CELL BIOLOGY

A balanced Oct4 interactome is crucial for maintaining pluripotency

Dong Han[1], Guangming Wu[1,2], Rui Chen[3], Hannes C. A. Drexler[4], Caitlin M. MacCarthy[1], Kee-Pyo Kim[1,5], Kenjiro Adachi[1], Daniela Gerovska[6,7], Lampros Mavrommatis[1], Ivan Bedzhov[3], Marcos J. Araúzo-Bravo[1,6,7], Hans R. Schöler[1]*

Oct4 collaborates primarily with other transcriptional factors or coregulators to maintain pluripotency. However, how Oct4 exerts its function is still unclear. Here, we show that the Oct4 linker interface mediates competing yet balanced Oct4 protein interactions that are crucial for maintaining pluripotency. Oct4 linker mutant embryonic stem cells (ESCs) show decreased expression of self-renewal genes and increased expression of differentiation genes, resulting in impaired ESC self-renewal and early embryonic development. The linker mutation interrupts the balanced Oct4 interactome. In mutant ESCs, the interaction between Oct4 and Klf5 is decreased. In contrast, interactions between Oct4 and Cbx1, Ctr9, and Cdc73 are increased, disrupting the epigenetic state of ESCs. Control of the expression level of Klf5, Cbx1, or Cdc73 rebalances the Oct4 interactome and rescues the pluripotency of linker mutant ESCs, indicating that such factors interact with Oct4 competitively. Thus, we provide previously unidentified molecular insights into how Oct4 maintains pluripotency.

INTRODUCTION

Pluripotent stem cells are characterized by their ability to self-renew indefinitely and give rise to all the cell types of an organism, except for extraembryonic tissues. The pluripotent state is maintained by a highly interconnected gene regulatory network centered on the core pluripotency transcription factors (TFs) Oct4, Sox2, and Nanog (*1*). Oct4 is considered to be a master transcriptional regulator of pluripotency. It is expressed in oocytes, blastomeres, inner cell mass (ICM), epiblast, and germ cells in vivo and in pluripotent cells in vitro (*2*, *3*). Oct4 deficiency causes preimplantation lethality due to failure of mouse embryos to form a pluripotent ICM (*4*). In embryonic stem cells (ESCs), Oct4 expression needs to be tightly regulated, as either the silencing of Oct4 or the overexpression of Oct4 causes ESCs to undergo differentiation (*5*). Our group has recently shown that while Oct4 is not required for the establishment of totipotency and initial stages of reprogramming, the TF is indispensable for maintenance of pluripotency (*6*, *7*).

Oct4 functions mainly through the activation of pluripotency- and self-renewal–associated genes and the simultaneous repression of differentiation-promoting genes (*8*). Genome-wide mapping of the binding sites of Oct4 and other ESC factors has led to the identification of sets of jointly regulated or bound targets, suggesting that Oct4 may function by collaborating primarily with other transcriptional factors or coregulators in regulating transcriptional programs to maintain ESC pluripotency (*9*, *10*). Therefore, several mass spectrometry (MS)–based studies were conducted in mouse ESCs to reveal the interaction partners of Oct4 (*11*–*14*). Not only TFs but also many components of multiple epigenetic regulatory complexes, such as NuRD, COMPASS-like, SWI/SNF, LSD1, and PAF1 complexes, are associated with Oct4, contributing to epigenetic modifications and maintenance of pluripotency in ESCs.

However, the mechanism by which Oct4 interacts with partner proteins to exert its function remains unclear. On the basis of the Oct4 crystal structure, we have shown that the highly evolutionary conserved part of the linker region between the two DNA binding domains is structured as an α helix and exposed on the protein's surface (*14*). The L80A point mutation in this region markedly decreased Oct4 reprogramming ability, suggesting that this interface is critical for Oct4 function in reprogramming (*14*). By showing that the Oct4 linker interface mediates a competing yet balanced Oct4 interactome, we provide previously unidentified molecular insights into the mechanism underlying the interaction of Oct4 with its partners to regulate pluripotency.

RESULTS

Oct4 linker is crucial for maintaining pluripotency in vitro

To investigate the role of the Oct4 linker in maintaining pluripotency, we replaced the mouse Oct4 linker with its *Xenopus*, zebrafish, or medaka orthologs and performed rescue experiments using ZHBTc4 ESCs (*5*). ZHBTc4 ESCs have had both endogenous Oct4 alleles disrupted; a doxycycline (Dox)–repressible transgene Oct4 maintains the cells as ESCs in the absence of Dox. In the presence of Dox, the transgene Oct4 is repressed, and ZHBTc4 ESCs differentiate into the trophectoderm lineage. By using the lentivirus overexpression system, wild-type (WT) or linker mutant Oct4 was transfected into ZHBTc4 ESCs. Simultaneously, the inducible transgene Oct4 was repressed by adding Dox, and only the transfected WT or linker mutant Oct4 was expressed. After 96 hours of Dox treatment, ESC colonies were obtained to delineate whether mutant Oct4 could rescue ESC identity (Fig. 1A). The chimeric protein with the *Xenopus* linker rescued stem cell renewal as effectively as WT Oct4. In contrast, chimeric proteins with either the medaka or zebrafish linker

[1]Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, 48149 Münster, Germany. [2]Guangzhou Regenerative Medicine and Health Guangdong Laboratory, 6 Luoxuan Avenue, Haizhu District, 510320 Guangzhou, P. R. China. [3]Embryonic Self-Organization Research Group, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, 48149 Münster, Germany. [4]Bioanalytical Mass Spectrometry, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, 48149 Münster, Germany. [5]Department of Medical Life Sciences, College of Medicine, The Catholic University of Korea, 222 Banpo-daero Seocho-gu, Seoul 06591, Republic of Korea. [6]Group of Computational Biology and Systems Biomedicine, Biodonostia Health Research Institute, San Sebastian 20014, Spain. [7]IKERBASQUE, Basque Foundation for Science, Bilbao 48011, Spain.
*Corresponding author. Email: office-schoeler@mpi-muenster.mpg.de

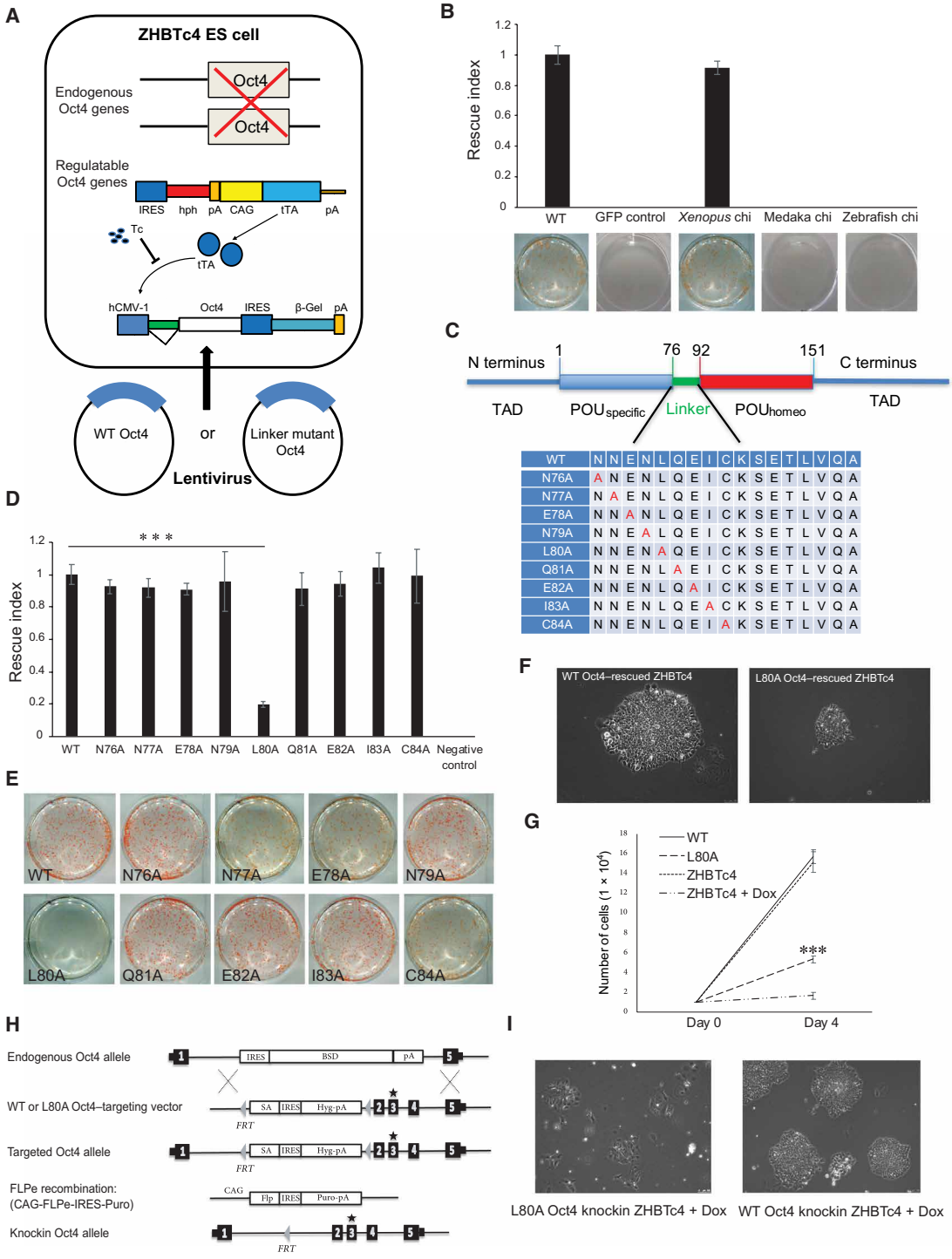

Fig. 1. Oct4 linker is crucial for maintaining pluripotency in vitro. (**A**) Schematic of complementation assay system. In ZHBTc4 ESCs, both *Oct4* alleles have been disrupted, and a Dox-suppressible transgene Oct4 maintains ESC pluripotency. The WT or linker mutant Oct4 was transfected into ZHBTc4 ESCs by a lentivirus overexpression system. (**B**) Rescue index of WT Oct4 and chimeric Oct4 with different species linker orthologs in the presence of Dox and AP staining of ZHBTc4 ESC colonies rescued by different Oct4 proteins. (**C**) Representation of mouse Oct4 protein (WT) and different alanine mutations on the linker segment. N-terminal transactivation domain (TAD), POU-specific domain, linker, POU homeodomain, and C-terminal TAD are shown in different colors. Amino acids that are replaced are colored in red. (**D**) Rescue index of WT Oct4 and mutant Oct4 with different alanine mutations on the linker segment. (**E**) AP staining of ZHBTc4 ESCs colonies rescued by different Oct4 mutants. (**F**) Morphology of WT and L80A Oct4–rescued ZHBTc4 ESC colony. (**G**) Proliferation rate of WT, L80A Oct4–rescued ZHBTc4 ESCs, and ZHBTc4 ESCs cultured without or with Dox. *** represents *P* value between WT and L80A Oct4–rescued ZHBTc4 ESCs. (**H**) Strategy of WT and L80A Oct4–KI ZHBTc4 ESCs expressing FLPe. Black boxes represent Oct4 exons. Star represents the position of L80 amino acid in the linker region. FLPe, FLP recombinase; SA, splice acceptor; BSD, blasticidin; Hyg, hygromycin; Puro, puromycin. (**I**) Morphology of L80A and WT Oct4–KI ZHBTc4 cells cultured with Dox.

sequences could not produce any stem cell colonies (Fig. 1B). Viral transcription levels were comparable (fig. S1A) and thus could not account for the failure to rescue pluripotency. This dataset demonstrates that the linker region is important for the biological function of Oct4 in maintaining pluripotency.

Sequence alignment of the linker regions in human, mouse, *Xenopus*, zebrafish, and medaka revealed that human, mouse, and *Xenopus* linkers share more conserved amino acids than do zebrafish and medaka (fig. S1B). The similarity is especially notable in the N-terminal part of the linker region, which is structured as an α helix and exposed on the protein's surface (*14*). Next, we performed an alanine scan on the conserved linker residues of α5 helix (Asn^{76}-Cys^{84}) to investigate whether these residues are required for the biological activity of Oct4 (Fig. 1C). At the same expression level (fig. S1C), the L80A mutant of Oct4 led to significantly fewer ESC colonies; other mutants did not affect the biological function of Oct4 in maintaining pluripotency (Fig. 1, D and E). L80A Oct4–rescued ZHBTc4 ESCs expressed pluripotency markers (fig. S1, D and E), and they could form embryoid bodies and differentiate into all three germ layers under differentiation culture conditions (fig. S1, F and G), suggesting that they are pluripotent. However, we observed that the cell proliferation of L80A Oct4–rescued ZHBTc4 ESCs was markedly impaired compared with that of WT Oct4–rescued ZHBTc4 ESCs, as shown by smaller colonies and poor cell growth (Fig. 1, F and G). Together, these results indicate that the L80A mutation impairs the rescue efficiency and the self-renewal capacity of pluripotent stem cells.

To further investigate the effect of the L80A mutation on Oct4 function, we replaced one allele of Oct4 in ZHBTc4 ESCs with either WT or L80A Oct4 (Fig. 1H). In the absence of Dox, one-allele WT or L80A Oct4–knockin (KI) ZHBTc4 ESCs expressed the original transgene Oct4 of ZHBTc4 ESCs and WT/L80A–KI Oct4 and could be maintained as ESCs. When these cells were cultured in the presence of Dox for 12 days, WT Oct4 was found to rescue ZHBTc4 ESCs; however, all the L80A Oct4–KI ZHBTc4 cells underwent differentiation (Fig. 1I and fig. S1H). Quantitative reverse transcription polymerase chain reaction (RT-qPCR) results showed that L80A Oct4 could not maintain the expression of the core pluripotency genes (fig. S1I). We examined the expression and localization of the L80A mutant Oct4 by sequencing and immunohistochemistry in L80A Oct4–KI ZHBTc4 ESCs cultured in the presence of Dox for 2 days and found that L80A Oct4 protein was expressed and localized correctly in the nucleus (fig. S1, J and K). As one-allele L80A Oct4–KI ZHBTc4 ESCs could not maintain pluripotency, many cells differentiated after 2 days of culture in the presence of Dox. Therefore, fewer than 50% cells expressed Oct4. In contrast, most one-allele WT Oct4–KI ZHBTc4 ESCs were pluripotent and positive for Oct4 (fig. S1K). These data confirm that the mutation in the linker region impairs the biological function of Oct4 in maintaining pluripotency in vitro.

Oct4 linker is crucial for maintaining pluripotency in vivo

To determine the effect of the Oct4 linker domain on development, an L80A mutant allele of Oct4 was generated via homologous recombination in ESCs. By using the tetraploid complementation method, we generated heterozygous male L80A Oct4–KI mice in one step. The F0 heterozygous male WT/L80A Oct4 mice were crossed with WT female mice to generate F1 heterozygous mice. The F1 heterozygous male and female mice grew normally without any evidence of developmental and growth (size) issues (fig. S2A); they were fertile and transmitted the mutant allele to approximately 50% of their progeny when they were crossed with WT mice (fig. S2B). We confirmed L80A Oct4 expression by checking the gonads of F1 heterozygous embryos on embryonic day 13.5 (E13.5) and found that they expressed a mixture of WT and L80A Oct4 (fig. S2C). However, when we intercrossed the F1 or subsequent generations of male and female WT/L80A Oct4 mice, of the 462 pups that were born, only 5 pups were homozygous (1.1%), indicating that L80A Oct4 portends a heavy embryonic impairment (Fig. 2A).

