2017

SMARCAD1 contributes to the regulation of naive pluripotency by interacting with histone citrullination

Gary Stormo
Washington University School of Medicine in St. Louis

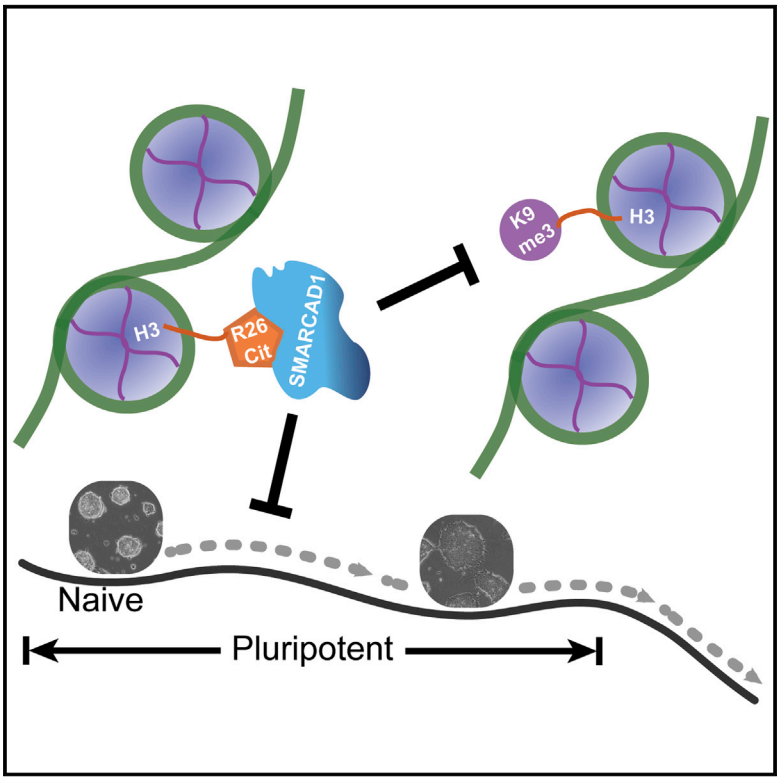
et al

Article

Cell Reports

SMARCAD1 Contributes to the Regulation of Naive Pluripotency by Interacting with Histone Citrullination

Graphical Abstract



Authors

Shu Xiao, Jia Lu, Bharat Sridhar, ..., Tetsuya S. Tanaka, Gary Stormo, Sheng Zhong

Correspondence

szhong@ucsd.edu

In Brief

Xiao et al. identify SMARCAD1 as a candidate reader protein of the H3R26Cit histone modification. Knockdown of Smarcad1 causes mouse ESCs to exit the naive pluripotent state and leads to increased H3K9me3 at SMARCAD1-binding regions. This suggests a model whereby SMARCAD1 guards naive pluripotency by interacting with H3R26Cit and suppressing H3K9me3.

Highlights

- SMARCAD1 preferentially binds histone peptides carrying the H3R26Cit modification
- Smarcad1 KD embryonic stem cells lose naive state features but remain pluripotent
- Suppression of Smarcad1 induces H3K9me3 at SMARCAD1-binding regions
- Inhibition of H3Cit induces H3K9me3 at SMARCAD1-binding regions

Accession Numbers

GSE45338

Xiao et al., 2017, Cell Reports *18*, 3117–3128
March 28, 2017
http://dx.doi.org/10.1016/j.celrep.2017.02.070

Cell Reports
Article

OPEN ACCESS
CellPress

SMARCAD1 Contributes to the Regulation of Naive Pluripotency by Interacting with Histone Citrullination

Shu Xiao,[1] Jia Lu,[1] Bharat Sridhar,[1,2] Xiaoyi Cao,[1] Pengfei Yu,[1] Tianyi Zhao,[1] Chieh-Chun Chen,[2] Darina McDee,[1] Laura Sloofman,[2] Yang Wang,[3] Marcelo Rivas-Astroza,[1] Bhanu Prakash V.L. Telugu,[4] Dana Levasseur,[5] Kang Zhang,[6] Han Liang,[7] Jing Crystal Zhao,[3] Tetsuya S. Tanaka,[2,9] Gary Stormo,[8] and Sheng Zhong[1,10,*]

[1]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA
[2]Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3]Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, USA
[4]Department of Animal and Avian Sciences, University of Maryland, College Park, MD 20742, USA
[5]Department of Internal Medicine, Carver College of Medicine, University of Iowa, Iowa City, IA 52242, USA
[6]Department of Ophthalmology, University of California, San Diego, La Jolla, CA 92093, USA
[7]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[8]Department of Genetics, Washington University at St. Louis, St. Louis, MO 63108, USA
[9]Present address: Departments of Biological Sciences and Chemical and Biomolecular Engineering, University of Notre Dame, Notre Dame, IN 46556, USA
[10]Lead Contact
*Correspondence: szhong@ucsd.edu
http://dx.doi.org/10.1016/j.celrep.2017.02.070

SUMMARY

Histone citrullination regulates diverse cellular processes. Here, we report that SMARCAD1 preferentially associates with H3 arginine 26 citrullination (H3R26Cit) peptides present on arrays composed of 384 histone peptides harboring distinct post-transcriptional modifications. Among ten histone modifications assayed by ChIP-seq, H3R26Cit exhibited the most extensive genomewide co-localization with SMARCAD1 binding. Increased Smarcad1 expression correlated with naive pluripotency in pre-implantation embryos. In the presence of LIF, Smarcad1 knockdown (KD) embryonic stem cells lost naive state phenotypes but remained pluripotent, as suggested by morphology, gene expression, histone modifications, alkaline phosphatase activity, energy metabolism, embryoid bodies, teratoma, and chimeras. The majority of H3R26Cit ChIP-seq peaks occupied by SMARCAD1 were associated with increased levels of H3K9me3 in Smarcad1 KD cells. Inhibition of H3Cit induced H3K9me3 at the overlapping regions of H3R26Cit peaks and SMARCAD1 peaks. These data suggest a model in which SMARCAD1 regulates naive pluripotency by interacting with H3R26Cit and suppressing heterochromatin formation.

INTRODUCTION

Histone citrullination, also called deimination, is a post-translational conversion of arginine into the amino acid citrulline (Klose and Zhang, 2007). The functional importance of H3 citrullination (H3Cit) includes regulation of naive pluripotency (Christophorou et al., 2014), gene regulation in cancer (Zhang et al., 2012), immune response (Neeli et al., 2008), and autoimmune disease (Sharma et al., 2012). However, the mechanisms underlying histone-citrullination-mediated regulation of cell physiology remain elusive. The available information includes that citrullination is antagonistic to arginine methylation (Cuthbert et al., 2004) and negatively regulates chromatin compaction (Christophorou et al., 2014), partially by weakening H3K9me3-mediated heterochromatin formation (Sharma et al., 2012). It is unknown whether any protein could read histone citrullination and relay its regulatory signal.

To search for proteins that may interact with citrullinated histones, we leveraged the discovery that H3Cit is associated with the establishment of naive pluripotency during embryonic development and cellular reprogramming (Christophorou et al., 2014). We searched for proteins that can be attached to chromatin and that are correlated with the establishments of inner cell mass (ICM) and naive-state pluripotent stem cells. This search identified two genes, with temporal expression patterns that peak at blastocyst stage and that are elevated in the ICM, that share sequence similarities with chromatin modification protein genes. These genes are Smarca4 (a.k.a. Brg1) and Smarcad1. Because the role of Smarca4 on embryonic stem cell (ESC) regulation has been characterized and the published data do not involve H3Cit (Kidder et al., 2009), we focused on the less studied Smarcad1.

Smarcad1, a.k.a. *Etl1* or *Hel1*, is one of the mammalian SNF2 family genes (Soininen et al., 1992). It contains DEAD/H ATP-binding domains and a bipartite nuclear localization signal (Schoor et al., 1993; Adra et al., 2000). It is expressed in all assayed embryonic and adult tissues (Adra et al., 2000), but its expression levels are exceptionally high in embryonal,

mammary, and lymphoid tumors (Okazaki et al., 2008). SMARCAD1 is attached to chromatin in both cancer and normal cells (Okazaki et al., 2008; Rowbotham et al., 2011).

RESULTS

Increased SMARCAD1 Expression Is Associated with Naive Pluripotency in Developing Embryos and in Cell Culture

We analyzed the association between Smarcad1 expression levels and pluripotent states in mouse and human preimplantation embryos. We generated RNA sequencing (RNA-seq) data from the ICM and the whole blastocyst of preimplantation mouse embryos. *Smarcad1* was expressed 2.7 times higher in ICM (RPKM [reads per kilobase per million] = 132.4) than in the whole blastocyst (RPKM = 49.4) (Figure 1A). Consistently, in a published RNA-seq dataset (Tang et al., 2010), *Smarcad1* exhibited 1.5 times higher expression in the ICM than in trophectoderm. Next, we re-analyzed our previously generated gene expression data from mouse preimplantation embryos at seven developmental stages with three biological replicates at each stage (Xie et al., 2010). *Smarcad1* mRNA was strongly and reproducibly induced at the eight-cell stage (Figure 1B), approximately 2 days prior to implantation. At the blastocyst stage, mouse SMARCAD1 protein expression is restricted to the ICM (see Figure 4 of Schoor et al. [1993]). In a single-cell RNA-seq dataset from human preimplantation embryos (GEO: GSE36552) (Yan et al., 2013), *SMARCAD1* expression increased from the two-cell stage to the blastocyst stage, peaking in some single cells in morula and in ICM (Figure 1C). All human mural trophectoderm cells except one had low *SMARCAD1* expression (Figure 1C). Human 32-cell morula is formed approximately 2 days prior to implantation (Cockburn and Rossant, 2010); therefore, the *SMARCAD1* mRNA peaked at approximately the same amount of time prior to implantation in mice and in humans.

