Context-Dependent Wiring of Sox2 Regulatory Networks for Self-Renewal of Embryonic and Trophoblast Stem Cells

Kenjiro Adachi,1,9,10,* Itoshi Nikaido,2,11 Hiroshi Ohta,3,12 Satoshi Ohtsuka,1 Hiroki Ura,1 Mitsutaka Kadota,5 Teruhiko Wakayama,3,6 Hiroki R. Ueda,2,4 and Hitoshi Niwa1,7,8,*

1Laboratory for Pluripotent Stem Cell Studies
2Functional Genomics Unit
3Laboratory for Genome Reprogramming
4Laboratory for Systems Biology
5Genome Resource and Analysis Unit
RIKEN Center for Developmental Biology, 2-2-3 Minatojima-minamimachi, Chuo-ku, Kobe 6500047, Japan
6Faculty of Life and Environmental Sciences, University of Yamanashi, Yamanashi 4008510, Japan
7Laboratory for Development and Regenerative Medicine, Kobe University Graduate School of Medicine, 7-5-1 Kusunokicho, Chuo-ku, Kobe 6500017, Japan
8JST, CREST, Sanbancho, Chiyoda-ku, Tokyo, 1020075, Japan
9These authors contributed equally to this work
10Present address: Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
11Present address: Bioinformatics Research Unit, Advanced Center for Computing and Communication, RIKEN, Wako, Saitama 3510198, Japan
12Present address: Department of Anatomy and Cell Biology, Graduate School of Medicine, Kyoto University, Sakyo-ku, Kyoto 6068501, Japan
*Correspondence: kenjiro.adachi@mpi-muenster.mpg.de (K.A.), niwa@cdb.riken.jp (H.N.)
http://dx.doi.org/10.1016/j.molcel.2013.09.002

SUMMARY

Sox2 is a transcription factor required for the maintenance of pluripotency. It also plays an essential role in different types of multipotent stem cells, raising the possibility that Sox2 governs the common stemness phenotype. Here we show that Sox2 is a critical downstream target of fibroblast growth factor (FGF) signaling, which mediates self-renewal of trophoblast stem cells (TSCs). Sustained expression of Sox2 together with Esrrb or Tfp2c can replace FGF dependency. By comparing genome-wide binding sites of Sox2 in embryonic stem cells (ESCs) and TSCs combined with inducible knockout systems, we found that, despite the common role in safeguarding the stem cell state, Sox2 regulates distinct sets of genes with unique functions in these two different yet developmentally related types of stem cells. Our findings provide insights into the functional versatility of transcription factors during embryogenesis, during which they can be recursively utilized in a variable manner within discrete network structures.

INTRODUCTION

The transcriptional output of a given cell type is controlled by unique combinations of transcription factors under the control of extrinsic signals that can modulate the expression and activity of transcription factors, forming a gene regulatory network that dictates a specific cellular phenotype. Tissue-specific transcription factors play deterministic roles in cell-type specification, which is manifested as lineage reprogramming by forced expression of such transcription factors (Graf and Enver, 2009; Zhou and Melton, 2008). Sox2 is one such transcription factor required for the maintenance of pluripotent stem cells in vivo (Avilion et al., 2003) and in vitro (Masi et al., 2007) and for the induction of pluripotency (Takahashi and Yamanaka, 2006). However, it is also preferentially expressed in neural, retinal, and trophoblast stem cells (TSCs) (Avilion et al., 2003; Pevny and Nicolis, 2010), suggesting a possible role for Sox2 in governing a common stemness phenotype.

In embryonic stem cells (ESCs), Sox2 forms a heterodimer with Oct3/4 (also known as Pou5f1) on DNA with the OCT-SOX composite motifs, and these factors cooperatively activate pluripotency-related target genes such as Nanog, Fgf4, Utf1, Lefty1, and Fbxf15, as well as their own expression (Nakatake et al., 2006, and references therein). Oct3/4-knockout ESCs are differentiated along the trophoblast lineage in a highly homogeneous manner (Niwa et al., 2000). In contrast, the loss of Sox2 causes differentiation of ESCs accompanied by upregulation of markers for trophoblast and embryonic germ layers, although artificial maintenance of Oct3/4 from the transgene can sustain self-renewal and pluripotency of Sox2-null ESCs (Masi et al., 2007), suggesting that the unique function of Sox2 may be to maintain Oct3/4 expression. These two core transcription factors, along with Nanog, form an interconnected and hierarchical network downstream of the leukemia inhibitory factor (LIF)-Stat3 and LIF-phosphatidylinositol 3-kinase (PI3K) signaling
pathways, which are connected together via signal-responsive transcription factors, such as Kitf4 and Tbx3 (Niwa et al., 2009). However, the expression of core transcription factors, including Sox2, is also maintained by shielding from ERK signaling even in the absence of LIF, so long as the GSK3 activity is simultaneously blocked (Ying et al., 2008), indicating that the ERK signaling, mainly activated by autocrine fibroblast growth factor (FGF) 4, negatively regulates the transcriptional network for pluripotency in ESCs.

