In this perspective, chromatin remodelers may play crucial roles to control the activity of downstream chromatin regulators as they fine-tune the structure of nucleosomal arrays.

Limitations of the study

Our data have unveiled the molecular basis for the establishment and synchronization of H3K27me3 and H2AK119ub1 at non-transcribed CGI-containing promoters at the peri-implantation stage. In this study, we initiated the transcriptomic analysis of PcG-associated genes and focused our functional and mechanistic studies on KDM2BLF. However, we do not exclude the other possibilities underlying these chromatin dynamics. Notably, in addition to the peri-implantation time window, PcG marks are also *de novo* acquired in somatic tissues (Jadhav et al., 2016; Weiner et al., 2016) and KDM2BLF expression may be transiently induced at specific contexts (Li et al., 2020). Thus, it will be interesting to further investigate whether our findings would have broader implications in general cell fate transitions, including even cell reprogramming or malignant transformation.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines and cell culture
  - Generation of Brg1 Knockdown Cell Lines
  - Generation of knock-in Cell Lines with CRISPR-Cas9
  - Protein tagging by BAC transgenesis
  - Mice
- METHOD DETAILS
  - CRISPR/dCas9 Assay
  - Real-Time Quantitative PCR
  - CUT&Tag
  - STAR ChIP-seq
  - Bulk ChIP-seq
  - Library preparation and sequencing strategy
  - ATAC-seq and ATAC-qPCR
  - Immunofluorescence
  - Immunoprecipitation
  - Complex purification coupled with LC-MS/MS analysis
  - AP staining
  - WGBS library preparation
  - ChIP-seq and CUT&Tag data processing
  - RNA-seq data processing
  - ATAC-seq data processing
  - Whole genome bisulfite sequencing (WGBS) data analysis
  - Genomic loci annotation
  - Datasets used in this study
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.01.027.

ACKNOWLEDGMENTS

We thank members of the Wu and Helin labs for insightful comments and discussions over the course of the study. We are grateful to Deqing Hu (TMU) for providing critical comments on the manuscript and for helpful discussions. This study was supported by the National Key Research and Development Program (2017YFA0504102 to X.W.), the National Natural Science Foundation of China (81772676 and 31970579 to X.W.; 32001025 to D.H.; 32030022 to Y.Z.), the Natural Science Foundation of Tianjin Municipal Science and Technology Commission (18JCJQJC48200 to X.W.), Key Research Project of Tianjin Education Commission (2020ZD13 to X.W.), an open grant from Chinese Academy of Medical Sciences (157-Z20-04 to X.W.), and the National Youth Talent Support Program to X.W. The work was also supported by an HHMI International Research Scholar grant (55008737) to G.L. and the Danish National Research Foundation (DNRF82) to K.H.

AUTHOR CONTRIBUTIONS

X.W. conceptualized and supervised the project and wrote the manuscript. Y.Z., K.H., G.L., and W.X. supervised the project. D.H. performed experiments, analyzed the data, prepared the figures, and wrote the manuscript. Z.Y. did the bioinformatics analysis. G.L. conducted WGBS assay. S.S. and O.J. conducted MS and data analysis. R.L., X.L., M.G., Y.H., Z.D., Y.K., and Y.Z. provided assistance in molecular cloning, tissue culture, animal assay, and library preparation. All authors commented on figures and read and edited the manuscript.

Figure 7. Kdm2b mutant mice show defects of post-implantation development and H3K27me3 establishment

(A) Morphology of E6.5 uterus from WT and MT.

(B) Quantification of implanted E6.5 embryos per mouse from WT and MT mice. Statistical tests were performed using two-tailed unpaired t test with Welch's correction. ****p value < 0.0001.

(C) Morphology of E6.5 embryos from WT and Kdm2b mutant (MT) mice. Scale bars, 200 μm.

(D) Heatmaps showing H3K36me2 STAR ChIP-seq, H3K27me3 CUT&Tag, and ATAC-seq signals in WT and MT E6.5 Epi at top 5,000 H3K27me3+ CGIs (H3K27me3+ CGIs were defined in the same way in Figure 1F). Top 5,000 H3K27me3+ CGIs were selected based on average H3K27me3 signals in CGIs. Colors represent RPM, and rows were ranked by H3K27me3 signals in WT E6.5 Epi.

(E) The UCSC genome browser view of H3K27me3, H3K36me2, and ATAC signals in WT and MT E6.5 Epi at *Hoxb* locus. CpG islands loci are shaded in gray, and signals represent ChIP-seq RPM.

(F) Boxplots showing H3K36me2, H3K27me3, and ATAC signals from wild type to mutant at E6.5 Epi stage at top 5,000 H3K27me3+ CGIs. ****p value < 0.0001.

(G) Upper panel, line charts showing average H3K27me3 signals at H3K27me3 peaks in CGIs or non-CGI regions in WT and MT E6.5 Epi. The x axis represents the distance to H3K27me3 peak center. Lower panel, line charts showing average ATAC signals around transcription start site (TSS) in CGIs or non-CGI regions in WT and MT E6.5 Epi. The x axis represents the distance to TSS.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: August 9, 2021
Revised: December 3, 2021
Accepted: January 27, 2022
Published: February 23, 2022

REFERENCES

Atlasi, Y., Megchelenbrink, W., Peng, T., Habibi, E., Joshi, O., Wang, S.Y., Wang, C., Logie, C., Poser, I., Marks, H., et al. (2019). Epigenetic modulation of a hardwired 3D chromatin landscape in two naive states of pluripotency. Nat. Cell Biol. *21*, 568–578.

Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell *143*, 470–484.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315–326.

Blackledge, N.P., Rose, N.R., and Klose, R.J. (2015). Targeting Polycomb systems to regulate gene expression: modifications to a complex story. Nat. Rev. Mol. Cell Biol. *16*, 643–649.

Boulard, M., Edwards, J.R., and Bestor, T.H. (2015). FBXL10 protects Polycomb-bound genes from hypermethylation. Nat. Genet. *47*, 479–485.

Bultman, S., Gebuhr, T., Yee, D., La Mantia, C., Nicholson, J., Gilliam, A., Randazzo, F., Metzger, D., Chambon, P., Crabtree, G., et al. (2000). A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Mol. Cell *6*, 1287–1295.

Chen, Z., Djekidel, M.N., and Zhang, Y. (2021). Distinct dynamics and functions of H2AK119ub1 and H3K27me3 in mouse preimplantation embryos. Nat. Genet. *53*, 551–563.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. Nat. Rev. Mol. Cell Biol. *18*, 407–422.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25*, 1010–1022.

Di Croce, L., and Helin, K. (2013). Transcriptional regulation by polycomb group proteins. Nat. Struct. Mol. Biol. *20*, 1147–1155.

Dorighi, K.M., and Tamkun, J.W. (2013). The trithorax group proteins Kismet and ASH1 promote H3K36 dimethylation to counteract Polycomb group repression in *Drosophila*. Development *140*, 4182–4192.