Furthermore, we tested the association between Smarcad1 expression and pluripotent states using various cell lines. SMARCAD1 protein levels were significantly higher in mouse ES cells (naive state) than in epiblast-derived stem cells (EpiSCs, primed state) (Figure 1D). In humans, *SMARCAD1* mRNA was more abundant in naive cells as compared to primed embryonic stem cells (ESCs) (RNA-seq data from ArrayExpress: E-MTAB-2857 [Takashima et al., 2014], $p = 0.012$, two-tailed t test). In addition, our western blots suggest that SMARCAD1 proteins were more abundant in pig naive ESCs (Telugu et al., 2011) than in the primed pig induced pluripotent stem cells (iPSCs) (Ezashi et al., 2009; Figure 1E). We do not have the data to directly compare *Smarcad1* expression levels in preimplantation ICM and postimplantation epiblasts. However, re-analysis of a published microarray dataset (Tesar et al., 2007) suggested higher *Smarcad1* mRNA expression in ESCs than in postimplantation epiblasts ($p = 0.00039$) or in EpiSCs ($p = 0.00015$) (Figure 1F). Taken together, higher Smarcad1 mRNA and protein levels are associated with naive pluripotent stem cells in developing embryos and in cell culture.

SMARCAD1 Possesses Binding Specificity to Histone Modification but Not to DNA Sequence

SMARCAD1 binds to chromatin (Okazaki et al., 2008; Rowbotham et al., 2011). To determine whether the specificity of SMARCAD1-chromatin interaction is achieved by SMARCAD1's recognition of specific DNA sequences, we carried out high-throughput SELEX (HT-SELEX) (Zhao et al., 2009) with two randomized sequence libraries, one with a 10-bp and the other with a 20-bp randomized region. HT-SELEX was performed for two rounds on each of the two libraries. Neither selection enriched for any sequence motif, suggesting that the SMARCAD1 protein used in this study does not recognize specific DNA sequences.

We then asked whether SMARCAD1 recognizes specific histone post-translational modifications. We incubated SMARCAD1 with two MODified Histone Peptide Arrays. Each array contained 384 19-mer histone peptides. Each peptide is an N-terminal tail of H2A, H2B, H3, or H4, with a unique combination of post-translational modifications (Active Motif, 2014). The two arrays exhibited reproducible binding signals, and they consistently identified the peptide with a single modification (H3R26Cit) as the strongest binding peptide (Figure 2A). We then compared the post-translationally modified peptides to those with the same amino-acid sequence without any modification by calculating the ratio of the binding signals between every modified peptide and its unmodified counterpart. H3R26Cit exhibited the largest binding increment (ratio) to the unmodified peptide, followed by H3K27ac; however, only the H3R26Cit peptide exhibited 3-fold or larger binding signals than the background peptides in both arrays (Figures 2B and S1). These data prioritize H3R26Cit followed by H3K27ac as histone modifications that SMARCAD1 may preferentially associate with in vitro.

SMARCAD1 Co-localizes with H3R26Cit on the Genome

The in-vitro-binding specificity led us to posit that SMARCAD1 binds to chromatin with H3R26Cit or H3K27ac in vivo. To test this, we carried out chromatin immunoprecipitation (ChIP) sequencing (ChIP-seq) on H3R26Cit (Abcam: ab19847) (Zhang et al., 2012) and SMARCAD1 (Abcam: ab67548) in mouse ESCs. These data were jointly analyzed with our published ChIP-seq datasets for nine other histone modifications from the same cell line, including H3K27ac, H2A.Z, H3K4me1, H3K4me2, H3K4me3, H3K9me3, H3K27me3, H3K36me3, and 5-mC (MeDIP-seq [methylated DNA immunoprecipitation followed by sequencing]) (Xiao et al., 2012; Figure S2A). The extent of H3K27ac-SMARCAD1 co-localization was 4.6-fold greater than expectation (odds ratio = 4.56; Figure 2C, blue bars). Among the ten histone modifications, H3R26Cit exhibited the strongest co-localization with SMARCAD1 (odds ratio = 9.94, $p < 10^{-20}$, chi-square test) (Figure 2C, blue bars). These data suggest a genomewide correlation of H3 citrullination and SMARCAD1 binding in ESCs.

We used OCT4 ChIP-seq (Xiao et al., 2012) as another control to further assess the specificity of the genomewide co-localization of SMARCAD1 binding with any the ten histone modifications. OCT4 and SMARCAD1 exhibited similar degrees of co-localization (odds ratio $\approx$ 1) with nine of the ten histone marks

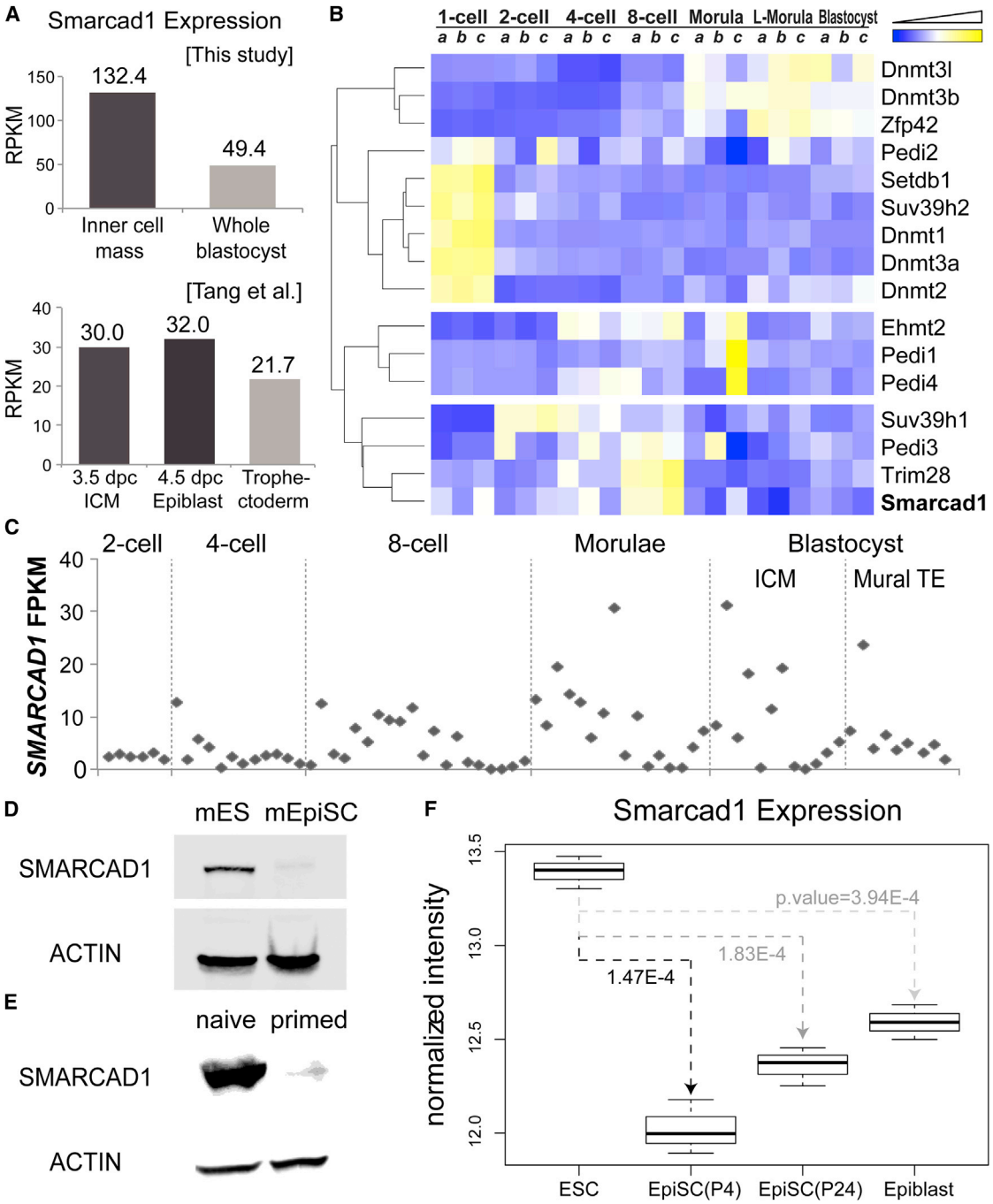


Figure 1. Smarcad1 Expression Patterns in Early Embryonic Development and in Pluripotent Stem Cells

(A) RNA-seq-derived *Smarcad1* mRNA levels in ICM, whole blastocyst, trophectoderm, and epiblast. dpc, days post-conception.

(B) Expression heatmap of genes related to DNA methylation, H3K9 methylation, and histone citrullination. Three biological replicates (a, b, and c) were analyzed in each of seven stages in mouse preimplantation development. L-Morula, late morula stage. Gene expression levels were normalized across samples, clustered (dendrogram), and visualized (yellow, high expression; blue, low expression).

(C) *SMARCAD1* expression (y axis) in single cells of human preimplantation embryos (GEO: GSE36552). Each column represents a single cell, and the cells were grouped by developmental stage (x axis). TE, trophectoderm.

(D and E) Western blots of SMARCAD1 and ACTIN in mouse ESCs (mES) and mouse EpiSCs (mEpiSC) (D) and in pig naive and primed pluripotent cells (E).

(F) Microarray-derived Smarcad1 expression levels in mouse ESC, EpiSC, and epiblast. P4 and P24 denotes passages 4 and 24, respectively.