The role of Sox2 as a crucial safeguard of stem cell state in ESCs could be phenotypically conserved in other types of stem cells. In neural stem cells and retinal progenitor cells, Sox2 can cooperate with other tissue-specific homeodomain transcription factors, such as Bm2 (also known as Pou3f2) and Pax6, respectively, to activate specific target genes (Inoue et al., 2007; Tanaka et al., 2004) and plays an essential role in stem cell maintenance (Favaro et al., 2009; Taranova et al., 2006). In TSCs, Sox2 is expressed predominantly in undifferentiated cells (Avilion et al., 2003). Although a previous study on Sox2-knockout mice revealed its critical role in the trophoblast lineage in vivo (Avilion et al., 2003), the mode of Sox2 function, as well as its possible partners, is unknown. TSCs are derived from embryos at a similar developmental stage as ESCs but originate from a distinct lineage and show completely different characteristics from ESCs. Self-renewal of TSCs is absolutely dependent on FGF-ERK signaling that supposedly activates the expression of Sox2, which implied the trophoblast-specific transcription factors such as Cdx2 and Eomes (Kunath et al., 2004; Tanaka et al., 1998), which is in contrast to the detrimental effect of ERK activation in ESCs. Therefore, it is intriguing to examine the extent to which Sox2 functions are conserved within discrete transcriptional networks in these well-characterized and developmentally similar stem cells.

Here, we performed a comprehensive analysis of Sox2 functions in ESCs and TSCs. First, we analyzed functions of transcription factors that are activated downstream of FGF signaling in TSCs and revealed the pivotal role of Sox2 in FGF-dependent self-renewal of TSCs. Then we compared the target genes regulated by Sox2 in ESCs and TSCs using transcriptome and chromatin immunoprecipitation analyses and found that Sox2 regulates quite different sets of genes in ESCs and TSCs. Finally, we identified Tfp2c as one of the TSC-specific Sox2 partners partly responsible for the recruitment of Sox2 to TSC-specific target genes. These data indicated that the function of Sox2 is highly flexible in the context of transcriptional networks controlled by different extrinsic signals.

RESULTS

FGF-ERK Signaling Converges on a Subset of Transcription Factors in TSCs

We have previously shown that trophoblast stem (TS)-like cells can be induced by forced repression of Oct3/4 in ESCs, which results in gradual but homogenous downregulation of ESC markers, like Nanog, and concomitant upregulation of TSC markers, like Cdx2, when exogenous FGF4 is supplied (Figure 1A and see Figure S1A available online; Niwa et al., 2000, 2005). After subcloning, these TS-like cells (hereafter called ZHBTc4-TSCs) were stably maintained as TSCs with a global expression profile similar to that of embryo-derived EGF-P-TS3.5 TSCs (eTSCs) (Tanaka et al., 1998; Figure 3A). This in vitro system allowed us to compare transcriptional networks in two distinct types of stem cells with identical genotypes. We first investigated the effect of FGF4 on the expression of key transcription factors in TSCs, which are less well understood than the ESC transcription factors regulated by LIF. Increased expression was found for transcription factors previously implicated in TSC maintenance, including Cdx2 (Chawengsaksophak et al., 1997; Strumpf et al., 2005), Eomes (Russ et al., 2000; Strumpf et al., 2005), Elf5 (Donnison et al., 2005), Esrb (Luo et al., 1997), and Sox2 (Avilion et al., 2003) (Figure 1B). In contrast, the expression of differentiated trophoblast markers, such as Ascl2 (also known as Mash2) and Cdkn1c (also known as p57Kip2), was repressed by FGF4 (Figure S1B). Other possible regulators of trophoblast development, such as Ets2, Tfp2c, and Tead4, did not appear to be activated by FGF4 (Figure S1B). The FGF responsiveness was confirmed by the rapid downregulation in ZHBTc4-TSCs and eTSCs after withdrawal of FGF4 or treatment with the FGF receptor inhibitor PD173074 (Figures 1C and S1C). Conversely, when TSCs were stimulated with FGF4 after starvation, the transcription factors were upregulated in 12 hr (Figure 1D). This effect was abrogated in the presence of the MEK inhibitor PD0325901, but not by the PI3K inhibitor LY294002 (Figure 1D), indicating that it was primarily mediated by the FGF-ERK signaling pathway. Consistent with this, these transcription factors showed largely overlapping expression with phosphorylated ERK in the extraembryonic ectoderm of postimplantation embryos (E5.5), although the expression of Esrb and Sox2 was even more highly restricted, and that of Eomes and Elf5 was extended (Figure 1E). It should be noted that the response of Esrb and Sox2 to FGF is TSC-specific, as the expression of these genes in ESCs did not depend on FGF signaling (Figure 1F). Instead, their expression was maintained by either activation of LIF signaling or intrinsic transcriptional circuitry without the counteracting FGF signal (Figure 1F; Niwa et al., 2009; Ying et al., 2008). Indeed, around the time of implantation (E3.5–E4.75), they are expressed exclusively in the epiblast without detectable phosphorylated ERK (Figure 1G; data not shown). These results also suggest that TSCs, with their FGF/ERK dependency, correspond to a population of trophoblasts in postimplantation, but not in preimplantation, embryos.