To determine the timing of homozygous L80A Oct4 embryo death, we examined the progression of embryonic development. Freshly isolated ~E3.5 mutant embryos normally expressed Oct4 and morphologically resembled normal blastocysts (Fig. 2B). Oct4–knockout (KO) embryos failed to produce primitive endoderm (PrE), which typically is clearly seen on day 4.5 (*15*). In contrast, on day 4.5, L80A/L80A Oct4 embryos exhibited a normal phenotype and formed an Oct4-positive epiblast and Gata6-positive PrE (Fig. 2C). The numbers of Oct4-positive and Gata6-positive cells were similar between WT/WT and L80A/L80A Oct4 ~E4.5 embryos (fig. S2D). However, at the molecular level, there were notable differences between the WT/WT and L80A/L80A Oct4 embryos. Although L80A/L80A Oct4 embryos showed unaltered expression of the core pluripotency marker genes *Oct4*, *Sox2*, *Nanog*, *Esrrb*, and *Rex1*, they showed lower expression of the growth-related genes *Utf1*, *Sall4*, and *Fgf4* and higher expression of the differentiation-related genes *Cdx2*, *Eomes*, *Sox17*, and *Fgf5* (Fig. 2D). To further evaluate the pluripotency of ~E4.5 L80A/L80A Oct4 embryos, we plated the embryos onto mouse embryonic fibroblasts (MEFs) to derive pluripotent ESCs. We derived 23 ESC lines from 46 embryos, of which 6 were homozygous for the L80A allele (Fig. 2E and fig. S2E). Whole-genome transcription analysis showed that L80A/L80A Oct4 ESCs were similar to WT/WT Oct4 ESCs (Fig. 2F). Among 25,601 detected genes, 408 genes were up-regulated, while 272 genes were down-regulated in L80A/L80A Oct4 ESCs compared to WT/WT Oct4 ESCs (Fig. 2G). Functional enrichment analysis showed that most of the differentially expressed genes were development-related genes (fig. S2, F and G). Although L80A/L80A Oct4 ESCs expressed pluripotent markers (fig. S2, H and I), they proliferated poorly (Fig. 2E), likely due to lower expression of genes involved in self-renewal (Fig. 2H). Whole-genome transcription showed that L80A/L80A Oct4 ESCs exhibited lower levels of self-renewal–related genes and higher levels of differentiation–related genes (Fig. 2F and fig. S2I). Whole-genome transcription analysis showed that *Lin28a* was significantly down-regulated in L80A/L80A Oct4 ESCs. Decreased *c-Myc*, *Utf1*, *Tcl1*, and *Fgf4* expression was also observed from the whole-genome transcription analysis in L80A/L80A Oct4 ESCs, although the down-regulation of these genes was not significant (Fig. 2F and fig. S2J). Similarly, gene ontology (GO) and RT-qPCR analysis of down-regulated genes also showed that L80A Oct4–rescued ZHBTc4 ESCs exhibited significantly reduced proliferation-related gene expression (fig. S2, K and L). These results demonstrate that L80A Oct4 impairs self-renewal and cannot efficiently repress differentiation genes in epiblasts.

To assess whether the up-regulation of differentiation-associated genes reflected a change in the proportion of undifferentiated to differentiated cells in L80A/L80A Oct4 ESCs, we stained WT/WT and L80A/L80A Oct4 ESCs with pluripotency and differentiation lineage markers. Most of the L80A/L80A Oct4 ESCs expressed pluripotency markers, but they did not express differentiation lineage

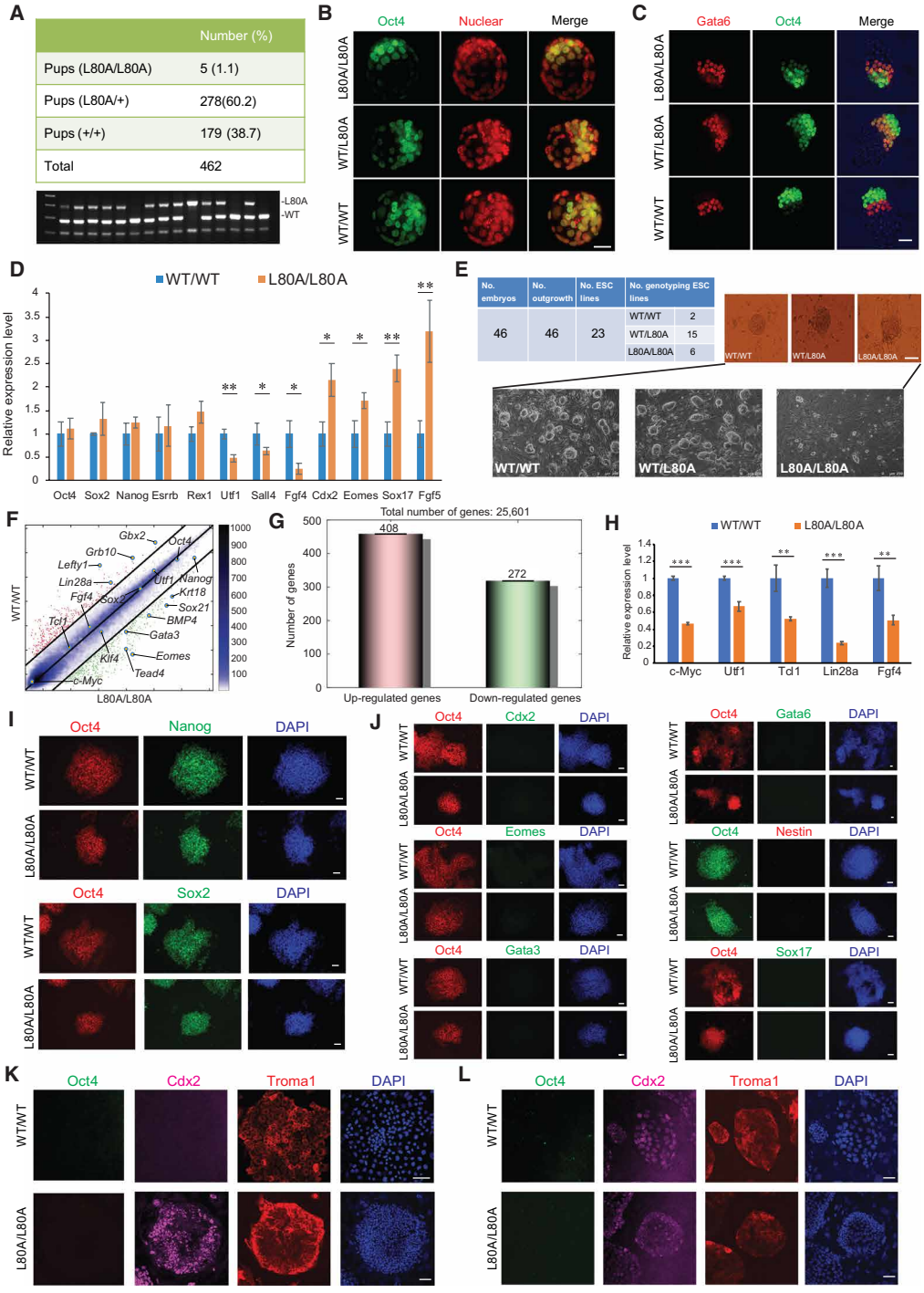

Fig. 2. Oct4 linker is crucial for maintaining pluripotency in preimplantation embryos. (**A**) Number of pups with different genotypes from intercross matings of WT/L80A Oct4 mice and representative genotyping gel picture. (**B**) Immunocytochemistry on ~E3.5 blastocysts using anti-Oct4 antibody and confocal microscopy. Scale bar, 30 μm. (**C**) Immunocytochemistry on ~E4.5 embryos using anti-Oct4 and anti-Gata6 antibodies and confocal microscopy. Scale bar, 30 μm. (**D**) RT-qPCR gene expression analysis of E4.5 embryos. All data are calibrated to the WT/WT embryos, whose expression is considered to be 1 for all genes. (**E**) Derivation of ESC line with different genotypes from E4.5 embryos. Upright pictures show outgrowths from embryos. Scale bar, 50 μm. Landscape pictures below show established ESCs lines with different genotypes. (**F**) Pairwise scatterplot analysis of the global gene expression profiles of L80A/L80A and WT/WT Oct4 ESCs obtained from the microarray analysis. (**G**) Bar plot of the number of differentially expressed genes between WT/WT and L80A/L80A Oct4 ESCs. (**H**) Relative mRNA expression of self-renewal–related genes in WT/WT and L80A/L80A Oct4 ESC lines. (**I** and **J**) Immunofluorescence microscopy images of WT/WT and L80A/L80A Oct4 ESCs using antibodies against Oct4, Sox2, and Nanog (I) and Oct4, Cdx2, Eomes, Gata3, Gata6, Nestin, and Sox17 (J). Scale bars, 50 μm. (**K** and **L**) Immunofluorescence microscopy images of WT/WT and L80A/L80A Oct4 ESCs expressing Cdx2-ERT2-RFP cultured under the TSC culture condition without (K) or with (L) tamoxifen for 3 weeks using antibodies against Oct4, Cdx2, and Troma1. Scale bars, 50 μm. DAPI, 4′,6-diamidino-2-phenylindole.

markers (Fig. 2, I and J). Only very rare spontaneously differentiated cells expressed lineage markers in WT/WT and L80A/L80A Oct4 ESCs (fig. S2M). The positive staining of lineage markers in ESCs, neural stem cells, and trophoblast stem cells (TSCs) demonstrated that the staining worked well (Fig. 2K and fig. S2, M and N). These data suggest that there is no change in the proportion of undifferentiated to differentiated cells in L80A/L80A Oct4 ESCs.

As the L80A mutation causes up-regulation of *Tead4*, *Gata3*, and *Eomes*, which is associated with the trophectoderm lineage in L80A/L80A Oct4 ESCs, we examined whether L80A/L80A Oct4 ESCs could differentiate into TSCs under Fgf4 plus feeder cell culture condition. We generated WT/WT and L80A/L80A Oct4 ESC lines that were stably transfected with a CAG promoter–driven Cdx2-ERT2-RFP (red fluorescent protein) coding plasmid for inducible overexpression of Cdx2. Under the ESC culture condition, these cells expressed Oct4, but they did not express the TSC markers Cdx2 and Troma1 (fig. S3, A and B). When these cells were cultured under the TSC culture condition (MEF medium containing Fgf4 plus feeder cells) without adding tamoxifen, some L80A/L80A Oct4 ESCs, but not WT/WT Oct4 ESCs, differentiated into Cdx2 and Troma1 double-positive TSCs (Fig. 2K and fig. S3C). When we added tamoxifen into the TSC culture medium to induce Cdx2 overexpression, both WT/WT and L80A/L80A Oct4 ESCs could differentiate into TSCs (Fig. 2L and fig. S3D). These data demonstrate that L80A/L80A Oct4 ESCs, but not WT/WT Oct4 ESCs, can spontaneously differentiate into TSCs under the TSC culture condition.

Next, we assessed the development of postimplantation L80A/L80A Oct4 embryos. Postimplantation L80A/L80A Oct4 embryos (~E5.5) showed normal morphology and normally expressed Oct4 and Cdx2 (Fig. 3A). However, L80A/L80A Oct4 embryos showed defective growth, with the onset of growth retardation observed at ~E6.5 (Fig. 3, B and C). At ~E7.75, although L80A/L80A Oct4 embryos underwent gastrulation, they grew much slower and were of smaller size than WT/WT Oct4 embryos (Fig. 3, D and E). L80A/L80A Oct4 embryos (~E9.5) showed diverse phenotypes. Some embryos died and degraded. While some embryos survived, they grew slower and were of smaller size than WT embryos (Fig. 3F and fig. S3E). A previous study has reported that postimplantation deletion of Oct4 ~E6.0 and ~E6.5 results in a fully penetrant phenotype, with the embryos being amorphous at ~E9.5. Deletion of Oct4 ~E7.0 results in a number of phenotypes, such as craniorachischisis, failed turning, defective somitogenesis, and posterior truncation observed at E9.5 (*16*). In contrast, a few ~E9.75 L80A/L80A Oct4 embryos survived, and they were much smaller than WT embryos, with normal turning, posterior extension reaching 21 to 29 somites, and normal neural tube closure, indicating that the phenotype of L80A/L80A Oct4 embryos is different from that of postimplantation Oct4 deletion embryos (Fig. 3G). These few smaller L80A/L80A Oct4 embryos that did not have obvious developmental defects might develop to term (Fig. 2A). At ~E12.5, we found that some L80A/L80A Oct4 embryos developed normal placenta (Fig. 3H), although the placentas were smaller and degraded in the case of earlier embryo death (Fig. 3I), suggesting that L80A Oct4 does not have a direct effect on extraembryonic tissue development. Dead L80A/L80A Oct4 embryos were not found after ~E13.5 (Fig. 3J), indicating that the death period of such embryos was from E9.5 to E13.5. All these results show that L80A Oct4 causes an impairment of pluripotent stem cells in embryos, leading, in most cases, to early embryo death.

Notably, two male and three female homozygous L80A Oct4 mice grew normally to adult age. We crossed the male L80A/L80A mice at the age of 3 months with WT female mice. Matings were successful, and 15 pups were born from two litters. As expected, all the pups were heterozygous for L80A Oct4 as shown by genotyping (fig. S3F). However, when we crossed the three female L80A/L80A Oct4 mice at the age of 2 months with WT male mice, pregnancy could not be supported for more than 3 months, indicating that the mice were infertile (fig. S3G). When we sacrificed the L80A/L80A Oct4 female mice and checked their ovaries, we found no developing follicles and oocytes (fig. S3H). These results suggest that L80A Oct4 impairs female germ cell development, causing female infertility, whereas male germ cell development is unaffected.