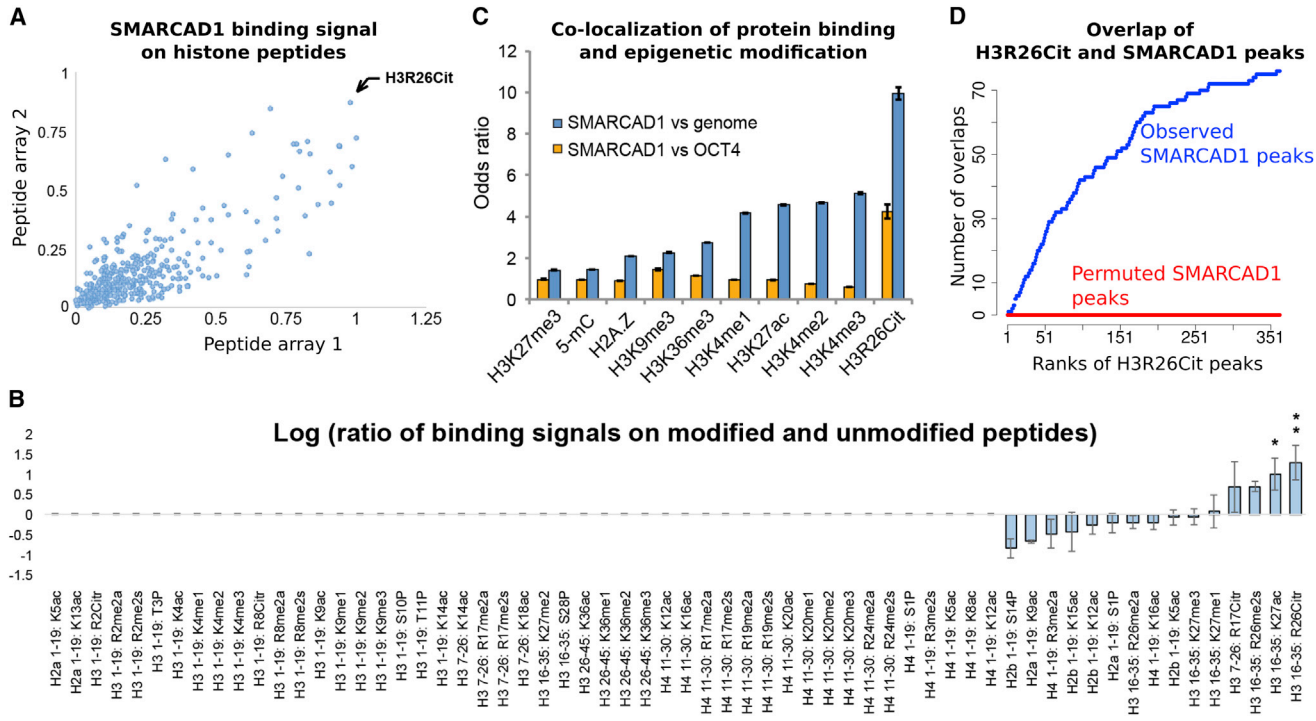


Figure 2. SMARCAD1 Recognizes H3R26Cit In Vitro and Co-localizes with H3R26Cit In Vivo

(A) Binding signals of SMARCAD1 on MODified Histone Peptide Arrays. Each dot represents the binding intensities to a peptide, which is modified with a unique combination of post-transcriptional modifications, on array 1 (x axis) and array 2 (y axis).

(B) Binding signals on post-translationally modified versus unmodified peptide (log ratio, y axis). All the peptides with a single modification and the unmodified peptides are shown (columns). If the raw binding signal to a modified peptide was smaller than the average binding signal to background peptides, the log ratio was assigned to 0 (non-informative, columns 1–50). If the raw binding signal was above background, this binding signal was divided by that of another peptide with identical amino-acid sequence without any modification (y axis, in log scale). One asterisks indicates signal >3-fold background in one array. Two asterisks indicate signal >3-fold background in both arrays.

(C) Relative levels (y axis) of co-localization of SMARCAD1 and each epigenetic modification (x axis), using the entire genome (blue bars) or OCT4 (orange bars) as the controls. Odds ratio >1 or <1 corresponds to an increased level or a decreased level, respectively, of co-localization.

(D) Cumulative counts of overlaps (y axis) of SMARCAD1 and H3R26Cit peaks, ordered by the significance (MACS-reported p value) of H3R26Cit peaks (x axis). Red curve shows the overlaps from a permutation analysis where SMARCAD1 peaks were randomly shifted to other genomic locations while keeping the size of each peak and the locations of the H3R26Cit peaks.

Error bars represent SEM.

including H3K27ac (Figure 2C, orange bars). In particular, the extent of H3K27ac-SMARCAD1 co-localization does not clearly exceed that of H3K27ac-OCT4 (odds ratio ≈ 1; orange bar in H3K27ac column of Figure 2C), suggesting that H3K27ac may generally enhance protein-chromatin interactions, but such an effect is not specific to SMARCAD1. The only exception among the ten analyzed histone marks was H3R26Cit, which exhibited stronger genomewide co-localization to SMARCAD1 than OCT4 (odds ratio = 4.22, $p < 10^{-20}$, chi-square test) (orange bar in H3R26Cit column, Figure 2C).

Using MACS (Model-based Analysis of ChIP-Seq; v. 1.4.0 beta) (Zhang et al., 2008) and immunoglobulin G (IgG) ChIP-seq as control, we identified a total of 363 H3R26Cit peaks in the genome. Among them, 22% (79 peaks) were bound by SMARCAD1, which correspond to an approximately 300-fold enrichment of overlap than what is expected from an independence model (1,279 SMARCAD1 peaks in the genome, covering ~1 million bases, approximately 1/1,800 of the effective genome); odds ratio > 300, $p < 10^{-10}$, chi-square test (Figure 2D). Moreover, 34%, 42%, and 48% out of the top 150, top 100, and top 50 H3R26Cit peaks, respectively, overlapped with SMARCAD1 peaks, suggesting that the more significant H3R26Cit peaks were more preferentially bound by SMARCAD1 (Figure 2D).

Phenotypic Differences between Smarcad1 Knockdown and Naive ESCs

We asked whether SMARCAD1 expression relates to naive-state pluripotency. To test this, we knocked down Smarcad1 (Smarcad1 knockdown [KD]) in naive ESCs with two different short hairpin RNA (shRNA) constructs. These constructs reduced Smarcad1 mRNA to 42% of the original level, on average (Figure 3E), and resulted in decreased protein levels (Figure 3F). We examined cell morphology, alkaline phosphatase (AP) activity, energy metabolism, contribution to ICM in chimeras, switch of culture condition, and the transcriptome of Smarcad1 KD cells.

Within 48 hr of shRNA transfection, mouse ESCs lost their typical morphology of compact, rounded colonies and exhibited

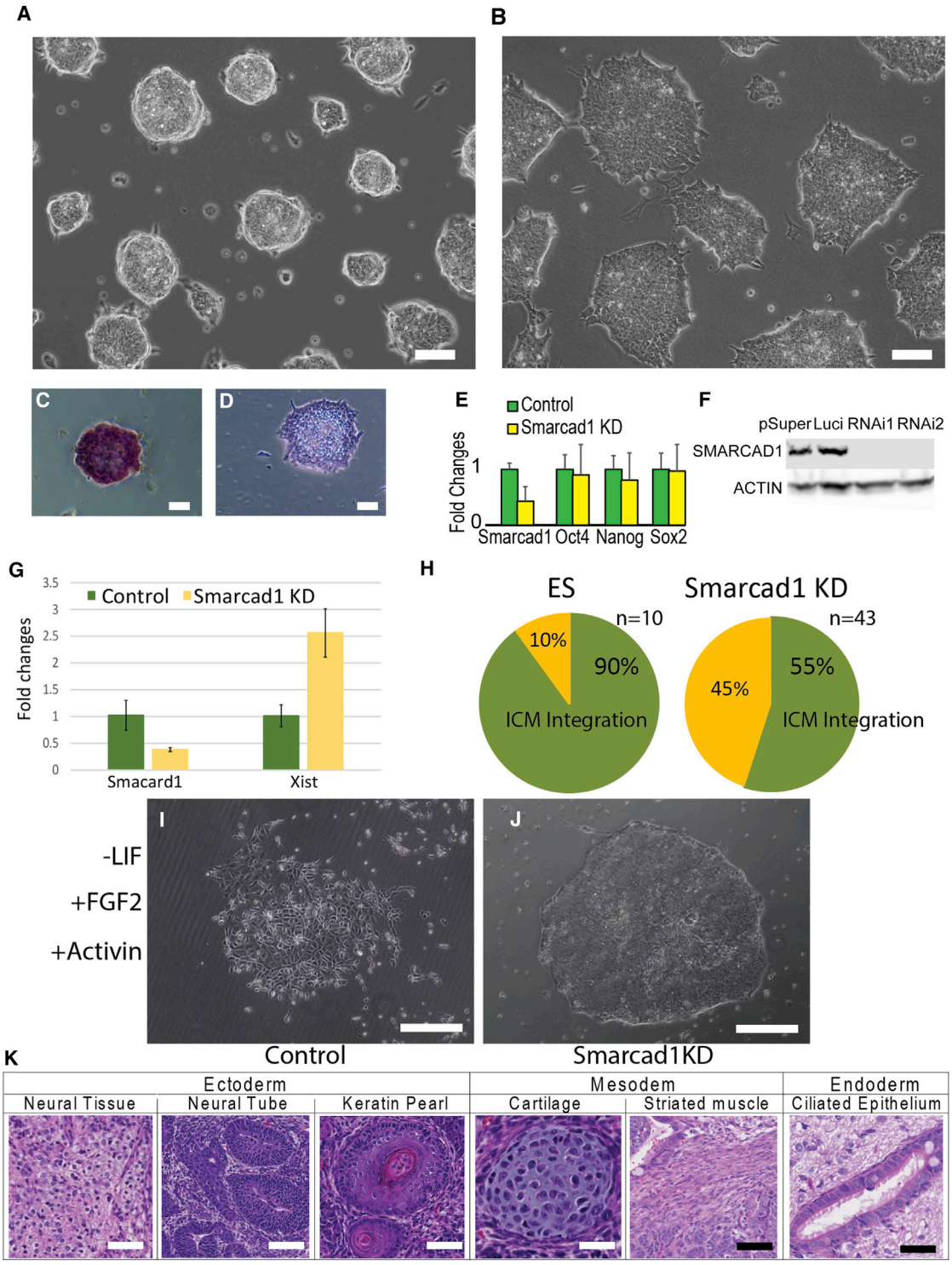


Figure 3. Smarcad1 KD in Mouse ESCs

(A and B) Mouse ESCs transfected with a control shRNA (A) and a *Smarcad1*-targeting shRNA (B).