Sox2 Is a Critical Downstream Target of FGF Signaling that Mediates Self-Renewal of TSCs

To evaluate the functional roles of these FGF-responsive transcription factors in self-renewal downstream of FGF-ERK signaling, we used a gain-of-function approach using eTSCs to test whether the requirement for FGF4 can be replaced by artificial maintenance of these transcription factors from transgene expression. Although sustained expression of a single factor alone did not confer independence from FGF4, combined expression of Sox2 and Esrb enabled prolonged self-renewal in the absence of FGF4 (Figures S2A and S2B). These results were confirmed using ZHBTc4-derived TS-like cells, although
Sox2 alone slightly promoted self-renewal (Figures S2C and S2D). We next assessed the multipotency of the transgenic TSCs by chimera formation assay. The stable transfection of an expression vector for Sox2 and Esrrb, linked by a self-cleaving 2A peptide sequence and flanked by loxP sites, enabled the self-renewal of eTSCs at a clonal density in the absence of FGF4 (Figure 2A). These cells maintained Cdx2 and Eomes expression even in the presence of PD173074, confirming their independence from FGF signaling (Figure 2B). After the culture without FGF4 for 4 weeks, the cells were transiently transfected
with Cre recombinase to excise the floxed transgene in the presence of FGF4, followed by negative selection against thymidine kinase with gancyclovir. The resulting transgene-deleted clones restored FGF dependency and readily differentiated after withdrawal of FGF4 (Figures 2C and S2E). When the reverted TSCs were injected into blastocysts, they contributed to all of the placental cell lineages in chimeras at the same efficiency as the parental line maintained in the presence of FGF4 (Figures 2D, 2E, and S2F), demonstrating that the expression of Sox2 and Esrrb is minimally sufficient to maintain the multipotency of TSCs in the absence of FGF4.

**Induced Knockout of Sox2 in ESCs and TSCs Causes Differentiation along Their Respective Developmental Pathways**

The results described above indicate that Sox2 plays a central role in regulating self-renewal of TSCs downstream of FGF signaling. The function of Sox2 in TSCs stands in contrast to that in ESCs, in which Sox2 acts as a core transcription factor downstream of LIF signaling to govern pluripotency (Niwa et al., 2009). In ESCs, Sox2 prevents trophoblast differentiation mainly by maintaining Oct3/4 expression or repressing Eomes expression and restricts the specification of the primary germ layers as well (Ivanova et al., 2006; Masui et al., 2007).

To compare the functional role of Sox2 in ESCs and TSCs in more detail, we performed whole-genome expression analysis after induced knockout of Sox2 in each cell type. For this, we first established a stable TSC line by transient induction of Cdx2 (Niwa et al., 2005) in the Tet-inducible Sox2-knockout ESC line, 2TS22C (Masui et al., 2007). Global gene expression profiles of ZHBTc4- and 2TS22C-derived TSCs, together with embryoid-derived EGFP-TS3.5, revealed that these TSCs are clustered together and separated from the ESCs along the primary principal component (PC1) axis (Figure 3A). Upon knockout of Sox2 in TSCs, expression of TSC markers, especially Esrrb, gradually decreased even in the presence of FGF4, concomitant with upregulation of differentiatied trophoblast markers (Figures 3B and S3A). These TSCs gradually lost their self-renewal capacity, manifesting flat morphological changes (Figure 3C). Global expression analysis identified 3,440 and 1,066 genes that were affected by Sox2 knockout in each cell type (Figures 3D, 3E, and S3B–S3F; Table S1). The smaller number of affected genes in TSCs, especially upregulated genes, could represent the restricted developmental capacity of TSCs compared to ESCs. Only a few genes were commonly downregulated in ESCs and TSCs, while larger numbers of genes, especially genes involved in placental development, were upregulated in both cell types; this makes sense, considering Sox2’s role in preventing trophoblast commitment in ESCs (Masui et al., 2007). Downregulation of pluripotency markers and concomitant upregulation of mesodermal and endodermal markers, but not neuroectodermal markers, was observed exclusively in ESCs (Figures 3F and 3G). In contrast, genes involved in cell-cycle regulation, especially in mitotic checkpoint control, were downregulated in TSCs (Figure 3G); this likely reflects the endoreduplication cycle, characteristic of differentiating trophoblast giant cells (Ullah et al., 2008). Gene ontology (GO) analysis using DAVID (Huang da et al., 2009) showed that GO terms related to placental function, like blood vessel development and negative regulation of immune system processes, are overrepresented in genes upregulated in TSCs and, to a lesser extent, in genes upregulated in ESCs (Figures 3G, S3G, and S3H; Table S2). Gene set enrichment analysis (GSEA; Subramanian et al., 2005) using published microarray data (Ralston et al., 2010; Schulz et al., 2009) further revealed that genes affected by Sox2 knockout are preferentially enriched in genes whose expression is similarly regulated during normal differentiation of the respective cell types (Figures 3H, 3I, and S3I). These results indicate that Sox2 regulates unique sets of genes to maintain stemness and prevent precocious differentiation along the possible developmental pathways in each stem cell type.

