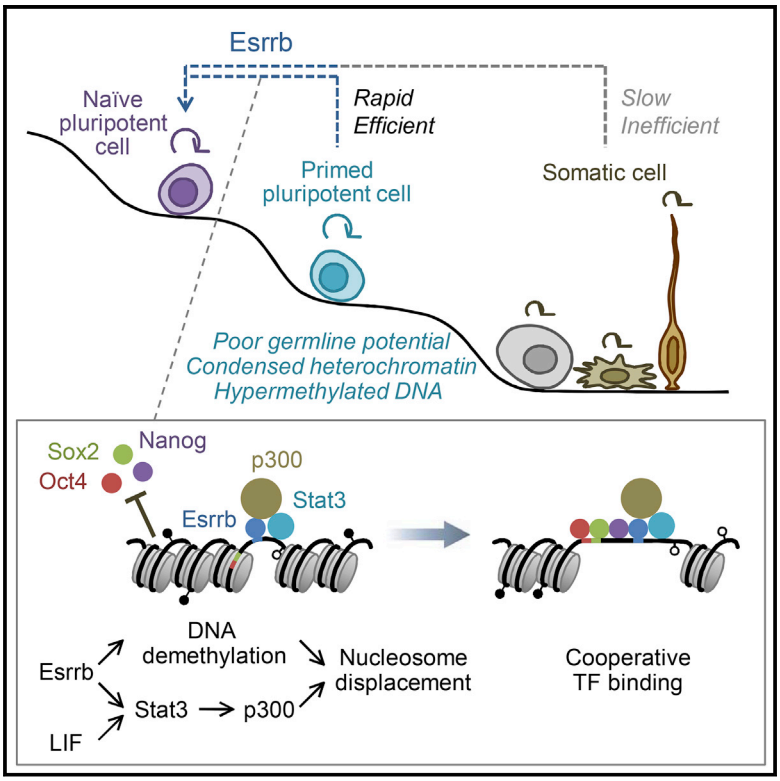
Short Article

Cell Stem Cell

# Esrrb Unlocks Silenced Enhancers for Reprogramming to Naive Pluripotency

## Graphical Abstract




## Authors

Kenjiro Adachi, Wolfgang Kopp, Guangming Wu, ..., Bernd Timmermann, Martin Vingron, Hans R. Schöler

## Correspondence

office@mpi-muenster.mpg.de

## In Brief

Rapid reprogramming of epiblast stem cells (EpiSCs) to induced pluripotent stem cells (iPSCs) allows a detailed analysis of epigenetic remodeling induced by transcription factors. The orphan nuclear receptor Esrrb plays a pioneering role in establishing a permissive chromatin environment for the recruitment of core pluripotency factors Oct4, Sox2, and Nanog.

## Highlights

- Loss of Esrrb leads to incomplete reprogramming of EpiSCs and MEFs to iPSCs
- Esrrb can bind to silenced enhancers containing stable nucleosomes and methylated DNA
- Esrrb and p300 establish permissive chromatin for core pluripotency TF recruitment
- Loss of DNA methylation facilitates LIF-driven nucleosome displacement

Adachi et al., 2018, Cell Stem Cell *23*, 266–275
August 2, 2018 
https://doi.org/10.1016/j.stem.2018.05.020

CellPress

CellPress

Cell Stem Cell
Short Article

# Esrrb Unlocks Silenced Enhancers for Reprogramming to Naive Pluripotency

Kenjiro Adachi,[1] Wolfgang Kopp,[2] Guangming Wu,[1] Sandra Heising,[1] Boris Greber,[3,4] Martin Stehling,[5] Marcos J. Araúzo-Bravo,[6,7] Stefan T. Boerno,[8] Bernd Timmermann,[8] Martin Vingron,[2] and Hans R. Schöler[1,9,10,*]
[1]Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
[2]Department of Computational Molecular Biology, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany
[3]Human Stem Cell Pluripotency Laboratory, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
[4]Chemical Genomics Centre of the Max Planck Society, 44227 Dortmund, Germany
[5]Flow Cytometry Unit, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
[6]Computational Biology and Systems Biomedicine, Biodonostia Health Research Institute, San Sebastián 20014, Spain
[7]IKERBASQUE, Basque Foundation for Science, Bilbao 48013, Spain
[8]Sequencing Core Facility, Max Planck Institute for Molecular Genetics, 14195 Berlin, Germany
[9]Medical Faculty, University of Münster, 48149 Münster, Germany
[10]Lead Contact
*Correspondence: office@mpi-muenster.mpg.de
https://doi.org/10.1016/j.stem.2018.05.020

## SUMMARY

**Transcription factor (TF)-mediated reprogramming to pluripotency is a slow and inefficient process, because most pluripotency TFs fail to access relevant target sites in a refractory chromatin environment. It is still unclear how TFs actually orchestrate the opening of repressive chromatin during the long latency period of reprogramming. Here, we show that the orphan nuclear receptor Esrrb plays a pioneering role in recruiting the core pluripotency factors Oct4, Sox2, and Nanog to inactive enhancers in closed chromatin during the reprogramming of epiblast stem cells. Esrrb binds to silenced enhancers containing stable nucleosomes and hypermethylated DNA, which are inaccessible to the core factors. Esrrb binding is accompanied by local loss of DNA methylation, LIF-dependent engagement of p300, and nucleosome displacement, leading to the recruitment of core factors within approximately 2 days. These results suggest that TFs can drive rapid remodeling of the local chromatin structure, highlighting the remarkable plasticity of stable epigenetic information.**

## INTRODUCTION

The pluripotent state of self-renewing embryonic stem cells (ESCs), so-called naive pluripotency, is maintained by a transcription factor (TF) network that is under the control of extrinsic signals (Ng and Surani, 2011). Mouse ESCs express the core pluripotency TFs Oct4, Sox2, and Nanog as well as auxiliary TFs such as Esrrb, Klf4, Klf2, and Tbx3, depending on the activation of LIF and WNT/beta-catenin signaling and the suppression of FGF signaling. The core TFs play a central role in maintaining a pluripotent state, as well as inducing pluripotency in somatic cells. Reprogramming to pluripotency by the ectopic expression of TFs is a slow and inefficient process involving multiple stochastic events. Although Oct4, Sox2, and Nanog co-occupy many pluripotency-specific enhancers in ESCs, access of these TFs to relevant target sites and subsequent activation of pluripotency genes are highly restricted in somatic cells (Chen et al., 2016; Chronis et al., 2017). As epigenetic remodeling enabling TF access occurs at the final step of reprogramming (Chen et al., 2016), the stochastic nature of reprogramming limits our understanding of the underlying molecular mechanisms.

Mouse epiblast stem cells (EpiSCs) are pluripotent stem cells originated from the postimplantation epiblast and exhibit distinct biological features from ESCs. EpiSCs express the core pluripotency TFs but not the aforementioned auxiliary TFs, analogous to intermediate cells of somatic cell reprogramming. EpiSCs exhibit a constrained developmental potential compared with ESCs—EpiSCs do not efficiently differentiate into germ cells *in vitro* (Hayashi et al., 2011) or *in vivo* (Ohtsuka et al., 2012). Electron microscopy studies have shown that ESCs and preimplantation epiblast cells have globally open chromatin structure characterized by uniformly dispersed chromatin, whereas EpiSCs and postimplantation epiblast cells have condensed heterochromatin domains (Ahmed et al., 2010). The preimplantation epiblast exhibits genome-wide DNA hypomethylation and two active X chromosomes in female cells, while the postimplantation epiblast undergoes *de novo* DNA methylation and X chromosome inactivation (Borgel et al., 2010). Consistent with the observed differences in chromatin status of EpiSCs and ESCs, EpiSCs rarely revert to a more primitive ESC state (Greber et al., 2010; Guo et al., 2009).

We have developed an efficient EpiSC reprogramming system to address how pluripotency TFs gain access to their target sites within repressive chromatin.

## RESULTS

### *Esrrb* Is a Crucial TF for Establishing Naive Pluripotency

We used a male EpiSC line containing a transgenic *Oct4-EGFP* reporter (E3; Greber et al., 2010) that was maintained under

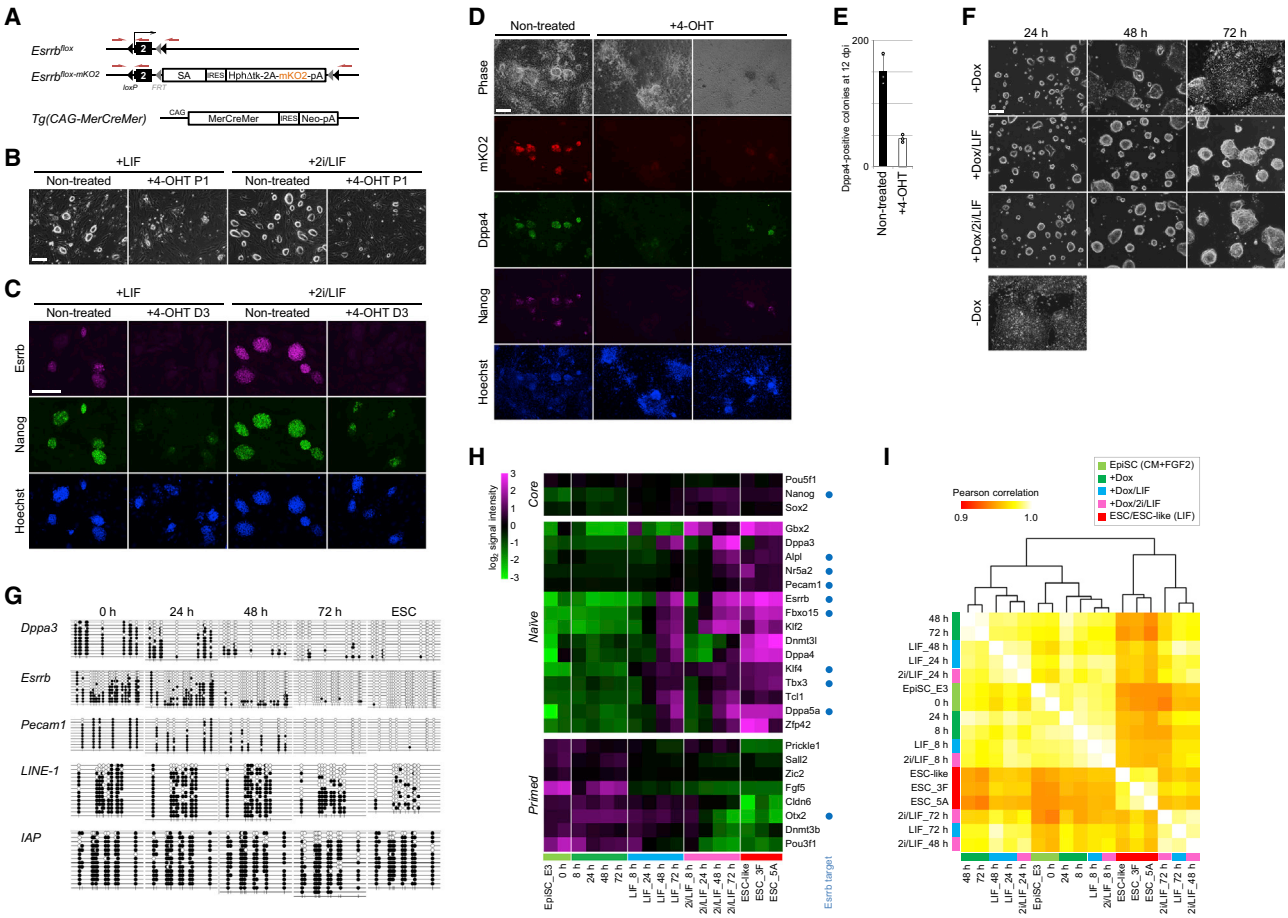



**Figure 1. The Crucial Role of Esrrb in Establishing Naive Pluripotency**

(A) Schematic representation of *Esrrb* conditional KO ESCs (*Esrrb*$^{flox/flox\text{-}mKO2}$) carrying the tamoxifen-inducible Cre (*MerCreMer*) transgene. The major transcription start site (TSS) is located upstream of exon 2 (Festuccia et al., 2012). Primer binding sites for genotyping in Figure S1E are indicated by red arrows. Tg, transgene; SA, splice acceptor; mKO2, monomeric Kusabira-Orange2.

