Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Distinct Enhancer Activity of *Oct4* in Naive and Primed Mouse Pluripotency

Hyun Woo Choi,[1,3] Jin Young Joo,[1,4] Yean Ju Hong,[1] Jong Soo Kim,[1] Hyuk Song,[1] Jeong Woong Lee,[2] Guangming Wu,[3] Hans R. Schöler,[3] and Jeong Tae Do[1,*]

[1]Department of Stem Cell and Regenerative Biology, College of Animal Bioscience and Technology, Konkuk University, Seoul 143-701, Republic of Korea
[2]Research Center of Integrative Cellulomics, Korea Research Institute of Bioscience and Biotechnology, Daejeon 305-806, Republic of Korea
[3]Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, 48149 Münster, Germany
[4]Dream-i Infertility Clinic, 45-17 Huimang-ro, 46 Beon-gil Baebang-eup, Asan-si 31470, Chungcheongnam-do, Republic of Korea
*Correspondence: dojt@konkuk.ac.kr
http://dx.doi.org/10.1016/j.stemcr.2016.09.012

SUMMARY

Naive and primed pluripotent stem cells (PSCs) and germ cells express the *Oct4* gene. The *Oct4* gene contains two *cis*-regulatory elements, the distal enhancer (DE) and proximal enhancer (PE), which differentially control *Oct4* expression in a cell-type-specific and stage-specific manner. Here, we generated double transgenic mice carrying both *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-tdTomato (RFP), enabling us to simultaneously monitor the activity of DE and PE. *Oct4* expression is stage-specifically regulated by DE and PE during embryonic and germ cell development. Using this dual reporter system, we successfully cultured pure populations of naive (GFP^+RFP^-) and primed (GFP^-RFP^+) PSCs. We found that GFP^+RFP^- cells were metastable (not naive) in serum-containing medium; stable naive pluripotent cells were observed in medium containing two inhibitors (Meki and GSKi) but lacked serum. Finally, we suggest that the activity of *Oct4* DE and PE is regulated by the repressive histone marks and DNA methylation in a cell-type-specific manner.

INTRODUCTION

Pluripotent stem cells (PSCs) can self-renew unlimitedly and have the potential to differentiate into all somatic and germ cell types. PSCs can be derived from the inner cell mass (ICM) of blastocyst or implanted epiblast cells (Hanna et al., 2010; Nichols and Smith, 2009). ICM and epiblast cells are pluripotent cells but not bona fide stem cells in vivo, as they differentiate to eventually establish the three germ layers in the gastrulating embryo. Both cell populations can be grown as PSCs when they are cultured in vitro with stem cell maintenance medium. Interestingly, the ICM of blastocysts forms "naive" pluripotent embryonic stem cells (ESCs) and epiblast cells form "primed" pluripotent epiblast stem cells (EpiSCs) (Hanna et al., 2010). Primed PSCs have limited differentiation potential in vivo; they barely contribute to chimeras by blastocyst injection analysis. Primed PSCs maintain stemness through basic fibroblast growth factor (bFGF) and Activin/Nodal signaling pathways (Brons et al., 2007; Greber et al., 2010; Vallier et al., 2009) but not by STAT3 and bone morphogenetic protein 4 pathways (Nichols and Smith, 2009). These two PSCs exhibit different molecular signatures but still share many important markers. One of the commonly expressed genes in these cells is *Oct4*, which is a PSC and germ cell marker (Scholer et al., 1990). In addition to maintenance of PSCs, *Oct4* alone can transform differentiated cells into PSCs, referred to as induced PSCs (iPSCs) (Kim et al., 2009). The *Oct4* gene contains three distinct *cis*-regulatory elements: the proximal promoter (PP), the distal enhancer (DE), and the proximal enhancer (PE), which differentially control *Oct4* expression during embryogenesis (Yeom et al., 1996). Although *Oct4* is expressed in both naive and primed PSCs, the regulatory mechanism of *Oct4* expression differs between these cell types; *Oct4* expression in naive and primed pluripotent cells is differentially controlled by DE and PE, respectively (Brons et al., 2007; Tesar et al., 2007; Yeom et al., 1996). Accordingly, enhancer activity is altered as primed PSCs are converted into naive PSCs through the induction of extrinsic signaling or genetic modification (Bao et al., 2009; Guo et al., 2009; Hanna et al., 2009).

Two recent reports used the *Oct4*-ΔPE-GFP marker to discriminate naive human PSCs from primed human PSCs (Gafni et al., 2013; Theunissen et al., 2014). However, as shown in flow cytometry data in Theunissen et al. (2014), *Oct4*-ΔPE-GFP reporter activity is not completely negative (including weak GFP activity) in primed human ESCs. Moreover, the *Oct4*-ΔPE-GFP^+ cells may still include ESCs utilizing *Oct4*-PE, since a mono-transgenic system cannot discriminate between cells using only *Oct4*-DE and cells using both *Oct4*-PE and *Oct4*-DE, which may constitute an impure population of naive PSCs. Therefore, in this study we established a dual reporter system for naive and primed mouse pluripotent cells, using two fluorescent reporters, GFP and tdTomato (RFP), controlled by the *cis*-regulatory elements DE and PE, respectively. We found that the expression of *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-RFP accurately represents the expression of naive and primed cells during the development of double transgenic mice. Thus, this double transgenic system can reproduce the in vivo *Oct4* regulatory system, providing a tool for studying the regulation of naive and primed pluripotency and

enabling the separation of pure populations of naive and primed PSCs.

RESULTS

Generation of Dual-Color Fluorescence Transgenic Mice Containing *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-RFP

Oct4 is expressed in both naive and primed PSCs. However, *Oct4* expression in naive and primed pluripotent cells is differentially controlled by two regulatory elements, DE and PE, respectively. We intended to understand how *Oct4* is regulated by DE and PE during development (Figure 1). Therefore, we generated double transgenic mice expressing GFP and RFP under the control of either DE or PE of *Oct4*, respectively. O4-DE-GFP mice carried the *Oct4*-ΔPE-GFP transgene, originally termed OG2 (Szabo et al., 2002), and O4-PE-RFP mice carried the *Oct4*-ΔDE-RFP transgene (Figure 1A). Five O4-PE-RFP founder transgenic mice (two male and three female) were generated (Figure S1). O4-PE-RFP mice were crossed with homozygous O4-DE-GFP mice, and subsequently O4-DE-GFP$^{+/-}$/O4-PE-RFP$^{+/+}$ double transgenic mice were derived (Figure 1A). O4-DE-GFP$^{+/-}$/O4-PE-RFP$^{+/-}$ embryos were obtained from wild-type female mice after crossing them with O4-DE-GFP$^{+/-}$/O4-PE-RFP$^{+/+}$ male mice. All transgenic animals that were studied for their expression had one allele of each transgene (O4-DE-GFP$^{+/-}$/O4-PE-RFP$^{+/-}$). In all transgenic animals both endogenous *Oct4* alleles were present.

O4-DE-GFP and O4-PE-RFP Recapitulate the Stage-Specific Expression of *Oct4* during Mouse Embryo Development

Two-cell-stage embryos did not express either GFP nor RFP (Figure 1B), in agreement with the zygotic genome not being active at this stage. GFP was initially detected in eight-cell embryos and was strongly expressed at the ICM of the blastocyst stage, whereas RFP was not detected even at the blastocyst stage (Figure 1B), indicating that PE is dispensable for *Oct4* expression in the pre-implantation embryo.