In ESCs, genomic binding and transactivation ability of L80A Oct4 are similar to those of WT Oct4

Next, we aimed to elucidate the molecular mechanism underlying the biological function of the impaired Oct4 linker. As a TF, Oct4 needs to bind to DNA to regulate gene expression. In vitro electrophoretic mobility shift assay (EMSA) analysis showed that L80A Oct4 bound to the octamer DNA probes W and PORE well (fig. S4A). However, L80A mutant Oct4 showed reduced DNA binding when bound alone on the sox-oct composite motif probes, Nanog and Utf1, compared to WT Oct4 (fig. S4B). The addition of Sox2 yielded similar amounts of the Oct4/Sox2 heterodimer on Nanog but not Utf1 (fig. S4B). This indicates that while L80A Oct4 binding alone appears to be impaired on certain motifs, the presence of Sox2 can augment L80A Oct4 binding on certain binding sites. Oct4 and Sox2 are likely to interact with other factors in ESCs and form multiprotein complexes (*11*–*14*). As we have previously shown for the POU factor Oct1, the coactivator OBF1 alleviates DNA sequence requirements of the Oct1 dimer on PORE-related sequences in vitro (*17*). Therefore, it can be assumed that in ESCs, binding of L80A Oct4 to sites like those in the *Utf1* gene may be enhanced by cofactors and may even be indistinguishable to binding of WT Oct4. We then assessed the genomic binding of WT and L80A Oct4 in WT/WT and L80A/L80A Oct4 ESCs derived from embryos by performing chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). L80A Oct4 bound similar genomic sites as WT Oct4 in ESCs (Fig. 4, A and B, and fig. S4, C and D). This was confirmed by Oct4 ChIP-qPCR analysis (Fig. 4C). *K*-means cluster analysis showed that the majority clusters of Oct4 binding sites were similar between WT and L80A Oct4 (fig. S4E). Among 20,165 WT and L80A Oct4 combined peaks, 2970 peaks (cluster 5) showed higher signals in WT Oct4, while 5311 peaks (cluster 7) showed higher signals in L80A Oct4 (fig. S4, F and G). As has been observed in genome-wide localization studies for Oct4, the sox-oct motif, which consists of neighboring sox (5′-CATTGTA-3′) and oct (5′-ATGCAAAT-3′) elements, was highly enriched across Oct4 binding sites in ESCs (*10*). Notably, the predominant motif found in our computational search of the top 1000 peaks in the L80A Oct4 dataset was a perfect match to the sox-oct composite element (Fig. 4D). GO analysis of the L80A Oct4 binding sites showed similar GO terms to those of the WT Oct4 binding sites (fig. S4, H and I). The transactivation activity of L80A Oct4 on most sox-oct reporters was not affected at all in ESCs (Fig. 4E). Only in the case of the 6W reporter, an oligomerized binding site of the immunoglobulin H (IgH) enhancer, there was a slight decrease in luciferase activity for L80A Oct4. These data demonstrate that the fundamental properties

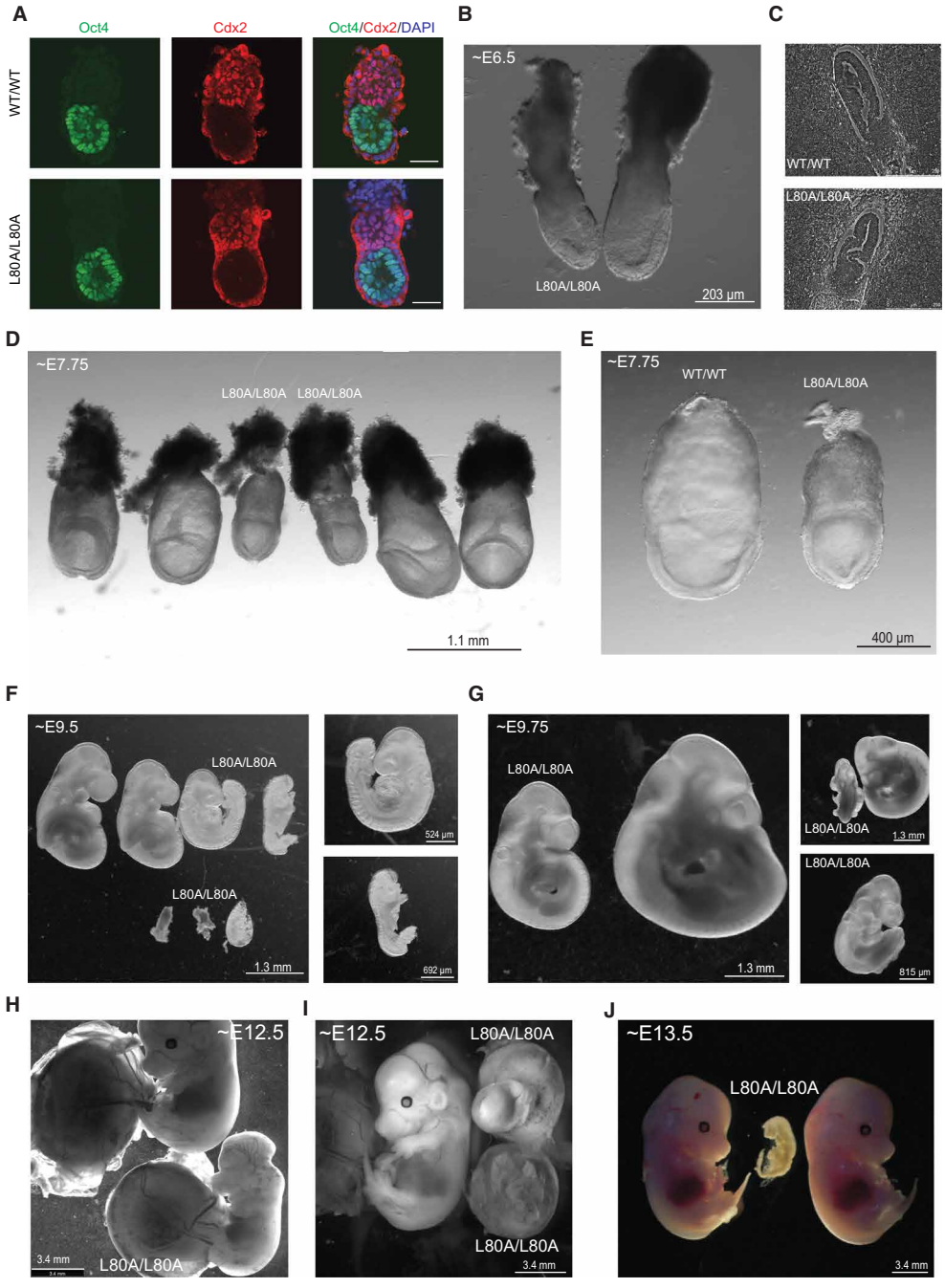

Fig. 3. Effect of L80A Oct4 on postimplantation embryo development. (**A**) Immunocytochemistry on ~E5.5 embryos using anti-Oct4 and anti-Cdx2 antibodies and confocal microscopy. Scale bars, 50 μm. (**B**) Morphology of embryos from intercross matings of WT/L80A Oct4 mice on ~E6.5. (**C**) Histological sections of ~E6.5 embryos from intercross matings of WT/L80A Oct4 mice showed growth defects in L80A/L80A Oct4 embryos. (**D** and **E**) Morphology of embryos from intercross matings of WT/L80A Oct4 mice on ~E7.75. (**F** and **G**) Morphology of embryos from intercross matings of WT/L80A Oct4 mice on ~E9.5 (F) and ~E9.75 (G). (**H** and **I**) Morphology of embryos with placentas from intercross matings of WT/L80A Oct4 mice on ~E12.5. (**J**) Morphology of embryos from intercross matings of WT/L80A Oct4 mice on ~E13.5.

of L80A Oct4 in genomic binding and transactivation are similar to those of WT Oct4 in ESCs.

L80A linker mutation interrupts the balanced Oct4 interactome

As the highly conserved N-terminal linker region is structured as an α helix and exposed on the surface of Oct4 protein, we proposed that the L80A mutation can potentially disturb this interface. In our previous study, we compared the interactome of the L80A mutant Oct4 with that of WT Oct4 by overexpressing strep-tagged WT and L80A Oct4 in WT ESCs (*14*). However, because of the dosage sensitivity of the Oct4 protein for ESC maintenance (*5*), we postulated that ESC-induced differentiation mediated by Oct4 overexpression does have an effect on Oct4 interactors. To solve this problem and

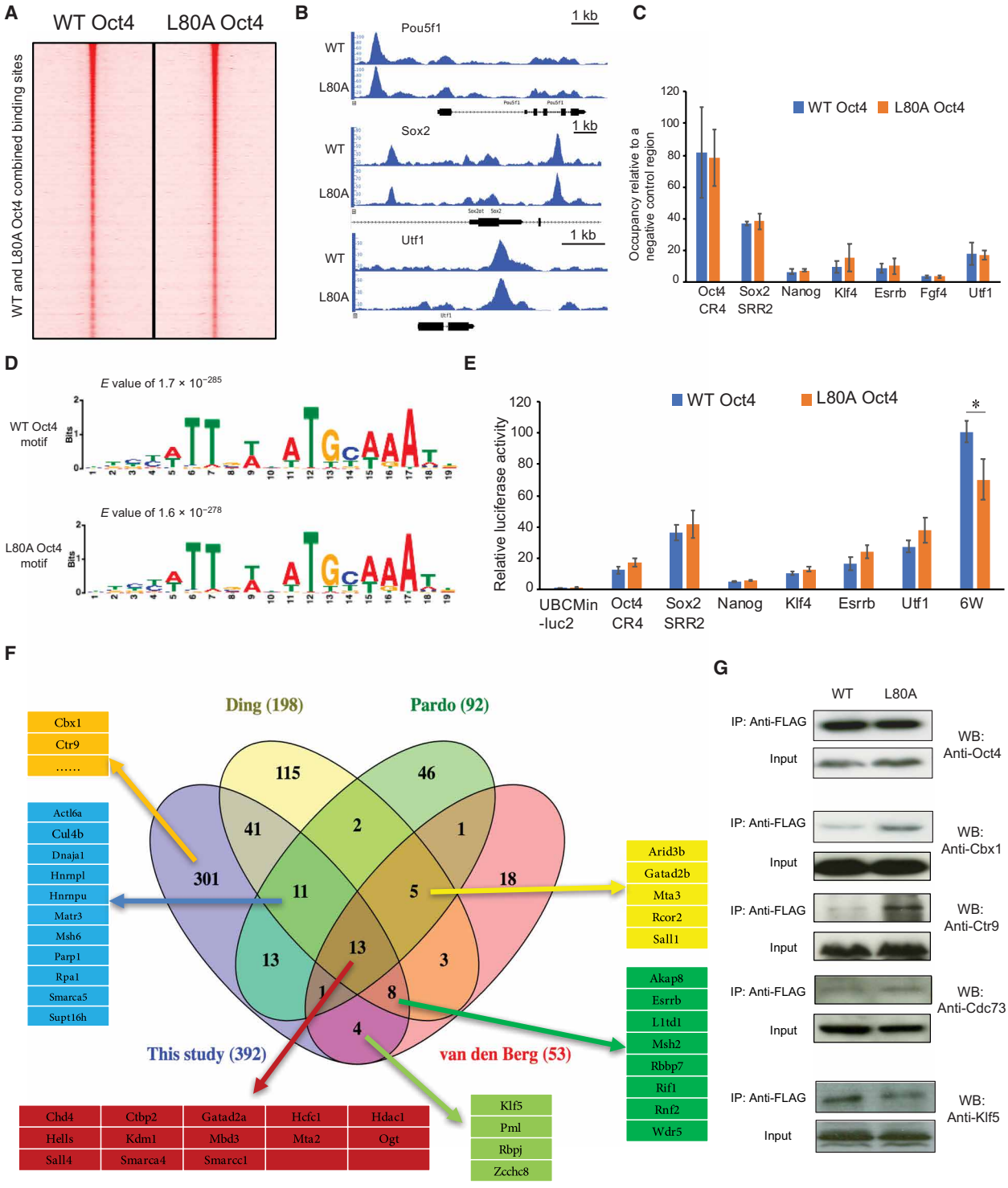

Fig. 4. Genomic binding and transactivation ability of L80A Oct4 are similar to those of WT Oct4 in ESCs, but L80A mutation interrupts the balanced Oct4 interactome. (**A**) Heatmaps of Oct4 ChIP-seq signals in WT and L80A Oct4 ESCs. (**B**) Representative Oct4 ChIP-seq tracks showing similar occupancy profiles of WT and L80A Oct4. (**C**) Oct4 ChIP-qPCR assay of WT and L80A Oct4 ESCs. Occupancy relative to a negative control region is shown. Error bars represent SD of biological triplicates. (**D**) The sox-oct motif is significantly enriched at the top 1000 WT and L80A Oct4 target sites. (**E**) The relative transcriptional activity levels of WT and L80A Oct4 were measured by luciferase reporter assay in WT/WT and L80A/L80A Oct4 ESCs. (**F**) Comparison of our Oct4 interactome with the three published Oct4 network studies. Overlapping proteins in different published datasets and novel proteins found in this study are listed in different color boxes. (**G**) Co-IP and Western blot (WB) showed decreased Oct4-Klf5 interaction and increased Oct4-Cbx1, Oct4-Ctr9, and Oct4-Cdc73 interactions in L80A Oct4 ESCs.

elucidate the Oct4 interactome underlying the physiological condition, we generated FLAG-tagged WT and FLAG-tagged L80A Oct4–rescued ZHBTc4 ESCs by using the lentivirus overexpression system. Western blots were conducted in the presence of Dox and showed that these ESCs expressed comparable levels of FLAG-WT and FLAG-L80A Oct4 in the absence of endogenous Oct4 (fig. S4J). To more acutely compare the interactome of the L80A mutant with that of the WT protein, we used the stable isotope labeling with amino acids in cell culture (SILAC) method to label FLAG-WT and FLAG-L80A Oct4 ZHBTc4 ESCs with light and heavy amino acids (fig. S4K).

Most in vivo protein-protein interactions are transient and in some cases are DNA dependent. Capturing or freezing these momentary instances of protein-DNA contact to study protein-protein interactions is meaningful for proteomics of cells. To this end, formaldehyde cross-linking of protein complexes combined with coimmunoprecipitation (Co-IP) and MS analysis is used to study the Oct4 interactome (*18*). During Co-IP, DNA was completely digested to eliminate protein interactions mediated indirectly by DNA bridging (fig. S4L). Co-IP and Western blot results showed that the positive Oct4 interactors, Oct4 and Nanog, were highly enriched in the FLAG-tagged Oct4 sample but not in the no-tag Oct4 sample (fig. S4M). With this method, we identified 392 Oct4-interacting proteins (excluding Oct4 itself) that include well-known Oct4 interactors, such as Sall4, Esrrb, Chd4, Mbd3, and Smarca4. Furthermore, we also uncovered many novel Oct4 interactors, such as Cbx1 and Ctr9 (Fig. 4F and table S1). When we compared our dataset with three other Oct4 interactome studies (*11*–*13*), we found a total of 91 previously found proteins. With our Oct4 interactome, we identified 13 of 18 (72%) overlapping Oct4-interacting proteins from the three prior studies (Fig. 4F). These results suggest that our study captures Oct4-interacting proteins with high confidence.

When we compared WT and L80A Oct4-interacting proteins, we found that most of the interacting proteins exhibited comparable intensity between WT and L80A Oct4, such as Esrrb (fig. S4N). However, five proteins (Hist1h1c, Sall4, Klf5, Rnf2, and Jade1) exhibited a significantly reduced intensity in the mutant interactome. Notably, four proteins (Cbx1, H2afy, Cdc73, and Ctr9) exhibited a significantly increased intensity in the mutant interactome (fig. S4N and table S1). To assess whether these Oct4-protein interaction changes are due to different protein expression levels in WT and mutant samples, we determined the expression levels of these nine proteins. Western blot results showed that Hist1h1c, Sall4, Rnf2, and Jade1 protein levels were decreased in the L80A mutant sample, while H2afy protein level was increased in the mutant sample (fig. S4O). Considering the critical role of Sall4 and Rnf2 in maintaining pluripotency, to completely rule out the differential interactions between Oct4 and Sall4/Rnf2, we cultured FLAG-WT and FLAG-L80A Oct4–rescued ZHBTc4 ESCs without adding Dox into the ESC medium. Under this culture condition, the original ZHBTc4 ESC transgene Oct4 was expressed together with FLAG-WT or FLAG-L80A Oct4. As the expression level of FLAG-WT or FLAG-L80A Oct4 was relatively low (fig. S4J), these ESCs did not differentiate. By using these ESCs, Western blot results showed that the protein levels of Sall4 and Rnf2 were similar between FLAG-WT and FLAG-L80A Oct4–rescued ZHBTc4 ESCs cultured without Dox. Co-IP and Western blot results showed the interactions between Oct4 and Sall4/Rnf2 were also similar in these ESCs (fig. S4P). These data suggest that the altered interactions between Oct4 and Hist1h1c, Sall4, Rnf2, Jade1, and H2afy in the mutant sample are due to the different protein expression levels. In contrast, the protein levels of Cbx1, Cdc73, Ctr9, and Klf5 were found to be comparable between the WT and L80A Oct4 samples (fig. S4O), and thus, expression levels of those proteins could not account for the observed interaction changes. The altered interactions between Oct4 and Cbx1, Ctr9, Cdc73, and Klf5 were further confirmed by Co-IP and Western blot (Fig. 4G). These data suggest that the Oct4 linker functions as a protein-protein interaction interface and that the L80A point mutation in the linker changes the Oct4 interactome to allow for more Oct4-Cbx1, Oct4-Ctr9, and Oct4-CDC73 interactions but fewer Oct4-Klf5 interactions.