(B) *Esrrb*$^{flox/flox\text{-}mKO2}$; *MerCreMer* ESCs were cultured in the presence of 4-hydroxytamoxifen (4-OHT) for one passage (P1). Scale bar, 200 μm.

(C) Immunostaining of *Esrrb* conditional KO ESCs treated with 4-OHT for 3 days. Scale bar, 200 μm.

(D) *Esrrb* conditional KO MEFs pretreated with 4-OHT were reprogrammed by Oct4, Sox2, and Klf4 for 12 days and immunostained for mKO2, Dppa4, and Nanog. Note that the colonies emerging from 4-OHT-pretreated MEFs contained many dead cells with fragmented nuclei. Scale bar, 200 μm.

(E) The number of Dppa4-positive colonies at 12 days post-infection (dpi). Open circles represent results from individual experiments in triplicate. Error bars represent SD of replicates.

(F) Tet-on Esrrb EpiSCs were cultured in the presence of Dox with or without LIF or 2i for up to 72 hr (h). Scale bar, 200 μm.

(G) Bisulfite sequencing analysis during reprogramming of EpiSCs. Genomic DNA was isolated from the cells cultured in 2i/LIF.

(H) Mean-centered expression values of pluripotency/epiblast markers obtained from microarray analysis. EpiSC_E3 represents the parental line and ESC-like represents reprogrammed Tet-on Esrrb EpiSCs cultured in the presence of LIF without Dox. ChIP-seq Esrrb targets during reprogramming (Figure 2) are indicated by filled blue circles.

(I) Pearson correlation coefficients for the microarray data.

See also Figures S1 and S2.

feeder-free conditions with minimal spontaneous differentiation (Figure S1A) and assessed the reprogramming activity of TFs involved in maintaining and inducing naive pluripotency. Gain-of-function studies identified the orphan nuclear receptor Esrrb as a potent reprogramming factor (Figure S1B), consistent with a previous report using ESC-derived EpiSCs (Festuccia et al., 2012). Furthermore, endogenous *Esrrb* was required for reprogramming by other TFs (Figure S1C). Esrrb plays an important role in the maintenance of ESCs (Martello et al., 2012) and in the generation of induced pluripotent stem cells (iPSCs) from mouse embryonic fibroblasts (MEFs; Feng et al., 2009). We confirmed that deletion of *Esrrb* compromised the self-renewal capacity of ESCs, showing decreased expression of Nanog and increased cell death (Figures 1A–1C). During fibroblast reprogramming by Oct4, Sox2, and Klf4, Esrrb was activated later than Sall4 and Nanog, concomitant with the expression of Dppa4 (Figures S1D–S1F). *Esrrb* knockout (KO) MEFs underwent reprogramming to form ESC-like colonies, but those colonies were often negative for Dppa4 and contained many dead cells (Figures 1D, 1E, S1E, and S1F).

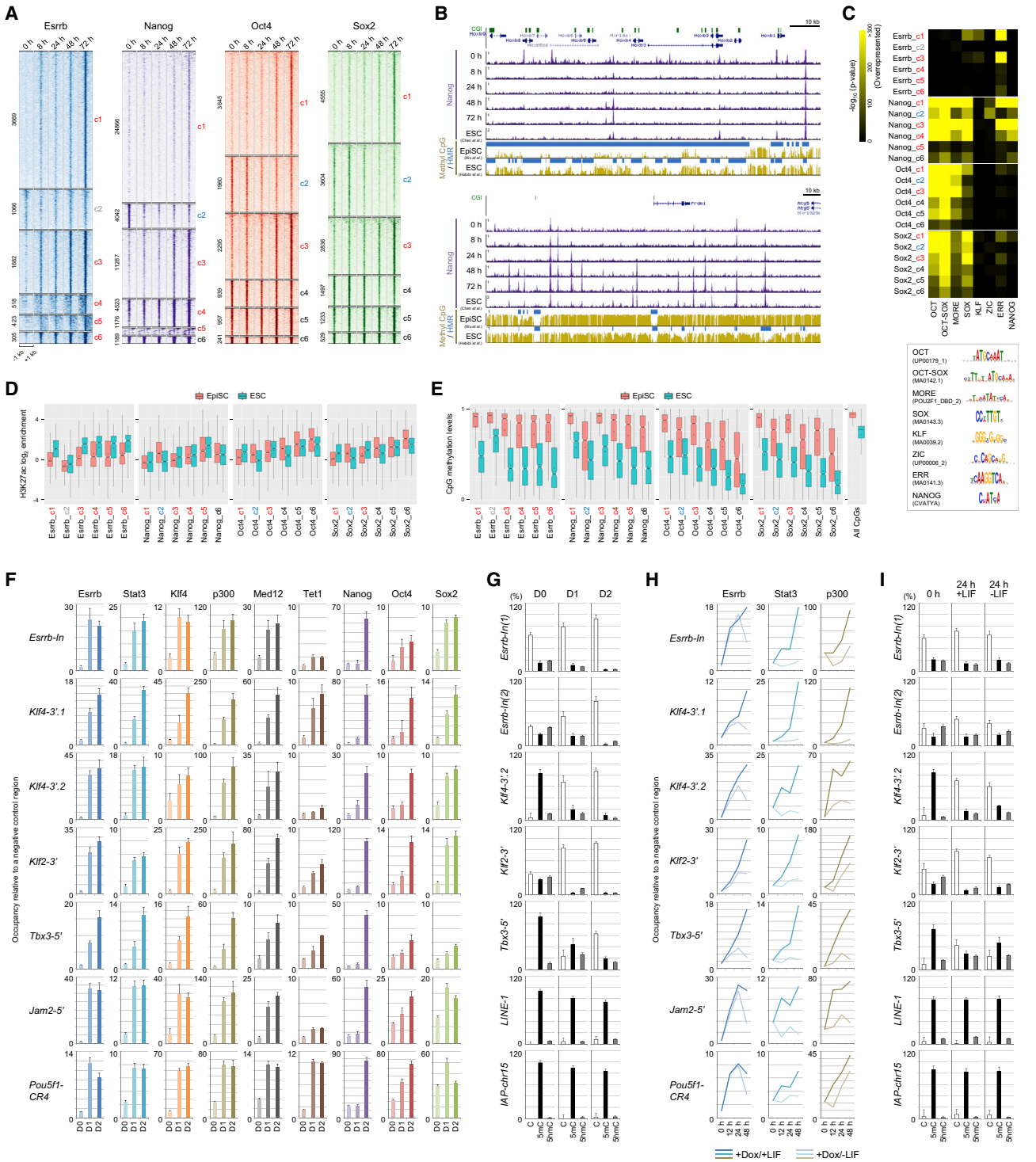



**Figure 2. Esrrb- and LIF-Dependent Remodeling of Core TF Occupancy during EpiSC Reprogramming**

(A) Heatmaps of ChIP-seq signals around the peaks that were grouped into six clusters (red, ESC specific; blue, EpiSC specific; black, common). Esrrb c2 peaks (gray), which do not show any increase in Esrrb occupancy or any DNA motif enrichment (Figure 2C), were considered as spurious. The number of peaks in each cluster is shown to the left of each heatmap.

(B) ChIP-seq tracks showing the EpiSC-specific (the *HoxB* cluster, top) and ESC-specific (the *Prdm1* locus, bottom) Nanog binding. CGI, CpG island.

(C) Enrichment of TF binding motifs in each peak cluster.

(D) Notched boxplots showing fold enrichment of H3K27ac (Factor et al., 2014; Whyte et al., 2012) relative to input around peaks in each cluster. Open circles in the boxplots represent means.

*(legend continued on next page)*

To elucidate the mechanistic role of Esrrb in establishing naive pluripotency, we developed a high-efficiency system to reprogram EpiSCs by the doxycycline (Dox)-induced expression of Esrrb (Figure S1G). Induction of Esrrb led to the rapid and homogeneous expression of the ESC markers Klf4 and Pecam1 (Figures S1H–S1J), and the acquisition of ESC-like colony morphology (Figure 1F). Most of the cells exhibited rapid loss of DNA methylation in the promoter regions of *Dppa3*, *Esrrb*, and *Pecam1* (Figure 1G). The cells treated with Dox and LIF for 3 days contributed to the somatic and germ cell lineages of chimeras (Figures S1K and S1L), demonstrating their unrestricted naive pluripotency. LIF was crucial for the induction of ESC markers (Figures 1H and S2A). The addition of 2i (a MEK inhibitor PD0325901 and a GSK3 inhibitor CHIR99021) (Ying et al., 2008) further enhanced the process, partly by suppressing the concomitantly induced primitive endoderm genes (Figures S2B–S2D). The global gene expression profiles of cells undergoing reprogramming in LIF for 72 hr and in 2i/LIF for 48–72 hr were similar to those of ESCs (Figure 1I). In contrast, the cells treated with Dox without LIF for 72 hr were still similar to EpiSCs. As *Klf4* is an important immediate early target of LIF signaling (Niwa et al., 2009), we examined whether Klf4 is required downstream of Esrrb and LIF for reprogramming. Although deletion of any single KLF TF (*Klf2*, *Klf4*, or *Klf5*) had a marginal effect on the formation of transgene-independent ESC-like colonies, EpiSCs deficient in both *Klf2* and *Klf4* were unable to complete Esrrb-mediated reprogramming and/or failed to maintain cells in a reprogrammed state (Figure S2E). Nanog, whose expression was moderately induced in the presence of LIF, was also required for reprogramming. These results indicate that Esrrb and LIF activate a self-reinforcing transcriptional circuitry including Klf4, Klf2, and Nanog, allowing for the rapid establishment of a self-sustaining transcriptional network for naive pluripotency.

## Esrrb and LIF Gradually Remodel Core TF Occupancy

We next analyzed the binding sites of Esrrb, Oct4, Sox2, and Nanog (EOSN) during reprogramming by chromatin immunoprecipitation sequencing (ChIP-seq). As expected, ChIP-seq signals at later stages of reprogramming showed high correlation with those in ESCs (Chen et al., 2008) (Figure S2F). Esrrb bound to target sites as early as 8 hr post-induction, and its occupancy increased monotonically throughout reprogramming (Figure 2A), despite the slight decrease in Esrrb expression after 24 hr (Figures S1G and S2A). In contrast, Nanog and, to a lesser extent, Oct4 and Sox2 showed both increased and decreased binding (Figures 2A and 2B). ESC-specific peaks of Esrrb and core TFs, particularly Nanog, are located preferentially close to genes expressed in the preimplantation epiblast, extra-embryonic lineages, and primordial germ cells (Figure S2G). In contrast, genes expressed later in development were enriched around EpiSC-specific Nanog peaks. Indeed, EpiSC-specific Nanog-bound genes were preferentially upregulated during embryoid body (EB) differentiation of EpiSCs (Figure S2H), representing poised genes. Considering the observed difference in developmental potential between ESCs and EpiSCs, these results suggest that differently wired pluripotency TF networks shape the enhancer landscapes poised for lineage specification in a developmental stage-specific manner to enable cells to respond properly to external stimuli.