Next, we observed the expression of O4-DE-GFP and O4-PE-RFP during the post-implantation stages (6.5–13.5 days post coitum [dpc]). The 5.5- and 6.5-dpc epiblasts were positive both for GFP and RFP (Figures 1C and S2). At 7.25 dpc the intensity of the GFP signal decreased, but the RFP signal remained strong in epiblast cells (Figure 1D). Primordial germ cells (PGCs) were not distinguishable at this stage. However, at 8.5 dpc, GFP-positive cells were localized to the posterior regions of the embryos where the PGCs form a cluster and begin migrating into the genital ridge (Figure 1E). Although RFP-positive cells were detected extensively at the posterior regions of the embryos, these cells did not overlap with the GFP-positive cells, indicating that early PGCs do not require PE for *Oct4* expression. At 9.5 dpc, GFP-positive cells were detected in the hindgut area (Figure 1F). RFP-positive cells disappeared from the soma; however, some cells in the hindgut expressed both RFP and GFP (approximately 34.7%), indicating that migratory PGCs at 9.5 dpc can be divided into two populations: GFP$^+$ and GFP$^+$/RFP$^+$ cells. At the 10.5-dpc stage, when getting close to the genital ridge, most PGCs expressed both *Oct4*-GFP and -RFP (Figure 1G). This was also the case once the PGCs arrived at and proliferated in the gonads (13.5 dpc; Figure 1H). These results demonstrate that the two regulatory regions, DE and PE, dynamically control *Oct4* expression during embryonic development and that founder PGCs use DE while migratory as well as post-migratory PGCs employ both DE and PE to drive *Oct4* expression.

Oct4 has been shown to be expressed in mitotically arrested prospermatogonia and type A spermatogonia, but is downregulated in type B spermatogonia and spermatocytes in adult testis (Pesce et al., 1998). Expression of both GFP and RFP was detected 7 days postpartum (dpp) in the seminiferous tubules of male transgenic testis (Figure 1I). Interestingly, although both GFP$^+$ and RFP$^+$ cells were detected in 4-week-old adult male mouse testis, only GFP$^+$ cells were localized to the periphery (near the basement membrane) of the seminiferous tubules while RFP$^+$ cells were detected at the center of the seminiferous tubules (Figure 1J). Immunohistochemistry analysis confirmed that the GFP$^+$ cells are present at the periphery (type A spermatogonia niche) of the seminiferous tubules and that the RFP$^+$ cells are detected at the center (differentiated germ cell niche) of the seminiferous tubules (Figure S3). These results indicate that type A spermatogonia at the pre-meiotic division stage (7 dpp) can use both DE and PE to express *Oct4* but that in adult testis only DE drives *Oct4* expression in type A spermatogonia of 4-week-old mice. Committed germ cells located near the lumen also express *Oct4*, which is controlled by PE.

In this study, we focused on spermatogenic cells since RFP was not detected in the transgenic ovary. We did not detect transgenic expression in somatic cells. For example, we detected neither a GFP nor an RFP signal in neural stem cells (NSCs) (Figure 1K), demonstrating that our transgenic system is specific for the germline, i.e., PSCs and germ cells, but not for somatic stem cells. Taken together, our results indicate that DE and PE activities stage-specifically regulate expression of *Oct4* in totipotent/pluripotent cells and germ cells in the developing mouse embryo.

Derivation of ESCs from Dual-Color Fluorescence Transgenic Blastocysts

Previous studies have suggested that ESCs grown under conventional ESC culture conditions are a heterogeneous

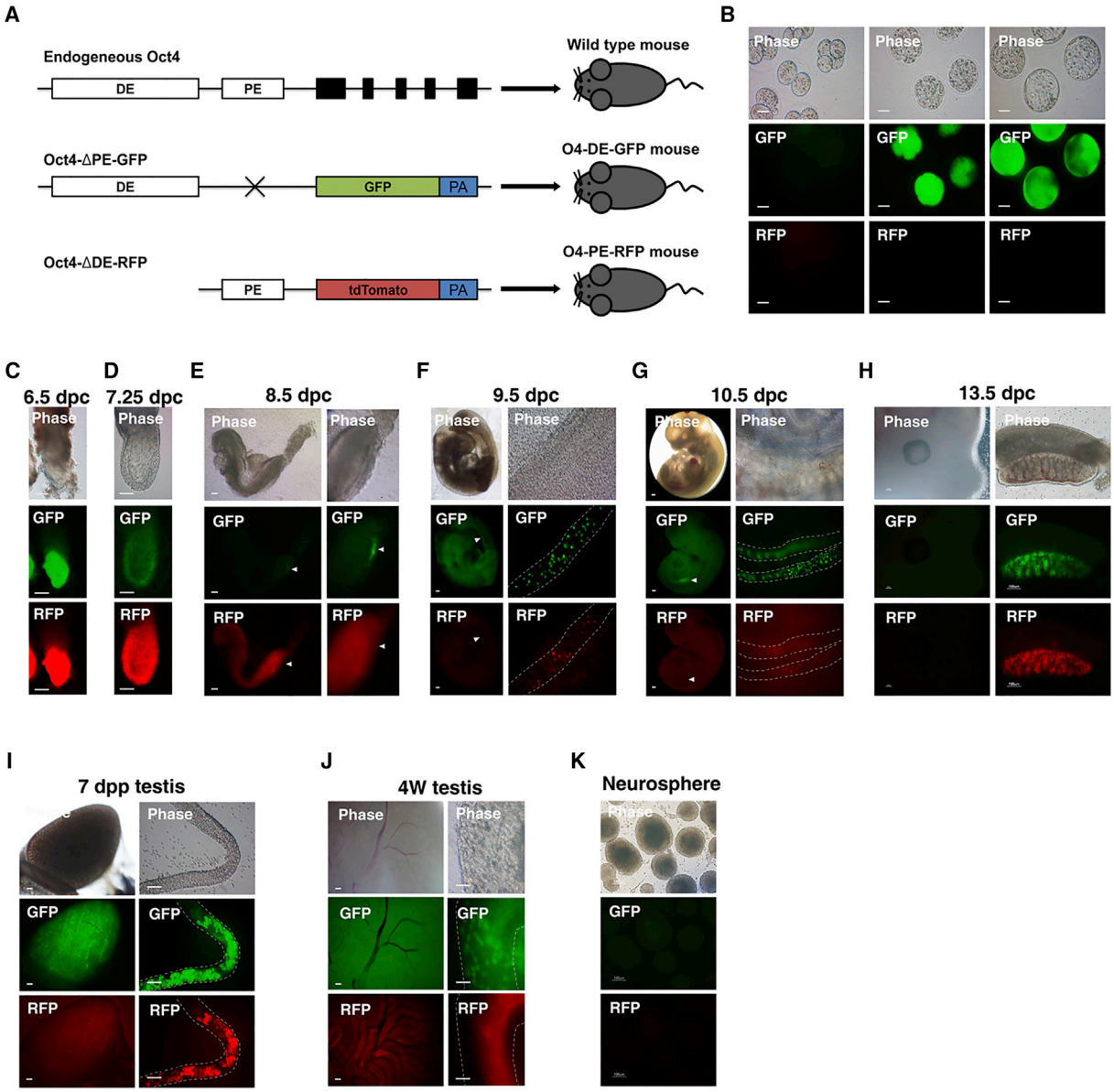

Figure 1. Generation of Dual Transgenic Mice (O4-DE-GFP/O4-PE-RFP) and the Distinct *Oct4* Regulatory Elements in the Totipotent Cycle

(A) Physical maps of wild-type endogenous *Oct4*, *Oct4*-ΔPE-GFP (O4-DE-GFP), and *Oct4*-ΔDE-RFP (O4-PE-RFP).

(B) The expression pattern of O4-DE-GFP and O4-PE-RFP in pre-implantation embryos (2C, blastocyst). The O4-DE-GFP/O4-PE-RFP pre-implantation embryos expressed only O4-DE-GFP. Scale bars, 50 μm.

(C–H) The expression pattern of O4-DE-GFP and O4-PE-RFP in post-implantation embryos. Phase and fluorescence images of (C) 6.5-dpc embryo (scale bar, 100 μm); (D) 7.25-dpc embryo (scale bar, 100 μm); (E) 8.5-dpc embryo (scale bar, 100 μm); (F) 9.5-dpc embryo and migrating PGCs (scale bar, 100 μm); (G) 10.5-dpc embryo and PGCs (scale bar, 100 μm); and (H) 13.5-dpc embryo and gonad (scale bar, 100 μm). Arrowheads indicate GFP-positive areas in the developing gonads.