Chromatin accessibility and H3K4me3 level are not affected, but H3K36me and H3K9me3 levels are increased in L80A/L80A Oct4 ESCs

Next, we investigated the effects of increased Oct4-Cbx1, Oct4-Ctr9, and Oct4-CDC73 interactions in ESCs. As Cbx1 is involved in forming heterochromatin and induces a more closed chromatin state (*19*), we assessed the chromatin state in WT/WT and L80A/L80A Oct4 ESCs derived from embryos by performing assay for transposase-accessible chromatin using sequencing (ATAC-seq). The chromatin accessibility in mutant ESCs was not affected by L80A Oct4 (Fig. 5, A to C). The position of nucleosome-free regions was the most enriched within a 5-kb region flanking the transcriptional start site, with no difference between WT and L80A Oct4 ESCs (Fig. 5D). In addition to being involved in heterochromatin formation, Cbx1 plays a role in regulating H3K9me3 (*20*). Ctr9 and Cdc73 belong to the Paf1 complex, which is important for regulating H3K4 (*21*) and H3K36 trimethylation (*22*). We next characterized the histone marks H3K4me3, H3K36me3, and H3K9me3 in WT/WT and L80A/L80A Oct4 ESCs derived from embryos (Fig. 5E). Unlike the methylation levels of H3K4me3, which were similar in WT and L80A Oct4 ESCs, the methylation levels of H3K9me3 and H3K36me3 were increased in L80A Oct4 ESCs, although the increase of H3K36me3 signals was slight (Fig. 5F and fig. S5, A to C). To confirm the H3K9me3 and H3K36me3 signal changes, we used the robust system MAnorm2 (*23*) to normalize and compare the H3K9me3 and H3K36me3 signals in WT/WT and L80A/L80A Oct4 ESCs. MA plots showed that H3K9me3 signals, but not H3K36me3 signals, were increased in L80A/L80A Oct4 ESCs (fig. S5D). However, MAnorm2 assumes that most shared peaks are not differentially bound. Thus, normalizing using MAnorm2 removes true binding differences in shared peak regions and may reduce the sensitivity of the H3K36me3 dataset. Therefore, we used another system, SICER (*24*), to normalize the data and do differential peak calling analysis. We detected 44,663 and 44,493 H3K9me3 peaks in WT/WT and L80A/L80A Oct4 ESCs, respectively. Differential peak calling analysis showed that 11,356 peaks were increased and only 3970 peaks were decreased in L80A/L80A Oct4 ESCs compared with WT/WT Oct4 ESCs. Similarly, 29,807 and 28,396 H3K36me3 peaks were detected in WT/WT and L80A/L80A Oct4 ESCs, respectively. A total of 11,614 peaks were increased and only 3466 peaks were decreased in L80A/L80A Oct4 ESCs compared with WT/WT Oct4 ESCs (fig. S5E). Western blot results showed that H3K4me3 protein level was not changed, but H3K36m3 and H3K9me3 protein levels were increased in L80A/L80A Oct4 ESCs (fig. S5, F and G). These results demonstrate that, although the increase of H3K36me3 signals is slight, H3K9me3 and H3K36me3 signals are increased in L80A/L80A Oct4 ESCs, which are consistent with higher interactions of

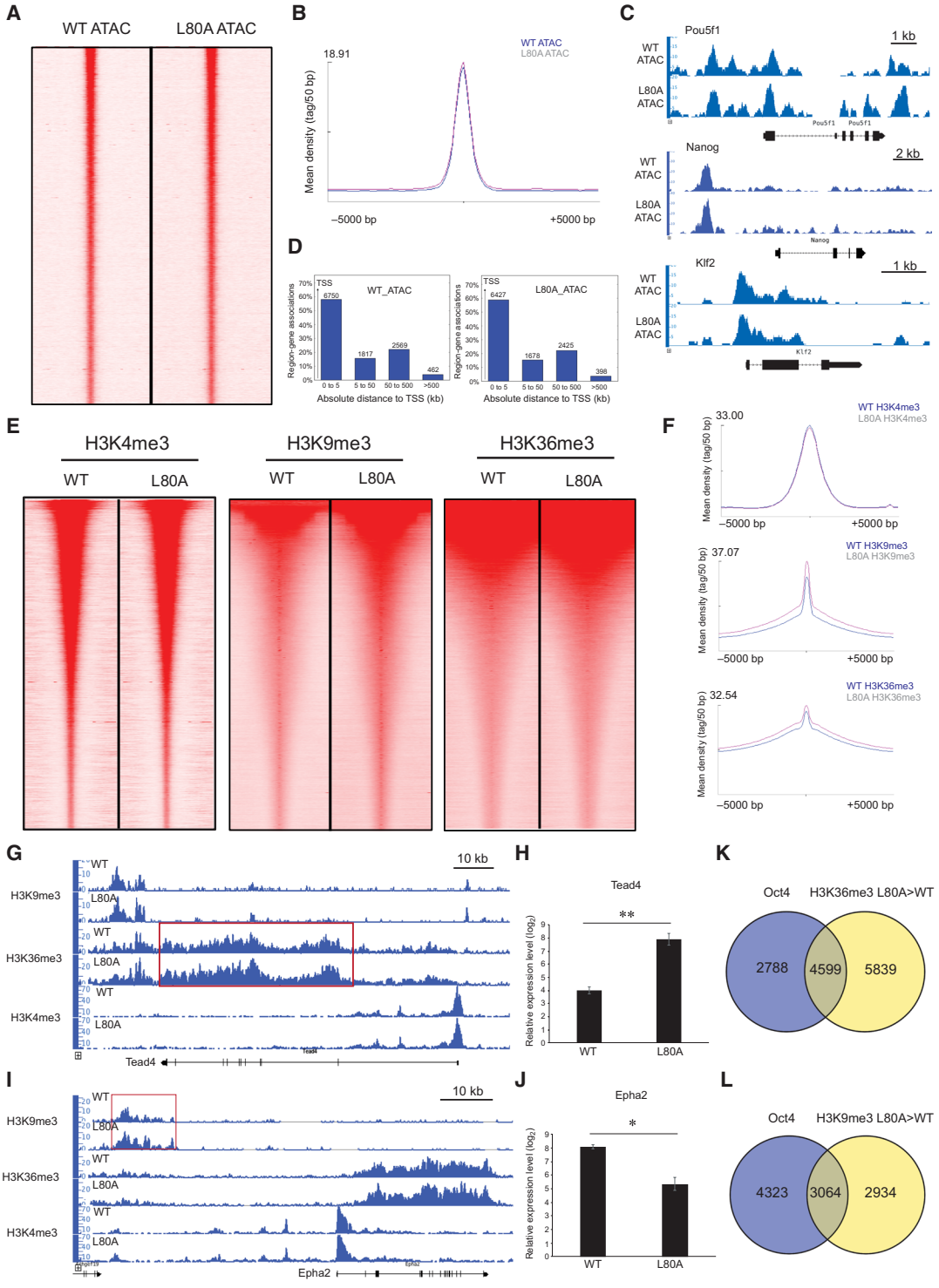

Fig. 5. Chromatin accessibility and H3K4me3 level are not affected, but H3K36me and H3K9me3 levels are increased in L80A/L80A Oct4 ESCs. (**A**) Heatmaps of ATAC-seq signals in WT and L80A Oct4 ESCs. (**B**) Average signal intensity of ATAC-seq in WT and L80A Oct4 ESCs. (**C**) Representative ATAC-seq tracks showing similar peak profile in WT and L80A Oct4 ESCs. (**D**) Genomic distribution of ATAC-seq signal in WT and L80A Oct4 ESCs. (**E**) Heatmaps of H3K4me3, H3K36me3, and H3K9me3 ChIP-seq profiles in WT and L80A Oct4 ESCs. (**F**) Average signal intensity of H3K4me3, H3K36me3, and H3K9me3 ChIP-seq in WT and L80A Oct4 ESCs. (**G**) H3K4me3, H3K36me3, and H3K9me3 ChIP-seq profiles in *Tead4* locus. Higher H3K36me3 peak signal in L80A Oct4 ESCs is shown in red box. (**H**) Relative expression level of *Tead4* in WT and L80A Oct4 ESCs from microarray data. Error bars represent SD of biological duplicates. (**I**) H3K4me3, H3K36me3, and H3K9me3 ChIP-seq profiles in *Epha2* locus. Higher H3K9me3 peak signal in L80A ESCs is shown in red box. (**J**) Relative expression level of *Epha2* in WT and L80A Oct4 ESCs from microarray data. Error bars represent SD of biological duplicates. (**K** and **L**) Venn diagram of Oct4 target genes overlapping with those genes that have higher H3K36me3 level (K) or H3K9me3 level (L) in L80A Oct4 ESCs.

L80A Oct4 and Cbx1, Ctr9, and Cdc73. We found that the altered histone markers were highly correlated with gene expression. About 40% of up-regulated genes and about 30% of down-regulated genes in L80A/L80A Oct4 ESCs were correlated with H3K36me3 and H3K9me3 alterations, respectively (fig. S5, H and I). Increased H3K36me3 levels resulted in higher expression levels of differentiation genes, such as *Tead4*, *BMP4*, *Htra1*, and *Krt8* (Fig. 5, G and H, and fig. S5, B and J), while increased H3K9me3 levels resulted in lower expression levels of self-renewal genes, such as *Epha2*, *Lin28a*, *Tcf15*, and *Sp5* (Fig. 5, I and J, and fig. S5, C and K). These results suggest that increased Oct4-Cbx1, Oct4-Ctr9, and Oct4-Cdc73 interactions elevate the H3K36me3 and H3K9me3 levels, resulting in abnormal gene expression in mutant ESCs. We found that the protein levels of H3K27me3 and H3K27ac were increased in L80A/L80A Oct4 ESCs (fig. S5, F and G), indicating that indirect epigenetic changing is induced by the unbalanced Oct4 interactome.

To address how the changed Oct4 interactome changes histone markers, we examined the H3K36me3 and H3K9me3 signals on Oct4 binding sites but did not find any overlap between Oct4 binding and histone marker signals (fig. S5, L and M). These results suggest that the histone marker changes did not occur on L80A Oct4 binding sites. However, when we examined the overlap between Oct4 target genes and genes that have higher H3K36me3 levels in L80A Oct4 ESCs, we found an overlap between more than three-fifths of Oct4 target genes and those genes that have increased H3K36me3 levels in mutant ESCs (Fig. 5K). Similarly, there was an overlap between roughly two-fifths of Oct4 target genes and those genes with increased H3K9me3 levels in L80A ESCs (Fig. 5L). Quantitative analysis of chromosome conformation capture assays (3C-qPCR) showed that the enhancer to promoter looping interactions were not affected by L80A Oct4 (fig. S5N). These data suggest that rather than by directly recruiting Cbx1, Ctr9, and Cdc73 to its own binding sites, L80A Oct4 recruits more Cbx1, Ctr9, and Cdc73 to downstream genomic regions where Cbx1, Ctr9, and Cdc73 change the histone marks in ESCs.

Overexpression of Klf5 rescues the phenotype of L80A Oct4 ZHBTc4 ESCs by rebalancing the Oct4 interactome

Interactions of L80A Oct4 and Klf5 were decreased. Therefore, we investigated whether overexpression of Klf5 could improve the pluripotency of L80A Oct4 ZHBTc4 ESCs. We coexpressed L80A Oct4 and Klf5 in ZHBTc4 ESCs by using the lentivirus overexpression system and found that the rescue efficiency was much higher than that of L80A Oct4 alone in the presence of Dox (Fig. 6A). Overexpression of Klf5 together with WT Oct4 did not significantly increase WT Oct4 rescue efficiency, suggesting that the effect of Klf5 overexpression is specific to the L80A mutation and is independent from the status of the Oct4 linker (Fig. 6A). The cell proliferation rate was also greatly improved when L80A Oct4 was coexpressed with Klf5 in ZHBTc4 ESCs (Fig. 6B). RT-qPCR results showed that the expression levels of self-renewal–related genes were increased in ZHBTc4 ESCs when L80A Oct4 and Klf5 were coexpressed (fig. S6A). These data demonstrate that overexpression of Klf5 improves the pluripotency of L80A Oct4 ZHBTc4 ESCs. Other Klf family members, Klf2 or Klf4, could not increase the L80A Oct4 rescue efficiency, indicating that the L80A linker mutation specifically affects the interaction between Oct4 and Klf5 (fig. S6B).

Next, we assessed whether overexpression of Klf5 rescues L80A Oct4 ZHBTc4 ESCs by enabling the recovery of the Oct4 interactome. Co-IP and Western blot results showed that the interaction between Oct4 and Klf5 was indeed increased when Klf5 was overexpressed in L80A ZHBTc4 ESCs (Fig. 6C). We also found that the interactions between L80A Oct4 and Cbx1, Ctr9, and Cdc73 were decreased (Fig. 6C). These data suggest that the different partners interact with Oct4 competitively and that the Oct4 interactions altered by the linker mutation can be rebalanced by controlling the expression levels of the Oct4 interactors. We also found that elevated H3K36me3 and H3K9me3 levels were decreased when Klf5 was overexpressed in L80A Oct4 ZHBTc4 ESCs (Fig. 6, D and E). Accordingly, the expression levels of genes that had more H3K36me3 in mutant ESCs were lower and the expression levels of genes that had more H3K9me3 in mutant ESCs were higher in L80A + Klf5 ZHBTc4 ESCs (fig. S6, C and D). These data suggest that rebalanced Oct4 interactions can also correct gene expression by reorganizing the epigenetic state of ESCs. Cbx1 or Cdc73 knockdown significantly increased the rescue efficiency of L80A Oct4 in ZHBTc4 ESCs (fig. S6, E and F). This observation suggests that control of the expression level of the Oct4-interacting proteins can overcome the detrimental effect of unbalanced Oct4-protein interactions caused by the linker mutation. However, due to a very high cellular death rate induced by Ctr9 knockdown, L80A Oct4 ZHBTc4 ESCs with Ctr9 knockdown could not be rescued (fig. S6, E and F).

To understand how overexpression of Klf5 rescues the phenotype of L80A Oct4 ZHBTc4 ESCs, we assessed the DNA binding of Klf5, Cbx1, Ctr9, and Cdc73 in WT, L80A Oct4–, and L80A Oct4 + Klf5–rescued ZHBTc4 ESCs. ChIP-qPCR results showed that Klf5 DNA binding was decreased, while the DNA binding of Cbx1, Ctr9, and Cdc73 was increased in L80A Oct4–rescued ZHBTc4 ESCs. Overexpression of Klf5 rescued the DNA binding of these factors (Fig. 6, F to I). These data suggest that the interaction changes between Oct4 and the cofactors affect the DNA binding of the cofactors, which can be rescued by rebalancing the interactions between Oct4 and the cofactors.