Farcas, A.M., Blackledge, N.P., Sudbery, I., Long, H.K., McGouran, J.F., Rose, N.R., Lee, S., Sims, D., Cerase, A., Sheahan, T.W., et al. (2012). KDM2B links the polycomb repressive complex 1 (PRC1) to recognition of CpG islands. Elife *1*, e00205.

Fenouil, R., Cauchy, P., Koch, F., Descostes, N., Cabeza, J.Z., Innocenti, C., Ferrier, P., Spicuglia, S., Gut, M., Gut, I., et al. (2012). CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. Genome Res. *22*, 2399–2408.

Fidalgo, M., Huang, X., Guallar, D., Sanchez-Priego, C., Valdes, V.J., Saunders, A., Ding, J., Wu, W.S., Clavel, C., and Wang, J. (2016). Zfp281 coordinates opposing functions of Tet1 and Tet2 in pluripotent states. Cell Stem Cell *19*, 355–369.

Gearhart, M.D., Corcoran, C.M., Wamstad, J.A., and Bardwell, V.J. (2006). Polycomb group and SCF ubiquitin ligases are found in a novel BCOR complex that is recruited to BCL6 targets. Mol. Cell. Biol. *26*, 6880–6889.

He, J., Kallin, E.M., Tsukada, Y., and Zhang, Y. (2008). The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15(Ink4b). Nat. Struct. Mol. Biol. *15*, 1169–1175.

He, J., Shen, L., Wan, M., Taranova, O., Wu, H., and Zhang, Y. (2013). Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. Nat. Cell Biol. *15*, 373–384.

Ho, L., Ronan, J.L., Wu, J., Staahl, B.T., Chen, L., Kuo, A., Lessard, J., Nesvizhskii, A.I., Ranish, J., and Crabtree, G.R. (2009). An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. Proc. Natl. Acad. Sci. USA *106*, 5181–5186.

Hofemeister, H., Ciotta, G., Fu, J., Seibert, P.M., Schulz, A., Maresca, M., Sarov, M., Anastassiadis, K., and Stewart, A.F. (2011). Recombineering, transfection, Western, IP and ChIP methods for protein tagging via gene targeting or BAC transgenesis. Methods *53*, 437–452.

Iacovino, M., Bosnakovski, D., Fey, H., Rux, D., Bajwa, G., Mahen, E., Mitanoska, A., Xu, Z., and Kyba, M. (2011). Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells. Stem Cells *29*, 1580–1588.

Jadhav, U., Nalapareddy, K., Saxena, M., O'Neill, N.K., Pinello, L., Yuan, G.C., Orkin, S.H., and Shivdasani, R.A. (2016). Acquired tissue-specific promoter bivalency is a basis for PRC2 necessity in adult cells. Cell *165*, 1389–1400.

Joshi, O., Wang, S.Y., Kuznetsova, T., Atlasi, Y., Peng, T., Fabre, P.J., Habibi, E., Shaik, J., Saeed, S., Handoko, L., et al. (2015). Dynamic reorganization of extremely long-range promoter-promoter interactions between two states of pluripotency. Cell Stem Cell *17*, 748–757.

Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat. Commun. *10*, 1930.

Kearns, N.A., Genga, R.M.J., Enuameh, M.S., Garber, M., Wolfe, S.A., and Maehr, R. (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. Development *141* (1), 219–223.

King, H.W., Fursova, N.A., Blackledge, N.P., and Klose, R.J. (2018). Polycomb repressive complex 1 shapes the nucleosome landscape but not accessibility at target genes. Genome Res. *28*, 1494–1507.

Klose, R.J., Cooper, S., Farcas, A.M., Blackledge, N.P., and Brockdorff, N. (2013). Chromatin sampling—an emerging perspective on targeting polycomb repressor proteins. PLoS Genet. *9*, e1003717.

Lagarou, A., Mohd-Sarip, A., Moshkin, Y.M., Chalkley, G.E., Bezstarosti, K., Demmers, J.A., and Verrijzer, C.P. (2008). dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. Genes Dev. *22*, 2799–2810.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Laugesen, A., and Helin, K. (2014). Chromatin repressive complexes in stem cells, development, and cancer. Cell Stem Cell *14*, 735–751.

Lee, C.H., Wu, J., and Li, B. (2013). Chromatin remodelers fine-tune H3K36me-directed deacetylation of neighbor nucleosomes by Rpd3S. Mol. Cell *52*, 255–263.

Lei, I., West, J., Yan, Z., Gao, X., Fang, P., Dennis, J.H., Gnatovskiy, L., Wang, W., Kingston, R.E., and Wang, Z. (2015). BAF250a protein regulates nucleosome occupancy and histone modifications in priming embryonic stem cell differentiation. J. Biol. Chem. *290*, 19343–19352.

Li, B., Jackson, J., Simon, M.D., Fleharty, B., Gogol, M., Seidel, C., Workman, J.L., and Shilatifard, A. (2009). Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. J. Biol. Chem. *284*, 7970–7976.

Li, H., Liefke, R., Jiang, J., Kurland, J.V., Tian, W., Deng, P., Zhang, W., He, Q., Patel, D.J., Bulyk, M.L., et al. (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. Nature *549*, 287–291.

Li, R., Xia, X., Wang, X., Sun, X., Dai, Z., Huo, D., Zheng, H., Xiong, H., He, A., and Wu, X. (2020). Generation and validation of versatile inducible CRISPRi embryonic stem cell and mouse model. PLoS Biol. *18*, e3000749.

Li, W., Shen, W., Zhang, B., Tian, K., Li, Y., Mu, L., Luo, Z., Zhong, X., Wu, X., Liu, Y., et al. (2020). Long non-coding RNA LncKdm2b regulates cortical neuronal differentiation by cis-activating Kdm2b. Protein Cell *11*, 161–186.

Liu, G., Wang, W., Hu, S., Wang, X., and Zhang, Y. (2018). Inherited DNA methylation primes the establishment of accessible chromatin during genome activation. Genome Res. *28*, 998–1007.

Liu, W., Liu, X., Wang, C., Gao, Y., Gao, R., Kou, X., Zhao, Y., Li, J., Wu, Y., Xiu, W., et al. (2016a). Identification of key factors conquering developmental arrest of somatic cell cloned embryos by combining embryo biopsy and single-cell sequencing. Cell Discov. *2*, 16010.

Liu, X., Wang, C., Liu, W., Li, J., Li, C., Kou, X., Chen, J., Zhao, Y., Gao, H., Wang, H., et al. (2016b). Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. Nature *537*, 558–562.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., Nichols, J., Kranz, A., Stewart, A.F., Smith, A., et al. (2012). The transcriptional and epigenomic foundations of ground state pluripotency. Cell *149*, 590–604.

Mazzoni, E.O., Mahony, S., Iacovino, M., Morrison, C.A., Mountoufaris, G., Closser, M., Whyte, W.A., Young, R.A., Kyba, M., Gifford, D.K., et al. (2011). Embryonic stem cell-based mapping of developmental transcriptional programs. Nat. Methods *8*, 1056–1058.