(I and J) The expression pattern of O4-DE-GFP and O4-PE-RFP in testis. Phase and fluorescence image of (I) 7-dpp testis and seminiferous tubules and (J) 4-week-old adult mouse testis and seminiferous tubules. Scale bars, 100 μm.

(K) The neurosphere did not express either O4-DE-GFP or O4-PE-RFP. Scale bar, 100 μm.

population (Hayashi et al., 2008; Martinez Arias and Brickman, 2011; Miyanari and Torres-Padilla, 2012). Thus, we attempted to verify the heterogeneity of ESCs and derive a pure population of naive pluripotent ESCs, using our double transgenic system. Blastocysts were plated on a feeder-layered dish in conventional mouse ESC medium (serum + leukemia inhibitory factor [LIF]) (Figures 2A and 2B). Initially the ICM of blastocysts expressed only GFP, whereas RFP was expressed after 3 days. As the ICM outgrowths expanded, the GFP$^+$RFP$^-$ cells became GFP$^+$RFP$^+$ cells (Figure 2B). This result supports a previous study demonstrating that pluripotent ESC derivation requires epiblast specification (Boroviak et al., 2014). This in vitro culture system indirectly suggests that during the 4.5- to 5.5-dpc embryonic development, GFP$^+$RFP$^-$ cells may be gradually changed into GFP$^+$RFP$^+$ cells in vivo. Trypsinized ICM outgrowth cells were transferred into a new dish with serum + LIF medium where two distinct ESC populations were cultured (GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$; Figure 2C). Fluorescence-activated cell sorting analysis showed that during the initial stage of ESC derivation in serum + LIF the GFP$^+$RFP$^-$ cells, presumably naive PSCs, constituted only 72.7%, while GFP$^+$RFP$^+$ and GFP$^-$RFP$^+$ cells constituted 22.1% and 0.49%, respectively, of the population (Figure 2D). When GFP$^+$RFP$^-$ cells were sorted and recultured in serum + LIF medium, GFP$^+$RFP$^-$ cells reverted to GFP$^+$RFP$^+$ subsequent to ten passages (Figure 2E). In contrast, when GFP$^+$RFP$^+$ cells were sorted and cultured in serum + LIF medium, GFP$^+$RFP$^+$ cells reverted to GFP$^+$RFP$^-$ cells subsequent to ten passages (Figure 2E). These results demonstrate that the ESCs in serum + LIF medium were heterogeneous; they contained two populations of cell types that were interconvertible. These GFP$^+$RFP$^+$ and GFP$^+$RFP$^-$ cell populations could not have been distinguished using an *Oct4*-ΔPE-GFP mono-transgenic system.

Next, we modified the culture conditions by adding two inhibitors (2i), ERK1/2 (PD0325901) and GSK3 (CHIR99021), into the serum + LIF medium. Previous reports have demonstrated that these 2i contributed to the establishment of naive PSCs and shielded PSCs from differentiation triggers (Ying et al., 2008). Under serum + LIF + 2i conditions, only 76.0% of cells were GFP$^+$RFP$^-$ and 23.0% were GFP$^+$RFP$^+$ (Figures 2F and 2G). When the GFP$^+$RFP$^-$ cells were sorted and cultured for ten passages under serum + LIF + 2i conditions, 3.45% became GFP$^+$RFP$^+$ (Figure 2H, left panel). In contrast, the sorted GFP$^+$RFP$^+$ cells also converted into GFP$^+$RFP$^-$ cells (70.1%) (Figure 2H, right panel). These findings indicate that serum + LIF + 2i culture conditions are insufficient for the maintenance of GFP$^+$RFP$^-$ cells. Thus, ESCs are metastable and interconvertible between GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells in serum-containing medium. Next, we used the new culture medium to obtain a pure population of naive PSCs. Serum was removed from the medium since mouse ESCs exhibited elevated levels of ERK1/2 phosphorylation, a key event in priming of ESCs for differentiation (Kunath et al., 2007; Wray et al., 2011; Yamaji et al., 2013), in serum-containing medium. Serum-free ESC culture medium (N2B27 medium) supplemented with 2i in combination with LIF (Marks et al., 2012; Silva et al., 2008; Ying et al., 2008) was used for ESC culturing. ESCs (from serum + LIF culture) were transferred into N2B27 + LIF + 2i medium. On day 2 after transfer to the N2B27 + LIF + 2i medium (Figure 2I), most GFP$^+$RFP$^+$ cells were converted into GFP$^+$RFP$^-$ cells (98.9%; Figure 2J). Subsequent to further culturing, nearly all cells were maintained as GFP$^+$RFP$^-$ cells (94.3%; Figure 2K). Next, the ESCs (from N2B27 + LIF + 2i medium) were transferred into serum + LIF medium. Subsequent to 2 days of culturing, nearly all of the GFP$^+$RFP$^-$ cells were converted into GFP$^+$RFP$^+$ cells (67.2%; Figure 2L). Taken together, our double transgenic system unequivocally shows that the ESCs in serum-containing medium were heterogeneous. This heterogeneous ESC population became homogeneous subsequent to culturing in serum-free medium supplemented with 2i and LIF. Therefore, serum-free culture condition is an essential requirement for culturing a pure population of naive PSCs.

Distinct Gene Expression Patterns in 2i-GFP-, GFP-, and GFP/RFP-Positive ESCs

Oct4 expression is controlled by the DE in naive PSCs, while the PE regulates *Oct4* expression in primed PSCs (Brons et al., 2007; Tesar et al., 2007; Yeom et al., 1996). Therefore, we investigated whether the expression of *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-RFP accurately represents the naive and primed PSC states in an in vitro culture system. We compared the global gene expression patterns of GFP$^+$RFP$^-$ (from N2B27 + LIF + 2i and serum + LIF medium) and GFP$^+$RFP$^+$ cells (from serum + LIF) using microarray analysis (Illumina MouseRef-8 v2 Expression BeadChip). Surprisingly, the gene expression profile of GFP$^+$RFP$^-$ cells cultured in N2B27 + LIF + 2i (2i-GFP$^+$RFP$^-$) was distinct from that of GFP$^+$RFP$^-$ cells cultured in serum + LIF (Figure 3A). The hierarchical clustering and multidimensional scaling (MDS) plot analyses showed that the gene expression pattern of GFP$^+$RFP$^-$ cells was more similar to GFP$^+$RFP$^+$ cells than to 2i-GFP$^+$RFP$^-$ cells (Figures 3B and 3C). Pairwise scatter-plot analyses also demonstrated that GFP$^+$RFP$^+$ cells were almost identical to GFP$^+$RFP$^-$ cells (0.998), which were distinct from 2i-GFP$^+$RFP$^-$ cells (Figures 3D–3F). Differentially expressed genes in GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells were largely associated with germline development; GFP$^+$RFP$^-$ cells highly expressed *Wnt3a*, *Rhox5*, *Rhox6*, *Rhox9*, *Nanos3*, and *Tcfap2c* (Figure 3G). We found that 249 genes were upregulated and 446 genes were downregulated in 2i-GFP$^+$RFP$^-$ versus GFP$^+$RFP$^-$ cells