We then examined changes in the expression of TF target genes predicted by ChIP-seq during reprogramming. Esrrb target genes were preferentially upregulated at the beginning of reprogramming even in the absence of LIF (Figure S2I, 8 versus 0 h). In contrast, increased OSN binding highly correlated with gene activation after 48 hr in the presence of LIF (Figure S2I, LIF_48 versus 0 h and 2i/LIF_48 versus 0 h), consistent with the crucial role played by Nanog in reprogramming (Figure S2E). Taken together these results demonstrate that the Esrrb- and LIF-dependent gradual remodeling of core TF occupancy is crucial for establishing naive pluripotency.

To understand the mechanisms underlying the remodeling of TF binding, we first analyzed the enrichment of DNA motifs around peak summits. The Esrrb binding sites (c1 and c3) were highly enriched for the cognate estrogen-related receptor (ERR) motif but only slightly enriched for secondary motifs that are potentially occupied by cooperative TFs (Figure 2C), suggesting that Esrrb binding is less dependent on binding of other TFs. In contrast, the ESC-specific (c1 and c3) binding sites of Nanog and, to a lesser extent, Sox2 were significantly enriched for the ERR motif in addition to their cognate motifs. To assess the dependency of OSN binding on Esrrb, we deleted the ERR motifs in OSN peaks near the *Klf4*, *Klf2*, and *Jam2* genes in ESCs (Figure S2J). Deletion of the ERR motifs resulted in specific reduction in both OSN binding (Figure S2J) and target gene expression (Figure S2K), contingent upon the juxtaposition of the ERR and OCT-SOX motifs, suggesting that Esrrb facilitates the recruitment of core TFs to these enhancer regions. ESC-specific OSN clusters were also more enriched for their own recognition sequences compared with EpiSC-specific clusters (e.g., NANOG motifs in Nanog c1 and c3), suggesting that the binding to ESC-specific peak regions is also driven by the intrinsic DNA binding affinity of the individual TFs.

We then analyzed the histone modification and DNA methylation states of peak regions in EpiSCs and ESCs using publicly available data (Factor et al., 2014; Habibi et al., 2013; Whyte et al., 2012; Wu et al., 2015). ESC-specific EOSN clusters showed lower levels of H3K27ac and H3K4me1 and higher levels of DNA methylation in EpiSCs compared with ESCs (Figures 2D, 2E, and S2L). We also found that many EpiSC- and ESC-specific peaks overlap with cell-type-specific broad hypomethylated regions (HMRs) (Figures 2B, S2M, and S3A). These data suggest that the accessibility of ESC-specific enhancers is limited in

(E) CpG methylation levels (Habibi et al., 2013; Wu et al., 2015) around peaks in each cluster and across all CpGs.
(F) ChIP assays in Tet-on Esrrb EpiSCs cultured with Dox and 2i/LIF. Occupancy relative to a negative control region is shown. Error bars represent SD of triplicates.
(G) GlucMS-PCR analysis on the same chromatin DNA as in (F). Error bars represent SD of triplicates.
(H) ChIP assays in Tet-on Esrrb EpiSCs cultured with Dox ± LIF without 2i.
(I) GlucMS-PCR analysis on the same chromatin DNA as in (H). Error bars represent SD of triplicates.
See also Figures S2 and S3.

EpiSCs and that the remodeling of TF binding was accompanied by the opening of repressive chromatin.

### Esrrb Binds to Silenced Enhancer Regions and Induces Chromatin Remodeling

To uncover the sequence of events leading to reactivation of silenced enhancers, we focused on enhancer regions of the *Esrrb*, *Klf4*, *Klf2*, *Tbx3*, and *Jam2* genes, which are co-occupied by EOSN during reprogramming (Figure S3A) and functional in ESCs (Figures S2J, S2K, and S3B). We first performed ChIP assays for Esrrb, Stat3 (a signal transducer of LIF signaling), Klf4, p300 (a histone acetyltransferase), Med12 (a Mediator component), and Tet1 (a methylcytosine dioxygenase) as well as for Oct4, Sox2, and Nanog (Figure 2F). Esrrb, Stat3, Klf4, p300, and Med12 were rapidly recruited to the silenced enhancers on day 1, and binding of these factors slightly increased on day 2. Increased Tet1 binding was observed at some of the enhancer regions. Among the core TFs, Sox2 was recruited earlier, whereas binding of Nanog and, less markedly, Oct4 increased only on day 2, suggesting that the assembly of enhancer complexes is a hierarchical process.

We then measured the levels of 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) at these enhancer regions by glucosylation-coupled methylation-sensitive PCR (GlucMS-PCR) assays (Figure 2G). The levels of 5mC decreased rapidly and inversely with enhancer occupancy by Esrrb. 5hmCs present at the *Esrrb* and *Klf2* enhancers in EpiSCs were also lost during reprogramming. The *Tbx3* enhancer transiently gained 5hmC, suggesting that active demethylation takes place. We confirmed the rapid recruitment of Esrrb and local loss of DNA methylation in the absence of 2i (Figures 2H and 2I). p300 and Stat3 were efficiently recruited only when LIF was present (Figure 2H), although the rapid loss of DNA methylation was still observed without LIF, albeit to a lesser extent (Figure 2I). Accordingly, in the absence of LIF, the initial binding of Esrrb was gradually decreased after 24 hr (Figure 2H). The genetic and pharmacological inhibition of p300 resulted in decreased expression of the Esrrb target genes in the presence of 2i/LIF (Figure S3C). These results suggest that chromatin remodeling of enhancer regions by p300 is critical for persistent binding of Esrrb and the subsequent activation of target genes. In contrast to the immediate engagement of Esrrb, Nanog and Klf4 were unable to access the silenced enhancers when overexpressed (Figures S3D and S3E). Instead, they transiently bound to accessible regions (*Nkx3-2-5′*, *Tal2-5′*, and *Hoxb1-5′*) or regions with relatively low levels of DNA methylation (*Esrrb-In* and *Jam2-5′*) in EpiSCs.

We next examined the changes in chromatin accessibility by nucleosome occupancy and methylation sequencing (NOMe-seq), a GpC methyltransferase-based assay combined with bisulfite sequencing analysis (Kelly et al., 2010). The regions around the putative Esrrb binding sites became rapidly accessible upon expression of Esrrb (Figure 3A), presumably following p300 binding and preceding OSN recruitment (Figure 2F). A rapid loss of CpG methylation was also observed, especially on nucleosome-free DNA. To test whether Esrrb binds to nucleosomal DNA, we performed ChIP assays on chromatin digested with micrococcal nuclease (MNase) and Exonuclease III (ExoIII) (Figure 3B). As expected, the ESC-specific enhancers were resistant to MNase digestion and enriched for histone H3 in EpiSCs at levels comparable to intergenic regions with no enhancer activity. Irrespective of the presence of 2i or LIF, exogenously expressed Flag-tagged Esrrb showed significant binding to MNase-resistant nucleosome-occupied chromatin fragments at some of these enhancer regions. We next performed ChIP assays followed by bisulfite sequencing analysis to test whether Esrrb binds to chromatin containing methylated DNA. The input DNA and DNA bound by Esrrb contained similar levels of 5mC plus 5hmC (Figure 3C). ChIP followed by GlucMS-PCR confirmed the binding of Esrrb to DNA containing 5mC (Figure 3D). Taken together, these results suggest that Esrrb accesses its target sites within inactive chromatin containing nucleosomes and high levels of 5mC. Subsequently, Esrrb, along with Stat3, recruit chromatin modifiers such as p300 and induce chromatin remodeling, which then allows for the cooperative binding of multiple TFs.

### Global Hypomethylation Caused by *Dnmt1* Deficiency Facilitates LIF-Driven Nucleosome Displacement

We next examined the role of DNA methylation in reprogramming by generating *Dnmt1*-deficient EpiSCs (Figure S4A). Loss of *Dnmt1* resulted in hypomethylation of CpG-rich regions of unique gene promoters and repetitive elements, as well as CpG-poor regions including the ESC-specific enhancers (Figures 4A and 4B). In addition, *Dnmt1* mutants showed derepression of the *Oct4-EGFP* transgene, which is aberrantly silenced in EpiSCs (Greber et al., 2010) (Figures 4C and 4D). In contrast to the essential role of DNMT1 in the survival of human ESCs (Liao et al., 2015), *Dnmt1* deficiency did not affect the ability of EpiSCs to self-renew. *Dnmt1*-deficient EpiSCs often formed dome-shaped colonies (Figures 4C and 4D), probably due to the LIF/IL-6 activity present in MEF-conditioned media (MEF-CM), as the addition of JAK inhibitor (JAK-i) led to the reversion of this phenotype (Figure S4B). When the cells were stimulated with LIF, rapid tyrosine phosphorylation of Stat3 was observed to a similar extent in *Dnmt1* mutants and parental cells (Figure S4C), suggesting that events downstream of Stat3 phosphorylation are responsible for the *Dnmt1* mutant phenotype. More strikingly, despite retaining LIF-independence and the molecular identity of EpiSCs, *Dnmt1* mutants showed rapid activation of ESC markers and concomitant suppression of EpiSC markers when cultured in the presence of 2i/LIF (Figures 4D, 4E, S4B, S4D, and S4E). LIF, but not 2i, was required for the induction of ESC markers, among which *Esrrb* was one of the most highly induced (Figure S4F). We observed a rapid and dramatic increase in the binding of Esrrb, Stat3, and p300 (Figure 4F) as well as the core TFs, especially Nanog (Figure 4G), to the ESC-specific enhancer regions, without significant differences in the protein levels of OSN compared with wild-type (Figure S4G). Consistently, the *Esrrb* and *Klf2* enhancers became rapidly accessible in *Dnmt1* mutants upon treatment with 2i/LIF (Figure 4H). These results indicate that global DNA hypomethylation provides the permissive chromatin environment, in which LIF and endogenous TFs can readily rewire TF binding to activate ESC-specific enhancers.