**Sox2 Occupies Distinct Sets of Genes Implicated in Later Differentiation in ESCs and TSCs**

To better understand the mechanisms underlying the versatility in transcriptional regulation by Sox2, we next analyzed genome-wide binding sites of Sox2 during the differentiation of ZHBTc4 ESCs into TS-like cells by chromatin immunoprecipitation, followed by high-throughput sequencing (ChiP-seq) (Figures 4A–4C and S4A–S4F; Table S3). During the differentiation, Sox2 binding sites changed dynamically. For instance, a conserved region 4 within a distal enhancer of Pou5f1 was preferentially occupied in ESCs (Figure 4A), while the Cdx2 locus was specifically occupied in differentiating TSCs (Figure 4B). Persistent binding was less frequently observed, as exemplified at the Col4a1 locus (Figure 4C). As a whole, the numbers of peaks (FDR < 0.05) that overlapped with those of other samples were quite limited (Figure S4G). K-means clustering of normalized ChIP-seq signals around the merged peak regions over the course of differentiation revealed that the vast majority of Sox2 binding sites are cell-type specific (Figure 4D). About half of the Sox2 binding sites in ESCs specific (cluster 2 [c2] and c5), while the other half are predominant in differentiating TSCs (c1, c3, and c4), which are further classified as either increased (c1), transiently increased (c3), or gained (c4) binding. The Sox2 binding sites in any clusters are predominantly found at transcription start site (TSS)-distal intronic or intergenic regions (Figure 4E). To gain insights into the underlying mechanisms for such a context-dependent Sox2 binding, we performed de novo motif analysis using MEME (Bailey et al., 2009) and found that different noncanonical motifs, besides the canonical SOX motif, were enriched in the peak centers of each cluster (Figure 4F). We then mapped these motifs to the peak regions (Figure 4G). As expected, overrepresentation of the OCT-SOX composite motif was observed only in ESC-specific clusters (c2 and c5), confirming the pivotal role of Oct3/4 as a recruiter of Sox2 to the ESC-specific target sites. Interestingly, the AP-2 binding motif was overrepresented only in the TSC-predominant clusters (c1, c3, and c4) and found together with the canonical SOX motif within a subset of peaks (Figure 4H), suggesting co-binding of Sox2 and AP-2 transcription factors. It also should be noted that although the canonical SOX motif was enriched in all clusters, it was more overrepresented in c1 and c3, where Sox2 binding is relatively persistent. Although the ESRRB motif was also overrepresented in Sox2 peak clusters, we could not
A. loxP

PB  CAG  Esrb  2A  Sox2  IRES  PacIΔk  pA  EGFP  pA  PB

EGFP-TS<sub>3.5</sub> → Transfection → Pick up (B) → Cre → Pick up (C) → Blastocyst inj. (D,E)

B. Cdx2  Eomes  Hoechst  Esrb (Tg)  Sox2 (Tg)  Hoechst

- FGF4

- FGF4/ +PD17

C. 628-1c2

+ FGF4  - FGF4

D. EGFP-TS<sub>3.5</sub>

628-1c2

E. de  sp  la

(legend on next page)
Distinct Sox2 Networks in ESCs and TSCs

find a significant difference of the motif enrichment in each peak cluster (data not shown). We also mapped these motifs to the peak summits of original samples, which were ranked by the -log₁₀(q-value) for peak call (Figure 4I). The more confident peaks in ESCs were highly enriched for the OCT- Sox motif compared to the less confident peaks. In contrast, the more confident peaks in differentiating TSCs were enriched for the canonical SOX motif, suggesting a stronger dependence of Sox2 binding on its own recognition motif. These results suggest that sequence features could, in part, explain the dynamic behavior of Sox2 binding.

To explore the functional significance of differential Sox2 binding, GO enrichment analysis was performed using GREAT (McLean et al., 2010), considering the predominance of TSS-distal binding sites (Figure 4J; Table S4). Genes involved in stem cell maintenance were enriched for the peaks of all clusters, but especially c5, while genes for early placental development were enriched for c3, c4, and c5. Enrichment of genes associated with placental functions, such as artery development, regulation of lymphocyte differentiation, and regulation of T cell activation, are observed in c1, c2, or c3. These results suggest that contextual Sox2 binding is related to specific developmental processes.

Next we investigated the functional consequence of Sox2 binding by comparing the ChiP-seq data with the microarray data. The peaks of each cluster were assigned to RefSeq genes whose TSSs are located within a 10 kb distance (Table S5). Enrichment of these genes in genes affected by Sox2 knockout in ESC or TSC was analyzed by GSEA (Figures 4K, 4L, and 4N). We found that genes associated with the ESC-specific clusters (c2 and c5) are significantly enriched in genes downregulated in Sox2-knockout ESCs, while genes near the TSC-predominant clusters (c1, c3, and c4) are enriched in genes downregulated in Sox2-knockout TSCs. GSEA using the microarray data of ZHBTc4 ESCs differentiating into TSCs showed that genes associated with the ESC-specific clusters (c2 and c5) and with the cluster showing gains of Sox2 binding in TSCs (c4) are significantly enriched in genes downregulated and upregulated during the trophoblast differentiation, respectively (Figures 4M and 4N). It should be noted that when more distant binding sites were assigned to genes, such a positive effect of Sox2 binding on expression of nearby genes was less evident (data not shown). These results suggest that Sox2 primarily contributes to activation of neighboring genes in both ESCs and TSCs. Nevertheless, this does not exclude a role of distal Sox2 binding in transcriptional regulation and a role of Sox2 binding in priming gene expression without significant consequences (see Discussion). Taken together, self-renewal of ESCs and TSCs is regulated by a context-dependent role of Sox2.