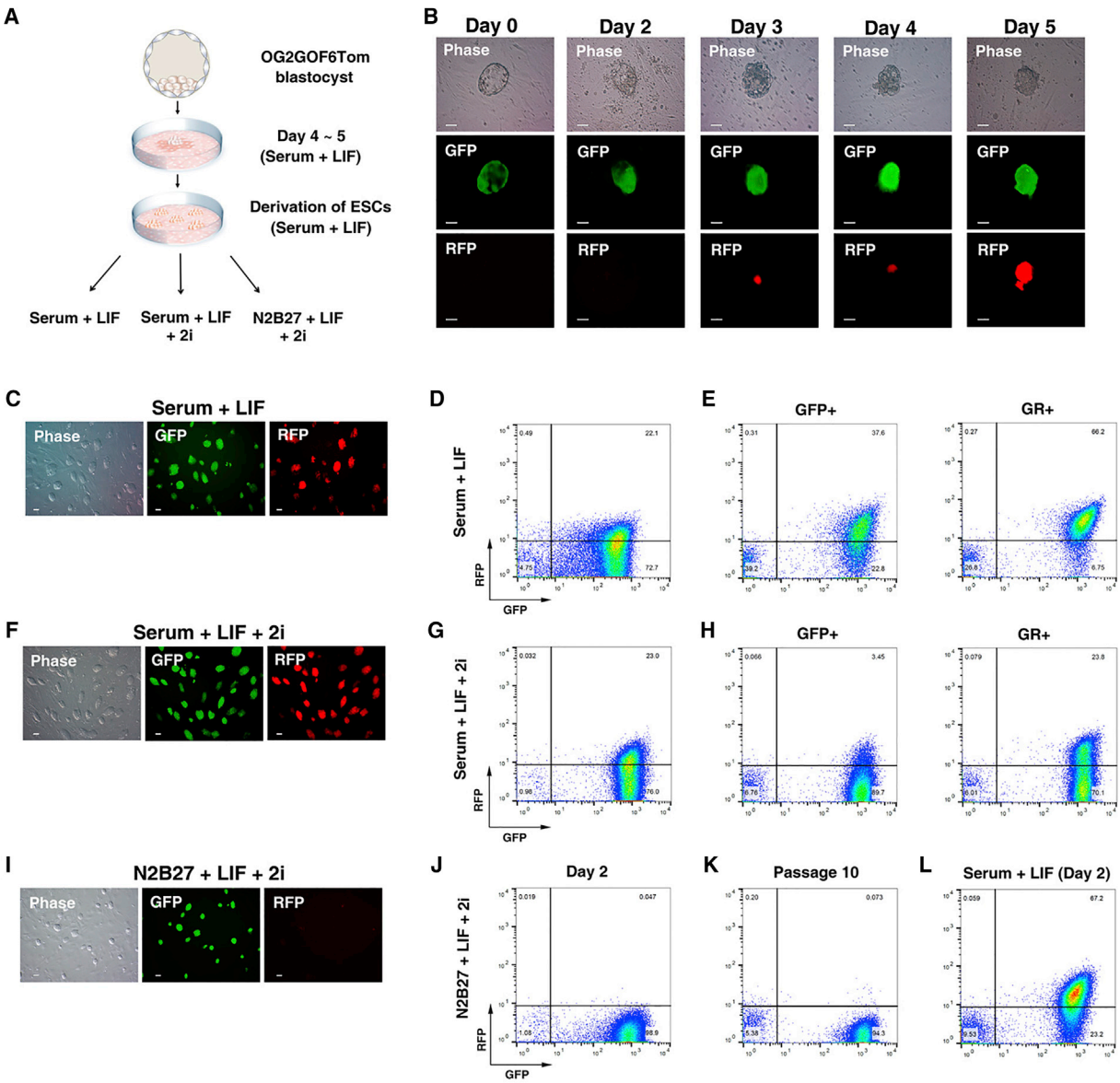

Figure 2. ESC Culture Conditions Influence the Activity of *Oct4* Enhancers

(A and B) Derivation of ESCs from 04-DE-GFP/04-PE-RFP blastocysts. 04-PE-RFP was initially expressed during the establishment of ESCs. Scale bars, 50 μm.

(C) Phase and fluorescence images of 04-DE-GFP/04-PE-RFP ESCs in serum + LIF medium. Scale bars, 50 μm.

(D) Flow cytometry analysis of the proportion of 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells in serum + LIF medium.

(E) Flow cytometry analysis of the proportion of 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells from sorted 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells cultured in serum + LIF medium for ten passages.

(F) Phase and fluorescence images of 04-DE-GFP/04-PE-RFP ESCs in serum + LIF + 2i medium. Scale bars, 50 μm.

(G) Flow cytometry analysis of the proportion of 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells in serum + LIF + 2i medium.

(H) Flow cytometry analysis of the proportion of 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells from sorted 04-DE-GFP$^+$only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells cultured in serum + LIF + 2i medium for ten passages.

(I) Phase and fluorescence images of 04-DE-GFP/04-PE-RFP ESCs in N2B27 + LIF + 2i medium. Scale bars, 50 μm.

(J and K) Flow cytometry analysis of the proportion of 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells cultured in N2B27 + LIF + 2i medium for 2 days (J) or ten passages (K).

(L) Flow cytometry analysis of the proportion of 04-DE-GFP$^+$ only cells or 04-DE-GFP$^+$ and 04-PE-RFP$^+$ cells cultured in serum + LIF medium for 2 days from N2B27 + LIF + 2i medium.

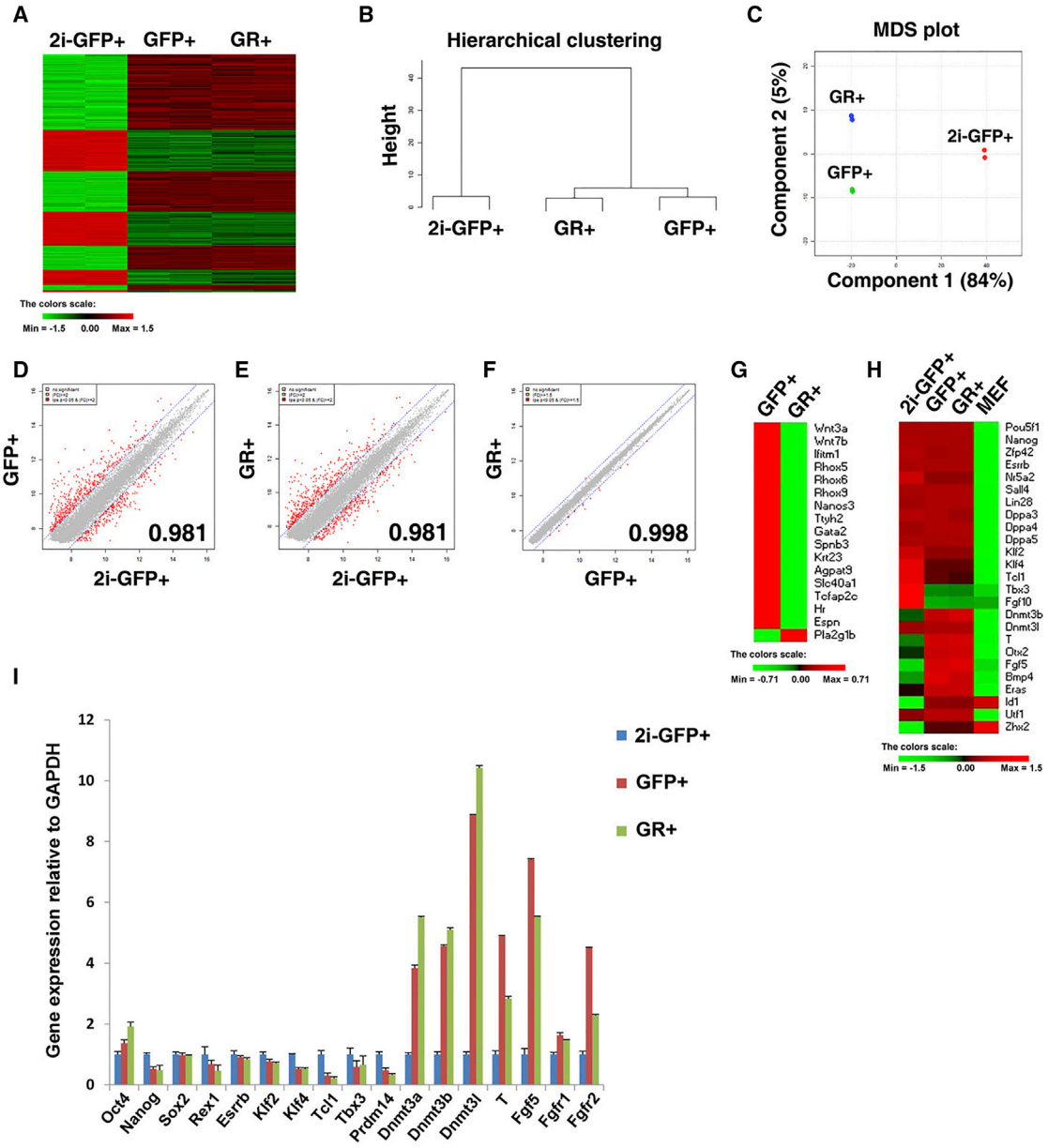

Figure 3. 2i-GFP-Positive Cells Cultivated in N2B27 + LIF + 2i Medium and GFP- or GR-Positive Cells Cultivated in Serum + LIF Medium Exhibit Distinct Gene Expression Patterns

(A) Heatmap of global gene expression patterns in 2i-GFP$^+$, GFP$^+$RFP$^-$ (GFP), and GFP$^+$RFP$^+$ (GR) cells.