Last, we explored whether the competing and balanced Oct4 interactome exists and functions in pluripotency maintenance and differentiation in WT Oct4 background ESCs. Since we did not incorporate FLAG tag in KI targeting vector, we used constitutive lentiviral FLAG-WT Oct4–rescued ZHBTc4 ESCs as WT Oct4 background ESCs. Co-IP and Western blot results showed that the interaction between Oct4 and Klf5 was markedly decreased when Klf5 was knocked down in FLAG-WT Oct4–rescued ZHBTc4 ESCs. Meanwhile, the interactions between Oct4 and Cbx1, Ctr9, and Cdc73 were increased (Fig. 6J and fig. S6G). These data suggest that the competing and balanced Oct4 interactome also exists in WT Oct4 background ESCs. To further confirm that the balanced Oct4 interactome exists in WT Oct4 ESCs, we overexpressed Cbx1, Cdc73, and Ctr9 in FLAG-WT Oct4–rescued ZHBTc4 ESCs. Because of a very high cellular death rate induced by the overexpression of Ctr9, we focused on the overexpression of Cbx1 and Cdc73 samples. Cbx1- or Cdc73-overexpressing FLAG-WT Oct4 ZHBTc4 ESCs had much higher Oct4 protein levels (fig. S6H). We speculate that only the ESCs that have higher Oct4 levels can tolerate the overexpression of Cbx1 or Cdc73. The higher Oct4 levels in Cbx1- or Cdc73-overexpressing ESCs resulted in more interactions between Oct4 and Klf5, Cbx1, Ctr9, and Cdc73, regardless of the protein expression levels of these factors (fig. S6H). Therefore, we could not assess the balanced Oct4 interactome in Cbx1- or Cdc73-overexpressing ESCs. Similarly, the protein levels of Oct4, Klf5, Cbx1, Cdc73,

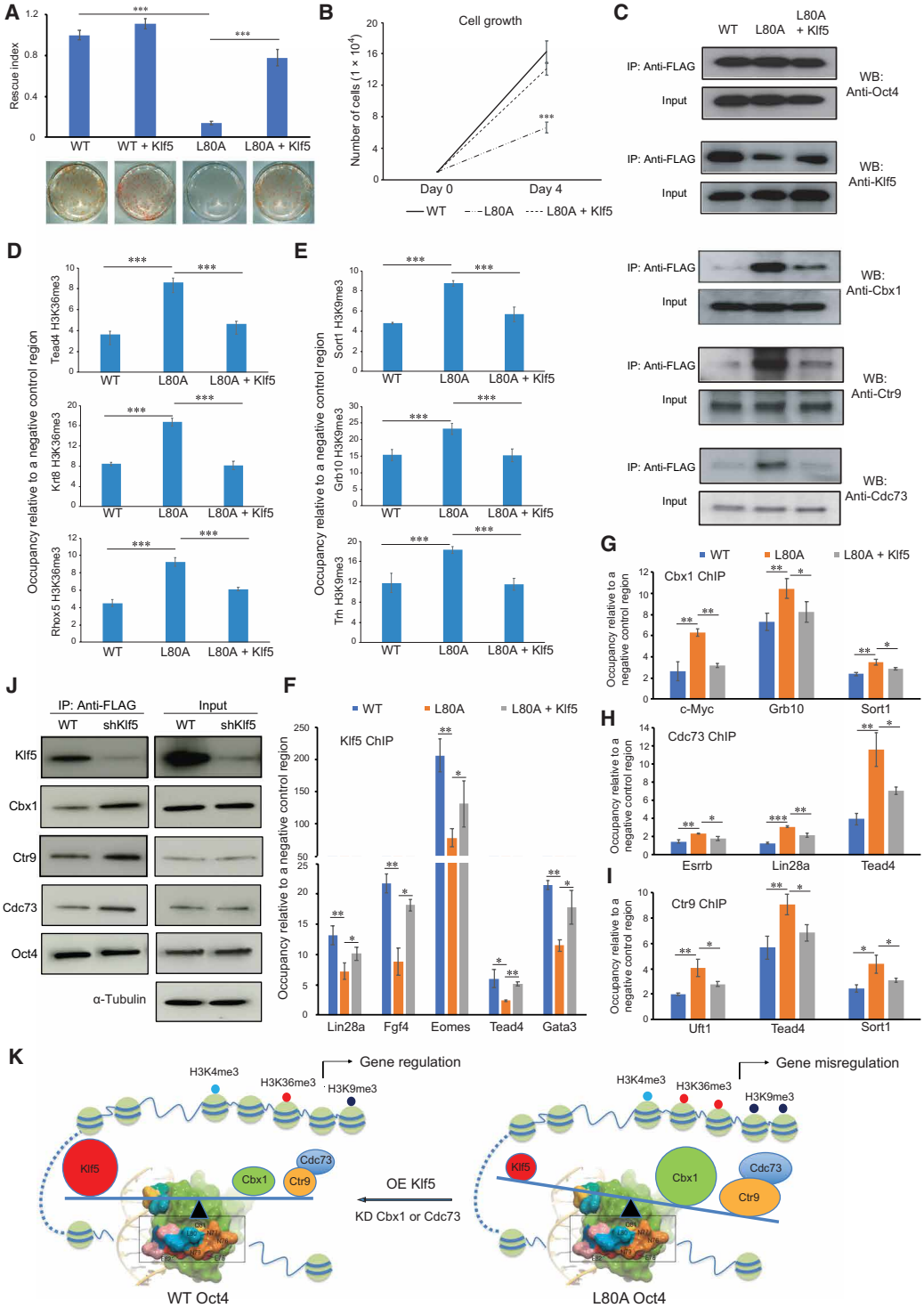

Fig. 6. Overexpression of Klf5 rescues the phenotype of L80A Oct4 ZHBTc4 ESCs by rebalancing the Oct4 interactome. (**A**) Overexpression of Klf5 increased the L80A Oct4 rescue efficiency in ZHBTc4 ESCs. Bottom: AP-positive colonies rescued by WT, L80A Oct4, and L80A Oct4 + Klf5. (**B**) Proliferation rate of WT, L80A Oct4–, and L80A Oct4 + Klf5–rescued ZHBTc4 ESCs. *** represents *P* values between L80A Oct4 and L80A Oct4 + Klf5–rescued ZHBTc4 ESCs. (**C**) Overexpression of Klf5 rebalanced L80A Oct4 interactions with Klf5, Cbx1, Ctr9, and Cdc73, as shown by Western blot after Co-IP with FLAG antibody. (**D** and **E**) ChIP-qPCR analysis of H3K36me3 (D) and H3K9me3 (E) in WT, L80A Oct4–, and L80A Oct4 + Klf5–rescued ZHBTc4 ESCs. (**F** to **I**) Klf5 (F), Cbx1 (G), Cdc73 (H), and Ctr9 (I) ChIP-qPCR assay of WT, L80A Oct4–, and L80A Oct4 + Klf5–rescued ZHBTc4 ESCs. Occupancy relative to a negative control region is shown. (**J**) Knockdown of Klf5 changed the interactions between WT Oct4 and Klf5, Cbx1, Ctr9, and Cdc73, as shown by Western blot after Co-IP with FLAG antibody. (**K**) Schematic diagram of the Oct4 linker interface, which mediates competing yet balanced Oct4-protein interactions. L80A Oct4 recruits less Klf5 but more Cbx1, Ctr9, and Cdc73 to Oct4 downstream genomic regions. The decreased Klf5 DNA binding and increased DNA binding of Cbx1, Ctr9, and Cdc73 result in dysregulated gene expression and mediate abnormal epigenetic changes at those specific loci.

and Ctr9 were decreased during FLAG-WT Oct4 ZHBTc4 ESC differentiation (fig. S6I). We also could not evaluate the balanced Oct4 interactome during ESC differentiation because of the changing protein levels. Similar to lentiviral FLAG-L80A Oct4–rescued ZHBTc4 ESCs, ESCs exhibited impaired proliferation by Klf5 knockdown or Cbx1/Cdc73 overexpression in lentiviral FLAG-WT Oct4–rescued ZHBTc4 ESCs (fig. S6J). Together, these results suggest that the competing and balanced Oct4 interactome exists in WT Oct4 background ESCs and plays a role in maintaining pluripotency.

DISCUSSION

Oct4, as a master regulator of pluripotency, collaborates primarily with other transcriptional factors or coregulators to maintain pluripotency. However, the precise mechanism by which Oct4 interacts with its partners and exerts its function remains unknown. In this study, we found that the Oct4 linker interface mediates a competing yet balanced Oct4 interactome that is crucial for the biological function of Oct4 in the maintenance of pluripotency in ESCs and embryos.

Oct4 contains a bipartite DNA binding domain, a POU-specific domain (POUS) and a POU homeodomain (POUHD) tethered by a linker region. In contrast to the unstructured linker of Oct1 and Oct6 (*25*, *26*), the highly conserved N-terminal linker region is structured as an α helix that is located on the surface of Oct4 (*14*). Here, we report that the Oct4 linker is crucial for the biological function of Oct4 in the maintenance of pluripotency in ESCs and in embryo development. Similarly, the Oct4 linker is also important for the biological function of Oct4 in reprogramming (*14*, *27*). It has recently been shown that heterochromatin loosening by the Oct4 linker region facilitates Klf4 binding and induced pluripotent stem cell (iPSC) reprogramming (*28*). Together, these data demonstrate the importance of the Oct4 linker on Oct4 biofunction.

Our in vitro EMSA data showed that L80A Oct4 alone could bind to oct motif probes, but the DNA binding of L80A Oct4 alone on sox-oct motif probes was affected. This might be due to the L80A mutation disrupting the helical conformation of the linker, thus leading to an impaired Oct4-DNA interaction. However, the decreased DNA binding of L80A Oct4 on the certain sox-oct motif could be recovered by the presence of Sox2. A study using single-molecule imaging showed that Sox2 binds to DNA first and subsequently recruits Oct4 to assemble the sox-oct heterodimer in ESCs (*29*). Oct4 is likely to be part of a multiprotein complex in ESCs, and Sox2 as well as other cofactors might further augment Oct4 genomic binding. Our ChIP-seq data showed that the L80A Oct4 bound similar genomic sites to WT Oct4 in ESCs. This was further confirmed by ChIP-qPCR. Similarly, Chen *et al.* (*28*) showed that the binding of exogenous Oct4 to target genes is not affected by the L80A mutation at the beginning of reprogramming. Our luciferase reporter assay showed that the transactivation activity of L80A Oct4 on most sox-oct reporters was not affected in ESCs. This is consistent with our previous report demonstrating that point mutations in the linker do not affect the transactivation ability of Oct4 (*14*). All these data suggest that the L80A mutation does not affect the fundamental properties of Oct4 in genomic binding and transactivation in ESCs. However, it will be of interest to determine whether the L80A mutation specifically affects the interface region that binds to interactors and/or whether it also has an indirect impact on the interactome by changing the structure of the N-terminal α helix. This indirect impact may be possible considering that the L80A mutation leads to increased binding to some proteins. For example, a structural change that disrupts the binding of certain proteins might allow for enhanced binding of other proteins, which may even be augmented by areas beyond the linker region (i.e., in the POUS or POUHD).

In previous studies, Oct4-interacting proteins were detected by the non–cross-linking Co-IP method (*11*–*13*). However, this classical Co-IP method has some drawbacks: It may not detect weak, transient protein interactions and those protein interactions that depend on DNA. Recent data have shown that interactions between many cooperative TFs are mediated by DNA (*30*). Here, we used formaldehyde cross-linking to capture protein interactions of protein bound to DNA and to stabilize the weak and transient protein interactions. Formaldehyde is the shortest available cross-linker in terms of spacer arm length, such that only proteins in the immediate vicinity can be cross-linked due to the small size of formaldehyde (*18*). By using this method, we reported an extended Oct4 interactome composed of a much larger repertoire of interacting proteins than had previously been reported. Many proteins are well-known Oct4 interactors, including TFs and multiple epigenetic regulators, such as Sall4, Esrrb, Chd4, Mbd3, and Smarca4. We also found many previously unidentified Oct4 interactors, such as Cbx1 and Ctr9. Cbx1, also called HP1β, is highly expressed in pluripotent cells and is important for maintaining pluripotency (*31*). Ctr9 belongs to the Paf1 complex (*32*) and is essential for ESC identity (*33*). Other studies, as well as our MS data, have shown that Oct4 also interacts with other Cbx family members and Paf1 complex components (*13*, *34*, *35*), suggesting that there are interactions between Oct4 and Cbx1 and Ctr9. Like in two previously published studies (*12*, *13*), in this study, we failed to uncover with high confidence the well-known Oct4 partner Sox2. The Sox2-Oct4 interaction may be too weak to be detected even with cross-linking. This would explain why the interaction between Oct4 and Sox2 was confirmed by only one previous study (*11*) but not by our study or others. This also suggests that we did not over-cross-link the samples, thus inducing indirect interactions. We conclude that our study captures most Oct4-interacting proteins with high confidence.

Our MS data showed that nine Oct4-interacting proteins were significantly affected in the L80A Oct4 interactome, although this means neither that all interacting proteins bind directly to the linker nor that directly binding proteins interact exclusively with the linker. As the linker region is relatively small, it follows that these proteins might interact with other areas of the POU domain. Western blot and Co-IP results confirmed that all these altered interactions were due to either altered protein expression levels or real interaction changes, strongly pointing to the high quality of our MS data. We found that the interaction between Oct4 and Klf5 was decreased. Klf5 can positively regulate the expression of many self-renewal–related genes and potentially repress the expression of lineage-specific genes (*36*, *37*). Compared with the Oct4-Klf5 interaction, which was decreased, Oct4-Cbx1, Oct4-Ctr9, and Oct4-Cdc73 interactions were increased, which coincided with elevated levels of H3K36me3 and H3K9me3 in mutant ESCs. Cbx1 plays a role in regulating H3K9me3 (*20*). Ctr9 and Cdc73 belong to the Paf1 complex, which is important for regulating H3K36 trimethylation (*22*, *38*). We found that the altered histone markers highly correlated with changes in gene expression. In mutant ESCs, increased H3K36me3 levels resulted in higher expression levels of differentiation genes, while increased H3K9me3 levels resulted in lower expression

levels of self-renewal genes. These results are consistent with the previous finding that Oct4 not only positively regulates pluripotency- and self-renewal–associated genes but also represses differentiation-promoting genes (*8*). We did not find any overlap between Oct4 binding and histone marker signals, suggesting that the histone marker changes did not occur on Oct4's own binding sites. 3C-qPCR results showed that the enhancer to promoter looping interactions were not affected by L80A Oct4. Furthermore, ChIP-qPCR results showed that the Klf5 DNA binding was decreased, while the DNA binding of Cbx1, Ctr9, and Cdc73 was increased in L80A Oct4 ZHBTc4 ESCs. All these data indicate that L80A Oct4 binds to enhancers and loops around to promoters, where less Klf5 but more Cbx1, Ctr9, and Cdc73 are recruited to those specific loci. The decreased Klf5 DNA binding and increased DNA binding of Cbx1, Ctr9, and Cdc73 result in dysregulated gene expression and mediate abnormal epigenetic changes at those specific loci. We found that overexpression of Klf5 rebalanced the Oct4 interactome, indicating that different partners interact with Oct4 competitively. These data suggest that the interactions between Oct4 and pluripotency TFs, such as Klf5, can prevent Oct4 from excessively interacting with epigenetic regulators that induce abnormal epigenetic modifications and consequently impaired pluripotency in ESCs.