McQuin, C., Goodman, A., Chernyshev, V., Kamentsky, L., Cimini, B.A., Karhohs, K.W., Doan, M., Ding, L., Rafelski, S.M., Thirstrup, D., Wiegraebe, W., Singh, S., Becker, T., Caicedo, J.C., and Carpenter, A.E. (2018). CellProfiler 3.0: Next-generation image processing for biology. PLoS Biol *16* (7), e2005970.

Mei, H., Kozuka, C., Hayashi, R., Kumon, M., Koseki, H., and Inoue, A. (2021). H2AK119ub1 guides maternal inheritance and zygotic deposition of H3K27me3 in mouse embryos. Nat. Genet. *53*, 539–550.

Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017). Intragenic DNA methylation prevents spurious transcription initiation. Nature *543*, 72–77.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol. Cell. Biol. *21*, 4330–4336.

Oksuz, O., Narendra, V., Lee, C.H., Descostes, N., LeRoy, G., Raviram, R., Blumenberg, L., Karch, K., Rocha, P.P., Garcia, B.A., et al. (2018). Capturing the onset of PRC2-mediated repressive domain formation. Mol. Cell *70*, 1149–1162.e5.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J. *23*, 4061–4071.

Peng, X., Wu, J., Brunmeir, R., Kim, S.Y., Zhang, Q., Ding, C., Han, W., Xie, W., and Xu, F. (2015). TELP, a sensitive and versatile library construction method for next-generation sequencing. Nucleic Acids Res. *43*, e35.

Perino, M., van Mierlo, G., Karemaker, I.D., van Genesen, S., Vermeulen, M., Marks, H., van Heeringen, S.J., and Veenstra, G.J.C. (2018). MTF2 recruits polycomb repressive complex 2 by helical-shape-selective DNA binding. Nat. Genet. *50*, 1002–1010.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842.

Riising, E.M., Comet, I., Leblanc, B., Wu, X., Johansen, J.V., and Helin, K. (2014). Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. Mol. Cell *55*, 347–360.

Rose, N.R., and Klose, R.J. (2014). Understanding the relationship between DNA methylation and histone lysine methylation. Biochim. Biophys. Acta *1839*, 1362–1372.

Sánchez, C., Sánchez, I., Demmers, J.A.A., Rodriguez, P., Strouboulis, J., and Vidal, M. (2007). Proteomic analysis of Ring1B/Rnf2 interactors identifies a novel complex with the Fbxl10/ Jmjd1B histone demethylase and the BcoR corepressor. Mol. Cell. Proteomics *6*, 820–834.

Schmitges, F.W., Prusty, A.B., Faty, M., Stützer, A., Lingaraju, G.M., Aiwazian, J., Sack, R., Hess, D., Li, L., Zhou, S., et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. Mol. Cell *42*, 330–341.

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

Shirane, K., Miura, F., Ito, T., and Lorincz, M.C. (2020). NSD1-deposited H3K36me2 directs de novo methylation in the mouse male germline and counteracts polycomb-associated silencing. Nat. Genet. *52*, 1088–1098.

Stadler, M.B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E.J., Gaidatzis, D., et al. (2011). DNA-binding factors shape the mouse methylome at distal regulatory regions. Nature *480*, 490–495.

Streubel, G., Watson, A., Jammula, S.G., Scelfo, A., Fitzpatrick, D.J., Oliviero, G., McCole, R., Conway, E., Glancy, E., Negri, G.L., et al. (2018). The H3K36me2 methyltransferase Nsd1 demarcates PRC2-mediated H3K27me2 and H3K27me3 domains in embryonic stem cells. Mol. Cell *70*, 371–379.e5.

Sun, D., Xi, Y., Rodriguez, B., Park, H.J., Tong, P., Meong, M., Goodell, M.A., and Li, W. (2014). MOABS: model based analysis of bisulfite sequencing data. Genome Biol. *15*, R38.

Tee, W.W., Shen, S.S., Oksuz, O., Narendra, V., and Reinberg, D. (2014). ERK1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs. Cell *156*, 678–690.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-seq. Bioinformatics *25*, 1105–1111.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. Nature *439*, 811–816.

van Mierlo, G., Dirks, R.A.M., De Clerck, L., Brinkman, A.B., Huth, M., Kloet, S.L., Saksouk, N., Kroeze, L.I., Willems, S., Farlik, M., et al. (2019). Integrative proteomic profiling reveals PRC2-dependent epigenetic crosstalk maintains ground-state pluripotency. Cell Stem Cell *24*, 123–137.e8.

Venkatesh, S., Smolle, M., Li, H., Gogol, M.M., Saint, M., Kumar, S., Natarajan, K., and Workman, J.L. (2012). Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. Nature *489*, 452–455.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc. Natl. Acad. Sci. USA *100*, 2468–2473.

Wang, C., Liu, X., Gao, Y., Yang, L., Li, C., Liu, W., Chen, C., Kou, X., Zhao, Y., Chen, J., et al. (2018). Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. Nat. Cell Biol. *20*, 620–631.

Wang, T., Chen, K., Zeng, X., Yang, J., Wu, Y., Shi, X., Qin, B., Zeng, L., Esteban, M.A., Pan, G., et al. (2011). The histone demethylases jhdm1a/1b enhance somatic cell reprogramming in a vitamin-C-dependent manner. Cell Stem Cell *9*, 575–587.

Wang, X., Paucek, R.D., Gooding, A.R., Brown, Z.Z., Ge, E.J., Muir, T.W., and Cech, T.R. (2017). Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. Nat. Struct. Mol. Biol. *24*, 1028–1038.

Wang, T., Wei, J.J., Sabatini, D.M., and Lander, E.S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. Science *343* (6166), 80–84.

Weber, C.M., Hafner, A., Kirkland, J.G., Braun, S.M.G., Stanton, B.Z., Boettiger, A.N., and Crabtree, G.R. (2021). mSWI/SNF promotes polycomb repression both directly and through genome-wide redistribution. Nat. Struct. Mol. Biol. *28*, 501–511.

Weinberg, D.N., Papillon-Cavanagh, S., Chen, H., Yue, Y., Chen, X., Rajagopalan, K.N., Horth, C., McGuire, J.T., Xu, X., Nikbakht, H., et al. (2019). The histone mark H3K36me2 recruits DNMT3A and shapes the intergenic DNA methylation landscape. Nature *573*, 281–286.

Weinberger, L., Ayyash, M., Novershtern, N., and Hanna, J.H. (2016). Dynamic stem cell states: naive to primed pluripotency in rodents and humans. Nat. Rev. Mol. Cell Biol. *17*, 155–169.

Weiner, A., Lara-Astiaso, D., Krupalnik, V., Gafni, O., David, E., Winter, D.R., Hanna, J.H., and Amit, I. (2016). Co-ChIP enables genome-wide mapping of histone mark co-occurrence at single-molecule resolution. Nat. Biotechnol. *34*, 953–961.

Wu, J., Huang, B., Chen, H., Yin, Q., Liu, Y., Xiang, Y., Zhang, B., Liu, B., Wang, Q., Xia, W., et al. (2016). The landscape of accessible chromatin in mammalian preimplantation embryos. Nature *534*, 652–657.