(B and C) Hierarchical clustering (B) and MDS plot (C) analysis showing that 2i-GFP-positive cells are distinct from GFP- or GR-positive cells.

(D–F) Scatter plots of global gene expression comparing (D) 2i-GFP- and GFP-positive cells, (E) 2i-GFP- and GR-positive cells, and (F) GFP- and GR-positive cells.

(G) Heatmap analysis shows that GFP-positive cells express germ cell-related genes including *Wnt3a*, *Rhox5*, *Rhox6*, *Rhox9*, *Nanos3*, and *Tcfap2c* at a higher level than do GR-positive cells.

(H) Heatmap analysis showing that 2i-GFP-positive cells highly express naive pluripotency-related genes and GFP- and GR-positive cells highly express differentiation markers.

(I) Quantitative gene expression analysis of pluripotency markers (*Oct4*, *Nanog*, *Sox2*, *Esrrb*), naive pluripotency markers (*Rex1*, *Klf2*, *Klf4*, *Tcl1*, *Tbx3*, *Prdm14*), and differentiation markers (*Dnmt1*, *Dnmt3a*, *Dnmt3b*, *T*, *Fgf5*). Data are presented as mean ± SEM for n = 3 independent experiments. *GAPDH* was used as a housekeeping gene.

(Figure S4A). The expression levels of pluripotency factors such as *Pou5f1* (*Oct4*), *Nanog*, *Zfp42* (*Rex1*), *Esrrb*, *Sall4*, and *Lin28* were almost identical in 2i-GFP$^+$RFP$^-$, GFP$^+$RFP$^-$, and GFP$^+$RFP$^+$ cells (Figure 3H). However, the naive pluripotency markers, *Klf4*, *Tcl1*, and *Tbx3*, were more highly expressed in 2i-GFP$^+$RFP$^-$ than in GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells. In contrast, differentiation-related genes such as *Dnmt3b*, *Brachyury* (*T*), *Otx2*, and *Fgf5* were more highly upregulated in GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells than in 2i-GFP$^+$RFP$^-$ cells. These gene expression patterns were further confirmed by qRT-PCR analysis (Figure 3I). The expression levels of the pluripotency markers (*Oct4*, *Nanog*, *Sox2*, *Rex1*, *Esrrb*, and *Klf2*) in 2i-GFP$^+$RFP$^-$ cells were similar to those in GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells (Figure 3I). However, while naive pluripotency markers were upregulated in 2i-GFP$^+$RFP$^-$ cells, the development-related genes were downregulated (Figure S4B). These findings suggest that although GFP$^+$RFP$^-$ cells display more naive characteristics (such as higher expression of germ cell markers) under serum + LIF culture conditions, they are in a metastable state and are readily converted into GFP$^+$RFP$^+$ cells. Stable naive PSCs could be maintained only in serum-free N2B27 + LIF + 2i medium, since 2i-GFP$^+$RFP$^-$ cells were found to be distinct from GFP$^+$RFP$^-$ cells cultured in serum-containing medium and expressed a greater number of naive pluripotency markers.

Functional annotation clustering of differentially expressed genes using gene ontology (GO) analysis revealed that the upregulated genes (>2-fold) in 2i-GFP$^+$RFP$^-$ cells were significantly enriched for GO terms linked to "sterol biosynthetic and metabolic process," "cholesterol biosynthetic and metabolic process," and "lipid biosynthetic process" (Marks et al., 2012) (Table S1). The upregulated genes in GFP$^+$RFP$^-$ or GFP$^+$RFP$^+$ cells were enriched for GO terms linked to "gland, tube development, and developmental growth," "collagen metabolic and catabolic process," "membrane organization," and "germ cell development" (Table S2).

GFP$^-$RFP$^+$ Cells Represent Primed PSCs

The results of our study show that GFP$^+$RFP$^-$ ESCs stably maintain a naive pluripotent state when cultured in serum-free N2B27 + LIF + 2i medium. Next, we sought to identify primed PSCs using the double transgenic system. Since the PE regulates *Oct4* expression in EpiSCs (Brons et al., 2007; Tesar et al., 2007), primed PSCs must be GFP$^-$RFP$^+$ cells and naive PSCs must be GFP$^+$RFP$^-$ cells. During EpiSC derivation from O4-DE-GFP$^{+/-}$/O4-PE-RFP$^{+/-}$ epiblasts, the initial expression of GFP decreased and only RFP$^+$ cells were expanded (Figures S5A and S5B). Next, we tried to differentiate GFP$^+$RFP$^+$ ESCs into GFP$^-$RFP$^+$ EpiSC-like cells (EpiLCs). The GFP$^+$RFP$^+$ ESCs changed morphologically to form flat colonies that became GFP$^-$RFP$^+$ cells (Figure 4A). Heatmap and hierarchical clustering analyses showed that the gene expression patterns of GFP$^-$RFP$^+$ cells were more similar to EpiSCs than to ESCs (Figures 4B and 4C). The GFP$^-$RFP$^+$ cells were more similar to EpiSCs than to 2i-GFP$^+$, GFP$^+$RFP$^-$, or GFP$^+$RFP$^+$ cells by MDS plot and scatter-plot analyses (Figures 4D–4G). These results demonstrate that transition of enhancer activity from DE to PE parallels the conversion from the naive to primed pluripotent state. Next, we examined the expression of core pluripotency-, naive pluripotency-, and differentiation-related genes in GFP$^-$RFP$^+$ EpiLCs (Figure 4H). *Oct4* and *Nanog* were highly expressed in EpiLCs, EpiSCs, and all ESC lines (Figure 4H). However, GFP$^-$RFP$^+$ EpiLCs expressed very low levels of the naive pluripotency-related genes *Klf2*, *Klf4*, *Klf5*, *Dppa3*, *Dppa4*, *Zfp42*, *Tbx3*, and *Tcl1*, as shown for the EpiSCs. However, the differentiation-related genes Krt18, *Fgf5*, *Fgf8*, *Otx2*, *T* (*Brachyury*), and *Nestin* were highly expressed in GFP$^-$RFP$^+$ EpiLCs and EpiSCs. qRT-PCR analysis confirmed the high expression of the core pluripotency genes (*Oct4* and *Nanog*) and EpiSC markers (*T* and *Fgf5*), and low expression of *Sox2*, *Klf2*, and *Klf4* in GFP$^-$RFP$^+$ EpiLCs and EpiSCs (Figure 4G). These data show that GFP$^-$RFP$^+$ EpiLCs are similar to EpiSCs in terms of morphology and gene expression profiles.