It has been shown that the interaction between Oct4 and Brg1 is decreased by the L80A mutation (*14*, *28*). However, we found that the interaction between L80A Oct4 and Brg1 is not affected in ESCs. Two reasons might explain the difference in observations between this study and other studies. First, in our previous study, WT and L80A Oct4 were overexpressed in ESCs to determine the interactomes of WT and L80A Oct4 (*14*). However, the differentiation of ESCs induced by WT or L80A Oct4 overexpression might affect the expression level of Brg1, which subsequently affects the interaction between Brg1 and Oct4. It was shown that Brg1 expression is lower in differentiated cells than in ESCs (*39*). Second, identical WT and L80A Oct4 expression levels are required to accurately compare the interactions of these proteins with Brg1. The lower expression level of L80A Oct4 may account for the decreased Brg1–L80A Oct4 interaction observed in the Co-IP experiment (*28*).

In conclusion, our results demonstrate that the Oct4 linker interface mediates a competing yet balanced Oct4 interaction network that is crucial for the biological function of Oct4 for pluripotency maintenance in ESCs in vitro and for embryo development in vivo. Mutation of the linker disturbs the interface and interrupts the balanced Oct4 interactome, inducing abnormal epigenetic changes and gene misregulation in ESCs. Control of the expression levels of Oct4-interacting proteins can rebalance the Oct4 interactome and enable the recovery of pluripotency, which has been impaired in mutant ESCs (Fig. 6K). Thus, by showing how Oct4 interacts with its partners, we provide previously unidentified molecular insights into explaining how Oct4 contributes to pluripotency. Furthermore, we also suggest a "competing yet balanced model" that may also apply to other protein interaction networks.

MATERIALS AND METHODS

Cell culture

MEF and human embryonic kidney (HEK) 293T cells were cultured in MEF medium: high-glucose Dulbecco's modified Eagle's medium (DMEM) containing 10% fetal bovine serum (Biochrom), 1% GlutaMAX, 1% penicillin-streptomycin, 1% nonessential amino acids (NEAAs), and 0.5% β-mercaptoethanol (all from Life Technologies). ZHBTc4 ESCs were grown on gelatin-coated dishes in ESC medium: KO DMEM medium (Life Technologies) supplemented with 10% KO serum replacement, 10% fetal calf serum, 1% GlutaMAX, 1% NEAA, 1 mM sodium pyruvate (Sigma-Aldrich), 1% penicillin-streptomycin, 1% β-mercaptoethanol, and human recombinant leukemia inhibitory factor (LIF; 20 ng/ml) (purified in-house). Mouse-derived ESCs were grown on irradiated C3H MEFs or gelatin-coated dishes in ESC medium. The cells were routinely tested for mycoplasma contamination.

Vector construction and virus production

Mouse WT Oct4 and different Oct4 mutants were amplified by PCR and pasted into pLVTHM (Addgene, #12247) using Pme I and Spe I restriction sites. HEK293T cells were seeded (density of 2.2×10^6 cells per 100-mm dish) in MEF medium. The following day, MEF medium was changed to ESC medium, and cells were transfected with 3 μg of lentiviral vector, 2 μg of PAX2, and 1 μg of vesicular stomatitis virus packaging plasmid using Fugene 6 (Roche) according to the manufacturer's instructions. Cells were incubated at 37°C with 5% CO_2. The virus-containing supernatant was collected and filtered (0.45 μm; Millex-HV, Millipore) 24 and 48 hours after infection. The supernatant was supplemented with protamine sulfate (6 μg/ml; Sigma-Aldrich) before infection.

ZHBTc4 ESC rescue assay

For the Oct4 virus expression rescue assay, 1×10^4 ZHBTc4 ESCs were seeded on gelatin-coated six-well plates in ESC medium. ZHBTc4 ESCs were infected with virus 6 hours later. Dox was added into the medium 18 hours after infection. After 96 hours of Dox treatment, the cells were either fixed with 4% paraformaldehyde (PFA), permeabilized, and then subjected to naphthol and Fast Red substrates for alkaline phosphatase (AP) staining or passaged and cultured in ESC medium with Dox to generate stable rescued cell lines. For the Oct4 KI rescue assay, ZHBTc4 ESCs were targeted with WT or L80A Oct4 constructs containing Oct4 genomic sequences and the *FRT-SA-IRES-Hyg-pA* cassette in the first intron of Oct4. Then, the targeted cells were transfected with an FLPe expression vector to remove the FRT cassette. The morphology of the WT and L80A Oct4–KI ZHBTc4 ESCs was assessed after 12 days of Dox treatment. To check the proliferation rate of WT Oct4, L80A Oct4, FLAG-WT Oct4, FLAG-L80A Oct4, FLAG-WT Oct4 + shKlf5, and FLAG-WT Oct4 + OE Cbx1/Cdc73–rescued ZHBTc4 ESCs and ZHBTc4 ESCs cultured without or with Dox, 10^4 cells were plated on gelatin-coated dishes, and the total cell number of cells was determined after 4 days in culture.

Knockdown and overexpression of Klf5, Cbx1, Ctr9, and Cdc73

Knockdown of Klf5, Cbx1, Ctr9, and Cdc73 was performed by using the lentiviral short hairpin RNA (shRNA) system. Lentiviruses containing different shRNAs were produced as mentioned above and used to infect ESCs. shRNAs used in this study are listed in table S2. Overexpression of Klf5 was performed by using the lentiviral overexpression system. Overexpression of Cbx1, Ctr9, and Cdc73 was performed by using the CAG promoter overexpression system. Constitutive lentiviral FLAG-WT Oct4–rescued ZHBTc4 was stably transfected with a CAG promoter–driven Cbx1/Ctr9/Cdc73-IRES-puromycin coding plasmid for constitutive overexpression. After selection with puromycin, FLAG-WT Oct4–rescued ZHBTc4-overexpressing Cbx1/Cdc73 was established.

RT-qPCR analysis

Total RNA was extracted using the RNeasy Mini Kit (QIAGEN). Complementary DNA (cDNA) synthesis was performed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Transcript levels were determined using iTaq SYBR Green Supermix with ROX (Bio-Rad). Relative gene expression levels were calculated by the $2^{-\Delta\Delta Ct}$ method, normalized to an endogenous control gene *GAPDH*, and presented as fold change over control samples. Primer sequences are listed in table S2.

Immunofluorescence of cells

The cells were fixed with 4% PFA (Sigma-Aldrich) for 15 min, incubated with 0.1% Triton X-100/phosphate-buffered saline (PBS) for 15 min, and blocked in 5% bovine serum albumin (BSA)/PBS for 1 hour. The cells were then incubated with the following primary antibodies overnight at 4°C: rat monoclonal anti-Nanog (1:1000; Thermo Fisher Scientific, 14-5761-80), goat anti-Oct4 (1:300; Santa Cruz Biotechnology, sc-8628), mouse monoclonal anti-Oct4 (1:1000; Santa Cruz Biotechnology, sc-5279), rabbit anti-Oct4 (1:1000; Cell Signaling Technology, 83932S), goat anti-Sox2 (1:1000; Santa Cruz Biotechnology, sc-17320), mouse anti-Cdx2 (1:300; Biogenex, MU392A-5UC), rat anti-Troma1 (1:200; home-made), rabbit anti-RFP (1:300; Biomol, 600-401-379), goat anti-Sox17 (1:300; R&D Systems, AF1924), rat anti-Nestin (1:1000; Sigma-Aldrich, MAB353C3), rabbit anti-Eomes (1:500; Abcam, ab23345), anti-Gata3 (1:1000; Cell Signaling Technology, #5852), and anti-Gata6 (1:300; Thermo Fisher Scientific, #PA1-104). The cells were then washed three times with PBS and incubated with the appropriate fluorescently labeled Alexa Fluor secondary antibodies (1:1000; Invitrogen) for 1 hour. The cells were then washed three times with PBS, incubated with 4′,6-diamidino-2-phenylindole (0.5 μg/ml; Molecular Probes) for 10 min, and washed once with PBS. Images were acquired using the Leica DMI6000B inverted fluorescence microscope equipped with the Orca-R2 charge-coupled device camera (Hamamatsu) and analyzed by the Leica application suite advanced fluorescence software.

Generation of L80A Oct4 mice by tetraploid embryo complementation

An L80A mutant allele of Oct4 was generated via homologous recombination in male ESCs derived from Acr–enhanced green fluorescent protein transgenic mice. By using the tetraploid complementation method, we generated L80A Oct4–KI heterozygous mice in one step. Briefly, two-cell embryos were flushed 20 hours post–human chorionic gonadotropin from the oviducts of B6C3F1 mice and fused with a peak pulse of 50 V for 35 μs in 0.3 M mannitol to generate tetraploid embryos. The tetraploid embryos were cultured overnight in $KSOM^{AA}$. Then, clumps of 15 to 20 trypsin-treated ESCs were transferred into individual depressions in drops of $KSOM^{AA}$ under mineral oil. Meanwhile, batches of 30 to 50 embryos were briefly incubated in acidified Tyrode's solution to remove the zonae pellucidae. Two embryos were placed on each ESC clump and cultured in an incubator with 5% CO_2 in air at 37°C. After 24 hours of culture, 10 to 12 embryos at the blastocyst stage were transferred into one uterine horn of a pseudopregnant recipient at 2.5 days post coitum (dpc) for full-term development. Tail biopsies of mice were performed for genotyping. Primer sequences are listed in table S2.

All mice used were bred and housed at the mouse facility of Max Planck Institute (MPI) in Muenster, and animal handling was in accordance with MPI animal protection guidelines. A protocol for animal handling and maintenance for this study was approved by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen under the supervision of a certified veterinarian in charge of the MPI animal facility.

Early developmental embryo collection and immunocytochemistry

Embryos were flushed out of the uteri of mice at 3.5 or 4.5 dpc in M2 medium to collect embryos at different stages of development as required for specific experiments. Briefly, embryos were fixed in 4% PFA (20 min), permeabilized with 0.1% Triton X-100 (Sigma-Aldrich) for 1 hour, and blocked with 3% BSA in PBS for 1 hour. This was followed with binding of primary antibodies (overnight at 4°C) and secondary antibodies at a 1:1000 dilution for 1 hour at room temperature. The following primary antibodies were used: goat anti-Oct4 (1:100; Santa Cruz Biotechnology, SC-8628), mouse anti-Oct4 (1:200, Santa Cruz Biotechnology, sc-5279), goat anti-Gata6 (1:200; R&D Systems, AF1700), and mouse anti-Cdx2 (1:300; Biogenex, MU392A-5UC). Samples were examined under a laser scanning confocal microscope (UltraVIEW, PerkinElmer Life Sciences) with 488- and 568-nm lasers.

Derivation of ESC lines

After removal of the zonae pellucidae, as described above, E4.5 embryos were cultured on irradiated C3H MEFs in ESC culture medium to establish ESC lines. All ESC lines newly derived in the present study tested negative for mycoplasma contamination. The ESCs were cultured on gelatin-coated dishes without feeder cells in ESC culture medium for several passages to establish feeder-free ESC lines. Feeder-free ESCs were used for further experiments.

TSC differentiation

ESCs were maintained on mitotically inactivated DR-4 feeders in ESC medium at 37°C and 5% CO_2 atmosphere in air. WT/WT and L80A/L80A Oct4 ESCs were stably transfected with a CAG promoter–driven Cdx2-ERT2-RFP coding plasmid for constitutive overexpression. After fluorescence-activated cell sorting–mediated selection of RFP-positive cells, WT/WT or L80A/L80A Oct4 ESCs overexpressing Cdx2-ERT2-RFP were established. To induce spontaneous TSC differentiation, the WT and mutant Oct4 ESCs overexpressing Cdx2-ERT2-RFP were cultured on MEFs in MEF medium supplemented with human fibroblast growth factor 4 (25 ng/ml; Peprotech, #100-31) and heparin (1 μg/ml; Sigma-Aldrich, #H3393) (TSC medium) for 3 weeks. To convert ESCs into TSCs by overexpressing Cdx2, conversion was induced by switching ESC medium to TSC medium and adding 0.5 μM 4-hydroxytamoxifen (Sigma-Aldrich, #H7904-5MG) for 3 weeks, which forced fusion construct nuclear translocation and subsequent TSC transcriptional program activation. Cells were split every 3 days using 0.25% trypsin-EDTA.

Histological sections

Deciduas from L80A Oct4 heterozygous matings were dissected between E6.5 and 8.5 dpc and were fixed overnight with 4% PFA in PBS at 4°C. Fixed deciduas were then dehydrated through a graded series of ethanol, incubated twice in xylene, and embedded in paraffin. Paraffin-embedded deciduas were then sectioned at a thickness of 10 μm.

Microarray and data analysis

Following procedures previously described in (*40*), RNA samples to be analyzed by microarrays were prepared using QIAGEN RNeasy columns with on-column DNA digestion. Total RNA (300 ng) per sample was used as input into a linear amplification protocol (Ambion) that involved synthesis of T7-linked double-stranded cDNA and 12 hours of in vitro transcription incorporating biotin-labeled nucleotides. Purified and labeled complementary RNA was then hybridized for 18 hours onto MouseRef-8 v2 expression BeadChips (Illumina) following the manufacturer's instructions. After washing as recommended, chips were stained with streptavidin-Cy3 (GE Healthcare) and scanned using the iScan reader (Illumina) and accompanying software. Samples were exclusively hybridized as biological replicates. The bead intensities were mapped to gene information using BeadStudio 3.2 (Illumina). Background correction was performed using the Affymetrix Robust Multiarray Analysis background correction model. Variance stabilization was performed using the $\log_2$ scaling, and gene expression normalization was calculated with the method implemented in the lumi package of R-Bioconductor. Data postprocessing and graphics were performed with in-house developed functions in MATLAB. Hierarchical clustering of genes and samples was performed with one minus correlation metric and the unweighted average distance (UPGMA) (also known as group average) linkage method. The molecular signatures were taken from the gene set collection C5 of version 3.0 of the Molecular Signatures Database. The significance of the gene set of the different expressed genes was analyzed using an enrichment approach based on the hypergeometric distribution. The significance (*P* value) of the gene set enrichment was calculated using the hypergeometric distribution. The multitest effect influence was corrected through controlling the false discovery rate (FDR) using the Benjamini-Hochberg correction at a significance level of $\alpha = 0.05$. In pairwise scatterplot figure, black lines indicate boundaries of twofold difference in gene expression levels. The bar to the right indicates the scattering density; the higher the scattering density, the darker the blue color. Gene expression levels are depicted in $\log_2$ scale.

Electrophoretic mobility shift assay

HEK293T cells were transformed with pLV-Oct4 or Sox2 for 3 days. PBS-washed cell pellets were resuspended in lysis buffer [20 mM Hepes-KOH (pH 7.8), 150 mM NaCl, 0.2 mM EDTA, 25% glycerol, 1 mM dithiothreitol (DTT), and cOmplete protease inhibitor cocktail] and lysed by freeze/thaw. Protein expression from the different constructs was analyzed by Western blot and adjusted in the EMSA for equal expression. Two to three micrograms of WT or mutant protein lysates were incubated with 1.6 pmol of a Cy5-labeled double-stranded DNA probe in the presence or absence of lysates containing Sox2. Binding reactions were incubated at 37°C for 1 hour in binding reaction buffer [25 mM Hepes (pH 8.0), 0.5 mM EDTA, 7.7 mM DTT, 10% glycerol, 50 mM NaCl, BSA (4 mg/ml), 0.07% Triton X-100, and 1 μg of poly(dI*dC)] and then loaded directly onto 6% native polyacrylamide gels and run at 10 mA (200 to 300 V) for 2.5 hours. The probe sequences are listed in table S2.