Wu, X., Bekker-Jensen, I.H., Christensen, J., Rasmussen, K.D., Sidoli, S., Qi, Y., Kong, Y., Wang, X., Cui, Y., Xiao, Z., et al. (2015). Tumor suppressor ASXL1 is essential for the activation of INK4B expression in response to oncogene activity and anti-proliferative signals. Cell Res. *25*, 1205–1218.

Wu, X., Johansen, J.V., and Helin, K. (2013). Fbxl10/Kdm2b recruits polycomb repressive complex 1 to CpG islands and regulates H2A ubiquitylation. Mol. Cell *49*, 1134–1146.

Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. BMC Bioinformatics *10*, 232.

Xiang, Y., Zhang, Y., Xu, Q., Zhou, C., Liu, B., Du, Z., Zhang, K., Zhang, B., Wang, X., Gayen, S., et al. (2020). Epigenomic analysis of gastrulation identifies a unique chromatin state for primed pluripotency. Nat. Genet. *52*, 95–105.

Xu, Q., Xiang, Y., Wang, Q., Wang, L., Brind'Amour, J., Bogutz, A.B., Zhang, Y., Zhang, B., Yu, G., Xia, W., et al. (2019). SETD2 regulates the maternal epigenome, genomic imprinting and embryonic development. Nat. Genet. *51*, 844–856.

Xu, W., Li, J., Rong, B., Zhao, B., Wang, M., Dai, R., Chen, Q., Liu, H., Gu, Z., Liu, S., et al. (2020). DNMT3A reads and connects histone H3K36me2 to DNA methylation. Protein Cell *11*, 150–154.

Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., Li, X., Xu, M., Zhang, Z., Niu, T., et al. (2012). Dense chromatin activates polycomb repressive complex 2 to regulate H3 lysine 27 methylation. Science *337*, 971–975.

Zhang, B., Zheng, H., Huang, B., Li, W., Xiang, Y., Peng, X., Ming, J., Wu, X., Zhang, Y., Xu, Q., et al. (2016). Allelic reprogramming of the histone modification H3K4me3 in early mammalian development. Nature *537*, 553–557.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Meyers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol *9* (9), R137.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137.

Zhang, Y., Xiang, Y., Yin, Q., Du, Z., Peng, X., Wang, Q., Fidalgo, M., Xia, W., Li, Y., Zhao, Z.A., et al. (2018). Dynamic epigenomic landscapes during early lineage specification in mouse embryos. Nat. Genet. *50*, 96–105.

Zheng, H., Huang, B., Zhang, B., Xiang, Y., Du, Z., Xu, Q., Li, Y., Wang, Q., Ma, J., Peng, X., et al. (2016). Resetting epigenetic memory by reprogramming of histone modifications in mammals. Mol. Cell *63*, 1066–1079.

Zhou, Q., Zhang, Y., Wang, B., Liu, X., and Zhan, Z. (2019). KDM2B epigenetically promotes IL-6 production and inflammatory responses through Brg1-mediated chromatin remodeling. Eur. J. Immunol. *49*, 110.

Zhou, Z., Yang, X., He, J., Liu, J., Wu, F., Yu, S., Liu, Y., Lin, R., Liu, H., Cui, Y., et al. (2017). Kdm2b regulates somatic reprogramming through variant PRC1 complex-dependent function. Cell Rep. *21*, 2160–2170.

Zhu, Y., Yu, J., Rong, Y., Wu, Y.-W., Li, Y., Zhang, L., Pan, Y., Fan, H.-Y., and Shen, L. (2021). Genomewide decoupling of H2AK119ub1 and H3K27me3 in early mouse development. Sci. Bull. *66*, 2489–2497.

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
KDM2B	home-made	N/A
RING1B	home-made	N/A
NSPc1	home-made	N/A
SUZ12	Cell Signaling Technology	#3737S
EZH2	Cell signaling technology	#5246
β-tublin	Abclonal	AC021
H3	Millipore	07-690
Gapdh	Abclonal	AC002
H3K27me3	Cell Signaling Technology	#9733
H3K27me3	Diagenode	pAb-069-050
H3K36me2	Cell Signaling Technology	#2901
H2AK119ub1	Cell Signaling Technology	#8240
H2AK119ub1	Millipore	05-678
BRG1	Abcam	ab110641
FLAG	Sigma	F1804
Chemicals, peptides, and recombinant proteins		
GSK inhibitor	Selleck Chemicals	S2924
MEK inhibitor	Selleck Chemicals	S1036
recombinant mouse lif	Sinobiological	50756-MNAH-20
mouse Activin A protein (His tag)	Sinobiological	50659-M08H-20
recombinant mouse FGF2 Protein	Sinobiological	50037-M07E-100
Puromycin	Sigma-Aldrich	P8833-25MG
Doxycycline hyclate	Sigma-Aldrich	D9891-10G
N2	Gibco	17502048
B27	Gibco	17504044
Critical commercial assays		
trueprep DNA library prep kit V2 for illuminated	Vazyme	TD501; TD503
Hyperactive pA-Tn5 Transposase for CUT&Tag	Vazyme	S603
Pico Methyl-seq Library Prep Kit	Zymo Research	D5456
Gateway LR Clonase II Enzyme mix	Thermo Fisher	Cat#11791020
Deposited data		
Sequencing data	This paper	EBI: ERP128231
Proteomics raw files	This paper	PRIDE: PXD031298
RNA-seq data in pre-implantation and post-implantation development	Liu et al., 2016a; Zhang et al., 2018	GEO: GSE70605, GSE76505
2i, SL, E3.5 ICM, E6.5 epiblast ATAC-seq data	Atlasi et al., 2019; Wu et al., 2016; Xiang et al., 2020	GEO: GSE92411, GSE66390, GSE125318
E6.5 Epiblast and E3.5 ICM H3K27me3 ChIP-seq data	Wang et al., 2018; Liu et al., 2016b	GEO: GSE98149, GSE73952
Original images	This paper	DOI: 10.17632/tkngw2xwjz.1

(*Continued on next page*)