We also performed ChiP-seq analysis of Esrrb using the same chromatin samples (Figures S4A–S4F). In stark contrast to Sox2, Esrrb bindings were less dynamically regulated during the trophoblast differentiation (Figure S4G). K-means clustering of normalized ChiP-seq signals around the merged peak regions identified ESC-specific binding peaks (c5) as a major cluster (Figure S4I). The remaining peaks were classified as relatively persistent clusters, with a TSC-predominant cluster not apparent. The Esrrb binding sites, especially of c2 and c3, were frequently found around the TSS (Figure S4J), consistent with the suggested role of Esrrb in collaboration with basal transcription machinery (van den Berg et al., 2010). The peaks of all clusters were enriched for the canonical ESRRB recognition motif (a motif for the nuclear receptor superfamily), and strong binding sites were associated with the occurrence of the ESRRB motifs (Figures S4L and S4M). These results suggest a common mechanism for Esrrb in regulating stemness in ESCs and TSCs. Although a minor fraction of Sox2 peak regions was also bound by Esrrb, we found only a slight enrichment of genes related to specific biological processes near these regions (Figures S4N–S4P), suggesting that Sox2 and Esrrb might play complementary roles via different sets of target genes.

Tfap2c Is a Cooperative Factor of Sox2 in TSCs

The motif analysis described above suggests a possible role of AP-2 transcription factors in cooperating with Sox2 in TSCs. Among the five members of the AP-2 family of transcription factors, Tfap2c (also known as Tfap2c or AP-2γ) is most abundantly expressed in TSCs (Figure 5A), and its expression is restricted to the trophoblast lineages in vitro and in vivo (Figures 5B, 5C, and S1B). The forced expression of Tfap2c rapidly suppressed Oct3/4 expression and induced trophoblast differentiation in ESCs (Figures S5A and S5B), confirming previous reports (Berg et al., 2011; Kuckenberg et al., 2010). When Tfap2c expression was induced in ESCs, the enhanced recruitment of
Figure 3. Microarray Analysis of Genes Affected by Sox2 Knockout in ESCs and TSCs
(A) Principal component analysis (PCA) of global gene expression profiles. The first two PCs explain 37% and 24%, respectively, of the total variance.
(B) Changes in gene expression after Tet-induced Sox2 knockout in ESCs and TSCs. Error bars indicate SD of three replicates.
(C) Photomicrographs of Tet-inducible Sox2-knockout ESC line 2TS22C and its derivative TSC line cultured in the absence or presence of Tet for 4 days. Scale bar, 200 μm.
(D) MA plots showing the log₂ ratio and average log₂ intensity of control and Sox2-knockout ESCs or TSCs. Whole-genome expression analysis was performed 4 days after Sox2 knockout. Blue lines indicate log₂ ratio cutoff of 0.25. Blue, red, or green dots indicate selected genes affected only in ESCs, in TSCs, or in both, respectively.
(E) The number of genes affected by Sox2 knockout in ESCs and TSCs. Cutoffs for q-value and log₂ ratio were set to 0.1 and 0.25, respectively. The numbers of genes upregulated or downregulated are shown in red or blue, respectively. The remaining genes in the intersection were affected in an opposite manner in ESCs and TSCs. Selected genes downregulated or upregulated in common are shown. Genes in bold were also considered significant when more stringent cutoff values (q-value < 0.05) were used (Figure S3 F).
(F) Changes in expression levels of lineage markers. Note that some markers for trophoblast lineage (asterisks) are also expressed in embryonic tissues.
(G) GO analysis of genes affected by Sox2 knockout using DAVID. The \log_{10}(p\text{ value}) for selected GO terms in the Biological Process category (GOTERM_BP_FAT) is shown in a heatmap.
(H and I) GSEA of changes in expression of genes affected by Sox2 knockout during the differentiation of ESCs (H) and TSCs (I). Genes are ordered according to their expression levels in differentiating ESCs (embryoid bodies) relative to undifferentiated ESCs (H) or levels in differentiating TSCs relative to undifferentiated TSCs (I). Normalized enrichment scores (NES) and FDR are shown.
See also Figure S3.
Figure 4. ChIP-Seq Analysis of Sox2 Occupancy during Differentiation of ESCs into TSCs

(A–C) Normalized Sox2 ChIP-seq signals at the Pou5f1 (A), Cdx2 (B), and Col4a1 (C) loci during the differentiation of ZHBTc4 ESCs into TS-like cells. ChIP-seq signals were normalized as levels per million mapped reads. The chromatin of Sox2 KO ESCs was used as a negative control. The Cdx2 region analyzed by ChIP-PCR in Figure 5 is shown (an arrowhead).