(C and D) AP staining of control (C) and Smarcad1 KD cells (D).

(E) RT-PCR-derived expression fold changes between Smarcad1 KD and control KD cells (Control). Error bars were derived from three biological replicates.

(F) Western blots. pSuper, empty vector control; Luci, control knockdown; RNAi1 and RNAi2, two shRNAs targeting different parts of *Smarcad1* mRNA.

(G) Xist expression measured by qPCR in control (luciferase KD) El16.6 (green), and Smarcad1 KD El16.6 (yellow) ESCs.

(H) Proportions of injected embryos with ICM integration. ES, ESCs.

(legend continued on next page)

a flattened shape (Figures 3A and 3B). These changes were consistently observed in the two shRNA transfections (Figures S3A–S3C). We passaged Smarcad1 KD cells for more than 80 passages and could not observe colonies with the appearance of differentiation, but flattened, monolayer colonies reminiscent of EpiSCs (primed state) were persistently observed (Figures S3A–S3C). A total of 27 out of 30 Smarcad1 KD colonies lost AP activity (Figure 3C), a marker of self-renewal that is absent in EpiSCs (Brons et al., 2007), whereas nine out of ten control (Luciferase KD) colonies retained AP activity (Figure 3D). Compared to control (luciferase KD), Smarcad1 KD cells exhibited a reduced oxygen consumption rate and an increased extracellular acidity rate (Figures S4A and S4B), suggesting greater dependence of glycolysis for energy metabolism (Zhou et al., 2012). In a female ESC line (EL16.6, established by the Jeannie Lee lab [Zhao et al., 2008]), Smarcad1 KD cells exhibited a 2.6-fold increase in Xist expression, as compared to luciferase KD control (Figure 3G).

We compared the frequencies of the ICM contribution of chimeric blastocysts using an Oct4-EGFP reporter ESC line (OGR1). While 90.0% of the embryos (n = 10) injected with OGR1 exhibited ESC integration to ICM, 55.8% of Smarcad1 KD OGR1-injected embryos (n = 43) exhibited ICM integration, suggesting a reduction of chimeric formation in Smarcad1 KD cells ($p = 0.04$, Fisher's exact test) (Figures 3H and S4C–S4E). In summary, Smarcad1 KD ESCs exhibited phenotypic differences compared to naive ESCs in cell morphology, AP activity, energy metabolism, and capacity of forming chimeras. Moreover, these phenotypic assays were carried out without transferring the cells to an EpiSC culture condition (with fibroblast growth factor 2 [FGF2] and ACTIVIN-A, without leukemia inhibitory factor [LIF]) (Wu et al., 2015), suggesting that it is unlikely that the observed phenotypic differences were due to environmental differences. Therefore, by controlling for the culture condition, Smarcad1 KD cells maintained in ESC culture may offer a unique opportunity to dissect the genetic factors responsible for naive pluripotency.

Finally, we moved cells from the ESC culture condition (with LIF) into the EpiSC culture condition (with FGF2 and ACTIVIN-A, without LIF) (Wu et al., 2015). After changing to the EpiSC culture condition, control (luciferase KD) E14 ESCs exhibited a differentiated cell morphology at passage 5 (Figure 3I). In comparison, Smarcad1 KD E14 ESCs kept their monolayer, human ESC-like morphology at the same passage (Figure 3J). This suggests that Smarcad1 KD and control ESCs respond differently to FGF2/ACTIVIN-A signaling. We note that all other characterizations of the Smarcad1 KD cells in the rest of this paper were carried out in the ESC culture condition (with LIF) (Li et al., 2011) for the purpose of teasing out the direct effects of Smarcad1 expression change.

Smarcad1 KD ESCs Retain Pluripotency Characteristics

We asked whether Smarcad1 KD ESCs lost pluripotency. Real-time qPCR analysis showed no expression differences of pluripotency markers *Oct4*, *Nanog*, and *Sox2* in Smarcad1 KD cells (passage 50) and in control ESCs (passage 50) (Figure 3E). Subjected to an embryoid body (EB) formation assay, Smarcad1 KD cells (passage 35) formed EBs that were morphologically indistinguishable from those formed by control cells (passage 35) (Figures S5D and S5E). Marker genes of all three germ layers, including *Mtap2*, *Nestin*, *Cd31*, *T*, *Flk-1*, *Gata4*, *Afp*, *Ihh*, and *Gata6*, were expressed 5 to 1,500 times higher in 16-day EBs than in ESCs but at similar levels in control EBs and in SMARCAD1 KD ESC-derived EBs (Figure S5F). No discernable difference in cell proliferation rates was identified between Smarcad1 KD cells and control (luciferase KD) cells (Figures S5B and S5C).

To test teratoma formation, we injected four mice with Smarcad1 KD cells (passage 55) and control ESCs (passage 55). The total numbers of tumors formed were the same (i.e., seven) for Smarcad1 KD and control cells (Figures S5G and S5H; Table S1). In the teratomas originating from Smarcad1 KD cells, we found characteristic cell types and tissues of all three germ layers, including neural cells, the neural tube, keratin pearl, cartilage, striated muscle, and ciliated epithelium (Figure 3K). Collectively, the expression of pluripotency markers and the formation of EBs and teratomas suggest that Smarcad1 KD ESCs maintained the capacity for pluripotent lineage specification.

Transcriptome of Smarcad1 KD Cells in ESC Culture

It remains impossible to dissect the genetic regulators of naive-to-primed transition from environmental regulators, because, in vitro, naive or primed cells have to be maintained in their respective culture conditions and, in vivo, these cells are exposed to different signals as well (Han et al., 2010). We assayed the transcriptomes of control (naive) and Smarcad1 KD ESCs cultured in ESC medium using microarrays and RNA-seq and compared them with published gene expression data of EpiSCs cultured in EpiSC medium and wild-type ESCs (naive) cultured in ESC medium (Tesar et al., 2007; Factor et al., 2014; Huang et al., 2014). Because our Smarcad1 KD cells were maintained in ESC medium, we did not expect the expression differences between Smarcad1 KD cells and ESCs to exhibit genomewide correlations to the expression differences between EpiSCs (in EpiSC culture) and ESCs (in ESC culture); however, the data exhibited a moderate genomewide correlation (Figures S3F and S3G) (Supplemental Analysis). In particular, Smarcad1 KD cells in ESC culture exhibited reduced expression of naive pluripotency markers *Klb*, *Tbx3*, *Bmp4*, *Tfcp2l1*, *Tet2*, *Piwil2*, *Klf4*, *Stra8*, *Fgf4*, *Wnt6*, *Esrrb*, *Pecam1*, and *Zfp42*, and they exhibited increased expression of primed-state markers *Smad3*, *Fgf8*, *Lefty1*, *Pitx2*, *Fgf5*, *Pim2*, and *Fabp7* (Figure S3D). In addition, nine genes in the ACTIVIN/TGF-β (transforming growth factor β) pathway were upregulated in Smarcad1 KD cells, making the TGF-β pathway the only signaling pathway that was statistically enriched with differentially expressed genes (Fisher's test, $p < 0.001$; Figure S3E). Activation of the ACTIVIN/TGF-β signal is another hallmark of the primed state, but previously reported changes in this pathway were compounded with changes of

(I and J) Control (luciferase KD) (I) and Smarcad1 KD (J) ESCs cultured under EpiSC culture condition.
(K) Tissues and cell types identified by histological staining of EBs derived from SMARCAD1 KD E14 ESCs.
Scale bars, 100 μm in (A)–(D) and (I)–(K).