## DISCUSSION

Most regions of the genome are occupied by nucleosomes that restrict the access of TFs to DNA. Nevertheless, at

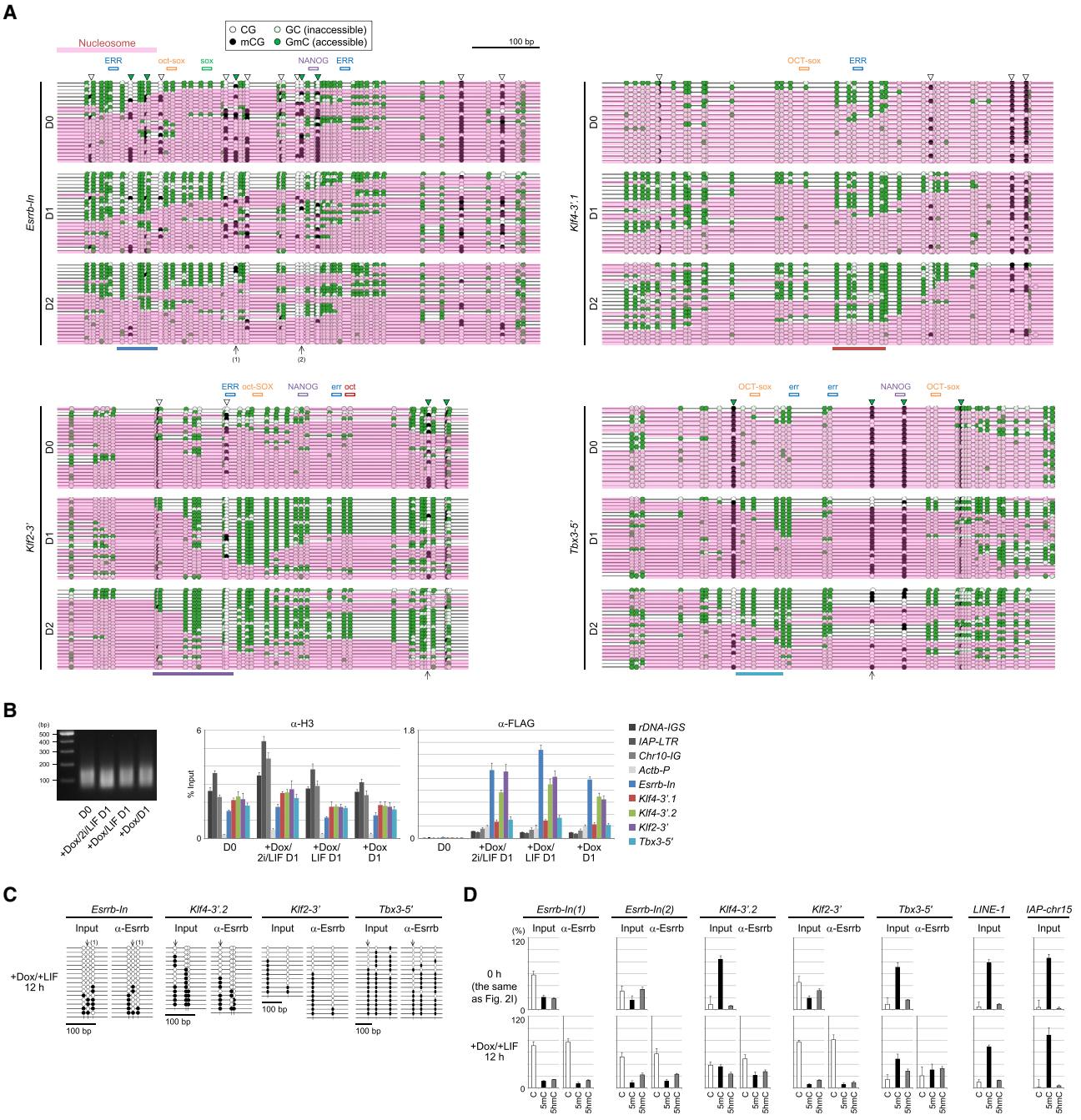



**Figure 3. Binding of Esrrb to Chromatin Containing Nucleosomes and Methylated DNA**

(A) NOMe-seq assays in Tet-on Esrrb EpiSCs cultured with Dox and 2i/LIF. M.CviPI-inaccessible regions that are at least 120 bp in length and therefore are likely to be occupied by nucleosomes are highlighted in pink. Open triangles indicate positions of CpG sites. Triangles filled in green represent CpGs in CCG trinucleotides that may be targeted by GpC methyltransferase at a low frequency. Color bars at the bottom of the plots indicate regions analyzed by MNase-ChIP in (B). Arrows indicate CpGs overlapping with MspI/HpaII restriction sites (CCGG), which were used for GlucMS-PCR in (C). Motifs in lowercase represent motifs that diverge from the consensus sequences.

(B) MNase-ChIP assays in Tet-on FLAG-Esrrb EpiSCs. Left, chromatin DNA digested with MNase and ExoIII. Right, ChIP assays for H3 and FLAG-tagged Esrrb. The percentage of input DNA is shown. Error bars represent SD of triplicates.

(C) ChIP-bisulfite sequencing assays on the same chromatin as in Figure 2H. Arrows indicate MspI/HpaII restriction sites used for GlucMS-PCR in (D).

(D) ChIP-GlucMS-PCR analysis. Only input DNA was analyzed for *LINE-1* and *IAP-chr15*. Error bars represent SD of triplicates.

See also Figure S3.

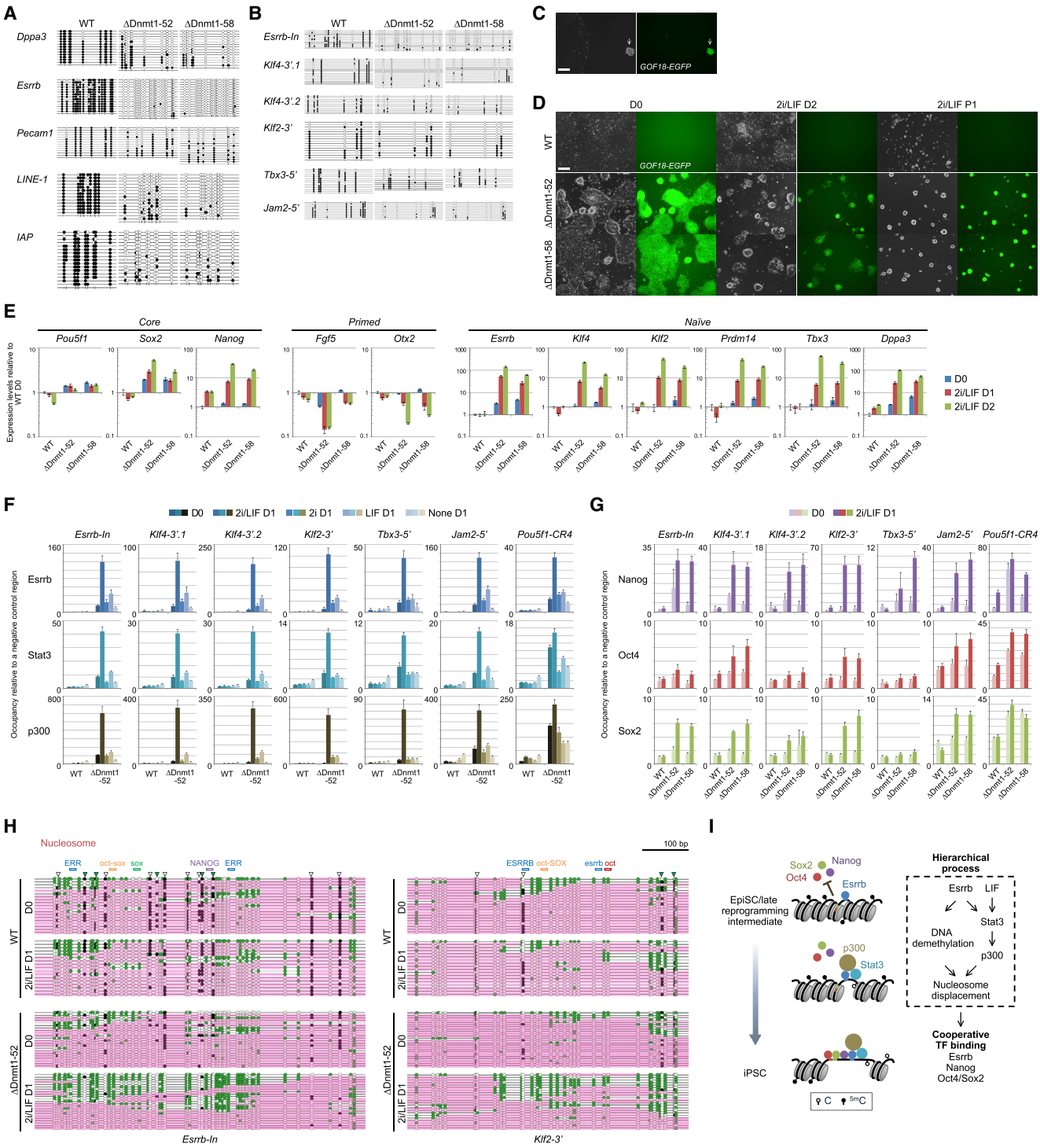



**Figure 4. Acute Remodeling of TF Binding in *Dnmt1*-Deficient EpiSCs Treated with 2i/LIF**

(A and B) Bisulfite sequencing analysis of promoters and repetitive elements (A) and promoter-distal regions (B) in *Dnmt1*-deficient EpiSCs.

(C) Representative colonies obtained after transfection of *Cas9* and *Dnmt1*-single guide RNA (sgRNA) expression constructs into E3 EpiSCs. A few compact colonies positive for the *Oct4-EGFP* reporter (*GOF18-EGFP*; arrows) were found to be mutants in most cases. Scale bar, 200 μm.

(D) Reprogramming of *Dnmt1*-deficient EpiSCs by 2i/LIF. Scale bar, 200 μm.

(E) RT-PCR analysis of *Dnmt1*-deficient EpiSCs treated with 2i/LIF. Error bars represent SD of triplicates.

(F) ChIP assays in *Dnmt1*-deficient EpiSCs cultured with or without LIF or 2i for 1 day. Error bars represent SD of triplicates.

*(legend continued on next page)*

some genomic regions TFs jointly compete with nucleosomes for DNA binding and influence the local chromatin structure by recruiting or repelling histone/DNA-modifying enzymes. Consequently, cell-type-specific sets of TFs and chromatin-modifying enzymes shape a chromatin landscape unique to each cell type. Such epigenetic information carried by histones and DNA acts to maintain transcriptional states over time and through cell division, ensuring that cell identity is stably maintained. This, however, is the major obstacle to TF-mediated reprogramming. When overexpressed in somatic cells, Oct4 and Sox2 are unable to access relevant target sites near pluripotency genes early in reprogramming, but predominantly bind to off-target sites accessible in the original cells (Chen et al., 2016; Chronis et al., 2017). On-target binding of Oct4 and activation of pluripotency genes are achieved at the last stage of reprogramming, roughly coinciding with the acquisition of H3K27ac and loss of DNA methylation around the Oct4 binding sites (Chen et al., 2016). These studies suggest that epigenetic reprogramming is not readily induced by the reprogramming TFs alone but is facilitated by secondary factors induced during reprogramming. We found that Esrrb acts as a pioneer reprogramming factor in EpiSCs (Figure 4I). Esrrb immediately binds to silenced enhancers containing nucleosomes and high levels of 5mC, where the access of other TFs and coactivators is limited. Subsequently, Esrrb induces chromatin reorganization including nucleosome displacement and DNA demethylation. Although it is well known that ERRs and their coactivators, such as SRC and PGC-1 family members, recruit various chromatin modifiers, p300 likely plays a key role in establishing an active chromatin state at enhancers, as forced tethering of p300 is sufficient to activate enhancers and target gene expression (Hilton et al., 2015). Interestingly, the recruitment of p300, but not the initial binding of Esrrb, depends upon LIF, consistent with the requirement of LIF for OSN target gene activation later in reprogramming (Figure S2I). In parallel, Esrrb binding induced a rapid loss of DNA methylation, especially on nucleosome-free DNA (Figure 3A). The initial loss of DNA methylation was only slightly affected in the absence of LIF. Rapid local demethylation presumably results from active mechanisms as suggested by the recruitment of Tet1 and a transient increase in 5hmC at subsets of the enhancer regions, although reprogramming of fibroblasts can be accomplished in the absence of TET family dioxygenases or thymine DNA glycosylase (Hu et al., 2014). In addition, passive mechanisms that depend on DNA replication might take place. TF binding may locally inhibit DNA methyltransferase activity, for instance, by displacing histones that can be used for targeting Dnmt1 and its partner Uhrf1 to chromatin (Nishiyama et al., 2013; Rothbart et al., 2012). Global hypomethylation caused by *Dnmt1* deficiency resulted in a dramatic increase in TF binding after treatment with LIF, suggesting that loss of DNA methylation is a rate-limiting step in enhancer activation. We speculate that loss of DNA methylation weakens the interactions between histones and DNA and increases nucleosome dynamics (Choy et al., 2010), allowing for TF-mediated nucleosome displacement (Figure 4H). Once a permissive chromatin environment is transiently created, a flexible set of TFs and coactivators can bind cooperatively to adjacent sites and reinforce the changes in enhancer structure. This final step is indeed crucial for enhancer activation as *Nanog* deficiency resulted in incomplete reprogramming.