(D) A heatmap of the ChIP-seq signals around the merged peak regions. ChIP-seq peaks of each sample (q-value < 0.05) were merged, and normalized ChIP-seq signals over ± 1 kb regions around the merged peak centers were clustered with the k-means algorithm into five clusters.

(E) Annotation of ChIP-seq clusters with genomic features. The merged peak centers of each cluster were annotated with genomic features using CEAS (Shin et al., 2009). Annotation of the whole genome is shown as a control.

(legend continued on next page)
Sox2 was observed at the Tfap2c binding sites, including the Cdx2 locus, as early as one day after induction, even with sustained expression of Oct3/4 by the transgene (Figures 5D and S5C). Conversely, knockdown of Tfap2c significantly decreased TSC-specific recruitment of Sox2 (Figures 5E and S5E). We confirmed that overexpression or knockdown of Tfap2c did not increase or decrease Sox2 protein levels, respectively (Figure S5F), which excludes the possibility that the observed changes in Sox2 binding are simply due to the modulation of Sox2 protein levels. To elucidate the mechanistic basis of cooperative binding, we performed coimmunoprecipitation experiments of endogenous Sox2 and Tfap2c proteins in TSCs. Tfap2c was coimmunoprecipitated with Sox2 (Figure 5F, left) and, reciprocally, Sox2 was coimmunoprecipitated with Tfap2c (Figure 5F, right), indicating that these proteins physically interact with each other. Importantly, overexpression of Tfap2c together with Sox2 also enabled the self-renewal of TSCs without FGF4 (Figures 5G and 5H), whereas knockdown of Tfap2c prevented the formation of undifferentiated TSC colonies (Figure S5D). Taken together, these results suggest that Tfap2c is a crucial cooperative factor for Sox2 to maintain self-renewal of TSCs.

To gain more global insights into the cooperativity between Sox2 and Tfap2c, we performed ChIP-seq analysis of these factors, along with Cdx2, in ZHBTc4-TSCs. We also analyzed the Sox2 binding sites in eTSCs, confirming a large overlap with those in ZHBTc4-TSCs (Figure S6). As expected, a number of previously identified TSC-predominant Sox2 binding sites, including the Esrrb, Eomes, and Atrx/Magt1 loci, were co-occupied by Sox2 and Tfap2c in ZHBTc4-TSCs (Figures 6A–6C and S6A–S6H). A significant number of Sox2 peaks overlapped with Tfap2c peaks, while Cdx2 peaks that overlapped with Sox2 or Tfap2c peaks were limited, irrespective of the relatively large numbers of peaks (Figure 6D). Interestingly, Tfap2c peaks were significantly enriched at proximal promoters of known genes, while Cdx2 binding sites were found exclusively at TSS-distal regions (Figure 6E). De novo motif discovery identified the canonical recognition motifs for each factor as primary motifs (Figure 6F). The motifs for AP-1 and AP-2, and TEAD and AP-1 transcription factors were also enriched in Sox2 and Tfap2c peaks, respectively. Motif mapping into the ranked peaks suggests that Tfap2c and Cdx2 bindings primarily depend on their recognition sequences, while the presence of canonical SOX motif is less predictive for Sox2 bindings in vivo, further emphasizing the contextual binding of Sox2 (Figure 6G). Finally, we found TSC-predominant Sox2 peak clusters (c1, c3, and c4), but not ESC-specific clusters (c2 and c5), shown in Figure 4D were significantly overlapped with the Sox2 and Tfap2c peaks in TSCs (Figure 6H). Taken together, these results suggest a global cooperativity of Sox2 and Tfap2c in regulating transcriptional network unique to TSCs.

**DISCUSSION**

Tissue-specific transcription factors are primarily responsible for directing transcriptional programs that confer unique cellular phenotypes. However, it is frequently found that the same transcription factors are recurrently expressed in different cell types at different stages during development. Sox2 is one such factor whose expression is dynamically regulated: inner cell mass at blastocyst stage; primitive ectoderm and extraembryonic ectoderm at egg-cylinder stage; neuroectoderm around gastrulation; and central nervous system, sensory placodes, and foregut endoderm thereafter (Collignon et al., 1996; Wood and Episkopou, 1999). In addition, Sox2 is also preferentially expressed in putative stem cell populations of various adult tissues (Arnold et al., 2011). Studies of germline or conditional knockout mice revealed that Sox2 is definitely required for the development and maintenance of these tissues (Arnold et al., 2011; Avilion et al., 2003; Favaro et al., 2009; Que et al., 2009; Taranova et al., 2006). Is such a pleiotropic role of Sox2 based on its common functions in different cell types or different modes of action in different contexts? Here, we have shown the different modes of Sox2 action in ESCs and TSCs underpinning the recursive deployment of Sox2 in different cell types during embryonic development.