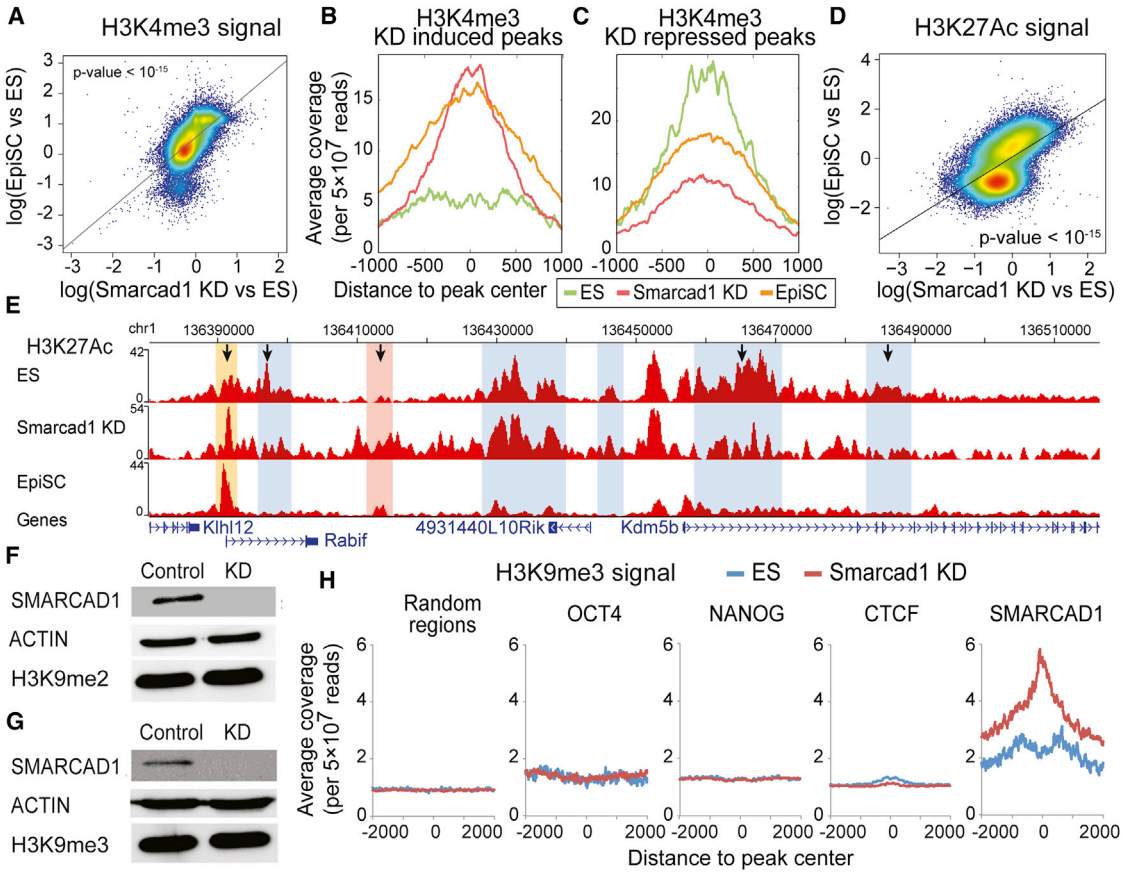


Figure 4. Epigenomic Difference between Smarcad1 KD Cells and ESCs

(A) H3K4me3 changes (log ratio) between Smarcad1 KD cells and control ESCs (ES) (x axis) versus H3K4me3 changes (log ratio) between EpiSCs and ESCs (y axis). Each dot represents an H3K4me3 peak identified in either EpiSCs or ESCs. 27,431 peaks (the union of 16,115 peaks in ESCs and 28,431 peaks in EpiSCs) are plotted.

(B and C) Average H3K4me3 ChIP-seq intensities (y axis) in Smarcad1 KD cells (red), ESCs (green), and EpiSCs (yellow) in a total of 565 Smarcad1 KD-induced peaks (B) and a total of 496 Smarcad1 KD-repressed peaks (C).

(D) H3K27ac differences between Smarcad1 KD cells and ESCs (x axis) versus H3K27ac changes between EpiSCs and ESCs (y axis) on the union of H3K27ac peaks (43,797) identified from EpiSCs and ESCs.

(E) H3K27ac distribution near the Kdm5b locus in ESCs, Smarcad1 KD cells, and EpiSCs, marked with previously (Factor et al., 2014) identified ESC-specific peaks (blue) and EpiSC-specific peaks (pink). H3K27ac in Smarcad1 KD exhibited reduced signals in three ESC-specific peaks and increased signals in the EpiSC-specific peak (marked with arrows). A new H3K27ac peak was identified (yellow), where both EpiSCs and Smarcad1 KD cells exhibited increased signals as compared to ESCs.

(F and G) Western blots of H3K9me2 (F) and H3K9me3 (G) in control (luciferase KD) and Smarcad1 KD (KD) mouse ESCs.

(H) Average H3K9me3 ChIP-Seq signals (read counts per 5 × 10^7 reads) in ESCs (blue) and Smarcad1 KD cells (red) are plotted against the distances to peak centers (x axis, peak center = 0) of nine chromatin-binding proteins (rows). A total of 10,000 random genomic locations are also included (Random regions).

external signals (Brons et al., 2007; Merrill, 2012). Finally, neither *Cdkn1a* (a.k.a. *P21*) nor *Cdkn2a* (a.k.a. *P16*) exhibited increased expression, which suggested that these cells were unlikely to have undergone senescence (Figure S5A).

Histone Modification Changes Induced by Smarcad1 KD

In order to characterize Smarcad1 KD-induced histone modification changes, we carried out ChIP-seq on H3K4me3 and H3K27ac in Smarcad1 KD cells (in ESC medium) and compared them with published ChIP-seq datasets in EpiSCs (in EpiSC medium) and ESCs (in ESC medium). Using published H3K4me3 ChIP-seq (GEO: GSM1382218), we identified a total of 27,431 peaks from EpiSCs or ESCs and calculated modification intensity in each peak. We compared the ratio of modification intensities between EpiSCs and ESCs (Figure 4A, y axis) to the ratio between Smarcad1 KD cells and ESCs (Figure 4A, x axis). The ratios between EpiSCs and ESCs are correlated with that between Smarcad1 KD cells and ESCs in a genomewide manner (Figure 4A, $p < 10^{-15}$). Next, we identified the genomic regions with strong induction of H3K4me3 in Smarcad1 KD (KD-induced peaks) and those with strong repression (KD-suppressed peaks). H3K4me3 in EpiSCs exhibited increased intensities in KD-induced peaks (Figure 4B) and reduced intensities in KD-suppressed peaks (Figure 4C). Taken together, without changing

the culture condition, suppression of Smarcad1 expression in ESCs makes the genomewide distribution of H3K4me3 shift toward that in EpiSCs.

We identified a total of 43,797 H3K27ac peaks in either EpiSCs or ESCs (GEO: GSM1382218). The ratio of H3K27ac intensities between EpiSCs and ESCs (Figure 4D, y axis) is correlated with that between Smarcad1 KD cells and ESCs in a genomewide manner (Figure 4D, x axis, $p < 10^{-15}$). In particular, we examined the Kdm5b gene locus, which was the main example where difference in H3K27ac was observed between EpiSCs and ESCs (see Figure 2 in Factor et al., 2014). Factor et al. (2014) identified one EpiSC-specific and five ESC-specific H3K27ac peaks. Smarcad1 KD cells exhibited H3K27ac decrease on three out of the five ESC-specific peaks (Figure 4E, blue regions) and H3K27ac increase on the EpiSC-specific peak (Figure 4E, pink region). These data suggest that Smarcad1 KD shared some, but not all, enhancer switches with EpiSCs, as indicated by H3K27ac changes. In addition, we identified a new EpiSC-specific peak at the Kdm5b locus (Figure 4E, yellow bar), which exhibits increased H3K27ac in both EpiSCs and Smarcad1 KD cells. Taken together, H3K27ac changes between Smarcad1 KD cells and ESCs correlate with the changes between EpiSCs (in EpiSC medium) and ESCs (in ESC medium).

SMARCAD1 Binding Modulates H3K9me3 at a Subset of SMARCAD1-Binding Sites

We explored the downstream components of the proposed SMARCAD1-chromatin interaction. ChIP-seq data suggested moderate co-localization of SMARCAD1 and H3K9me3 (Figure 2C, H3K9me3 columns). To assess the biological significance, we used SETDB1-H3K9me3 co-localization as a positive control, because H3K9 methyltransferase SETDB1 was thought to co-localize with H3K9me3 in mouse ESCs (Schultz et al., 2002; Karimi et al., 2011). Compared to SETDB1, SMARCAD1 ChIP-seq peaks exhibited larger overlaps with H3K9me3 peaks (Figures S2C and S2D), suggesting that the extent of SMARCAD1-H3K9me3 co-localization was not trivial. Considering that SMARCAD1 does not recognize H3 tail with H3K9me3 in vitro (Figure 2B, column 17), we posited that H3K9me3 is downstream to SMARCAD1 binding to chromatin.

To test this, we started by comparing the total amounts of H3K9me2 and H3K9me3 in wild-type and Smarcad1 KD ESCs, which revealed no discernable difference (Figures 4F, 4G, and S2G). Next, we carried out H3K9me3 ChIP-seq in Smarcad1 KD cells and compared it to H3K9me3 ChIP-seq data in ESCs. There was no discernable difference of H3K9me3 modification levels in 10,000 randomly selected genomic regions (Figure 4H), which is consistent with the western blot data (Figures 4F and 4G). In addition, there was no discernable difference in the ChIP-seq peaks of OCT4, NANOG, and CTCF (Figure 4H). However, in the 1,279 SMARCAD1 peaks (called with ChIP-seq in ESCs using MACS v. 1.4.0beta) (Zhang et al., 2008), H3K9me3 modification levels were substantially increased in Smarcad1 KD cells compared to ESCs ($p < 10^{-15}$, two-tailed t test) (Figure 4H). Furthermore, the largest changes of H3K9me3 precisely appeared at the centers of the Smarcad1 peaks (Figure 4H, x = 0). These data suggest that, although there are no global changes of H3K9me3 in Smarcad1 KD cells, there is an increase of H3K9me3 in SMARCAD1-binding regions. These data are incompatible with an alternative model, where Smarcad1 KD caused differentiation, which, in turn, resulted in global chromatin condensation and global H3K9me3 increases.