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental models: Cell lines		
WT (E14) mESC	Kristian lab	N/A
Kdm2b H283A mutant mESC	This paper	N/A
293FT cells	Kristian lab	N/A
A2Loxcre mESC	Iacovino et al., 2011	N/A
artificial CGI knock in mESC	This paper	N/A
Experimental models: Organisms/strains		
Mouse: Kdm2b H283A mutant	NBRI, Nanjing, China	C57BL/6J
Oligonucleotides		
Brg1 Knock down#1: 5'-CGCCCGACACATTATTGAGAA-3'	This paper	N/A
Brg1 Knock down#2: 5'-CGGCTCAAGAAGGAAGTTGAA-3'	This paper	N/A
Kdm2b-H283A sgRNA: 5'-CACTCAGGTTGGATCCATG-3'	This paper	N/A
artificial CGI knock in sgRNA: 5'-CTTATAGGCATGTCCGTAAG-3'	This paper	N/A
sgRNA target site Forward: CCCCCATTCCAGTGATGCTA	This paper	N/A
sgRNA target site Reversed: CAAACTGCACTTGTGGGGAA	This paper	N/A
Intergenic negative control Forward: CCTACTTACAGTCACTTTTA	This paper	N/A
Intergenic negative control Reversed: TCTCAAGAGCTTTCCACCACT	This paper	N/A
Sox2 promoter Forward: TCATGCAAAACCCTCTGGCG	This paper	N/A
Sox2 promoter Reversed: GGAATAAATGGGTTTCCGGC	This paper	N/A
Fgf3 promoter Forward: CAGATCAGGCCCATCCCG	This paper	N/A
Fgf3 promoter Reversed: GCGTGTGCTCCCAGCG	This paper	N/A
Wnt1 promoter Forward: TGCGCCCTGGTGCTTTTAGTGC	This paper	N/A
Wnt1 promoter Reversed: GCGGGCCGCAGGCAGCATG	This paper	N/A
Hoxb9 promoter Forward: CCATCCTGGATATGGAATGC	This paper	N/A
Hoxb9 promoter Reversed: GGAGCTCCTTGCAAAATGAT	This paper	N/A
Recombinant DNA		
pHAGE TRE dCas9-KRAB	Kearns et al., 2014	Addgene #50917
pLX-sgRNA	Wang et al., 2014	Addgene#50662
p2Lox-FLAG	Mazzoni et al., 2011	N/A
p2Lox-FLAG-dCas9-KDM2B WT or MT	This paper	N/A
Software and algorithms		
CellProfiler	McQuin et al., 2018	https://cellprofiler.org
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml

(*Continued on next page*)

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MACS2 (v2.1.1.20160309)	Zhang et al., 2008	https://pypi.org/project/MACS2/
TopHat (v2.1.1)	Trapnell et al., 2009	https://ccb.jhu.edu/software/tophat/index.shtml
Mcall	Sun et al., 2014	https://github.com/sunnyisgalaxy/moabs
IntersectBed	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/content/tools/intersect.html
Bsmap	Xi and Li, 2009	https://github.com/sunnyisgalaxy/moabs

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Xudong Wu (wuxudong@tmu.edu.cn).

Materials Availability

Constructs, ESC and mouse lines generated in this study are available from the lead contact upon request.

Data and code availability

- ChIP-seq, RNA-seq, ATAC-seq, CUT&Tag, and WGBS data that support the findings of this study have been deposited in the EBI under accession code ERP128231 and will be publicly available as of the date of publication. Proteomics raw files were deposited to PRIDE with the accession code PXD031298. Original images have been deposited to Mendeley data and are available at Mendeley Data: https://doi:10.17632/tkngw2xwjz.1.
- This paper does not report original code.
- Any additional information required to re-analyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines and cell culture

293FT cells were grown in DMEM medium with 10% FBS (Lonsera). Mouse ES cells were cultured on 0.1% gelatin coated plates. SL culture medium was DMEM medium (GIBCO) containing 15% FBS (Hyclone), 1% NEAA (GIBCO), 1% sodium pyruvate (GIBCO), 1% glutamax (GIBCO), 1% penicillin/streptomycin (GIBCO) and 0.1mM β-mercaptoethanol (sigma), 5 ng/mL LIF (Sino Biological). 2i culture medium was DMEM/F12 (GIBCO): Neurobasal medium (GIBCO)=1:1, containing N2 (Invitrogen), B27 (Invitrogen), 1% NEAA (GIBCO), 1% sodium pyruvate (GIBCO), 1% glutamax (GIBCO), 1% penicillin/streptomycin (GIBCO), 0.1 mM β-mercaptoethanol (sigma), 5 ng/mL LIF (Sino Biological), 3 nM GSK inhibitor CHIR99021 (Selleck) and 1 nM MEK inhibitor PD0325901 (Selleck). FA culture medium was DMEM/F12 (GIBCO): Neurobasal medium (GIBCO)=1:1 containing N2 (Invitrogen), B27 (Invitrogen), 1% NEAA (GIBCO), 1% sodium pyruvate (GIBCO), 1% glutamax (GIBCO), 1% penicillin/streptomycin (GIBCO), 0.1mM β-mercaptoethanol (sigma), 20 ng/mL Fgf (Sino Biological) and 8 ng/mL activin A (Sino Biological).

Generation of Brg1 Knockdown Cell Lines

Two pairs of oligos were designed according to specific gene sequence: 5'-CGCCCGACACATTATTGAGAA-3' and 5'-CGGCTCAAGAAGGAAGTTGAA-3' for Brg1, the synthesized oligos were cloned into the pLKO.1 TRC cloning vector using BamHI/KpnI sites. The lentiviruses were generated by co-transfection with pAX8 and pCMV-VSVG into 293FT cells. Virus supernatants were collected 2 days later. Then ES cells were transduced, followed by selection with puromycin (2 μg/ml) for 48 hrs. The Brg1 expression levels were tested by western-blot analysis.

Generation of knock-in Cell Lines with CRISPR-Cas9

Specific single guide RNAs (sgRNAs) were designed by synthego (https://www.synthego.com/products/bioinformatics/crispr-design-tool). The sgRNA fragments were cloned into the pX459 vector according to Zhang lab general targeting sequence cloning protocol.

To generate *Kdm2b*-H283A mutant ES cells, we used a single sgRNA (5'-CACTCAGGTTGGATCCATG-3') and a donor plasmid which contains H283A mutation and a synonymous mutation disrupting the PAM sequence. To integrate the artificial CGI, we chose a gene desert area with H3K36me2 positive, while H3K27me3, H3K4me3, H3K27ac, H3K4me1 and H3K9me3 negative. The guide

RNA sequences were as followed: 5'-CTTATAGGCATGTCCGTAAG-3' and the donor plasmid contained 500 bp of CG-rich DNA sequences. Totally 10 μg plasmid was transfected into ES cells in 6 cm dish using Lipofectamine 3000 transfection reagent. The individual positive clones were identified by PCR analysis using specific primers.

Protein tagging by BAC transgenesis

We generated the targeting vectors and Kdm2b C-terminal tagging stable ES cell line as described (Hofemeister et al., 2011). Briefly, we ordered BAC clone RP23-94P12 from Children's Hospital Oakland Research Institute (CHORI), http://bacpac.chori.org. Proper oligo sequences were designed including the homology arms to the BAC DNA sequence and attached by PCR to the adaptor cassette using the plasmid pR6K-2Ty1-rpsl-BSD-3FLAG as template DNA (Kdm2b_CTag_left: TGAGTGTCCAATTTGGGCAAGTGG AAGAGAAACTCCTGCAAAAACTAAGT GAAGTGCATACCAATCAGGACCCGC; *Kdm2b*_CTag_right: TCCTGGTTGTATTCCCCCCA ACCCTCCATATTTACATACTTGTCCTGGAA GTTCCTTCTTCACTGTCCCTTATTC). After introduction of the recombineering expression plasmid pSC101-BAD-Red$\alpha\beta\gamma$, the adaptor PCR product was electroporated and inserted into the BAC by recombineering, replacing the stop codon of the targeting gene. Then the tagging cassette generated by linearization of plasmid R6K-2Ty1-2PreS-mVenus-Biotin-T2A-gb3-neo was electroporated into the recombined bacterial stock. By the second recombineering, the tagging cassette replaces the adaptor cassette. The purified plasmid DNA was then used to transfected into mouse embryonic stem cell line E14 by lipofectamine 2000 (Life technologies). The neomycin-resistant clones were selected for the detection of Ty1 or GFP expression. The first round of positive clones was then transfected with Cre and Dre-expressing constructs to remove mVenus and neo cassette, thus the second round of subclones are neomycin sensitive and express the 3xFLAG tag.