First we found that Sox2 is both necessary and sufficient, if combined with Esrrb or Tfap2c, to mediate the effect of FGF on self-renewal of TSCs. Although some transcription factors are considered to be involved in this process, no factors have been identified that can replace FGF dependency. Although Sox2-deficient mice die shortly after implantation due to defective epiblast development (Avilion et al., 2003), chimeric analysis using Sox2-null blastocyst injected with wild-type ESCs suggests a critical role of Sox2 in the trophoblast lineage around E7.5, when Erk2 (also known as Mapk1)- or Frs2-deficient mice...
also result in lethality (Gotoh et al., 2005; Hadari et al., 2001; Saba-El-Leil et al., 2003). The function of Esrrb is also essential but auxiliary, since, unlike Sox2, Esrrb alone did not affect self-renewal of TSCs in the absence of FGF4 (Figure S2C). It should be noted that, as exemplified by the FGF-ERK dependency, extraembryonic ectoderm in postimplantation embryos is phenotypically and molecularly distinct from the trophectoderm in preimplantation embryos, in which Tead4 and Cdx2 play central roles. Thus, Sox2 is deployed for FGF-dependent self-renewal of TSCs after implantation (Figures 1E and 1G).

In two types of stem cells, Sox2 can connect different extrinsic signals, LIF-STAT and FGF-ERK, to different sets of transcription factors, enabling cell-type-specific responses to environmental stimuli (Figure 6I). We found that one of the mechanisms for such context-dependent functions of Sox2, with intrinsic low-affinity DNA-binding activity (Kamachi et al., 2000), is attributable to the cell-type-specific transcription factors that assist in its recruitment or binding to the target sites: Oct3/4 in ESCs and Tfap2c in TSCs. Mechanistically, Sox2 physically interacts with Tfap2c in TSCs, which may facilitate the efficient binding of Sox2 to the targets co-occupied with Tfap2c. The cooperative role of Sox2 and Tfap2c in vivo is supported by the evidence that Tfap2c-knockout embryos die between E7.5 and E8.5 due to a defect in trophoblast-derived tissues (Auman et al., 2002; Werling and Schorle, 2002), as is the case in Sox2-null embryos rescued by wild-type ESCs. Interestingly, such cooperative action of Sox2 and Tfap2c in TSCs was found on Cdx2, which is the most important regulator of trophoblast development. It should be noted that although Sox2 activates Cdx2 in TSCs, it is expressed in a highly reciprocal manner to Cdx2 from E3.0 to E5.5, in which Sox2 begins to be expressed in trophoblast lineages as well (Figures 1E and 1G; data not shown). Although epigenetic constraints imposed on the chromatin might be involved in different accessibility of target genes in different
cellular contexts, its contribution would be minimal, since the epigenetic state of ESCs is considered as a tabula rasa (Mura-kami et al., 2011; Niwa, 2007); this was confirmed by the dynamic change of the Sox2 binding sites immediately after elimination of Oct3/4.

However, even in a different mode of action with different partners, Sox2 plays a conceptually common role in ESCs and TSCs: the maintenance of self-renewal. This role is achieved by the activation of genes involved in maintaining the undifferentiated states and, albeit less commonly, restricts the expression of genes promoting differentiation. Furthermore, although a large fraction of the Sox2-bound genes whose functions are related to later developmental processes were not significantly affected by Sox2 knockout, Sox2 may prime these genes for subsequent activation (Wegner, 2011). A flexible mode of regulation by Sox2 fits well with the idea that Sox2 is one of the putative stemness genes that is predominantly expressed and serves as a molecular hub to respond to diverse environmental cues for stem cell maintenance in different types of stem cells. This is in sharp contrast to the case of Esrrb, as shown here, and Myc, which regulates a common set of genes involved in cell proliferation and metabolism in wide variety of stem or progenitor cells (Kim et al., 2010). The versatile functions of Sox2 provide a clear example of how tissue-specific transcription factors have maximum flexibility to play divergent roles in different networks. Considering that the sequence specificities of individual transcription factors are generally limited, and only a few of the potential recognition sites are bound in vivo, the context-dependent redeployment of these transcription factors to drive tissue-specific transcriptional programs, as shown here, may be prevailing. Such mechanisms should contribute to maximizing the complexity and diversity of cellular functions with a minimal set of transcription factors. Reorganization of transcription factor networks might occur extensively during

Figure 6. Global Cooperativity of Sox2 and Tfap2c in TSCs
(A–C) Normalized Sox2, Tfap2c, and Cdx2 ChIP-seq signals at the Esrrb (A), Eomes (B), and Atrx/Magt1 (C) loci in ZHBTc4-TSCs. The Esrrb region analyzed by ChIP-PCR in Figure 5 is shown (an arrowhead).

(D) A Venn diagram showing the number of overlapping peaks in each data set. Peak regions were defined as ± 100 bp regions around peak summits.

(E) Annotation of ChIP-seq peaks with genomic features using CEAS.

(F) Motifs found by MEME to be overrepresented in the top 1,000 ranked peaks (summits ± 100 bp).

(G) Occurrence of the motifs in the ranked peaks. The cumulative percentage of peaks (summits ± 200 bp) containing at least one sequence for a given motif are plotted against ranks of peaks.

(H) A heatmap representing the fractions of overlapping peaks of Sox2 ChIP-seq clusters found in Figure 4D with Sox2, Tcfap2c, and Cdx2 peaks in TSCs. The numbers of overlapping peaks are shown in the grid.