To further assess whether the H3K9me3 changes in SMARCAD1 peaks were due to reduced SMARCAD1 binding in the Smarcad1 KD condition, we conducted SMARCAD1 ChIP-seq in Smarcad1 KD E14 ESCs. Out of the 1,279 SMACARD1 peaks detected in ESCs, 1,207 (94.4%) exhibited a decrease of ChIP-seq signal in Smarcad1 KD cells, and 897 (70%) exhibited a strong decrease (KD sensitive). Among these 897 KD-sensitive SMARCAD1 peaks, 166 (18.5%) exhibited an increase of H3K9me3 levels in Smarcad1 KD, 3 (0.33%) exhibited decrease, and the rest (728) did not exhibit a significant change of H3K9me3 levels. Therefore, although the total amount of H3K9me3 did not appear to change with Smarcad1 KD, Smarcad1 KD is associated with increased H3K9me3 in a subset of SMARCAD1-binding regions.

H3Cit Is Negatively Associated with H3K9me3 in SMARCAD1-Binding Regions

The data led us to speculate a model that SMARCAD1 binding prevents the formation of extremely dense H3K9me3 regions. Considering the in vitro binding (Figures 2A and 2B) and in vivo (Figures 2C and 2D) correlation of SMARCAD1 and H3Cit, the aforementioned model is consistent with the idea that H3Cit interferes with H3K9me3-mediated heterochromatin formation (Sharma et al., 2012), and the latter is a characteristic of ESC differentiation. However, the data do not exclude H3Cit-independent mechanisms of SMARCAD1 recruitment to chromatin.

The aforementioned model would predict a negative correlation of H3Cit and H3K9me3 in SMARCAD1-binding regions. Among the 79 H3R26Cit peaks bound by SMARCAD1 in the entire genome, 73 (92.4%) exhibited at least a 1.5-fold increase of H3K9me3 in Smarcad1 KD cells (Figure 5E). In particular, the genomic loci of naive-state markers *Klf4* and *Wnt6* exhibited 2-fold more increases of H3K9me3 in Smarcad1 KD cells (Figures S6A and S6B).

To further test the speculated model, we inhibited H3Cit by Cl-amidine treatment (Christophorou et al., 2014) and carried out SMARCAD1 ChIP-seq and H3K9me3 ChIP-seq in Cl-amidine-treated (Cl+) cells. Compared to untreated ESCs, SMARCAD1 ChIP-seq signals exhibited mild decreases near H3R26Cit peaks (H3R26Cit peaks were defined by ChIP-seq in untreated ESCs) (Figure 5A). Next, we compared H3K9me3 ChIP-seq signals between Cl+ and untreated ESCs. No discernable H3K9me3 changes were identified in 10,000 randomly selected regions (Figure 5B). In addition, H3K9me3 ChIP-seq barely exhibited any changes in OCT4, NANOG, CTCF, H3K4me1, and H3K4me3 peaks (defined in untreated ESCs) either (Figure 5C). However, H3K9me3 exhibited strong increases in H3R26Cit peaks (defined in untreated ESCs) and in SMARCAD1 peaks (defined in untreated ESCs), lending additional support to the hypothesized roles of H3Cit and SMARCAD1 (Figure 5D and 5E). Furthermore, in the overlapped peaks of H3R26Cit and SMARCAD1 (defined in untreated ESCs), H3K9me3 exhibited even greater increases than those in H3R26Cit peaks or in

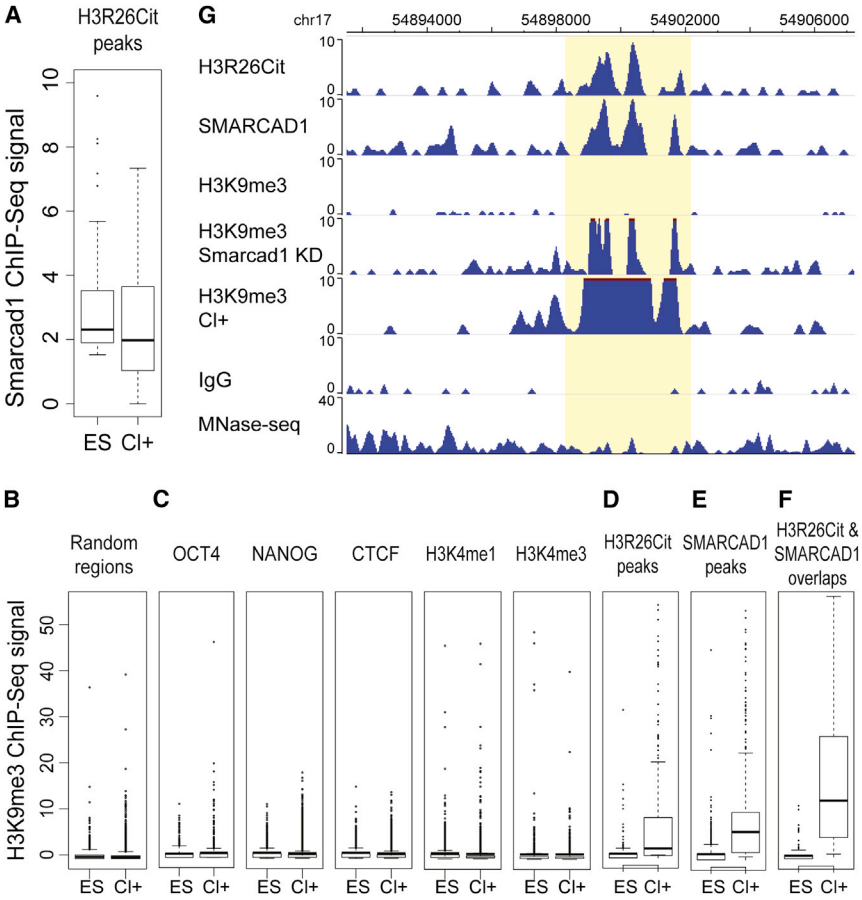


Figure 5. SMARCAD1 and H3K9me3 Changes in CI-Amidine Treatment

(A) SMARCAD1 ChIP-seq intensities in CI-amidine-treated (CI+) cells and untreated ESCs (ES).

(B–F) H3K9me3 ChIP-seq intensities in CI+ cells and ESCs in 10,000 random genomic regions (B); OCT4, NANOG, CTCF, H3K4me3, and H3K4me1 ChIP-seq peaks (C); H3R26Cit peaks (D); SMARCAD1 peaks (E); and the overlaps of H3R26Cit and SMARCAD1 peaks (F). All peaks were defined by ChIP-seq in ESCs.

(G) A genome browser view of H3R26Cit, SMARCAD1, H3K9me3 ChIP-seq in ESCs; H3K9me3 in Smarcad1 KD cells; H3K9me3 in CI+ ESCs; and IgG ChIP-seq and MNase-seq (sequencing of input DNA fragmented by micrococcal nuclease [MNase]) in ESCs.

SMARCAD1 peaks (Figure 5F). The specific H3K9me3 changes in SMARCAD1 peaks are also incompatible with the alternative model mentioned earlier.

Reproducibility of Genomewide SMARCAD1 Binding in Male and Female ESCs

To assess the variation of ChIP-seq data in different ESC lines, we generated SMARCAD1 ChIP-seq from a female ESC line derived by Wu et al. (Wu1) from C57BL mice (Wu et al., 2015) and compared it to E14 (derived from a male 129X1 mouse) data. In order to make an unbiased comparison, we fragmented the mouse genome (mm9) into 500-nt bins. After removing the bins with low mappability, we retained 4,854,116 bins. We calculated the ratio between SMARCAD1 and IgG ChIP-seq reads (SMARCAD1/IgG) in every bin and compared these ratios between the two ChIP-seq datasets. SMARCAD1 ChIP-seq in E14 and Wu1 cells exhibited a degree of correlation (Pearson correlation = 0.378) comparable to that of two H3K27me3 ChIP experiments conducted in the same cell line (Pearson correlation = 0.359) (Kaneko et al., 2013; Xiao et al., 2012) and that of two ESC ESET (H3K9 methyltransferase, ERG-associated protein with SET domain) ChIP-seq experiments (Pearson correlation = 0.344) (Yuan et al., 2009; Bilodeau et al., 2009; Figure 6). A caveat of this analysis is that Wu1 ESCs were cultured in the 2i condition (Wu et al., 2015). We transferred Wu1 cells from 2i medium into LIF-based culture condition (with LIF, and without 2i, as described in Li et al. [2011]), cultured them for 48 hours, and then performed SMARCAD1 ChIP-seq.

DISCUSSION

A Candidate System for Dissecting Genetic and Environmental Regulators of Naive-To-Primed Transition

It has been 10 years since the initial derivation and characterization of cultured primed-state pluripotent stem cells (Brons et al., 2007; Tesar et al., 2007). It remains extremely difficult to analyze the earliest genetic factors that underscore the naive-to-primed transition, because naive and primed cells are exposed to different signals either in vivo or in vitro (Han et al., 2010; Weinberger et al., 2016).

H3Cit was reported to correlate with the naive pluripotent state (Christophorou et al., 2014), but inhibition of H3Cit may not be an ideal system for studying naive-to-primed transition, because it resulted in cell differentiation as seen in global chromatin compaction, induction of differentiation markers, and reduction of Nanog mRNA expression to one-third that of the control ESCs (Christophorou et al., 2014). Compared to H3Cit inhibition, Smarcad1 KD phenotypes were more delicate. On one hand, Smarcad1 KD cells remain capable of pluripotent lineage specification; on the other hand, Smarcad1 KD ESCs lost representative cellular features of the naive state, even when they remain in ESC culture medium. This unique property makes Smarcad1 KD ESCs a candidate for the desired in vitro system.