Next, we investigated the DNA methylation status at the promoter regions of *Nanog*, *Stella* (*Dppa3*), and *Dppa5* in GFP$^+$RFP$^-$, GFP$^+$RFP$^+$, 2i-GFP$^+$RFP$^-$, and GFP$^-$RFP$^+$ EpiLCs (Figures 4J–4L). The *Nanog* promoter regions were hypomethylated in all samples (Figure 4J). However, the promoter regions of *Stella* and *Dppa5* were hypermethylated in GFP$^-$RFP$^+$ EpiLCs (Figures 4K and 4L). *Long interspersed nuclear element 1* (*LINE1*) and *Intracisternal A particle* (*IAP*) have been shown to be hypermethylated in primed PSCs (Yamaji et al., 2013). Thus, we examined the DNA methylation status of *LINE1* and *IAP* in GFP$^-$RFP$^+$ EpiLCs compared with GFP$^+$RFP$^-$, GFP$^+$RFP$^+$, and 2i-GFP$^+$RFP$^-$ cells (Figures 4M and 4N). The *LINE1* regions were more hypermethylated in GFP$^-$RFP$^+$ EpiLCs than in GFP$^+$ (GFP$^+$RFP$^-$, GFP$^+$RFP$^+$, and 2i-GFP$^+$RFP$^-$) cells (Figure 4M). The *IAP* regions were more methylated in RFP$^+$ cells (GFP$^+$RFP$^+$ and GFP$^-$RFP$^+$ EpiLCs) than in GFP$^+$RFP$^-$ and 2i-GFP$^+$RFP$^-$ cells (Figure 4N). Taken together, these results indicate that the *Oct4*-ΔPE-RFP reporter represents an applicable primed PSC marker. Thus, the conversion of ESCs into EpiLCs entails a shift from GFP$^+$ to RFP$^+$ (i.e., from DE to PE regulation), which also coincides with a change in transcriptome and epigenetic status specific for the primed pluripotent state.

Developmental Potential of GFP$^-$RFP$^+$ Cells

The ability of chimera formation is a strict criterion that distinguishes naive PSCs from primed PSCs (Chenoweth et al., 2010; Nichols and Smith, 2009, 2011). Naive

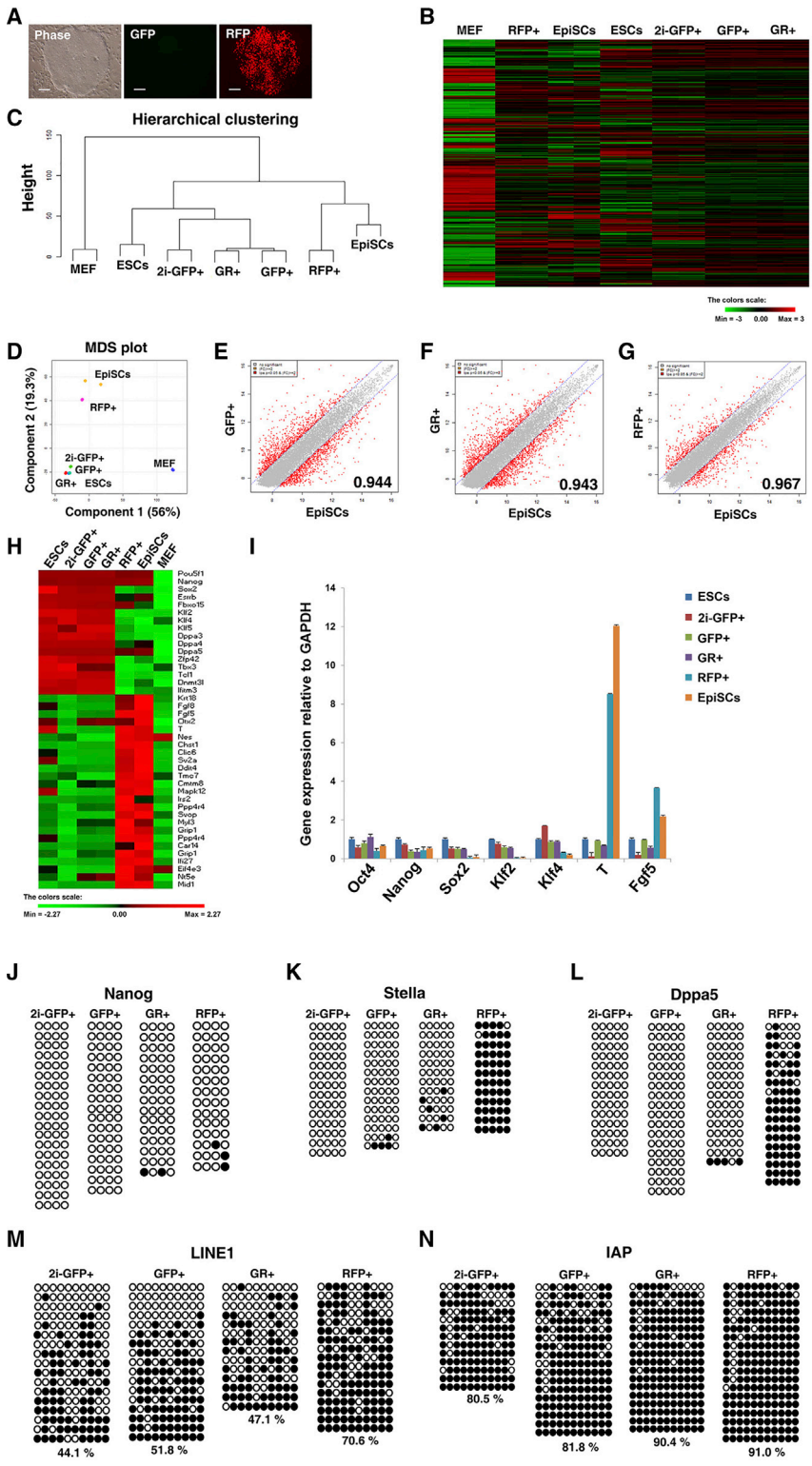

Figure 4. *Oct4*-ΔDE-RFP-Positive Cells Constitute Primed Pluripotent Stem Cells, as Determined by Gene Expression Patterns and Epigenetic Status

(A) Phase and fluorescence images of EpiSC-like cells (EpiLCs) from 04-DE-GFP/04-PE-RFP ESCs. These EpiLCs express *Oct4*-ΔDE-RFP but not *Oct4*-ΔPE-GFP. Scale bars, 100 μm.

(B) Heatmap of global gene expression patterns in 2i-GFP$^+$, GFP$^+$RFP$^-$ (GFP$^+$), GFP$^+$RFP$^+$ (GR), GFP$^-$RFP$^+$ (RFP) cells, ESCs, EpiSCs, and mouse embryonic fibroblasts (MEF).

(C and D) Hierarchical clustering and MDS plot analyses showing that RFP-positive cells are more similar to EpiSCs than to ESCs, 2i-GFP$^+$, GFP$^+$, or GR$^+$ cells.

(E–G) Scatter plots of global gene expression comparing (E) EpiSCs and GFP-positive cells, (F) EpiSCs and GR-positive cells, and (G) EpiSCs and RFP-positive cells.

(H and I) Heatmap (H) and quantitative gene expression (I) analysis show that RFP-positive cells do not express naive pluripotency-related genes but highly express primed pluripotency-related genes. Data are presented as mean ± SEM for n = 3 independent experiments. *GAPDH* was used as a housekeeping gene.