Previous studies suggested that exogenous Esrrb facilitates somatic cell reprogramming (Chronis et al., 2017). Interestingly, when ectopically expressed in somatic cells, Esrrb is unable to bind to ESC-specific enhancers but preferentially binds to accessible chromatin and activates genes associated with oxidative phosphorylation. Other contexts, such as higher-order chromatin structure and cooperatively binding factors, may be required for Esrrb to access ESC-specific enhancers in somatic cells. Such contextual prerequisites are also important for binding of other pioneer factors, such as Foxa1 (Lupien et al., 2008) and Ascl1 (Wapinski et al., 2013). Endogenous Esrrb plays an important role in establishing pluripotency in the final step of somatic cell reprogramming by Oct4, Sox2, and Klf4 (Figures 1D and 1E). Recent findings show that most ESC-specific open chromatin regions gain accessibility at the end of reprogramming and that those regions are highly enriched for the ERR motif (Li et al., 2017), suggesting that Esrrb plays a pioneering role similar to that observed in EpiSC reprogramming.

## STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell lines
  - Mice
- METHOD DETAILS
  - Reprogramming
  - RT-PCR
  - Immunofluorescence
  - Western Blot
  - Flow Cytometry
  - Bisulfite Sequencing
  - ChIP
  - NOMe-seq
  - GlucMS-PCR
  - Luciferase Reporter Assays
  - Microarray and Data Analysis
  - ChIP-seq and Data Analysis
  - RNA-seq and Data Analysis
  - Analysis of Publicly Available Data
- DATA AND SOFTWARE AVAILABILITY

(G) ChIP assays in *Dnmt1*-deficient EpiSCs cultured with 2i/LIF for 1 day. Error bars represent SD of triplicates.
(H) NOMe-seq assays in *Dnmt1*-deficient EpiSCs cultured with 2i/LIF for 1 day.
(I) Schematic diagram of Esrrb-mediated chromatin opening. See the Discussion for details.
See also Figure S4.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and four tables and can be found with this article online at https://doi.org/10.1016/j.stem.2018.05.020.

## ACKNOWLEDGMENTS

We thank Ingrid Gelker and Claudia Ortmeier for technical assistance; Martina Sinn for microarray analysis; Phuc-Loi Luu and Dirk Richter for assistance in computational work; Shinya Aramaki, Saya Kagiwada, Sergiy Velychko, Katsuhiro Kato, Ho-Ryun Chung, and Stefan Haas for helpful discussions; and Areti Malapetsas for final editing. The p300/CBP inhibitor A-485 and its inactive analog A-486 were supplied by the Structural Genomics Consortium under an Open Science Trust Agreement: https://www.thesgc.org/click-trust. This work was supported by the Max Planck Society.

## AUTHOR CONTRIBUTIONS

K.A. designed the study, performed the experiments and bioinformatic analysis, and wrote the manuscript; W.K. contributed to the sequencing data analysis; G.W. performed the mice experiments; S.H. assisted with the experiments; B.G. provided advice on the experiments; M.S. performed the flow cytometry experiments; M.J.A.-B. contributed to the microarray data analysis; S.T.B. and B.T. performed the sequencing analysis; M.V. supervised the sequencing data analysis; and H.R.S. supervised the study and edited the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 11, 2017
Revised: April 3, 2018
Accepted: May 21, 2018
Published: June 14, 2018

## REFERENCES

Adachi, K., Nikaido, I., Ohta, H., Ohtsuka, S., Ura, H., Kadota, M., Wakayama, T., Ueda, H.R., and Niwa, H. (2013). Context-dependent wiring of Sox2 regulatory networks for self-renewal of embryonic and trophoblast stem cells. Mol. Cell *52*, 380–392.

Ahmed, K., Dehghani, H., Rugg-Gunn, P., Fussner, E., Rossant, J., and Bazett-Jones, D.P. (2010). Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. PLoS ONE *5*, e10531.

Borgel, J., Guibert, S., Li, Y., Chiba, H., Schübeler, D., Sasaki, H., Forné, T., and Weber, M. (2010). Targets and dynamics of promoter DNA methylation during early mouse development. Nat. Genet. *42*, 1093–1100.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell *133*, 1106–1117.

Chen, J., Chen, X., Li, M., Liu, X., Gao, Y., Kou, X., Zhao, Y., Zheng, W., Zhang, X., Huo, Y., et al. (2016). Hierarchical Oct4 Binding in Concert with Primed Epigenetic Rearrangements during Somatic Cell Reprogramming. Cell Rep. *14*, 1540–1554.

Choy, J.S., Wei, S., Lee, J.Y., Tan, S., Chu, S., and Lee, T.H. (2010). DNA methylation increases nucleosome compaction and rigidity. J. Am. Chem. Soc. *132*, 1782–1783.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative binding of transcription factors orchestrates reprogramming. Cell *168*, 442–459.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Dowen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schujiers, J., Lee, T.I., Zhao, K., and Young, R.A. (2014). Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. Cell *159*, 374–387.

Du, P., Kibbe, W.A., and Lin, S.M. (2008). lumi: A pipeline for processing Illumina microarray. Bioinformatics *24*, 1547–1548.

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.

Feng, B., Jiang, J., Kraus, P., Ng, J.H., Heng, J.C., Chan, Y.S., Yaw, L.P., Zhang, W., Loh, Y.H., Han, J., et al. (2009). Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. Nat. Cell Biol. *11*, 197–203.

Festuccia, N., Osorno, R., Halbritter, F., Karwacki-Neisius, V., Navarro, P., Colby, D., Wong, F., Yates, A., Tomlinson, S.R., and Chambers, I. (2012). Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. Cell Stem Cell *11*, 477–490.

Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P., and Schöler, H.R. (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. Cell Stem Cell *6*, 215–226.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., and Smith, A. (2009). Klf4 reverts developmentally programmed restriction of ground state pluripotency. Development *136*, 1063–1069.

Habibi, E., Brinkman, A.B., Arand, J., Kroeze, L.I., Kerstens, H.H., Matarese, F., Lepikhov, K., Gut, M., Brun-Heath, I., Hubner, N.C., et al. (2013). Whole-genome bisulfite sequencing of two distinct interconvertible DNA methylomes of mouse embryonic stem cells. Cell Stem Cell *13*, 360–369.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. Cell *146*, 519–532.

Hilton, I.B., D'Ippolito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nat. Biotechnol. *33*, 510–517.

Hu, X., Zhang, L., Mao, S.Q., Li, Z., Chen, J., Zhang, R.R., Wu, H.P., Gao, J., Guo, F., Liu, W., et al. (2014). Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. Cell Stem Cell *14*, 512–522.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Kelly, T.K., Miranda, T.B., Liang, G., Berman, B.P., Lin, J.C., Tanay, A., and Jones, P.A. (2010). H2A.Z maintenance during mitosis reveals nucleosome shifting on mitotically silenced genes. Mol. Cell *39*, 901–911.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kumaki, Y., Oda, M., and Okano, M. (2008). QUMA: Quantification tool for methylation analysis. Nucleic Acids Res. *36*, W170–W175.

Langfelder, P., and Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics *9*, 559.

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

Lasko, L.M., Jakob, C.G., Edalji, R.P., Qiu, W., Montgomery, D., Digiammarino, E.L., Hansen, T.M., Risi, R.M., Frey, R., Manaves, V., et al. (2017). Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. Nature *550*, 128–132.

Li, D., Liu, J., Yang, X., Zhou, C., Guo, J., Wu, C., Qin, Y., Guo, L., He, J., Yu, S., et al. (2017). Chromatin accessibility dynamics during iPSC reprogramming. Cell Stem Cell *21*, 819–833.

Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C., et al. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat. Genet. *47*, 469–478.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell *132*, 958–970.

Machanick, P., and Bailey, T.L. (2011). MEME-ChIP: Motif analysis of large DNA datasets. Bioinformatics *27*, 1696–1697.

Mali, P., Aach, J., Stranges, P.B., Esvelt, K.M., Moosburner, M., Kosuri, S., Yang, L., and Church, G.M. (2013). CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nat. Biotechnol. *31*, 833–838.

Mammana, A., and Helmuth, J. (2016). bamsignals: Extract read count signals from bam files. R package version 160.

Martello, G., Sugimoto, T., Diamanti, E., Joshi, A., Hannah, R., Ohtsuka, S., Göttgens, B., Niwa, H., and Smith, A. (2012). Esrrb is a pivotal target of the Gsk3/Tcf3 axis regulating embryonic stem cell self-renewal. Cell Stem Cell *11*, 491–504.

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nat. Biotechnol. *28*, 495–501.

Nakanishi, T., Ikawa, M., Yamada, S., Parvinen, M., Baba, T., Nishimune, Y., and Okabe, M. (1999). Real-time observation of acrosomal dispersal from mouse sperm using GFP as a marker protein. FEBS Lett. *449*, 277–283.

Ng, H.H., and Surani, M.A. (2011). The transcriptional and signalling networks of pluripotency. Nat. Cell Biol. *13*, 490–496.

Nishiyama, A., Yamaguchi, L., Sharif, J., Johmura, Y., Kawamura, T., Nakanishi, K., Shimamura, S., Arita, K., Kodama, T., Ishikawa, F., et al. (2013). Uhrf1-dependent H3K23 ubiquitylation couples maintenance DNA methylation and replication. Nature *502*, 249–253.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. *24*, 372–376.

Niwa, H., Masui, S., Chambers, I., Smith, A.G., and Miyazaki, J. (2002). Phenotypic complementation establishes requirements for specific POU domain and generic transactivation function of Oct-3/4 in embryonic stem cells. Mol. Cell. Biol. *22*, 1526–1536.

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature *460*, 118–122.

Ohtsuka, S., Nishikawa-Torikai, S., and Niwa, H. (2012). E-cadherin promotes incorporation of mouse epiblast stem cells into normal development. PLoS ONE *7*, e45220.

Ramírez, F., Dündar, F., Diehl, S., Grüning, B.A., and Manke, T. (2014). deepTools: A flexible platform for exploring deep-sequencing data. Nucleic Acids Res. *42*, W187–W191.

Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-Lovejoy, D., Martinez, J.Y., Bedford, M.T., Fuchs, S.M., et al. (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. Nat. Struct. Mol. Biol. *19*, 1155–1160.

Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science *351*, 84–88.

Song, Q., Decato, B., Hong, E.E., Zhou, M., Fang, F., Qu, J., Garvin, T., Kessler, M., Zhou, J., and Smith, A.D. (2013). A reference methylome database and analysis pipeline to facilitate integrative and comparative epigenomics. PLoS ONE *8*, e81148.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., and Mesirov, J.P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. USA *102*, 15545–15550.

Taberlay, P.C., Kelly, T.K., Liu, C.C., You, J.S., De Carvalho, D.D., Miranda, T.B., Zhou, X.J., Liang, G., and Jones, P.A. (2011). Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell *147*, 1283–1294.