ChIP followed by qPCR or sequencing

Following procedures previously described in (*41*), for ChIP experiments, WT/WT or L80A/L80A Oct4 ESCs 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, and 0.25% Triton X-100] containing protease inhibitor cocktail (Sigma-Aldrich) 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, and 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, and 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 to 300 base pairs (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, and 1.25% Triton X-100] overnight at 4°C. The amounts of chromatin, beads, and antibodies for ChIP assay are as follows: chromatin corresponding to 50 μg of DNA + 25 μl of beads + 2 μg of primary antibody. The following primary antibodies were used: goat anti-Oct4 (Santa Cruz Biotechnology, sc-8628), rabbit anti-H3K4me3 (Abcam, ab8580), rabbit anti-H3K36me3 (Abcam, ab9050), rabbit anti-H3K9me3 (Abcam, ab8898), rabbit anti-Cbx1 (Abcam, ab10478), rabbit anti-Ctr9 (Bio-Techne, NB100-68205), rabbit anti-Cdc73 (Parafibromin) (Bethyl Laboratories, A300-170A-M), and rabbit anti-Klf5 (Sigma-Aldrich, 09-822). Beads were washed once with low-salt buffer [20 mM tris-HCl (pH 8.0), 150 mM NaCl, 2 mM EDTA, 0.1% SDS, and 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, and 1% Triton X-100], twice with radioimmunoprecipitation assay buffer [50 mM Hepes-KOH (pH 7.6), 250 mM LiCl, 1 mM EDTA, 1% IGEPAL CA630, and 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, and 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-qPCR experiments, reverse cross-linked and purified DNA was analyzed by real-time PCR. The fold enrichment was calculated by ddCt method and normalized to the values obtained at a negative control region within the intergenic spacer of ribosomal DNA, which gave reliable amplification due to its multiple copies in the genome. Primer sequences are listed in table S2.

Sequencing libraries were prepared according to the instruction of the KAPA Hyper Prep Kit (Roche). Amplified libraries were purified with 1.0 volume of AMPure XP beads and size-selected on a 2% agarose gel (250 to 450 bp). Sequencing was performed on the Illumina NextSeq 500 (Illumina) as single-end 75-bp reads. The sequencing data were uploaded to the Galaxy web platform, and we used the public server at usegalaxy.org to analyze the data (*42*). Sequence reads were aligned to the mm10 mouse reference genome using Map with Bowtie for Illumina with default parameters except for suppressing all alignments for a read if more than one reportable alignment exist. Peak calling was performed using MACS with the following parameters: Do not build the shifting model and shift size 100. Input DNA from each sample was used as a control. The Wig files generated before were converted into bigWig files for visualization. Motif analysis was performed using the MEME-ChIP. Calling differential ChIP-seq signals between samples was performed by using MAnorm2 (*23*) or SICER (*24*) with default parameters.

Heatmaps and *k*-means cluster analysis were generated using SEQMINER (*43*) on combined peaks in WT and L80A samples. GO analysis and annotations of the nearby genes associated with peaks were performed using GREAT (*44*). Peak overlapping analysis was performed using ChIPpeakAnno (*45*). Because of the potential for overlap between one peak from one dataset and two peaks from another dataset, the total peak number may slightly vary.

ATAC-seq

Briefly, WT/WT or L80A/L80A Oct4 ESCs were harvested and washed once with 1 ml of cold PBS buffer (5 min, 500*g*, 4°C). Cells were then resuspended in cold PBS buffer, and the number of cells was counted. An appropriate volume of cell suspension (with 50,000 cells) was transferred into a 1.5-ml tube and centrifuged for 5 min at 500*g* and 4°C. The cell pellet was resuspended in 50 μl of 1× cold lysis buffer (0.15% IGEPAL CA630 + 0.1% Tween 20) and lysed on ice for 10 min. Immediately after lysis, nuclei were spun at 500*g* for 5 min to remove the supernatant. Nuclei were then resuspended in 50 μl of transposition reaction mixture by pipetting up and down and then incubated at 37°C for 30 min in a water bath. The transposition reaction mixture contained 25 μl of TD (2× reaction buffer from a Nextera kit; Illumina, catalog no. FC-121-1030), 2.5 μl of TDE1 (Nextera Tn5 Transposase from a Nextera kit; Illumina, catalog no. FC-121-1030), and 22.5 μl of nuclease-free H_2O. Immediately following transposition, the reaction mixture was purified using the QIAGEN MinElute PCR Purification Kit. Transposed DNA was placed in 20 μl of elution buffer [10 mM tris-HCl (pH 8.0)]. PCR was performed to amplify the library for 14 cycles using the following PCR conditions: 72°C for 5 min; 98°C for 30 s; and thermocycling at 98°C for 10 s, 63°C for 30 s, and 72°C for 1 min; followed by 72°C for 5 min. After the PCR reaction, libraries were purified with the 1.2× AMPure (Beckman) beads two times. Sequencing was performed on the Illumina NextSeq 500 (Illumina) as single-end 75-bp reads. The sequencing data were uploaded to the Galaxy web platform, and we used the public server at usegalaxy.org to analyze the data (*42*). Sequence reads were aligned to the mm10 mouse reference genome using Map with Bowtie for Illumina with default parameters except for suppressing all alignments for a read if more than one reportable alignment exist. Peak calling was performed using MACS with the following parameters: Do not build the shifting model and shift size 100. The Wig files generated before were converted into bigWig files for visualization. Heatmaps and mean density of signals were generated using SEQMINER on combined peaks in WT and L80A samples (*43*). Region-gene associations were performed using GREAT (*44*).

Luciferase assay

A 6W enhancer containing six copies of oligonucleotides with octamer-binding motif from the mouse Ig heavy-chain gene enhancer and genomic regions amplified by PCR were cloned into a construct containing the human UBC minimal promoter and firefly luciferase (luc2). WT and L80A Oct4 homozygous ESCs derived from embryos were transfected with pGL4.75[hRluc/CMV], together with the reporter constructs. Twenty-four hours after transfection, luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega). The relative luciferase activity (luc2/hRluc) was further normalized to an empty vector control. Each experiment was performed in biological triplicate. Primer sequences used for amplification are listed in table S2.

SILAC labeling

FLAG-WT and FLAG-L80A Oct4 ZHBTc4 ESCs were cultured in SILAC DMEM medium without lysine and arginine (with glutamine) (Thermo Fisher Scientific, 88364), containing 10% dialyzed serum (Sigma-Aldrich, F0392), light lysine and arginine or heavy lysine and arginine (13C and 15N) (Silantes), 1% NEAA, 1 mM sodium pyruvate (Sigma-Aldrich), 1% penicillin-streptomycin, 1% β-mercaptoethanol, and human recombinant LIF (20 ng/ml) (purified in-house) for six passages. Then, the labeling efficiency was determined by MS.

Western blot

The protein extracts were separated by 12% SDS–polyacrylamide gel electrophoresis and transferred onto nitrocellulose membranes (GE Healthcare). The membranes were blocked in 5% skim milk in 0.1% Tween 20/PBS (PBST) for 1 hour and incubated with the following primary antibodies overnight at 4°C: mouse monoclonal anti–α-tubulin (Sigma-Aldrich, T6199), mouse monoclonal anti-FLAG (Sigma-Aldrich, F1804), rat monoclonal anti-Nanog (Thermo Fisher Scientific, 14-5761-80), mouse monoclonal anti-Oct4 (Santa Cruz Biotechnology, sc-5279), rabbit polyclonal anti-Cbx1 (Abcam, ab10478), rabbit polyclonal anti-Ctr9 (Bio-Techne, NB100-68205), rabbit polyclonal anti-Cdc73 (Parafibromin) (Bethyl Laboratories, A300-170A-M), rabbit polyclonal anti-Klf5 (Abcam, ab137676), rabbit monoclonal anti-histone H1.2 (Abcam, ab181973), rabbit polyclonal anti-Sall4 (Abcam, ab29112), rabbit polyclonal anti-Jade1 (Bethyl Laboratories, A302-814A-M), rabbit monoclonal anti-Rnf2 (RING1B) (Cell Signaling Technology, #5694), and rabbit polyclonal anti-H2afy (macroH2A.1) (Abcam, ab37264). The membranes were then washed three times with PBST and incubated with the following horseradish peroxidase (HRP)–conjugated secondary antibodies for 1 hour: donkey polyclonal anti-rabbit IgG HRP (GE Healthcare, NA934), goat polyclonal anti-mouse IgG + IgM (H + L) HRP (Dianova, 115-035-044), and goat polyclonal anti-rat IgG HRP (Santa Cruz Biotechnology, sc-2032). The membranes were washed three times with PBST, incubated with enhanced chemiluminescence solution (GE Healthcare), and exposed to x-ray films (GE Healthcare). Western blots were quantified by using ImageJ software.

Cross-linking Co-IP

Cells 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, and 0.25% Triton X-100] containing protease inhibitor cocktail (Sigma-Aldrich, P8340) 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, and 0.5 mM EGTA]. The nuclear pellet was resuspended in IP buffer [50 mM tris (pH 7.4), 150 mM NaCl, 1 mM EDTA, 20% glycerol, 0.1% NP-40, and 1.5 mM $MgCl_2$] containing protease inhibitor cocktail, followed by mild sonication using Diagenode Bioruptor. Insoluble proteins were separated by spinning for 30 min at 13,000 rpm at 4°C. The extract was incubated with anti-FLAG M2 agarose beads (Sigma-Aldrich, A2220) and benzonase (Merck, 71205-3) overnight on a rotating wheel at 4°C. The beads were washed in IP buffer for 5 min at 4°C with rotating. The wash step was repeated five times. The beads were eluted by boiling in SDS sample buffer for 5 min.

Mass spectrometry

Proteins eluted from the beads were first acetone-precipitated to remove detergent and salts and were then subjected to an in-solution digest. Briefly, following solubilization of the precipitate in 50 µl of UB [6 M urea, 2 M thiourea, and 20 mM Hepes (pH 7.5)], proteins were first reduced by addition of DTT (10 mM final concentration) for 1 hour at 37°C, alkylated with iodoacetamide (55 mM final concentration) at room temperature in the dark for 30 min, and then predigested using LysC (1 µg per sample) for 2 to 3 hours at 37°C. Following dilution of the digest with 250 µl of 50 mM ammonium bicarbonate buffer, trypsin was added (0.5 µg per sample), and the digestion was allowed to complete by overnight incubation of the protein mixture at 37°C. Digests were then acidified by addition of trifluoroacetic acid and subjected to stage tipping to remove salts and contaminants. Tryptic digests were stored on-tip at 4°C until further analyzed by MS.

All MS measurements were performed on a Q Exactive HF mass spectrometer (Thermo Fisher Scientific) online coupled to an EASY nLC 1200 nano-HPLC via a Nanospray Flex ion source and a 30-cm-long in-house packed reversed-phase fused silica capillary column with integrated emitter tip (nanoseparations; 360 µm, outer diameter, 75µm, inside diameter; Reprosil-Pur 120 C18-AQ, 1.9 µm by Dr. Maisch). Samples were dissolved in 0.1% formic acid (solvent A) and separated using a multilinear gradient running from 2 to 40% solvent B (0.1% formic acid and 80% acetonitrile) in 220 min, from 40 to 60% B in 20 min, followed by a washing step at 98% solvent B and reequilibration at starting conditions. MS was performed in data-dependent mode [spray voltage, 2.1 kV; column temperature maintained at 45°C using a PRSO-V1 column oven (Sonation, Biberach, Germany)]. MS1 scan resolution was set to 60,000 at mass/charge ratio (m/z) of 200 and the mass range to m/z of 300 to 1750. Advanced gain control (AGC) target value was 3×10^6 with a maximum fill time of 100 ms. Fragmentation of peptides was achieved by higher-energy collisional dissociation using a top17 method (MS2 scan resolution of 15,000 at 200 m/z; AGC target value of 1×10^5; maximum fill time of 50 ms; isolation width of 1.6 m/z; normalized collision energy of 27). Dynamic exclusion of previously identified peptides was allowed and set to 20 s; singly charged compounds and peptides assigned with a charge of 8 or more were excluded from the analysis. Data were recorded using Xcalibur software (Thermo Fisher Scientific).

Raw MS files were processed using the MaxQuant computational platform (version 1.5.3.8). Identification of peptides and proteins was enabled by the built-in Andromeda search engine by querying the concatenated forward and reverse mouse UniProt database (UP000000589_10090.fasta; version from December 2015) including common laboratory contaminants. Allowed initial mass deviation was set to 20 parts per million and 0.5 Da, respectively, in the search for precursor and fragment ions. Trypsin with full enzyme specificity and peptides with a minimum length of seven amino acids only were selected. A maximum of two missed cleavages were allowed; the "match between runs" option was turned on. Arg^{10} and Lys^8 were defined as the heavy SILAC amino acids, carbamidomethylation (Cys) was set as fixed modification, while oxidation (Met) and N-acetylation at the protein N terminus were defined as variable modifications. For peptide and protein identifications, a minimum FDR of 1% was required. All MS proteomic data were deposited to the ProteomeXchange Consortium via the PRIDE (*46*) partner repository with the dataset identifier PXD017456.

The list of identified proteins was first filtered, and potential contaminants, reverse hits derived from the target-decoy search, and proteins that were identified by a single modified peptide only were removed. To determine which of the proteins were enriched in the Co-IP experiments, significance B was calculated in Perseus (version 1.5.6.0) on $\log_2$-transformed H/L ratios using a Benjamini-Hochberg FDR set to 0.05 for truncation. As the Co-IP Oct4 levels in WT-heavy and L80A-light samples were significantly different, we used only WT-light and L80A-heavy Co-IP sample data for further candidate searching and analysis. We removed proteins with documented membrane, cytoplasmic, ribosomal, or mitochondrial localization. Those proteins identified in both duplicates were considered as Oct4-interacting proteins. Only those proteins determined to be significant outliers in both duplicates were considered for further candidate evaluation.

Chromosome conformation capture qPCR

WT/WT or L80A/L80A Oct4 ESCs (~2 to 3×10^7) 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 [10 mM tris-HCl (pH 7.5), 10 mM NaCl, 5 mM $MgCl_2$, and 0.1 mM EGTA] containing protease inhibitor cocktail (Sigma-Aldrich) for 10 min on ice and centrifuged for 5 min at 400*g* at 4°C. The pelleted nuclei were resuspended in 0.5 ml of 1.2× restriction enzyme buffer containing 0.3% SDS and incubated for 1 hour at 37°C while shaking at 900 rpm. Triton X-100 was added to a final concentration of 2%, and the nuclei were further incubated for 1 hour at 37°C while shaking at 900 rpm. The cross-linked chromatin was digested overnight at 37°C with 400 U of Hind III enzyme while shaking at 900 rpm. The restriction enzyme was heat-inactivated (25 min at 65°C). After addition of 6.5 ml of 1.15× ligation buffer containing 1% Triton X-100 and 100 U of T4 ligase, the chromatin was ligated for 4 hours at 16°C, followed by 30 min at room temperature. Proteinase K was added, and samples were incubated at 65°C overnight to decross-link the sample. The following day, samples were incubated with ribonuclease at 37°C for 30 to 45 min, and the DNA was purified by phenol-chloroform extraction and ethanol precipitation using glycogen as a carrier. qPCR was performed with PrimeTime PCR Master Mix and PrimeTime qPCR probe assays (Integrated DNA Technologies) on an Applied Biosystens Quantstudio 3 real-time PCR system, using the following cycling conditions: 94°C for 15 min and 44 cycles of 15 s at 94°C and 90 s at 60°C. Relative cross-linking frequency was calculated by the $2^{-\Delta\Delta Ct}$ method and normalized to β-globin control. Probe and primer sequences are listed in table S2.