H3Cit was thought to prevent heterochromatin formation by weakening H3K9me3 (Sharma et al., 2012); however, the mechanism was unknown. The histone modification changes in Smarcad1 repression and H3Cit inhibition conditions—especially the stepwise increases of H3K9me3 from other genomic regions to H3R26Cit peaks and SMARCAD1 peaks and then to

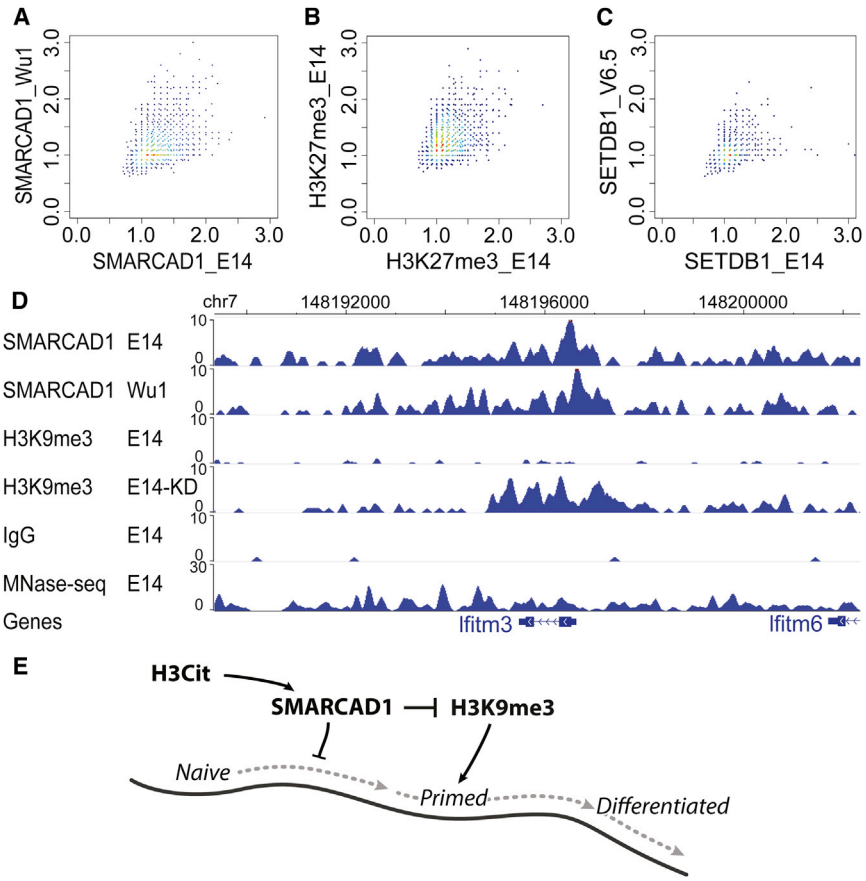


Figure 6. Variation of SMARCAD1 ChIP-Seq in Male and Female ESCs

(A–C) The mappable portion of the genome (mm9) was separated into a total of 4,854,116 non-overlapping 500-bp bins. A subset of 5,000 bins was drawn at random for plotting.

(A) Scatterplot of the ratio of SMARCAD1 and IgG ChIP-seq reads in 500-bp bins in E14 (male, x axis) and in Wu1 (female, y axis).

(B) Scatterplot between two biological replicates (GSM1199184 on the y axis, and GSM2111307 on the x axis). Each dot is a ratio of H3K27me3 and IgG ChIP-seq reads in a 500-bp bin.

(C) Scatterplot of ratio between ESET ChIP-seq and IgG ChIP-seq reads in E14 (GSM440256) (x axis) and V6.5 (GSM459273) (y axis) ESCs.

(D) Genome browser view of SMARCAD1 ChIP-seq in E14 and Wu1 cells, and H3K9me3 in E14 and in Smarcad1 KD cells (E14-KD). MNase-seq, sequencing of input DNA fragmented by MNase.

(E) A speculated model for the role of SMARCAD1 in regulating the naive pluripotent state. Naive pluripotent, primed pluripotent, and differentiated cells are situated on an epigenetic landscape (dark curve). In this model, SMARCAD1 contributes to keep cells at the highest position on the epigenetic landscape by translating H3Cit into an inhibitory signal of H3K9me3.

H3R26Cit/SMARCAD1 overlapping peaks—suggest SMARCAD1 could be a missing link of H3Cit-mediated suppression of H3K9me3 (Figure 6E).

Other Possible Mechanisms for SMARCAD1 Recruitment to Chromatin

The aforementioned data do not rule out H3Cit-independent mechanisms for SMARCAD1 recruitment to chromatin. The lack of DNA-binding specificity in vitro does not rule out the scenario where SMARCAD1 is tethered to chromatin by co-factors. LTR (long terminal repeat) retrotransposons and satellite repeats appear more frequently in SMARCAD1 peaks than in the entire genome (see Supplemental Analysis). In addition, SMARCAD1 may interact with histone deacetylase HDAC1 (Rowbotham et al., 2011), which may explain why SMARCAD1 binds to H3K27ac in vitro but SMARCAD1 binding only moderately correlates with H3K27ac in vivo. Furthermore, the peptide array data suggest that H3 S28 phosphorylation (H3S28P) inhibits SMARCAD1 binding to H3 in vitro (Figure S1), which suggests that combinatorial histone modifications could affect SMARCAD1 interaction with chromatin. Finally, we did not find a correlation between nucleosome positions (Chen et al., 2013) and SMARCAD1 binding. Nucleosome phase lengths were slightly reduced in SMARCAD1 peaks, but they did not exhibit clear differences between SMARCAD1-associated H3R26Cit peaks and the entire genome (Figure S2E).

Open-Chromatin versus Closed-Chromatin Hypotheses

H3Cit was thought to interfere with H3K9me3-mediated heterochromatin formation (open-chromatin hypothesis) (Sharma et al., 2012), whereas SMARCAD1 was thought to associate with and promote H3K9me3, a mark of closed chromatin (closed-chromatin hypothesis) (Rowbotham et al., 2011). At a first glance, the opposite effects do not seem to allow a simple model that directly puts H3K9me3 downstream to the proposed H3R26Cit/SMARCAD1 pathway. The increase of H3K9me3 in Smarcad1 KD cells on H3R26Cit peaks is better aligned with the open-chromatin hypothesis. In addition, the closed-chromatin hypothesis primarily relied on the observation of strong decreases of the total amounts of H3K9me2 and H3K9me3 in the Smarcad1 KD condition (Rowbotham et al., 2011). We carried out the same test in ESCs. The total amounts of H3K9me2 and H3K9me3 did not decrease in Smarcad1 KD ESCs, as compared to control ESCs, in multiple independent experiments (Figures 4F and 4G). These different results are potentially attributable to cell-type differences. In summary, at least in naive ESCs and on H3R26Cit peaks, SMARCAD1 binding may function as a checkpoint, which prevents the formation of overly dense H3K9me3 regions.

H3K9me3 changes in the genomic loci of naive pluripotency markers may not be the only plausible explanation of Smarcad1 KD induced change of pluripotent state. An non-exclusive alternative hypothesis is that SMARCAD1 binding mediated suppression of H3K9me3 keeps certain retrotransposons transcriptionally active, which is a feature of increased developmental

potency and is correlated to an earlier developmental stage (Macfarlan et al., 2012). Consistent to this hypothesis, LTR retrotransposons were enriched in SMARCAD1-binding regions and also enriched in Smarcad1-KD-induced H3K9me3 peaks (Supplemental Analysis). Interestingly, deletion of an H3K9 methyltransferase upregulated a set of LTR retrotransposons and their nearby genes (Karimi et al., 2011).

EXPERIMENTAL PROCEDURES

Institutional Permission and Oversight

The analysis of mouse embryos was approved by the University of California, San Diego, Institutional Animal Care and Use Committee (IACUC) and the University of Illinois Urbana-Champaign IACUC. All analyzed embryos were at blastocyst stage. The genders were not identified. Additional mouse gene expression data were retrieved from public records (GEO: GSE20187 [Tang et al., 2010] and GEO: GSE18290 [Xie et al., 2010]). All human data were retrieved from public records (GEO: GSE36552 [Yan et al., 2013]). The ESC analyses were approved by University of California, San Diego, Human Research Protection Program.

Change of Culture Condition

Under the ESC culture condition with LIF (Li et al., 2011), we carried out shRNA-mediated Smarcad1 KD and Luciferase KD (control). After 6 days of puromycin selection in ESC culture, the cells were transferred onto Matrigel-coated plates with EpiSC culture medium as previously described (Wu et al., 2015). Briefly, the EpiSC medium is composed of N2B27 basal medium and 20% knockout serum replacement (KSR), 12 ng/mL FGF2, and 2 ng/mL ACTIVIN-A.

Quantifying Genomewide Co-localization of Two Histone Modifications

Odds ratio was used to measure the extent of co-localization between two histone modifications, as previously described (Xiao et al., 2012). Briefly, the genome was split into 200-nt non-overlapping bins. Each histone modification was judged as either present or absent on each bin based on ChIP-seq data. A contingency table was built for the distribution of genomic bins, reflecting the presence and absence of each histone modification. Odds ratio was calculated based on this contingency table. The odds ratio larger or smaller than 1 reflects more or fewer overlaps of the two histone modifications than random expectation. The odds ratio between a chromatin-binding protein and a histone modification was calculated by the same approach, where one ChIP-seq dataset of histone modification was replaced by that of the chromatin-binding protein.

Cl-Amidine Treatment

Cl-amidine treatment of ESCs was carried out as previously described (Christophorou et al., 2014). Briefly, Cl-amidine (200 μM) was added to ESC culture medium, and E14 ESCs were maintained in this treatment condition for 48 hours. Approximately 10 million treated cells were crosslinked for each ChIP-seq experiment. Approximately 3 million treated cells were obtained at the same time for western blot experiments.

Additional materials and methods are available in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for all sequencing data reported in this paper is GEO: GSE45338.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, Supplemental Analysis, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2017.02.070.