Wapinski, O.L., Vierbuchen, T., Qu, K., Lee, Q.Y., Chanda, S., Fuentes, D.R., Giresi, P.G., Ng, Y.H., Marro, S., Neff, N.F., et al. (2013). Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell *155*, 621–635.

Whyte, W.A., Bilodeau, S., Orlando, D.A., Hoke, H.A., Frampton, G.M., Foster, C.T., Cowley, S.M., and Young, R.A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. Nature *482*, 221–225.

Wu, G., Han, D., Gong, Y., Sebastiano, V., Gentile, L., Singhal, N., Adachi, K., Fischedick, G., Ortmeier, C., Sinn, M., et al. (2013). Establishment of totipotency does not depend on Oct4A. Nat. Cell Biol. *15*, 1089–1097.

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.

Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519–523.

Zambelli, F., Pesole, G., and Pavesi, G. (2013). PscanChIP: Finding over-represented transcription factor-binding site motifs and their correlations in sequences from ChIP-seq experiments. Nucleic Acids Res. *41*, W535–W543.

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.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Antibodies | | |
| Mouse monoclonal anti-Esrrb | R&D Systems | Cat# PP-H6705-00; RRID: AB_1964232 |
| Rat monoclonal anti-Nanog | Thermo Fisher Scientific | Cat# 14-5761-80; RRID:AB_763613 |
| Rabbit polyclonal anti-mKO2 | MBL | Cat# PM051; RRID:AB_10597258 |
| Goat polyclonal anti-Dppa4 | R&D Systems | Cat# AF3730; RRID:AB_2094166 |
| Mouse monoclonal anti-Sall4 | Santa Cruz | Cat# sc-101147; RRID:AB_1129262 |
| Goat polyclonal anti-Klf4 | R&D Systems | Cat# AF3158; RRID: AB_2130245 |
| Mouse monoclonal anti-Oct4 | Santa Cruz | Cat# sc-5279; RRID: AB_628051 |
| Goat polyclonal anti-Sox17 | R&D Systems | Cat# AF1924; RRID: AB_355060 |
| Rabbit monoclonal anti-Dnmt1 | Cell Signaling Technology | Cat# 5032; RRID: AB_10828695 |
| Mouse monoclonal anti-α-tubulin | Sigma | Cat# T6199; RRID: AB_477583 |
| Rabbit polyclonal anti-Stat3 | Santa Cruz | Cat# sc-482; RRID: AB_632440 |
| Rabbit monoclonal anti-phospho-Stat3 (Tyr705) | Cell Signaling Technology | Cat# 9145; RRID: AB_2491009 |
| Rabbit polyclonal anti-Nanog | Bethyl Laboratories | Cat# A300-397A; RRID: AB_386108 |
| Goat polyclonal anti-Oct4 | Santa Cruz | Cat# sc-8628; RRID: AB_653551 |
| Goat polyclonal anti-Sox2 | Neuromics | Cat# GT15098; RRID: AB_2195800 |
| Rabbit polyclonal anti-H3 | Abcam | Cat# ab1791; RRID: AB_302613 |
| Rabbit polyclonal anti-p300 | Santa Cruz | Cat# sc-585; RRID: AB_2231120 |
| Rabbit polyclonal anti-Med12 | Bethyl Laboratories | Cat# A300-774A; RRID: AB_669756 |
| Rabbit polyclonal anti-Tet1 | GeneTex | Cat# GTX124207; RRID:AB_11176491 |
| Mouse monoclonal anti-FLAG | Sigma | Cat# F1804; RRID:AB_262044 |
| Rat monoclonal anti-mouse Pecam1/CD31 antibody, conjugated with APC | BioLegend | Cat# 102509; RRID:AB_312916 |
| Chemicals, Peptides, and Recombinant Proteins | | |
| Human FGF2 | PeproTech | Cat# 100-18B |
| Human Activin A | eBioscience | Cat# 34-8993-85 |
| Human LIF | Prepared in-house | N/A |
| PD0325901 | Cayman Chemical | Cat# 13034 |
| CHIR99021 | Tocris | Cat# 4423 |
| Y-27632 | Abcam | Cat# ab120129 |
| JAK Inhibitor I | Calbiochem | Cat# 420097 |
| A-485 | Lasko et al., 2017 | Structural Genomics Consortium |
| A-486 | Lasko et al., 2017 | Structural Genomics Consortium |
| L-ascorbic acid 2-phosphate | Sigma | Cat# A8960 |
| 4-hydroxytamoxifen (4-OHT) | Sigma | Cat# H7904 |
| Doxycycline (Dox) | Sigma | Cat# D9891 |
| Di(N-succinimidyl) glutarate (DSG) | Santa Cruz | Cat# sc-285455 |
| M.CviPI (GpC methyltransferase) | NEB | Cat# M0227 |
| T4 phage β-glucosyltransferase (T4-BGT) | NEB | Cat# M0357 |
| Micrococcal nuclease (MNase) | NEB | Cat# M0247 |
| Exonuclease III (ExoIII) | NEB | Cat# M0206 |
| Poly(vinyl alcohol) | Sigma | Cat# P8136 |
| Critical Commercial Assays | | |
| MouseRef-8 v2.0 Expression BeadChips | Illumina | Cat# BD-202-0602 |
| TruSeq ChIP Library Preparation Kit | Illumina | Cat# IP-202-1012 |
| TruSeq RNA Library Prep Kit v2 | Illumina | Cat# RS-122-2001 |

(*Continued on next page*)

*Continued*

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Deposited Data | | |
| Microarray data | This paper | ArrayExpress: E-MTAB-5341 |
| ChIP-seq data | This paper | ArrayExpress: E-MTAB-5342 |
| RNA-seq data | This paper | ArrayExpress: E-MTAB-5343 |
| ChIP-seq data in ESCs | Chen et al., 2008 | GEO: GSE11431 |
| ChIP-seq data in EpiSCs | Factor et al., 2014 | GEO: GSE57407 |
| ChIP-seq data in ESCs | Whyte et al., 2012 | GEO: GSE27841, GSE11724 |
| WGBS-seq data in EpiSCs | Wu et al., 2015 | GEO: GSE63568 |
| WGBS-seq data in ESCs | Habibi et al., 2013 | GEO: GSE41923 |
| Experimental Models: Cell Lines | | |
| Mouse: EpiSC_E3: Tg(GOF18-EGFP) | Greber et al., 2010 | N/A |
| Mouse: ESC_GOF18-3F/5A: Pou5f1$^{+/-}$; Tg(GOF18-EGFP) | This paper | N/A |
| Mouse: ESC_EB3: Pou5f1$^{+/BSD}$ | Niwa et al., 2002 | N/A |
| Mouse: ESC_ZHBTc4: Pou5f1$^{BSD/Zeo}$; Tg(CAG-tTA); Tg(TetO-Pou5f1) | Niwa et al., 2000 | N/A |
| Mouse: ESC_Esrrb$^{flox/flox-mKO2}$: Esrrb$^{flox/flox-mKO2}$; Tg(CAG-MerCreMer); Tg(Acr-EGFP) | This paper | N/A |
| Mouse: MEF_Esrrb$^{flox/flox-mKO2}$: Esrrb$^{flox/flox-mKO2}$; Tg(CAG-MerCreMer); Tg(Acr-EGFP) | This paper | N/A |
| Oligonucleotides | | |
| See Table S1 for oligonucleotides used in this paper | This paper | N/A |
| Recombinant DNA | | |
| gRNA_Cloning Vector | Mali et al., 2013 | Addgene Plasmid #41824 |
| pX330-U6-Chimeric_BB-CBh-hSpCas9 | Cong et al., 2013 | Addgene Plasmid #42230 |
| eSpCas9(1.1) | Slaymaker et al., 2016 | Addgene Plasmid #71814 |
| Software and Algorithms | | |
| GenomeStudio | Illumina | https://www.illumina.com/techniques/microarrays/array-data-analysis-experimental-design/genomestudio.html |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Hans R. Schöler (office@mpi-muenster.mpg.de).

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Cell lines

E3 GOF18-EGFP male EpiSCs (F1 of homozygous GOF18-EGFP [C57BL/6 and DBA/2 mixed background] and 129/Sv) and their derivatives were maintained as previously described (Greber et al., 2010) in MEF-CM containing 20% KnockOut Serum Replacement (KSR) supplemented with 10 ng/ml FGF2 (CM + FGF2) or in a 1:1 mixture of CM + FGF2 and N2B27 (50% DMEM/F12, 50% Neurobasal, 0.5 × N2, 0.5 × B27, 25 μg/ml BSA fraction V) supplemented with 10 ng/ml FGF2 and 10 ng/ml Activin A (CM/N2B27 + FA) on fetal bovine serum (FBS)-coated dishes without feeders. GOF18-3F and GOF18-5A male ESCs were derived from embryos carrying the *GOF18-EGFP* transgene and a heterozygous deletion of *Pou5f1* (Wu et al., 2013) in a mixed background and maintained in DMEM/F12 containing 10% KSR (K10) and LIF (prepared in house) on MEF feeders. For microarray analysis, ESCs were adapted to feeder-free conditions in K10 + LIF. For bisulfite sequencing, GOF18-3F ESCs were further adapted to 2i (1 μM PD0325901 and 3 μM CHIR99021) (Ying et al., 2008) plus LIF. For enhancer deletion experiments depicted in Figures S2J and S2K, EB3 ESCs (Niwa et al., 2002) cultured in a feeder-free condition with 2i/LIF were used. For luciferase reporter assays, ZHBTc4 ESCs (Niwa et al., 2000) cultured in a feeder-free condition with 2i/LIF were used.

For *Esrrb* conditional KO, male ESCs derived from *Acr-EGFP* transgenic mice (Nakanishi et al., 1999) were targeted with the *flox-mKO2* construct containing floxed *Esrrb* exon 2 followed by a *FRT-SA-IRES-Hph-HSVΔtk-P2A-mKO2-pA* cassette, transiently

transfected with a FLPe expression vector to remove the FRT cassette, and targeted with the same *flox-mKO2* construct to obtain *Esrrb*$^{flox/flox\text{-}mKO2}$ ESC lines, which were then stably transfected with a vector for constitutive expression of *MerCreMer-IRES-Neo*.

*Esrrb* conditional KO MEFs were obtained from E12.5 chimeric embryos generated by aggregation of *Esrrb*$^{flox/flox\text{-}mKO2}$; *MerCreMer* ESCs with tetraploid host embryos. Primary MEFs were split, treated with or without 0.5 μM 4-OHT for 6 days, and plated for reprogramming in the absence of 4-OHT. 900 μg/ml G418 was added to both groups to eliminate host cells.