(J–N) Bisulfite genomic sequencing of the *Nanog* (J), *Stella* (K), *Dppa5* (L), *LINE* (M), and *IAP* (N) regions in 2i-GFP$^+$, GFP$^+$, GR$^+$, and RFP$^+$ cells. Black and white circles represent methylated and unmethylated CpGs, respectively.

pluripotent ESCs contribute to chimeras; however, primed pluripotent EpiSCs barely contribute to chimeric embryos subsequent to blastocyst injection followed by transfer to a surrogate mother (Brons et al., 2007). GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells as well as GFP$^-$RFP$^+$ EpiLCs were aggregated with wild-type eight-cell embryos and further cultured

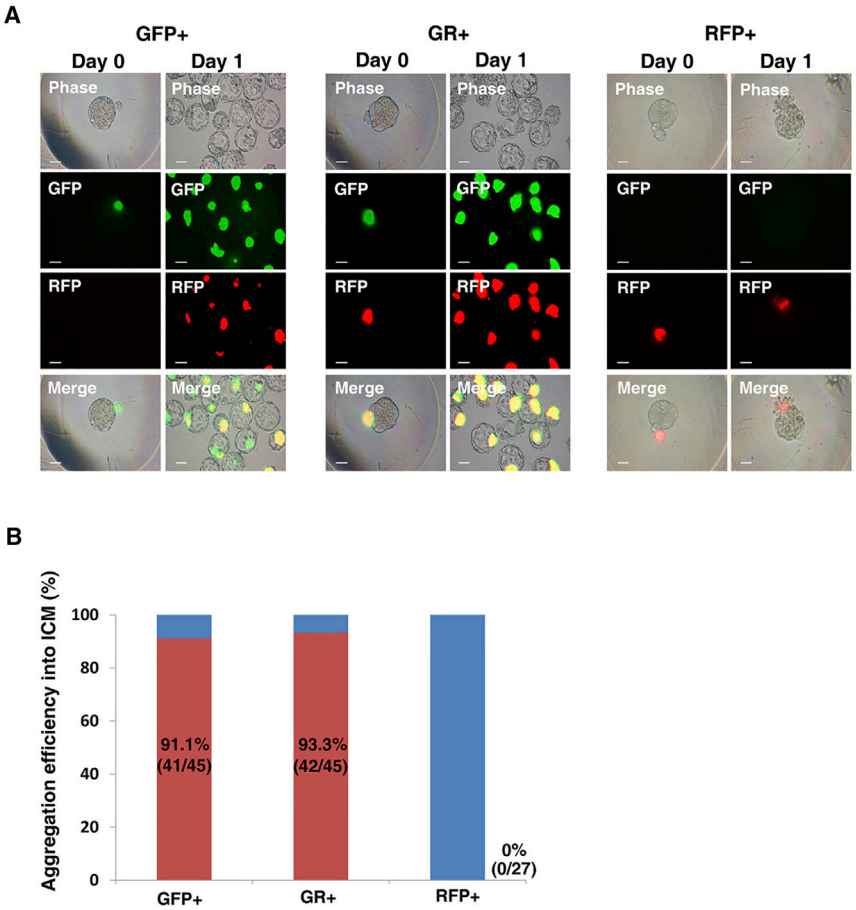

Figure 5. In Vivo Development Potential of *Oct4*-ΔDE-RFP-Positive Cells

(A) Aggregation of GFP-, GR-, and RFP-positive cells with normal embryo. RFP-positive cells did not incorporate into the embryos. Scale bars, 50 μm.

(B) Aggregation efficiency of GFP-, GR-, and RFP-positive cells.

until the blastocyst stage. The GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells were efficiently incorporated into the ICM of blastocysts (Figures 5A and 5B). Interestingly, at day 1 post aggregation, incorporated GFP$^+$RFP$^-$ cells co-expressed RFP, similar to GFP$^+$RFP$^+$ cells (Figure 5A). The aggregation efficiency of GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells was indistinguishable at 91.1% (41/45) and 93.3% (42/45), respectively (Figure 5B). However, basically all GFP$^-$RFP$^+$ EpiLCs failed to incorporate into the ICM of blastocysts, the aggregation efficiency being 0% (0/27) (Figures 5A and 5B). These results indicate that GFP$^+$ cells exhibit characteristics of naive pluripotency and GFP$^-$RFP$^+$ EpiLCs demonstrate those of primed pluripotency-like EpiSCs.

Control of the DE and PE Elements in Naive and Primed PSCs

The activity of the *Oct4* enhancers in naive and primed PSCs was evaluated in luciferase assays showing that 2i-GFP$^+$ cells mainly use DE for *Oct4* expression while PE was completely inactive; PE is slightly active in the control of ESCs cultured in serum + LIF medium (Figure 6A). GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells in serum + LIF medium also mainly used DE, but the activity of PE was higher than in 2i-GFP$^+$ cells. In contrast, the GFP$^-$RFP$^+$ cells mainly utilized PE for *Oct4* expression, similar to EpiSCs and P19 embryonic carcinoma cells (Figure 6A). These results confirm that the expression of *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-RFP accurately represents the naive and primed PSC state; thus, the double transgenic reporter system can be used as a tool for purifying populations of naive (2i-GFP$^+$ cells) and primed (GFP$^-$RFP$^+$ cells) PSCs, since the regulatory mechanisms underlying the two different pluripotent states can be precisely defined.

Recent studies have shown that histone modifications are closely associated with the activity of enhancers (Creyghton et al., 2010; Favaedi et al., 2012). Acetylation of histone H3 lysine 27 (H3K27) is an indicator of active enhancers (Bonn et al., 2012), and deacetylation of H3K27 is associated with decreased gene expression or poised enhancers. H3K27me3 and H3K9me3 have also been suggested to be poised enhancer marks (Creyghton et al., 2010; Favaedi et al., 2012). H3K9me3 can distinguish poised from active enhancers independently of H3K27me3 (Zentner et al., 2011). Thus, we investigated the chromatin enrichment state of H3K27ac, H3K27me3, and H3K9me3 in 2i-GFP$^+$, GFP$^+$RFP$^-$, GFP$^+$RFP$^+$, and