Mice

To generate a C57BL/6 mouse model with point mutation (H283A) at mouse *Kdm2b* locus by CRISPR/Cas9-mediated genome engineering (NBRI, Nanjing, China), Kdm2b-H283A (CAT to GCT) specific sgRNA oligos were designed by the CRISPR website (http://crispr.mit.edu/) targeting sequence at H283 locus (5'- TGGACTCTCTGGTGTTCGGC-3'). Cas9 mRNA, sgRNA generated by *in vitro* transcription and donor oligo were co-injected into fertilized eggs. The sgRNA directed Cas9 endonuclease cleavage in intron 8∼9 and on exon 9 and created a DSB (double-strand break). Such breaks were repaired and resulted in H283A mutation on exon9 respectively by homologous recombination. The target region of mouse *Kdm2b* locus was sequenced to confirm targeting.

8∼10 weeks old WT and H283A mutant mice were used for intercourse matings and the mice were maintained under specific pathogen-free (SPF) conditions. In order to obtain E6.5 epiblast, naturally mated female mice were dissected to expose the uterus. After taking out of the embryos, we placed them in DMEM containing 10% FBS, peeled off the decidua and reichert's membrane, then digested with pancreatic and trypsin enzyme for 5 minutes (min) before stopping with DMEM containing 10% FBS. Then a thin glass needle was used to peel off the visceral endoderm, and isolate the epiblast from the extraembryonic tissue.

METHOD DETAILS

CRISPR/dCas9 Assay

The inducible epigenome editing system was established similarly to our previous inducible CRISPRI system (Li et al., 2020). Briefly, the fragment of dCas9 was amplified from the plasmid (Addgene #50917). And the cDNA sequences of KDM2B-WT or H283A mutant (MT) were ligated with dCas9 fragment by overlap PCR. After purification, the dCas9-KDM2B WT or MT fragment were introduced into Gateway Entry vector pCR8/GW/TOPO (Invitrogen) following the manufacturer's protocol. Then the fragments of interest were subcloned into the destination vector p2Lox-FLAG(Mazzoni et al., 2011) through Gateway technology (Invitrogen). After induction by Doxycycline for 24 hrs, A2Loxcre mESCs (Iacovino et al., 2011) were transfected with p2Lox-FLAG-dCas9-KDM2B WT or MT plasmids using Lipofectamine 3000 (Thermo). After 3 days selection by G418 (100 μg/ml), the positive clones were identified by PCR analysis using specific primers (F: 5'-TTACCACTCCCTATCAGTGATAG-3', R: 5'-AGGAAGCTCTCTTCCAGCCTATG-3'). And the inducible expression of FLAG-dCas9-KDM2B protein was confirmed by Western Blot (WB) assay.

In order to direct the fusion protein to the target genomic region, we transduced the cells with sgRNAs expressed by a pLX-sgRNA vector from Addgene (#50662). The lentiviruses were generated by co-transfection with pAX8 and pCMV-VSVG into 293FT cells as described above.

Real-Time Quantitative PCR

For gene expression analysis, total RNA was extracted with TRIzol (15596026, invitrogen) and reverse transcribed with a revertAid first strand cDNA synthesis kit (#K1622, Thermo). Quantitative PCR analyses of cDNA or ChIP-DNA were performed with ChamQ Universal SYBR qPCR Master Mix (Q711, Vazyme) on Roche LightCycler® 480 system. The primer sequences were listed in key resources table

CUT&Tag

The CUT&Tag assay was performed as described previously (Kaya-Okur et al., 2019). In brief, 10,000-50,000 cells were harvested and washed twice in 1 ml wash buffer (20 mM HEPES PH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 1×Protease Inhibitor Cocktail (PIC, Bimake), and then resuspended in 300 μl wash buffer. Concanavalin A coated beads were activated by washing twice in binding

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

buffer (20 mM HEPES pH 7.5, 10 mM KCl, 1mM $MnCl_2$, 1mM $CaCl_2$). Then 10 μl activated beads were added to the resuspended cells and incubated for 30 min at room temperature (RT). The supernatant was removed and beads were resuspended in 50 μL antibody buffer (20 mM HEPES pH 7.5,150 mM NaCl, 0.5 mM Spermidine, 0.0125% Digitonin, 2 mM EDTA, 0.1% BSA, 1×PIC). Then 0.5 μL primary antibody was added and incubated for 2 hours (hrs) at RT. After the supernatant was gently removed by placing the tubes on magnet, the beads were resuspended in 100 μl Dig wash buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 0.0125% Digitonin, 1×PIC) which contains 1 μl Guinea Pig anti-Rabbit IgG antibody and incubated for 1 hr at RT. After being washed three times in 800 μl Dig wash buffer, the beads were resuspended in 100 μl Dig 300 wash buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5mM Spermidine, 0.0125% Digitonin, 1×PIC), and 0.04 μM pA-Tn5 adapter complex (Vazyme S603) was added and incubated for 1 hr at RT. Beads were washed three times in 800 μl Dig 300 wash buffer and resuspended in 300 μL tagmentation buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM Spermidine, 0.0125% Digitonin, 10 mM $MgCl_2$, 1×PIC) and incubated for 1 hr at 37°C. Then 10 μl of 0.5 M EDTA, 3 μl of 10% SDS and 1 μl of 20 mg/ml Proteinase K were added and incubated at 55°C for 1 hr to stop tagmentation before DNA extraction by PCI. The purified DNA was amplified with i5 and i7 primer, KAPA 2x PCR mix (KM2602), under the cycling program with heated lid, 72°C, 5 min; 98°C, 30 s; 17 cycles for 98°C, 10s; 60°C, 30 s; 72°C, 30 s; and 72°C for 5 min, hold at 4°C. After the program was completed, 0.5 volume (25 μl) Ampure XP beads were added to remove fragments above 1 kb and another 0.5 volume (25 μl) Ampure XP beads were added to enrich the 200-1,000bp fragments for sequencing. The constructed libraries were sequenced with PE150 on NovaSeq platform.