Cells with stable integration of Tet-on transgenes were generated using the *piggyBac* transposon system (Guo et al., 2009) as previously described (Adachi et al., 2013). Cells were co-transfected with a Dox-inducible expression vector containing *IRES-Venus* or *IRES-H2BmKO2* reporters and a vector for constitutive expression of *rtTAM2-IRES-Neo*, followed by selection with G418. The clonal Tet-on Esrrb-IRES-Venus EpiSC line was isolated and further modified by introducing a *Pou5f1* (3.5 kb upstream of the TSS)-*IRES-Puro* cassette to facilitate enrichment of undifferentiated cells. For the Tet-on FLAG-Esrrb construct, a 3 × FLAG tag was fused to the N terminus of *Esrrb*. KO EpiSC lines were generated using the CRISPR/Cas system (Cong et al., 2013; Mali et al., 2013; Slaymaker et al., 2016). sgRNA oligonucleotides were cloned downstream of the human U6 promoter of pBRBII-hU6-BsaI-gRNA (modified from Addgene #41824, a gift from George Church), pX330A-GFPT2APuro (modified from Addgene #42230, a gift from Feng Zhang), or eSpCas9-PuroP2AmCherry (modified from Addgene #71814, a gift from Feng Zhang). For *Esrrb*, *Nanog*, and *Dnmt1* KOs in EpiSCs, two sgRNAs were designed for each gene to delete critical exons, resulting in frameshift mutations. For *Ep300* mutants, two sgRNAs were designed to generate an in-frame deletion of the catalytic core domain. Clones carrying deletions in both alleles were identified by sequencing of the genomic PCR products and immunostaining. For *Klf2*, *Klf4*, *Klf2/4*, and *Klf5* KOs, single sgRNAs targeting the exons were expressed. Clones carrying frameshift mutations in both alleles were identified by sequencing of the genomic PCR products and immunostaining. For enhancer deletion, two sgRNAs were designed to delete putative enhancer regions. Clones carrying deletions in both alleles were identified by sequencing of the genomic PCR products. sgRNA oligonucleotide sequences are listed in Table S1.

### Mice

To generate chimeric mice, cells were aggregated with zona pellucida–free eight-cell–stage embryos obtained by mating either B6C3F1 or CFW 6- to 8-week-old female mice with CD1 male mice aged 3 months or older. After one day of *in vitro* culture in KSOM supplemented with 10% FBS, the chimeric blastocysts were transferred into the uteri of 2.5 days post-coitum pseudopregnant CD1 female mice aged 2 to 4 months. Animals were maintained under a 14-hour light/10-hour dark cycle with free access to food and water. Female mice were housed in groups of up to 5 per cage and male mice were housed individually. A protocol for animal handling and maintenance for this study was approved by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen (LANUV NRW) under the supervision of a certified veterinarian in charge of the Max Planck Institute animal facility.

## METHOD DETAILS

### Reprogramming

For the reprogramming of MEFs, retroviruses were prepared by transient transfection of HEK293T cells with pMXs retroviral plasmids encoding *Oct4*, *Sox2*, or *Klf4*, together with the packaging plasmid pCL-Eco. MEFs were plated at 2 × $10^4$ cells per well of a 12-well plate one day before transduction and infected with a mixture of viral supernatants for one day in the presence of 4 μg/ml protamine sulfate. The next day (1 dpi), the medium was replaced by K15 medium supplemented with LIF and 50 μg/ml L-ascorbic acid 2-phosphate.

For reprogramming of EpiSCs, cells dissociated into single cells using Accutase were plated in EpiSC media supplemented with 5 μM Y-27632 on FBS-coated dishes. After overnight culture, the medium was replaced by K10 medium with or without LIF, 2i or 1 μg/ml Dox.

### RT-PCR

Total RNA was isolated using the RNeasy Mini Kit (QIAGEN). First-strand cDNA was synthesized using oligo(dT) primers and M-MLV Reverse Transcriptase (USB). Real-time PCR was performed using iTaq Universal SYBR Green Supermix (Bio-Rad). Relative expression values were calculated by the standard curve method, normalized to *Gapdh*, and presented as fold change over control samples. Each experiment was performed in technical triplicate. Error bars indicate SD of triplicates. Primer sequences are listed in Table S1.

### Immunofluorescence

Immunofluorescence was performed using the following primary antibodies with dilutions: mouse monoclonal anti-Esrrb (PP-H6705-00, R&D Systems), 1:1000; rat monoclonal anti-Nanog (14-5761-80, Thermo Fisher Scientific), 1:1000; rabbit anti-mKO2 (PM051, MBL), 1:2000; goat anti-Dppa4 (AF3730, R&D Systems), 1:500 of 0.2 mg/ml; mouse monoclonal anti-Sall4 (sc-101147, Santa Cruz), 1:300; goat anti-Klf4 (AF3158, R&D Systems), 1:250 of 0.2 mg/ml; mouse monoclonal anti-Oct4 (sc-5279, Santa Cruz), 1:1000; goat anti-Sox17 (AF1924, R&D Systems), 1:1000 of 0.2 mg/ml.

### Western Blot

Western blotting was performed according to standard procedures using the following primary antibodies: rabbit monoclonal anti-Dnmt1 (5032, Cell Signaling Technology); mouse monoclonal anti-α-tubulin (T6199, Sigma); rabbit anti-Stat3 (sc-482, Santa Cruz);

rabbit monoclonal anti-phospho-Stat3 (Tyr705) (9145, Cell Signaling Technology); rabbit anti-Nanog (A300-397A, Bethyl Laboratories); goat anti-Oct4 (sc-8628, Santa Cruz); goat anti-Sox2 (GT15098, Neuromics); rabbit anti-H3 (ab1791, Abcam).

### Flow Cytometry

Cells were harvested using non-enzymatic cell dissociation solution (C-5914, Sigma) and stained with 4 μg/ml rat monoclonal anti-Pecam1 antibody conjugated with APC (102509, BioLegend) followed by DAPI. Fluorescence was measured using a FACSAria IIu cell sorter (BD Biosciences). Cells were gated based on FSC and SSC to exclude cellular debris and doublets, and on DAPI staining to exclude dead cells. Data analysis was done using FlowJo software (LLC).

### Bisulfite Sequencing

Bisulfite treatment was performed using the EZ DNA Methylation Kit (Zymo Research). PCR was performed at an annealing temperature of 50°C using EpiMark Hot Start Taq DNA Polymerase (NEB). Sequencing data was analyzed using QUMA (Kumaki et al., 2008). For ChIP-bisulfite sequencing assays, the EZ DNA Methylation-Direct Kit (Zymo Research) was used. Primer sequences are listed in Table S1.

### ChIP

For ChIP-seq experiments, Tet-on Esrrb EpiSCs undergoing reprogramming were cross-linked with 1% formaldehyde in PBS for 10 min at room temperature followed by quenching with 125 mM glycine. Cells were lysed in lysis buffer 1 (50 mM HEPES-KOH [pH 7.5], 140 mM NaCl, 1 mM 0.5 M EDTA, 10% glycerol, 0.5% IGEPAL CA630, 0.25% Triton X-100) containing protease inhibitor cocktail (Sigma) for 30 min at 4°C and washed in lysis buffer 2 (10 mM Tris-HCl [pH 8.0], 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA). Chromatin was sheared using Diagenode Bioruptor (high power, 40 cycles of 30 s on and 30 s off) equipped with a water-cooling system (4°C) in 1 mL of sonication buffer (50 mM Tris-HCl [pH 8.0], 10 mM EDTA, 0.5% SDS) per 1 × $10^7$ cells. Small aliquots of sheared chromatin were reverse cross-linked, purified, and analyzed for DNA concentration and fragment size. Sheared chromatin with an average fragment size of 200–300 bp was incubated with Dynabeads protein G coupled to primary antibodies in 4 × volume of ChIP dilution buffer (10 mM Tris-HCl [pH 8.0], 125 mM NaCl, 0.125% sodium deoxycholate, 1.25% Triton X-100) at 4°C overnight. The amount of chromatin, antibodies, and beads for each ChIP assay are as follows: chromatin corresponding to 400 μg DNA + 24 μg mouse anti-Esrrb (PP-H6705-00, R&D Systems) + 200 μL beads; chromatin corresponding to 500 μg DNA + 30 μg rabbit anti-Nanog (A300-397A, Bethyl Laboratories) + 300 μL beads; chromatin corresponding to 200 μg DNA + 8 μg goat anti-Oct4 (sc-8628, Santa Cruz) + 100 μL beads; chromatin corresponding to 500 μg DNA + 16 μg goat anti-Sox2 (GT15098, Neuromics) + 150 μL beads. Beads were washed once with low-salt buffer (20 mM, Tris-HCl [pH 8.0], 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), twice with high-salt buffer (20 mM, Tris-HCl [pH 8.0], 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), twice with RIPA buffer (50 mM, HEPES-KOH [pH 7.6], 250 mM LiCl, 1 mM EDTA, 1% IGEPAL CA630, 0.7% sodium deoxycholate), and once with TE containing 50 mM NaCl. After elution with elution buffer (10 mM, Tris-HCl [pH 8.0], 300 mM NaCl, 5 mM EDTA, 0.5% SDS) for 15 min at 65°C, immunoprecipitated chromatin, together with input chromatin, was reverse cross-linked and DNA was purified using the QIAquick PCR Purification Kit (QIAGEN).

For ChIP-PCR experiments, cells were cross-linked with formaldehyde for 8 min (for Nanog, Oct4, and Sox2), or cross-linked with 2 mM Di(N-succinimidyl) glutarate (DSG) in PBS for 5 min at room temperature, washed with 50 mM Tris-HCl (pH 7.4) for 5 min, and then cross-linked with formaldehyde for 8 min (for Esrrb, Stat3, Klf4, p300, Med12, and Tet1). Chromatin was sonicated for 45 cycles of 30 s on and 30 s off. Chromatin corresponding to approximately 25 μg DNA was incubated with 25 μL Dynabeads protein G coupled to 2 μg of the following primary antibodies: mouse anti-Esrrb (PP-H6705-00, R&D Systems); rabbit anti-Stat3 (sc-482, Santa Cruz); goat anti-Klf4 (AF3158, R&D Systems); rabbit anti-p300 (sc-585, Santa Cruz); rabbit anti-Med12 (A300-774A, Bethyl Laboratories); rabbit anti-Tet1 (GTX124207, GeneTex); rabbit anti-Nanog (A300-397A, Bethyl Laboratories); goat anti-Oct4 (sc-8628, Santa Cruz); goat anti-Sox2 (GT15098, Neuromics). Reverse cross-linked and purified DNA was analyzed by real-time PCR. The percentage of input was calculated by the standard curve method and normalized to the values obtained at a negative control region within the intergenic spacer (IGS) of ribosomal DNA (rDNA), which gave reliable amplification due to its multiple copies in the genome. Each experiment was performed in technical triplicate. Error bars indicate SD of three replicates.

For MNase-ChIP, Tet-on 3 × FLAG-Esrrb EpiSCs were cross-linked with DSG and formaldehyde, lysed in lysis buffer 1 containing protease inhibitor cocktail, washed in MNase buffer (50 mM Tris-HCl [pH 8.0], 5 mM $CaCl_2$, 0.25% IGEPAL CA-630), and treated with 10,000 gels units MNase (NEB) and 100 units ExoIII (NEB) in 0.5 mL MNase buffer supplemented with 0.1 mg/ml BSA per 1 × $10^7$ cells for 15 min at 37°C. After addition of EDTA to 45 mM to terminate the reaction, samples were sonicated for 15 cycles of 30 s on and 30 s off in the presence of 0.5% SDS to break nuclear membrane. The following primary antibodies were used for ChIP: rabbit anti-H3 (ab1791, Abcam); mouse monoclonal anti-FLAG (F1804, Sigma).

For ChIP-bisulfite sequencing and ChIP-GlucMS-PCR assays, chromatin corresponding to 150 μg DNA was incubated with 100 μL Dynabeads protein G coupled to 12 μg anti-Esrrb antibody. Chromatin was reverse cross-linked and DNA was extracted once with phenol-chloroform and once with chloroform, and precipitated with isopropanol. One-sixth of ChIP DNA, together with 100 ng input DNA, was used for bisulfite treatment, and the remaining ChIP DNA, together with 1 μg input DNA, was used for GlucMS-PCR. Primer sequences are listed in Table S1.