Statistics

Throughout the paper, P values were calculated with analysis of variance (ANOVA): single factor; not significant, $P > 0.05$; $^*P < 0.05$; $^{**}P < 0.01$; and $^{***}P < 0.001$. All data are presented as means ± SD of biological replicates. Error bars represent SD of biological triplicates, unless otherwise stated.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at https://science.org/doi/10.1126/sciadv.abe4375

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. M. Li, J. C. Belmonte, Ground rules of the pluripotency gene regulatory network. *Nat. Rev. Genet.* **18**, 180–191 (2017).

2. H. R. Schöler, R. Balling, A. K. Hatzopoulos, N. Suzuki, P. Gruss, Octamer binding proteins confer transcriptional activity in early mouse embryogenesis. *EMBO J.* **8**, 2551–2557 (1989).
3. M. H. Rosner, M. A. Vigano, K. Ozato, P. M. Timmons, F. Poirier, P. W. Rigby, L. M. Staudt, A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. *Nature* **345**, 686–692 (1990).
4. J. Nichols, B. Zevnik, K. Anastassiadis, H. Niwa, D. Klewe-Nebenius, I. Chambers, H. Schöler, A. Smith, Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95**, 379–391 (1998).
5. H. Niwa, J. Miyazaki, A. G. Smith, Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat. Genet.* **24**, 372–376 (2000).
6. G. Wu, D. Han, Y. Gong, V. Sebastiano, L. Gentile, N. Singhal, K. Adachi, G. Fischedick, C. Ortmeier, M. Sinn, M. Radstaak, A. Tomilin, H. R. Schöler, Establishment of totipotency does not depend on Oct4A. *Nat. Cell Biol.* **15**, 1089–1097 (2013).
7. S. Velychko, K. Adachi, K. P. Kim, Y. Hou, C. M. MacCarthy, G. Wu, H. R. Schöler, Excluding Oct4 from Yamanaka cocktail unleashes the developmental potential of iPSCs. *Cell Stem Cell* **25**, 737–753.e4 (2019).
8. L. A. Boyer, T. I. Lee, M. F. Cole, S. E. Johnstone, S. S. Levine, J. P. Zucker, M. G. Guenther, R. M. Kumar, H. L. Murray, R. G. Jenner, D. K. Gifford, D. A. Melton, R. Jaenisch, R. A. Young, Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
9. X. Chen, H. Xu, P. Yuan, F. Fang, M. Huss, V. B. Vega, E. Wong, Y. L. Orlov, W. Zhang, J. Jiang, Y. H. Loh, H. C. Yeo, Z. X. Yeo, V. Narang, K. R. Govindarajan, B. Leong, A. Shahab, Y. Ruan, G. Bourque, W. K. Sung, N. D. Clarke, C. L. Wei, H. H. Ng, Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133**, 1106–1117 (2008).
10. Y. H. Loh, Q. Wu, J. L. Chew, V. B. Vega, W. Zhang, X. Chen, G. Bourque, J. George, B. Leong, J. Liu, K. Y. Wong, K. W. Sung, C. W. Lee, X. D. Zhao, K. P. Chiu, L. Lipovich, V. A. Kuznetsov, P. Robson, L. W. Stanton, C. L. Wei, Y. Ruan, B. Lim, H. H. Ng, The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat. Genet.* **38**, 431–440 (2006).
11. D. L. van den Berg, T. Snoek, N. P. Mullin, A. Yates, K. Bezstarosti, J. Demmers, I. Chambers, R. A. Poot, An Oct4-centered protein interaction network in embryonic stem cells. *Cell Stem Cell* **6**, 369–381 (2010).
12. M. Pardo, B. Lang, L. Yu, H. Prosser, A. Bradley, M. M. Babu, J. Choudhary, An expanded Oct4 interaction network: Implications for stem cell biology, development, and disease. *Cell Stem Cell* **6**, 382–395 (2010).
13. J. Ding, H. Xu, F. Faiola, A. Ma'ayan, J. Wang, Oct4 links multiple epigenetic pathways to the pluripotency network. *Cell Res.* **22**, 155–167 (2012).
14. D. Esch, J. Vahokoski, M. R. Groves, V. Pogenberg, V. Cojocaru, H. Vom Bruch, D. Han, H. C. Drexler, M. J. Araúzo-Bravo, C. K. Ng, R. Jauch, M. Wilmanns, H. R. Schöler, A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nat. Cell Biol.* **15**, 295–301 (2013).
15. G. C. Le Bin, S. Muñoz-Descalzo, A. Kurowski, H. Leitch, X. Lou, W. Mansfield, C. Etienne-Dumeau, N. Grabole, C. Mulas, H. Niwa, A. K. Hadjantonakis, J. Nichols, Oct4 is required for lineage priming in the developing inner cell mass of the mouse blastocyst. *Development* **141**, 1001–1010 (2014).
16. B. DeVeale, I. Brokhman, P. Mohseni, T. Babak, C. Yoon, A. Lin, K. Onishi, A. Tomilin, L. Pevny, P. W. Zandstra, A. Nagy, D. van der Kooy, Oct4 is required ~E7.5 for proliferation in the primitive streak. *PLOS Genet.* **9**, e1003957 (2013).
17. K. Lins, A. Reményi, A. Tomilin, S. Massa, M. Wilmanns, P. Matthias, H. R. Schöler, OBF1 enhances transcriptional potential of Oct1. *EMBO J.* **22**, 2188–2198 (2003).
18. C. Klockenbusch, J. Kast, Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin β1. *J. Biomed. Biotechnol.* **2010**, 927585 (2010).
19. T. Cheutin, A. J. McNairn, T. Jenuwein, D. M. Gilbert, P. B. Singh, T. Misteli, Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**, 721–725 (2003).
20. P. R. Nielsen, D. Nietlispach, H. R. Mott, J. Callaghan, A. Bannister, T. Kouzarides, A. G. Murzin, N. V. Murzina, E. D. Laue, Structure of the HP1 chromodomain bound to histone H3 methylated at lysine 9. *Nature* **416**, 103–107 (2002).
21. N. J. Krogan, J. Dover, A. Wood, J. Schneider, J. Heidt, M. A. Boateng, K. Dean, O. W. Ryan, A. Golshani, M. Johnston, J. F. Greenblatt, A. Shilatifard, The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: Linking transcriptional elongation to histone methylation. *Mol. Cell* **11**, 721–729 (2003).
22. Y. Chu, R. Simic, M. H. Warner, K. M. Arndt, G. Prelich, Regulation of histone modification and cryptic transcription by the Bur1 and Paf1 complexes. *EMBO J.* **26**, 4646–4656 (2007).
23. S. Tu, M. Li, H. Chen, F. Tan, J. Xu, D. J. Waxman, Y. Zhang, Z. Shao, MAnorm2 for quantitatively comparing groups of ChIP-seq samples. *Genome Res.* **31**, 131–145 (2021).
24. C. Zang, D. E. Schones, C. Zeng, K. Cui, K. Zhao, W. Peng, A clustering approach for identification of enriched domains from histone modification ChIP-seq data. *Bioinformatics* **25**, 1952–1958 (2009).
25. J. D. Klemm, M. A. Rould, R. Aurora, W. Herr, C. O. Pabo, Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**, 21–32 (1994).
26. R. Jauch, S. H. Choo, C. K. Ng, P. R. Kolatkar, Crystal structure of the dimeric Oct6 (POU3f1) POU domain bound to palindromic MORE DNA. *Proteins* **79**, 674–677 (2011).
27. S. Jerabek, C. K. Ng, G. Wu, M. J. Arauzo-Bravo, K. P. Kim, D. Esch, V. Malik, Y. Chen, S. Velychko, C. M. MacCarthy, X. Yang, V. Cojocaru, H. R. Schöler, R. Jauch, Changing POU dimerization preferences converts Oct6 into a pluripotency inducer. *EMBO Rep.* **18**, 319–333 (2017).
28. K. Chen, Q. Long, G. Xing, T. Wang, Y. Wu, L. Li, J. Qi, Y. Zhou, B. Ma, H. R. Schöler, J. Nie, D. Pei, X. Liu, Heterochromatin loosening by the Oct4 linker region facilitates Klf4 binding and iPSC reprogramming. *EMBO J.* **39**, e99165 (2020).
29. J. Chen, Z. Zhang, L. Li, B. C. Chen, A. Revyakin, B. Hajj, W. Legant, M. Dahan, T. Lionnet, E. Betzig, R. Tjian, Z. Liu, Single-molecule dynamics of enhanceosome assembly in embryonic stem cells. *Cell* **156**, 1274–1285 (2014).
30. A. Jolma, Y. Yin, K. R. Nitta, K. Dave, A. Popov, M. Taipale, M. Enge, T. Kivioja, E. Morgunova, J. Taipale, DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527**, 384–388 (2015).
31. A. Mattout, Y. Aaronson, B. S. Sailaja, E. V. Raghu Ram, A. Harikumar, J. P. Mallm, K. H. Sim, M. Nissim-Rafinia, E. Supper, P. B. Singh, S. K. Sze, S. M. Gasser, K. Rippe, E. Meshorer, Heterochromatin protein 1β (HP1β) has distinct functions and distinct nuclear distribution in pluripotent versus differentiated cells. *Genome Biol.* **16**, 213 (2015).
32. C. L. Mueller, J. A. Jaehning, Ctr9, Rtf1, and Leo1 are components of the Paf1/RNA polymerase II complex. *Mol. Cell. Biol.* **22**, 1971–1980 (2002).
33. L. Ding, M. Paszkowski-Rogacz, A. Nitzsche, M. M. Slabicki, A. K. Heninger, I. de Vries, R. Kittler, M. Junqueira, A. Shevchenko, H. Schulz, N. Hubner, M. X. Doss, A. Sachinidis, J. Hescheler, R. Iacone, K. Anastassiadis, A. F. Stewart, M. T. Pisabarro, A. Caldarelli, I. Poser, M. Theis, F. Buchholz, A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell* **4**, 403–415 (2009).
34. N. Z. Zaidan, K. J. Walker, J. E. Brown, L. V. Schaffer, M. Scalf, M. R. Shortreed, G. Iyer, L. M. Smith, R. Sridharan, Compartmentalization of HP1 proteins in pluripotency acquisition and maintenance. *Stem Cell Rep.* **10**, 627–641 (2018).
35. M. P. Ponnusamy, S. Deb, P. Dey, S. Chakraborty, S. Rachagani, S. Senapati, S. K. Batra, RNA polymerase II associated factor 1/PD2 maintains self-renewal by its interaction with Oct3/4 in mouse embryonic stem cells. *Stem Cells* **27**, 3001–3011 (2009).
36. S. Parisi, F. Passaro, L. Aloia, I. Manabe, R. Nagai, L. Pastore, T. Russo, Klf5 is involved in self-renewal of mouse embryonic stem cells. *J. Cell Sci.* **121**, 2629–2634 (2008).
37. J. Jiang, Y.-S. Chan, Y.-H. Loh, J. Cai, G.-Q. Tong, C.-A. Lim, P. Robson, S. Zhong, H.-H. Ng, A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nat. Cell Biol.* **10**, 353–360 (2008).
38. K. Zhang, J. M. Haversat, J. Mager, CTR9/PAF1c regulates molecular lineage identity, histone H3K36 trimethylation and genomic imprinting during preimplantation development. *Dev. Biol.* **383**, 15–27 (2013).
39. B. L. Kidder, S. Palmer, J. G. Knott, SWI/SNF-Brg1 regulates self-renewal and occupies core pluripotency-related genes in embryonic stem cells. *Stem Cells* **27**, 317–328 (2009).
40. E. van der Wal, A. J. Bergsma, T. J. M. van Gestel, S. L. M. In't Groen, H. Zaehres, M. J. Araúzo-Bravo, H. R. Schöler, A. T. van der Ploeg, W. W. M. P. Pijnappel, GAA deficiency in pompe disease is alleviated by exon inclusion in iPSC-derived skeletal muscle cells. *Mol. Ther. Nucleic Acids* **7**, 101–115 (2017).
41. K. Adachi, W. Kopp, G. Wu, S. Heising, B. Greber, M. Stehling, M. J. Araúzo-Bravo, S. T. Boerno, B. Timmermann, M. Vingron, H. R. Schöler, Esrrb unlocks silenced enhancers for reprogramming to naive pluripotency. *Cell Stem Cell* **23**, 900–904 (2018).
42. E. Afgan, D. Baker, M. van den Beek, D. Blankenberg, D. Bouvier, M. Čech, J. Chilton, D. Clements, N. Coraor, C. Eberhard, B. Grüning, A. Guerler, J. Hillman-Jackson, G. Von Kuster, E. Rasche, N. Soranzo, N. Turaga, J. Taylor, A. Nekrutenko, J. Goecks, The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res.* **44**, W3–W10 (2016).
43. T. Ye, A. R. Krebs, M. A. Choukrallah, C. Keime, F. Plewniak, I. Davidson, L. Tora, seqMINER: An integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* **39**, e35 (2011).
44. C. Y. McLean, D. Bristor, M. Hiller, S. L. Clarke, B. T. Schaar, C. B. Lowe, A. M. Wenger, G. Bejerano, GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).
45. L. J. Zhu, C. Gazin, N. D. Lawson, H. Pages, S. M. Lin, D. S. Lapointe, M. R. Green, ChIPpeakAnno: A Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**, 237 (2010).
46. Y. Perez-Riverol, A. Csordas, J. Bai, M. Bernal-Llinares, S. Hewapathirana, D. J. Kundu, A. Inuganti, J. Griss, G. Mayer, M. Eisenacher, E. Pérez, J. Uszkoreit, J. Pfeuffer, T. Sachsenberg, S. Yilmaz, S. Tiwary, J. Cox, E. Audain, M. Walzer, A. F. Jarnuczak,

T. Ternent, A. Brazma, J. A. Vizcaíno, The PRIDE database and related tools and resources in 2019: Improving support for quantification data. *Nucleic Acids Ress* **47**, D442–D450 (2019).

Acknowledgments: We are grateful to V. Cojocaru, S. Jerabek, and D. Esch for helpful discussions. We thank M. Sinn for histology and microarray sample preparation. We thank Q. Chen for help with the confocal microscope. We thank A. Büchsenschütz for help with MS measurements. We thank A. Witten and the Core Facility Genomics of the Medical Faculty Münster for help with sequencing. Last, we thank A. Malapetsas for proofreading. **Funding:** This work was supported by the Max Planck Society. **Author contributions:** D.H. designed and executed experiments and wrote the manuscript. G.W., C.M.M., K.-P.K., K.A., and L.M. executed the experiments. H.C.A.D. performed the MS experiments and analyzed the proteomic datasets. R.C. and I.B. performed TSC differentiation experiments. M.J.A.-B. and D.G. analyzed the microarray datasets. H.R.S. provided the study concept and funding and edited the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The accession numbers for the microarray, ChIP-seq, and ATAC-seq are available from the Gene Expression Omnibus under accession numbers GEO: GSE144305 and GSE144876, respectively. All MS proteomic data were deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD017456.

Submitted 21 August 2020
Accepted 21 December 2021
Published 16 February 2022
10.1126/sciadv.abe4375