(I) Self-renewal networks regulated by transcription factors and extrinsic signals in ESCs (Niwa et al., 2009) and TSCs. Context-dependent role of Sox2 is in part mediated by the lineage-specific and mutually antagonistic transcription factors Oct3/4 and Tfap2c.
embryogenesis, in which networks transit from one state to another along the developmental trajectory. This strategy also should be beneficial in evolution, since it allows establishment of novel networks, such as those for placental formation, only from existing transcription factors in the genome. This idea is supported by the evidence that very few transcription factors are expressed only in the sole cell type, and their combinatorial usage is quite flexible. We expect that the concept found in this study will emphasize the significance of the entire transcription factor network for the understanding of refined strategies in evolution and development.

EXPERIMENTAL PROCEDURES

Cell Culture and Cell Lines

ESCs were maintained in GMEM supplemented with 10% FBS, 1 x nonessential amino acids, 1 mM sodium pyruvate, 0.1 mM 2-mercaptoethanol, and 1,000 U/ml LIF on gelatin-coated dishes. TSCs were induced and maintained in GMEM/10% FBS supplemented with 50 ng/ml human FG4, 2 μg/ml heparin, 10 ng/ml human Activin A, and 2 mM L-glutamine on mitomycin C treated mouse embryonic fibroblasts (MEFs). Tet-inducible Oct3/4-knockout ESC line ZHBTc4 and its derivative (4S2IP14) carrying an IRES-pac cassette knocked in downstream of the Sox2 coding region for selection of Sox2-positive cells were maintained as described previously (Niwa et al., 2000, 2005). Tet-inducible Sox2-knockout ESC line 2TS22C and its TSC derivative (22CEROH2) carrying a 4-hydroxytamoxifen (4-OHT)-inducible Cdx2-ERTM transgene and a hygromycin-thymidine kinase fusion gene targeted into the Pou5f1 locus were maintained as described previously (Masui et al., 2007; Niwa et al., 2005). ESC-derived TSC lines were established from ZHBTc4 in TSC medium supplemented with Tet on MEF feeders followed by subcloning, or from 22CEROH2 treated with 4-OHT followed by culture without 4-OHT and with GANC to eliminate unindifferentiated ESCs. Inducers and inhibitors were used at the following concentrations: Tet, 1 μg/ml; doxycycline (Dox), 1 μg/ml; 4-OHT, 200 ng/ml; GANC, 1 μM; PD173074, 100 nM; PD0325901, 1 μM; L1294020, 20 μM.

ChIP

ChIP assays were performed using the ChIP-IT Express Enzymatic Kit (Active Motif) following the manufacturer’s instructions. Briefly, cells were crosslinked with 1% formaldehyde for 10 min at room temperature, and chromatin was enzymatically sheared into fragments with an average size of 300–500 bp. The chromatin fragments were immunoprecipitated with the antibodies against Sox2 (raised by us), Esrb (H6705, Perseus Proteomics), Tfap2c (sc-8977, Santa Cruz Biotechnology), and Cdx2 (CDX2-88, BioGenex). Precipitated DNA was analyzed by real-time PCR. For each experiment, the percentage of input was determined and normalized to the value obtained at the 28S treated DNA was analyzed by real-time PCR. For each experiment, the percentage of input was determined and normalized to the value obtained at the 28S.

Mice and Embryos

All animal experiments were performed according to the guidelines for animal experiments of RIKEN Center for Developmental Biology and approved by the Animal Experiment Committee of the RIKEN Kobe Institute.

ACCESSION NUMBERS

Microarray and ChIP-seq data have been deposited at the Gene Expression Omnibus (GEO) database under accession number GSE28455.

SUPPLEMENTAL INFORMATION

Supplemental Information for this article includes six figures, six tables, and Supplemental Experimental Procedures and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2013.09.002.

AUTHOR CONTRIBUTIONS

K.A. and H.N. designed the study; K.A. performed experiments and analyzed the data; K.A., I.N., and H.R.U. performed bioinformatic analysis; H.O. and T.W. performed blastocyst injection; S.O. performed coimmunoprecipitation experiments; K.A., H.U., and M.K. performed ChIP-seq experiments; and K.A. and H.N. interpreted the data and wrote the paper.

ACKNOWLEDGMENTS

We thank A. Smith and G. Guo (University of Cambridge, Cambridge) for piggyBac transposon vectors, S. Tanaka (University of Tokyo, Tokyo) and S. Kuraku (Genome Resource and Analysis Unit, RIKEN, Kobe) for the EGFP-TS23.5 line, the RIKEN Omics Science Center (Yokohama) for high-throughput sequencing analysis, K. Uno (Functional Genomics Unit, CDB, RIKEN, Kobe) for microarray analysis, H. Schöler (Max Planck Institute for Molecular Biomedicine, Münster) for generous support, and Y. Nakai-Futasugi and D. Sipp (CDB, RIKEN, Kobe) for their critical reading of the manuscript. This work was supported by a RIKEN research grant and by the Cell Innovation Program from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) to H.N. and H.R.U. It was also supported by the CREST Program, from the Japan Science and Technology Agency (JST) to H.N.

Received: March 1, 2013
Revised: July 8, 2013
Accepted: August 29, 2013
Published: October 10, 2013

REFERENCES