AUTHOR CONTRIBUTIONS

S.X., B.S., and S.Z. designed the experiments. S.X., B.S., D.M., L.S., Y.W., B.P.V.L.T., J.C.Z., T.S.T., and G.S. performed the experiments. All authors analyzed the data and wrote the manuscript.

ACKNOWLEDGMENTS

We thank Dr. Azim Surani for providing EpiSCs, Drs. Jun Wu and Juan Carlos Izpisua Belmonte for providing Wu1 ESCs. We thank Drs. Sua Myong, Taekjip Ha, Wei Wang, Fernando H. Biase, David Granas, Ankur Jain, Xue Zou, Fabian Hertel, Jin Zhang, Lucie Hebert, and Norman Huang for carrying out related work and comments to the manuscript. This work is supported by NIH grants DP1HD087990 and DP2OD007417 to S.Z., R01EY025090 to K.Z., and R01HG008135 to S.Z. and K.Z.

Received: April 3, 2015
Revised: February 8, 2017
Accepted: February 23, 2017
Published: March 21, 2017

REFERENCES

Active Motif. (2014). MODified histone peptide array [Manual]. http://www.activemotif.com/documents/1691.pdf.

Adra, C.N., Donato, J.L., Badovinac, R., Syed, F., Kheraj, R., Cai, H., Moran, C., Kolker, M.T., Turner, H., Weremowicz, S., et al. (2000). SMARCAD1, a novel human helicase family-defining member associated with genetic instability: cloning, expression, and mapping to 4q22-q23, a band rich in breakpoints and deletion mutants involved in several human diseases. Genomics *69*, 162–173.

Bilodeau, S., Kagey, M.H., Frampton, G.M., Rahl, P.B., and Young, R.A. (2009). SetDB1 contributes to repression of genes encoding developmental regulators and maintenance of ES cell state. Genes Dev. *23*, 2484–2489.

Brons, I.G., Smithers, L.E., Trotter, M.W., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R.A., and Vallier, L. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. Nature *448*, 191–195.

Chen, K., Xi, Y., Pan, X., Li, Z., Kaestner, K., Tyler, J., Dent, S., He, X., and Li, W. (2013). DANPOS: dynamic analysis of nucleosome position and occupancy by sequencing. Genome Res. *23*, 341–351.

Christophorou, M.A., Castelo-Branco, G., Halley-Stott, R.P., Oliveira, C.S., Loos, R., Radzisheuskaya, A., Mowen, K.A., Bertone, P., Silva, J.C., Zernicka-Goetz, M., et al. (2014). Citrullination regulates pluripotency and histone H1 binding to chromatin. Nature *507*, 104–108.

Cockburn, K., and Rossant, J. (2010). Making the blastocyst: lessons from the mouse. J. Clin. Invest. *120*, 995–1003.

Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J., and Kouzarides, T. (2004). Histone deimination antagonizes arginine methylation. Cell *118*, 545–553.

Ezashi, T., Telugu, B.P., Alexenko, A.P., Sachdev, S., Sinha, S., and Roberts, R.M. (2009). Derivation of induced pluripotent stem cells from pig somatic cells. Proc. Natl. Acad. Sci. USA *106*, 10993–10998.

Factor, D.C., Corradin, O., Zentner, G.E., Saiakhova, A., Song, L., Chenoweth, J.G., McKay, R.D., Crawford, G.E., Scacheri, P.C., and Tesar, P.J. (2014). Epigenomic comparison reveals activation of "seed" enhancers during transition from naive to primed pluripotency. Cell Stem Cell *14*, 854–863.

Han, D.W., Tapia, N., Joo, J.Y., Greber, B., Araúzo-Bravo, M.J., Bernemann, C., Ko, K., Wu, G., Stehling, M., Do, J.T., and Schöler, H.R. (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. Cell *143*, 617–627.

Huang, K., Maruyama, T., and Fan, G. (2014). The naive state of human pluripotent stem cells: a synthesis of stem cell and preimplantation embryo transcriptome analyses. Cell Stem Cell *15*, 410–415.

Kaneko, S., Son, J., Shen, S.S., Reinberg, D., and Bonasio, R. (2013). PRC2 binds active promoters and contacts nascent RNAs in embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1258–1264.

Karimi, M.M., Goyal, P., Maksakova, I.A., Bilenky, M., Leung, D., Tang, J.X., Shinkai, Y., Mager, D.L., Jones, S., Hirst, M., and Lorincz, M.C. (2011). DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs. Cell Stem Cell *8*, 676–687.

Kidder, B.L., Palmer, S., and Knott, J.G. (2009). SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. Stem Cells *27*, 317–328.

Klose, R.J., and Zhang, Y. (2007). Regulation of histone methylation by demethylimination and demethylation. Nat. Rev. Mol. Cell Biol. *8*, 307–318.

Li, Y., Yokohama-Tamaki, T., and Tanaka, T.S. (2011). Short-term serum-free culture reveals that inhibition of Gsk3β induces the tumor-like growth of mouse embryonic stem cells. PLoS ONE *6*, e21355.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63.

Merrill, B.J. (2012). Wnt pathway regulation of embryonic stem cell self-renewal. Cold Spring Harb. Perspect. Biol. *4*, a007971.

Neeli, I., Khan, S.N., and Radic, M. (2008). Histone deimination as a response to inflammatory stimuli in neutrophils. J. Immunol. *180*, 1895–1902.

Okazaki, N., Ikeda, S., Ohara, R., Shimada, K., Yanagawa, T., Nagase, T., Ohara, O., and Koga, H. (2008). The novel protein complex with SMARCAD1/KIAA1122 binds to the vicinity of TSS. J. Mol. Biol. *382*, 257–265.

Rowbotham, S.P., Barki, L., Neves-Costa, A., Santos, F., Dean, W., Hawkes, N., Choudhary, P., Will, W.R., Webster, J., Oxley, D., et al. (2011). Maintenance of silent chromatin through replication requires SWI/SNF-like chromatin remodeler SMARCAD1. Mol. Cell *42*, 285–296.

Schoor, M., Schuster-Gossler, K., and Gossler, A. (1993). The Etl-1 gene encodes a nuclear protein differentially expressed during early mouse development. Dev. Dyn. *197*, 227–237.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G., and Rauscher, F.J., 3rd. (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. Genes Dev. *16*, 919–932.

Sharma, P., Azebi, S., England, P., Christensen, T., Møller-Larsen, A., Petersen, T., Batsché, E., and Muchardt, C. (2012). Citrullination of histone H3 interferes with HP1-mediated transcriptional repression. PLoS Genet. *8*, e1002934.

Soininen, R., Schoor, M., Henseling, U., Tepe, C., Kisters-Woike, B., Rossant, J., and Gossler, A. (1992). The mouse Enhancer trap locus 1 (Etl-1): a novel mammalian gene related to Drosophila and yeast transcriptional regulator genes. Mech. Dev. *39*, 111–123.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., Mansfield, W., et al. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254–1269.

Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M.A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-seq analysis. Cell Stem Cell *6*, 468–478.

Telugu, B.P., Ezashi, T., Sinha, S., Alexenko, A.P., Spate, L., Prather, R.S., and Roberts, R.M. (2011). Leukemia inhibitory factor (LIF)-dependent, pluripotent stem cells established from inner cell mass of porcine embryos. J. Biol. Chem. *286*, 28948–28953.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196–199.

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.

Wu, J., Okamura, D., Li, M., Suzuki, K., Luo, C., Ma, L., He, Y., Li, Z., Benner, C., Tamura, I., et al. (2015). An alternative pluripotent state confers interspecies chimaeric competency. Nature *521*, 316–321.

Xiao, S., Xie, D., Cao, X., Yu, P., Xing, X., Chen, C.C., Musselman, M., Xie, M., West, F.D., Lewin, H.A., et al. (2012). Comparative epigenomic annotation of regulatory DNA. Cell *149*, 1381–1392.

Xie, D., Chen, C.C., Ptaszek, L.M., Xiao, S., Cao, X., Fang, F., Ng, H.H., Lewin, H.A., Cowan, C., and Zhong, S. (2010). Rewirable gene regulatory networks in the preimplantation embryonic development of three mammalian species. Genome Res. *20*, 804–815.

Yan, L., Yang, M., Guo, H., Yang, L., Wu, J., Li, R., Liu, P., Lian, Y., Zheng, X., Yan, J., et al. (2013). Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1131–1139.

Yuan, P., Han, J., Guo, G., Orlov, Y.L., Huss, M., Loh, Y.H., Yaw, L.P., Robson, P., Lim, B., and Ng, H.H. (2009). Eset partners with Oct4 to restrict extraembryonic trophoblast lineage potential in embryonic stem cells. Genes Dev. *23*, 2507–2520.

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

Zhang, X., Bolt, M., Guertin, M.J., Chen, W., Zhang, S., Cherrington, B.D., Slade, D.J., Dreyton, C.J., Subramanian, V., Bicker, K.L., et al. (2012). Peptidylarginine deiminase 2-catalyzed histone H3 arginine 26 citrullination facilitates estrogen receptor α target gene activation. Proc. Natl. Acad. Sci. USA *109*, 13331–13336.

Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., and Lee, J.T. (2008). Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. Science *322*, 750–756.

Zhao, Y., Granas, D., and Stormo, G.D. (2009). Inferring binding energies from selected binding sites. PLoS Comput. Biol. *5*, e1000590.

Zhou, W., Choi, M., Margineantu, D., Margaretha, L., Hesson, J., Cavanaugh, C., Blau, C.A., Horwitz, M.S., Hockenbery, D., Ware, C., and Ruohola-Baker, H. (2012). HIF1α induced switch from bivalent to exclusively glycolytic metabolism during ESC-to-EpiSC/hESC transition. EMBO J. *31*, 2103–2116.