STAR ChIP-seq

STAR ChIP-seq was performed as described previously (Zhang et al., 2016). Briefly, embryo samples were lysed in a modified lysis buffer (0.5% NP-40, 0.5% Tween-20, 0.1% SDS and proteinase inhibitor). 0.02 unit of MNase was used for chromatin fragmentation. The reaction was terminated by adding 5 μl Stop buffer (110mM Tris-HCl pH8.0, 55mM EDTA), followed by adding another 45 μl cold "2xRIPA buffer" (1% Triton X-100, 280 mM NaCl, 0.1% SDS, 0.2% DOC, 5 mM EGTA supplemented with proteinase inhibitor). After being spinned at max speed in 4°C for 15 min, the supernatant was transferred to a new tube. Then the sample was supplemented with another 40 μl RIPA buffer (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.1% DOC, 1 mM EDTA). IP sample was then incubated with 1∼1.5 μg antibody overnight with rotation at 4°C. On the next day, the sample was incubated with Protein A dynabeads (Life technologies) for 2 hrs with rotation at 4°C. To elute DNA, the beads were resuspended with 27 μl ddH2O and 1 μl 10×Ex-Taq buffer (TaKaRa), 1 μl proteinase K (Roche, 10910000) and incubated at 55°C for 90 min. The supernatant was transferred to a new tube and the proteinase K was inactivated at 72°C for 40 min. The DNA sample was then treated with 1 μl rSAP (NEB, M0371) at 37°C for 1 hr, followed by inactivation at 65°C for 10 min. The resulting sample was subjected to TELP library preparation starting from poly-C tailing as described previously in full detail (Peng et al., 2015) with a slight modification. A modified anchor primer was used for better performance ([BIO]GTTCAGACGTGTGCTCTTCCGATCTGGGGGGGGG, where [BIO] refers to 5'-end biotin). The constructed libraries were sequenced with PE150 on NovaSeq platform.

Bulk ChIP-seq

ES cells cultured in 2i or SL medium were crosslinked for 10 min with 1% formaldehyde at RT, followed by adding glycine to 0.125M at final concentration and incubation for another 5 min to stop fixation. Then cells were rinsed twice with PBS and cell pellets were resuspended in SDS buffer (1% SDS, 5 mM EDTA, 100 mM NaCl, 0.02% NaN_3, 50 mM Tris-HCl, pH 8.1, 1×PIC and 1×PMSF). After the lysis, nuclei were obtained by spinning at 1200 rpm for 6 min and resuspended in appropriate volume of ice-cold IP Buffer (SDS Buffer: Triton Dilution Buffer (100 mM Tris-HCl, pH 8.6, 100 mM NaCl, 5 mM EDTA, 5% Triton X-100, 0.02% NaN_3, 1×PIC and 1×PMSF) =1:1) for sonication. The sonication was optimized to enrich chromatin fragments with an average length of 200-500 bp. If an optimal sonication was achieved, the chromatin samples were diluted with IP buffer to a desired concentration. Then primary antibodies were added and incubated overnight at 4°C rotating. On the next day, 30-40 μl of protein A/G magnetic beads were add and incubate 2-4 hrs at 4°C. Then the beads were washed one with 150 mM wash buffer (1% Triton X-100, 0.1% SDS, 150 mM NaCl, 2 mM EDTA, pH 8.0, 20 mM Tris-HCl, pH 8.0) and three times with 500 mM wash buffer (1% Triton X-100, 0.1% SDS, 500 mM NaCl, 2 mM EDTA, pH 8.0, 20 mM Tris-HCl, pH 8.0). After the washes, the beads were added 120 μl of de-crosslink buffer and incubated 3 hrs up to overnight at 65°C (shaking at 1,100 rpm) to elute DNA from beads. Then the input and ChIP DNA was purified with QIAGEN PCR purification kit. The enrichment was analyzed by qPCR or high-throughput sequencing.

For H2AK119ub1 ChIP, a 5-min pre-extraction step with detergents was performed before fixation. And after sonication, IP buffer was replaced with PBS through dialysis for antibody incubation.

Library preparation and sequencing strategy

The DNA Library of ChIP were prepared with a TruePrep DNA Library Prep Kit according to the manufacturer's instructions (TD503, Vazyme). The prepared libraries were sequenced with PE150 on NovaSeq platform. The RNA library preparation and sequencing were cooperated with Novogene Company.

ATAC-seq and ATAC-qPCR

50,000 ESCs or ∼20 E6.5 Epi were used for ATAC-seq or qPCR. After being washed by PBS, the cells or epiblasts were resuspended in lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.1% (v/v) igepal CA-630 and 1×PIC). After 10 min incubation on

Please cite this article in press as: Huo et al., CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency, Molecular Cell (2022), https://doi.org/10.1016/j.molcel.2022.01.027

ice, cells were spinned down by centrifugation at 500g, 4°C. Then the pellets were resuspended in transposome mix (37 μl ddH2O, 10 μl 5×TTBL, 3 μl TTE mixV50 (TD501, Vazyme)) and gently flicked. The reaction was incubated at 37°C for 0.5 h in a PCR cycler, before 1 μl 0.5M EDTA, 1 μl proteinase K (20 mg/ml) and 5 μl 10% SDS were added to stop the reaction. After another incubation at 50°C for 1 h, DNA was purified with Phenol-Chloroform-Isoamylol (ThermoFisher) and resuspended with 30 μl ddH2O. 15 μl DNA were amplified with i5 and i7 primers (Vazyme, TD202) and KAPA 2×PCR mix (KM2602) following the cycling program with heated lid, 72°C, 5min; 98°C, 30s; 10 cycles for 98°C, 10 s; 60°C, 30 s; 72°C, 30 s; and 72°C for 5 min, and hold at 4°C. Upon the program completed, 0.5 volume (25 μl) Ampure XP beads were added to remove fragments above 1 kb and another 0.5 volume (25 μl) Ampure XP beads were added to collect the 200-1,000bp fragments for sequencing. And the rest 15 μl DNA was used as input. For ATAC-seq, the libraries were sequenced with PE150 on NovaSeq platform. For ATAC-qPCR analysis, the library and input DNA samples were diluted 30 times for Quantitative PCR analysis.

Immunofluorescence

ES cells cultured in 2i or SL medium were fixed in cold methanol for 3 min. After being washed twice with PBS, the cells were blocked with 0.8% BSA (A80209, solarbio) for 10 min. NANOG (A300-397A, Bethyl) or OCT4 (ab184665, abcam) antibodies were diluted 1:100 with PBS and then incubated with cells at 37°C for 1 hr. After being washed twice with PBS for 5 min, secondary antibodies were incubated for another hour. Images were acquired using a laser scanning confocal microscope. Images were analyzed with CellProfiler image analysis software.

Immunoprecipitation

$1x10^7$ cells were harvested, washed three times with PBS, resuspended with cell lysis buffer (50 mM Tris-HCl pH=7.4, 0.2 mM EDTA, 150 mM NaCl, 0.3% NP40, 1×PIC) followed by rotation at 4°C for 30 min. After centrifugation at 12,000 rpm and 4°C, the supernatant was incubated with 5 μg antibody overnight at 4°C rotating. 5% of lysates were used as input. On the next day, 40 μl protein A/G magnetic beads were added and incubated for 2∼4 hrs. The beads were then washed 3 times with cell lysis buffer and boiled in 40 μl 1×SDS loading buffer, analyzed by western blot assays.