### NOMe-seq

NOMe-seq assays were performed as previously described (Taberlay et al., 2011) with some modifications. $1 \times 10^6$ cells were lysed in nuclei lysis buffer (10 mM Tris-HCl [pH 7.4], 10 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA, 0.5% IGEPAL CA-630) containing protease inhibitor cocktail for 10 min on ice and aliquots containing $2 \times 10^5$ cells were treated with 100 units M.CviPI (NEB) in the presence of 160 μM S-adenosylmethionine for 15 min at 37°C. Purified DNA was used for bisulfite treatment. Sequencing data was analyzed using methylcircleplot (https://doi.org/10.6084/m9.figshare.842634). GCG trinucleotides were excluded from the analysis as they are targets of both CpG and GpC methyltransferases. If at least 2 of 3 consecutive GpCs were methylated, the region between methylated GpCs was considered as accessible. Then regions with at least 120 bp of inaccessible chromatin were considered as nucleosome-occupied. Primer sequences are listed in Table S1.

### GlucMS-PCR

500 ng DNA isolated from sheared chromatin was treated with or without T4 phage β-glucosyltransferase (T4-BGT, NEB) in the presence of uridine diphosphoglucose for 12–16 hours at 37°C. DNA was then divided into three aliquots and treated with MspI (which is sensitive to glycosylated 5hmC at a CCGG site), HpaII (which cleaves only unmodified cytosine residues), or no enzyme for 8 hours at 37°C, followed by proteinase K treatment for 30 min at 40°C. The amount of undigested DNA was determined by real-time PCR. Relative abundance values were calculated by the standard curve method and normalized to the values obtained at a negative control region (IGS of rDNA) that does not contain a CCGG sequence. The percentage of each cytosine derivative was calculated as follows: %5hmC = [(T4-BGT_MspI) × (No T4-BGT_No enzyme) / (T4-BGT_No enzyme) − (No T4-BGT_MspI)] / (No T4-BGT_No enzyme) × 100 (%); %5mC = [(No T4-BGT_HpaII) − (T4-BGT_MspI) × (No T4-BGT_No enzyme) / (T4-BGT_No enzyme)] / (No T4-BGT_No enzyme) × 100 (%); %C = [(No T4-BGT_No enzyme) − (No T4-BGT_HpaII)] / (No T4-BGT_No enzyme) × 100 (%). Each experiment was performed in technical triplicate. Error bars indicate SD of three replicates. Primer sequences are listed in Table S1.

### Luciferase Reporter Assays

Genomic regions amplified by PCR were cloned into a construct containing the human UBC minimal promoter and firefly luciferase (luc2). ZHBTc4 ESCs were transfected with pGL4.75[hRluc/CMV] together with the reporter constructs. Twenty-four hours after transfection, luciferase activity was measured using Dual-Luciferase Reporter Assay System (Promega). The relative luciferase activity (luc2/hRluc) was further normalized to that of an empty vector control. Each experiment was performed in biological triplicate. Error bars indicate SD of three replicates. Primer sequences used for amplification are listed in Table S1.

### Microarray and Data Analysis

Total RNA was isolated from Tet-on Esrrb EpiSCs undergoing reprogramming, transgene-independent ESC-like cells that were reprogrammed and maintained in K10 + LIF, parental E3 EpiSCs that were cultured in CM + FGF2, and GOF18-3F and GOF18-5A ESCs that were adapted to feeder-free conditions in K10 + LIF. For each dataset, the experiments were performed in biological duplicate. Microarray analysis was performed using MouseRef-8 v2.0 Expression BeadChips (Illumina) as previously described (Greber et al., 2010). Background corrected probe-level data exported from the GenomeStudio software (Illumina) were variance-stabilizing transformed and robust-spline normalized using the Bioconductor lumi package (Du et al., 2008) in R (Table S2, lumi_exprs). Expression values from duplicate samples were averaged, and Pearson's correlation coefficients were computed based on linear scale data. For weighted gene co-expression network analysis (WGCNA; Langfelder and Horvath, 2008), probes with detection p values of $\leq 0.01$ (reported in the Group Probe Profile table in the GenomeStudio software) in at least one sample were selected (14284 of 25697 probes). Multiple probes with the same Ensembl gene ID were collapsed using the collapseRows function in the R package WGCNA (Langfelder and Horvath, 2008) by selecting those with the highest mean values across the samples (9965 genes) and only genes with $\log_2$ fold change relative to 0 h of $\geq 0.5$ in at least one sample were used (4434 genes) (Table S2, mean_detected_collapsed_lfc0.5). A signed weighted correlation network was constructed using the blockwiseModules function with a soft-thresholding power of 20, a minimum module size of 10, and a dendrogram merge cut height of 0.25. The most representative gene expression profiles of each module (module eigengenes) were calculated and correlations between individual gene expression and module eigengene (module membership values) were computed to identify hub genes that have high intramodular connectivity (module membership $\geq 0.8$) (Table S2, WGCNA_module membership). Functional annotation of hub genes was performed using DAVID 6.8 (Huang et al., 2009). Representative terms in enriched annotation clusters identified by the Functional Annotation Clustering tool, together with their Group Enrichment Scores, are presented.

### ChIP-seq and Data Analysis

Sequencing libraries were prepared according to the TruSeq ChIP Sample Preparation protocol (Illumina). Amplified libraries were purified with 1.0 volume of AMPure XP beads and size-selected on a 2% agarose gel (250-450 bp). Sequencing was performed on the HiSeq 2500 (Illumina) as single-end 50-bp reads. Sequence reads were aligned to the mm10 mouse reference genome using Bowtie2 (Langmead and Salzberg, 2012) with default parameters. Fragment sizes were estimated using phantompeakqualtools (https://code.google.com/archive/p/phantompeakqualtools/). Peak calling was performed using MACS2 (Zhang et al., 2008) with the following parameters: -g mm -B –SPMR -q 0.01 –nomodel –shiftsize 80 (Table S3, summits). Input DNA from each sample was used as controls. Reads per million (RPM)-normalized pileup signals in the bedGraph format were converted into bigWig files

using the UCSC bedGraphToBigWig tool. Peaks that overlap with the blacklisted regions (https://sites.google.com/site/anshulkundaje/projects/blacklists), satellite repeats (RepeatMasker) or regions with abnormally high signals in all samples (Table S3, blacklist_GOF18) were removed as probable mapping artifacts. Pairwise Pearson correlation coefficients between two bigWig files were calculated using the UCSC bigWigCorrelate tool based on mean values over merged peak regions, in which ± 100-bp regions around the peak summits from all datasets were combined. ChIP-seq peaks in time-series samples were merged separately for each TF and RPM-normalized pileup signals in 10-bp bins over ± 1-kb regions around the peak centers were grouped into 6 clusters using deepTools (Ramírez et al., 2014) (Table S3, cluster). Gene set enrichment analysis (GSEA; Subramanian et al., 2005) was performed to assess enrichment of the gene sets assigned to peaks. Peaks were assigned to the nearest Ensembl protein-coding genes whose TSSs are located within 10 kb of peak summits or merged peak centers (Table S3, genes_assigned). For GSEA on microarray data on reprogramming EpiSCs, all detected and collapsed probes were used. For GSEA on RNA-seq data on differentiating EpiSCs, enrichment of the gene sets in differentially expressed protein-coding genes based on rlog-transformed values was determined. FDR was estimated by gene set permutation tests. Heatmaps were generated from normalized enrichment scores (NESs) for gene sets with FDR $\leq$ 0.1. Motif analysis was performed using the MEME-ChIP (Machanick and Bailey, 2011). Sequences of ± 100-bp regions around the merged peak centers were masked for repetitive elements using RepeatMasker. One thousand randomly selected sequences were used for *de novo* motif discovery by MEME. Significantly enriched motifs found by *de novo* discovery (MEME and DREME) and local motif enrichment analysis (CentriMo in local mode) were used for PscanChIP (Zambelli et al., 2013) to scan the whole set of unmasked sequences. Heatmaps were generated from p values for global enrichment. Functional annotation of gene sets associated with peaks in each cluster was performed using GREAT (McLean et al., 2010). Because a large number of peaks can result in saturation of a gene-based hypergeometric test, only the top 10,000 peaks were used for Nanog c1 and c3. Binomial raw p values for ontology terms that were significant by both the region-based binomial test and the gene-based hypergeometric test are presented.

### RNA-seq and Data Analysis

For EB formation, E3 EpiSCs that were maintained in CM/N2B27 + FA were seeded into 96-well V-bottom plates at $3 \times 10^3$ cells per well in DMEM containing 10% FBS, 0.1% poly(vinyl alcohol), and 5 μM Y-27632. On days 1−5, 16 EBs were pooled and total RNA was isolated. Sequencing libraries were prepared according to the TruSeq RNA Sample Preparation protocol (Illumina) and sequenced on the HiSeq 2500 (Illumina) as paired-end 50-bp reads. Sequence reads were aligned to the mm10 mouse reference genome using TopHat2 (Kim et al., 2013) with default parameters. The number of fragments mapped to each Ensembl gene was counted using the Bioconductor bamsignals package (Mammana and Helmuth, 2016). Raw count data was normalized and transformed to the $\log_2$ scale using the rlog function of the Bioconductor DESeq2 package (Love et al., 2014) (Table S4).

### Analysis of Publicly Available Data

The publicly available datasets used in this study are listed in Table S1. For ChIP-seq data, sequence reads in the FASTQ format were downloaded from the European Nucleotide Archive (https://www.ebi.ac.uk/ena) and aligned to the mm10 mouse reference genome. Fragment sizes were estimated using phantompeakqualtools. Peak calling was performed using MACS2 (the settings are provided in Table S1). Input DNA or mock IP samples from each experiment were used as controls. Histone modification states in EpiSCs (Factor et al., 2014) and ESCs (Whyte et al., 2012) around EOSN peaks were analyzed using the UCSC bigWigAverageOverBed tool. To compare ChIP-seq data from different resources, fold enrichment over matching input DNA was calculated using the MACS2 bdgcmp function. Average fold enrichment values over ± 500-bp regions centered on the top 1,000 peaks in each EOSN cluster were calculated. For whole genome bisulfite sequencing (WGBS-seq) data (Habibi et al., 2013; Wu et al., 2015), bigWig files containing methylation levels and coverage information at each CpG site (.meth.bw and .read.bw files, respectively) were downloaded from MethBase (Song et al., 2013) via the UCSC Genome Browser. CpGs covered by at least 10 reads were selected, and average methylation levels over ± 500-bp regions centered on the top 1,000 peaks in each EOSN cluster were calculated using the bigWigAverageOverBed tool. The regions containing no CpGs were excluded. Global methylation levels were calculated by averaging methylation levels over 1-kb windows containing at least 5 CpGs (coverage $\geq$ 10). EpiSC- and ESC-specific HMRs were identified using the dmr tool in MethPipe (Song et al., 2013). Cell type–specific HMRs spanning at least 10 CpGs and containing at least 5 differentially methylated CpGs were selected. Overlap between two regions (200-bp regions around the centers) was calculated and its significance was assessed by Fisher's exact test using the R function fisher.test. Smc1a ChIA-PET data (Dowen et al., 2014) showing cohesin-mediated chromatin interactions was obtained from GEO (GSM1397342) and visualized using the WashU epigenome browser.

## DATA AND SOFTWARE AVAILABILITY

The accession numbers for the microarray, ChIP-seq, and RNA-seq datasets reported in this paper are ArrayExpress: E-MTAB-5341, E-MTAB-5342, and E-MTAB-5343, respectively (https://www.ebi.ac.uk/arrayexpress).