Complex purification coupled with LC-MS/MS analysis

In order to isolate Kdm2b-containing protein complexes, affinity purification was performed followed by MS analysis. Nuclear extracts (60 mg) from BAC engineered stable Kdm2b-2Ty1-3FLAG cell line was prepared as described (Wu et al., 2015) and precleared with protein G Sepharose (GE Healthcare) for 2 hrs and incubated with a 80 μl packed volume of anti-FLAG-beads (Anti-FLAG® M2-agarose, Sigma) overnight at 4 °C with rotation. The beads were collected by centrifugation at 700g for 5 min and washed 6 times with 40×resin bed volume of washing buffer (20 mM Tris-HCl, pH 8.0, 300 mM NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, 10% glycerol, 0.2 mM PMSF, 1 mM DTT, 1 μg/ml aprotinin and 1 μg/ml leupeptin). The beads were then boiled in SDS loading buffer and run shortly into a SDS gel. A gel slice containing the purified proteins was isolated for MS analysis. The extracted MS/MS data were processed using Mascot search engine. Tandem mass spectra were searched against SwissProt-Mouse database concatenated with reverse decoy database. Proteins containing unique mapped peptide with ion score meeting significant *P value* were considered.

AP staining

To test the pluripotent state of ESCs, appropriate number of cells were plated in the 6-well plate in 2i or SL medium and then cultured for 5 days. After fixation with 4% Paraformaldehyde for 20 min, ESCs were stained with an AP staining kit according to the manufacturer's instructions (C3206, Beyotime).

WGBS library preparation

WGBS libraries were generated using the Pico Methyl-seq Library Prep Kit (Zymo Research, D5456). Briefly, input DNA was mixed with 0.1% unmethylated lambda DNA (w/w) (Promega, D1521), followed by bisulfite conversion. The bisulfite-converted DNA was then annealed with random primers for initial amplification, followed by adaptor ligation and final amplification. Libraries were purified by the Genomic DNA Clean & ConcentratorTM (Zymo Research, D4010), and the concentration was measured by Qubit® 2.0 Fluorometer with the qubit dsDNA High-sensitivity assay kit (Invitrogen, Q32854). Then the libraries were sequenced by Illumina HiSeq X Ten sequencing.

ChIP-seq and CUT&Tag data processing

ChIP-seq and CUT&Tag reads (as listed in Table S1) were first processed using TrimGalore (v0.4.4_dev) to trim adaptor and low-quality reads. Trimmed reads were then aligned to the mouse genome build mm10 using bowtie2 (v2.4.2) (Langmead and Salzberg, 2012) with parameters –no-mixed –no-discordant. Low mapping quality reads (mapping quality < 30) and duplicates were discarded. Then biological replicates passed quality control were pooled together. Signal tracks for each sample were generated using the MACS2 (v2.1.1.20160309) (Zhang et al., 2008) pile-up function and were normalized to 1 million reads (RPM). ChIP-seq peaks were identified by MACS14 (v1.4.2.20120305) (Zhang et al., 2008) with parameters –nomodel –nolambda –shiftsize=73.

As H3K36me2 is depleted at CGIs, ChIP-seq signals at those loci are rather low. Here we utilized local normalization signals rather than ChIP-seq RPMs to represent relative binding levels of H3K36me2 when we compared the difference between MT and WT. Local

normalization signal at each locus was defined as Z-score transformation of its RPM where RPMs of its neighbouring genomic loci were used as the population. For each of interested regions, we first calculated its ChIP-seq RPM x and the mean (μ) and standard deviation (σ) of the population (all ChIP-seq RPMs at ±100 kb surrounding center of the given region), then local normalization signal (Z) can be calculated with the following formula.

$$Z = \frac{(x - \mu)}{\sigma}$$

RNA-seq data processing

RNA-seq reads (as listed in Table S1) were aligned to mouse genome build mm10 using TopHat (v2.1.1) (Trapnell et al., 2009). Expression levels for all RefSeq transcripts were quantified to fragment per kilobase million (FPKM) using Cufflinks (v2.2.1). Signal tracks for each stage were generated using reads density and normalized to 1 million reads (RPM).

ATAC-seq data processing

ATAC-seq reads (as listed in Table S1) were processed as previously described (Liu et al., 2018). In general, reads were trimmed using TrimGalore (v0.4.4_dev) to remove adaptor and aligned to mouse genome build mm10. After remove low quality reads (mapping quality < 30) and duplicates, only reads pairs with fragment length $\leq$ 100 bp were used as effective fragments for the following analysis. Signal tracks for each sample were generated using reads density and normalized to 1 million reads (RPM).

Whole genome bisulfite sequencing (WGBS) data analysis

WGBS reads (as listed in Table S1) were first processed using TrimGalore (v0.4.4_dev) to trim adaptor and low-quality reads and then aligned to mm10 genome using bsmap (Xi and Li, 2009). The DNA methylation level of each CpG site was calculated as methylated C/total C ratio and estimated using mcall (Sun et al., 2014). Signal tracks for each sample were generated using DNA methylation level at each CpG site.

Genomic loci annotation

Mouse genome were classified into 5 groups of loci in this manuscript, CGIs in promoter and non-promoter, promoters and gene body regions outside CGIs, and others. 23,784 CGIs were downloaded from UCSC genome browser database. Promoters were defined as ± 2 kb around transcription start sites (TSS) of RefSeq genes. CGIs in promoters were defined as CGIs that overlap with promoters and are extended to ± 2 kb around CGI centers. CGIs in non-promoter were defined as CGIs that do not overlap with promoters and are also extended to ± 2 kb around CGI centers. Promoters and gene body regions outside CGIs were defined as such regions that do not overlap with CGIs. Overlap events between CGIs and other regions were identified by using intersectBed (v2.27.1) (Quinlan and Hall, 2010) with parameters -e -f 0.5 -F 0.5. Gene body regions broader than 20 Kb were discarded in analysis for calculating fold change of signals. H3K27me3+/- genomic loci were defined as such genome loci with or without H3K27me3 peaks.

Datasets used in this study

ChIP-Seq, RNA-Seq, ATAC-Seq, CUT&Tag and WGBS data that support the findings of this study have been deposited in the EBI under accession codes ERP128231.

RNA-seq data in pre-implantation and post-implantation development were from previous publication (GSE70605; Liu et al., 2016a and GSE76505; Zhang et al., 2018) with quality control filtration. Public wild type 2i, SL, E3.5 ICM, E6.5 epiblast ATAC-seq data were from previous publication GSE92411 (Atlasi et al., 2019), GSE66390 (Wu et al., 2016) and GSE125318 (Xiang et al., 2020). E6.5 Epiblast and E3.5 ICM H3K27me3 ChIP-seq data were from previous publication (GSE98149; Wang et al., 2018) and (GSE73952; Liu et al., 2016b). E6.5 Epiblast and E3.5 ICM H2AK119ub1 CUT&RUN data were from previous publication (GSE153531; Chen et al., 2021).

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analysis are described in each figure legends and method details section.