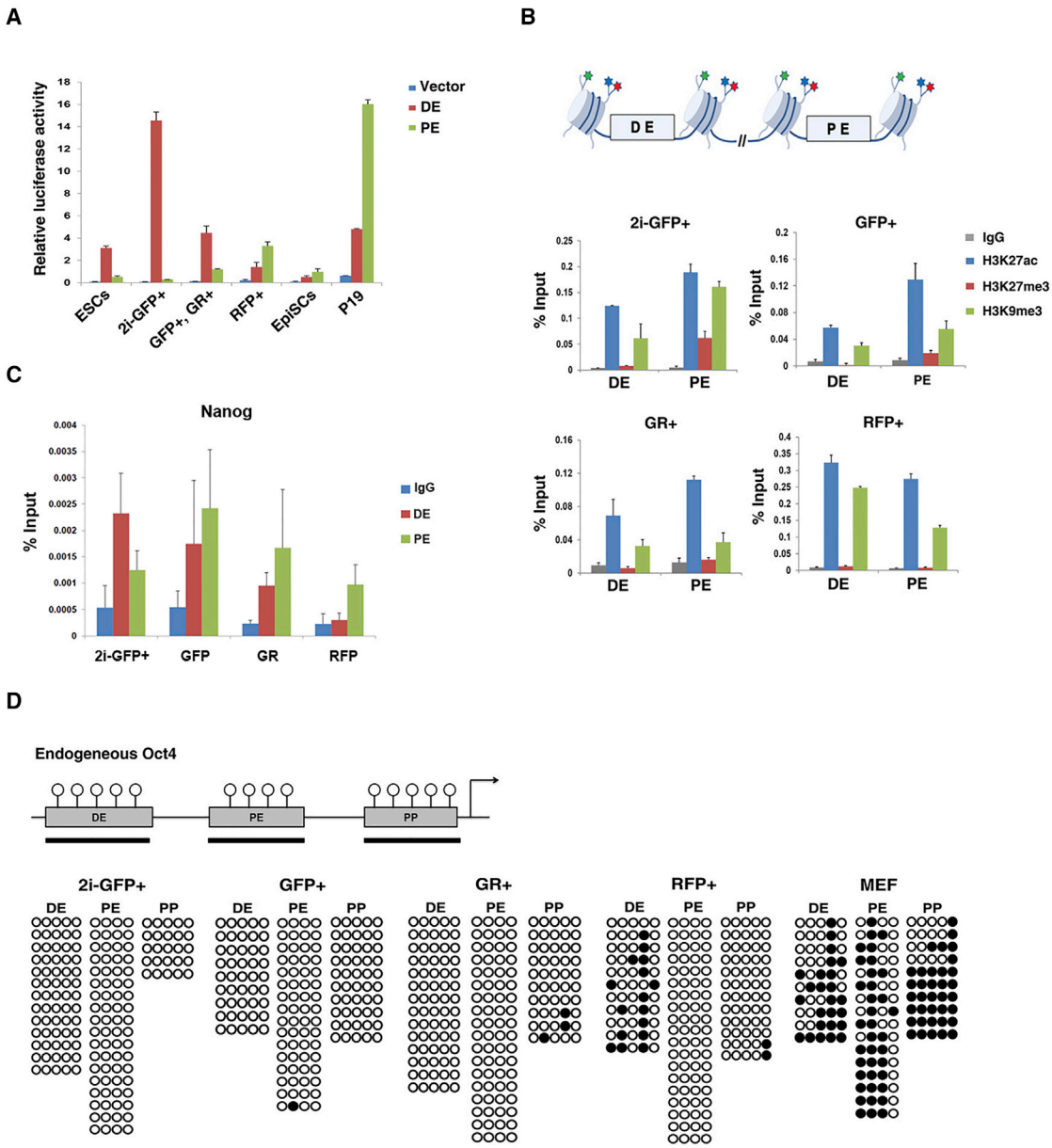

Figure 6. Epigenetic Status of *Oct4* Regulatory Elements in Naive and Primed Pluripotent Stem Cells

(A) Analysis of *Oct4* enhancer activity in ESCs, 2i-GFP$^+$ cells, GFP$^+$ GR$^+$ cells, RFP$^+$ cells, EpiSCs, and P19 embryonic carcinoma cells. Relative luciferase activity was normalized to the activity of the empty vector. Data are presented as mean ± SEM for n = 3 independent experiments.

(B) ChIP-qPCR analysis to determine immunoglobulin G (IgG), H3K27ac, H3K27me3, and H3K9me3 enrichment on the *Oct4* distal and proximal enhancers. Data are presented as mean ± SEM for n = 3 independent experiments. Coloured stars indicate histone tails, such as H3K9 and H3K27.

(C) ChIP-qPCR analysis to determine IgG and Nanog enrichment on the *Oct4* distal and proximal enhancers. Data are presented as mean ± SEM for n = 3 independent experiments.

(D) Bisulfite genomic sequencing of the regions of the *Oct4* DE, PE, and PP in 2i-GFP$^+$, GFP$^+$, GR$^+$, and RFP$^+$ cells, and MEF. Black and white circles represent methylated and unmethylated CpGs, respectively.

GFP$^-$RFP$^+$ cells (Figure 6B) using chromatin immunoprecipitation (ChIP)-qPCR analysis. The results of the ChIP-qPCR analysis showed that H3K27ac was highly enriched on the DE and PE of *Oct4* in all *Oct4*-expressing cells, naive (2i-GFP$^+$), metastable (GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$), and primed (GFP$^-$RFP$^+$) PSCs (Figure 6B). The poised enhancer marks H3K27me3 and H3K9me3 were enriched on PE rather than on DE in 2i-GFP$^+$ cells. The level of H3K27me3 on the PE was slightly higher than that on DE in metastable cells (GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$), but the level of H3K9me3 on the DE and PE did not differ significantly. Therefore, in the metastable state, epigenetic marks seem to fluctuate and the activity of PE and DE cannot be distinguished by histone marks. In primed pluripotent GFP$^-$RFP$^+$ cells, H3K9me3 was enriched on DE rather than on PE while the level of H3K27me3 was not different. These results suggest that enrichment of H3K9me3 and H3K27me3 on PE indicates the naive pluripotent state and that enrichment of H3K9me3 on DE marks the primed pluripotent state.

Transcription factors can regulate *Oct4* expression by binding to DE and PE (Wu and Scholer, 2014). Galonska et al. (2015) showed that 2i condition alters the binding enrichment of Nanog on PE in ESCs. Thus we next checked the binding profiles of Nanog on PE and DE of *Oct4* in different cell populations. ChIP-qPCR analysis showed that Nanog was highly enriched on the DE rather than on PE in 2i-GFP$^+$ cells, naive PSCs (Figure 6C). On the other hand, Nanog was more enriched on the PE in GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells. These results suggest that the differential activity of DE and PE is closely related to binding affinity of transcription factors.

Since both DNA methylation and histone modification regulate gene expression and affect each other (Cedar and Bergman, 2009), we next examined the DNA methylation pattern of the *Oct4* regulatory elements. The DNA methylation status of DE, PE, and PP of *Oct4* were analyzed by bisulfite DNA sequencing analysis (Figure 6D). The DE, PE, and PP of *Oct4* in naive PSCs (2i-GFP$^+$), GFP$^+$RFP$^-$, and GFP$^+$RFP$^+$ cells were completely unmethylated. However, primed PSCs (GFP$^-$RFP$^+$) showed relatively hypermethylated patterns in the DE of *Oct4*, indicating that the DE of *Oct4* in primed PSCs is regulated by DNA methylation as well as H3K9me3 (Figures 6B–6D).

DISCUSSION

In this study, we show that DE and PE coordinately regulate *Oct4* expression during embryonic and germ cell development in vivo and naive and primed pluripotency in vitro. The cell-specific and temporal control of DE and PE could be precisely determined using the dual reporter system. Approximately two decades ago, we first determined the activity of the genomic fragment containing *Oct4* using a β-galactosidase reporter and found that DE activity was ESC- and germ cell-specific and that PE was epiblast- and embryonic carcinoma cell (P19)-specific (Yeom et al., 1996). *Oct4*-GFP transgenic mice have been generated and used for the detection of PGCs and PSCs and for monitoring reprogramming through nuclear transfer, cell fusion, and the transduction of reprogramming factors (Boiani et al., 2002; Choi et al., 2014; Do and Schöler, 2004). However, these existing transgenic systems cannot concurrently discriminate between the activity of DE and PE. In this study, a developed dual reporter system allowed us to monitor the DE, PE, and DE + PE activity simultaneously.

Enhancers play a central role in cell-type-specific and stage-specific regulation of gene expression and are capable of acting over a distance ranging from several to hundreds, and in rare cases even thousands of kilobases from their target genes (Bulger and Groudine, 2011; Ong and Corces, 2011, 2012). In developing mouse embryos, *Oct4* gene expression is regulated spatially and temporally by at least two enhancers, DE and PE (Yeom et al., 1996). Here we have elucidated the control of *Oct4* expression by its DE and PE elements during the totipotent cycle (Figure 7A). The *Oct4* gene is expressed under the control of the DE during the pre-implantation stage, but its expression is controlled by PE in the epiblast subsequent to implantation (Figure 7A; Scholer, 1991). Expression of the *Oct4* gene is downregulated in epiblasts and expressed exclusively in the developing germline. In founder PGCs, *Oct4* is initially controlled by DE but subsequently by both DE and PE (Figures 1E–1J and 7A). Interestingly, we found that *Oct4* was expressed not only in spermatogonia but also in differentiating germ cells; in adult mice, *Oct4* expression was regulated by DE in spermatogonia but by PE in differentiating germ cells. It is well known that PGCs first emerge at the proximal region of the epiblast adjacent to the extraembryonic ectoderm and then migrate to form the PGC cluster at 7.25 dpc (Ginsburg et al., 1990; Saitou et al., 2002). However, while our dual reporter system could not distinguish the PGC population at 7.25 dpc, *Oct4*-ΔPE-GFP PGCs were apparent in 8.5-dpc embryos. Intriguingly, migrating PGCs in 9.5- and 10.5-dpc embryos contained two subpopulations, GFP$^+$RFP$^-$ and GFP$^+$RFP$^+$ cells, and the GFP$^+$RFP$^+$ cell population increased during PGC migration into the genital ridge. Since genome-wide epigenetic modification, including imprinting erasure, occurs in migrating PGCs (Sasaki and Matsui, 2008), further experiments are required to examine the differences between these two migrating PGC populations and whether they can be distinguished using genome-wide epigenetic markers.

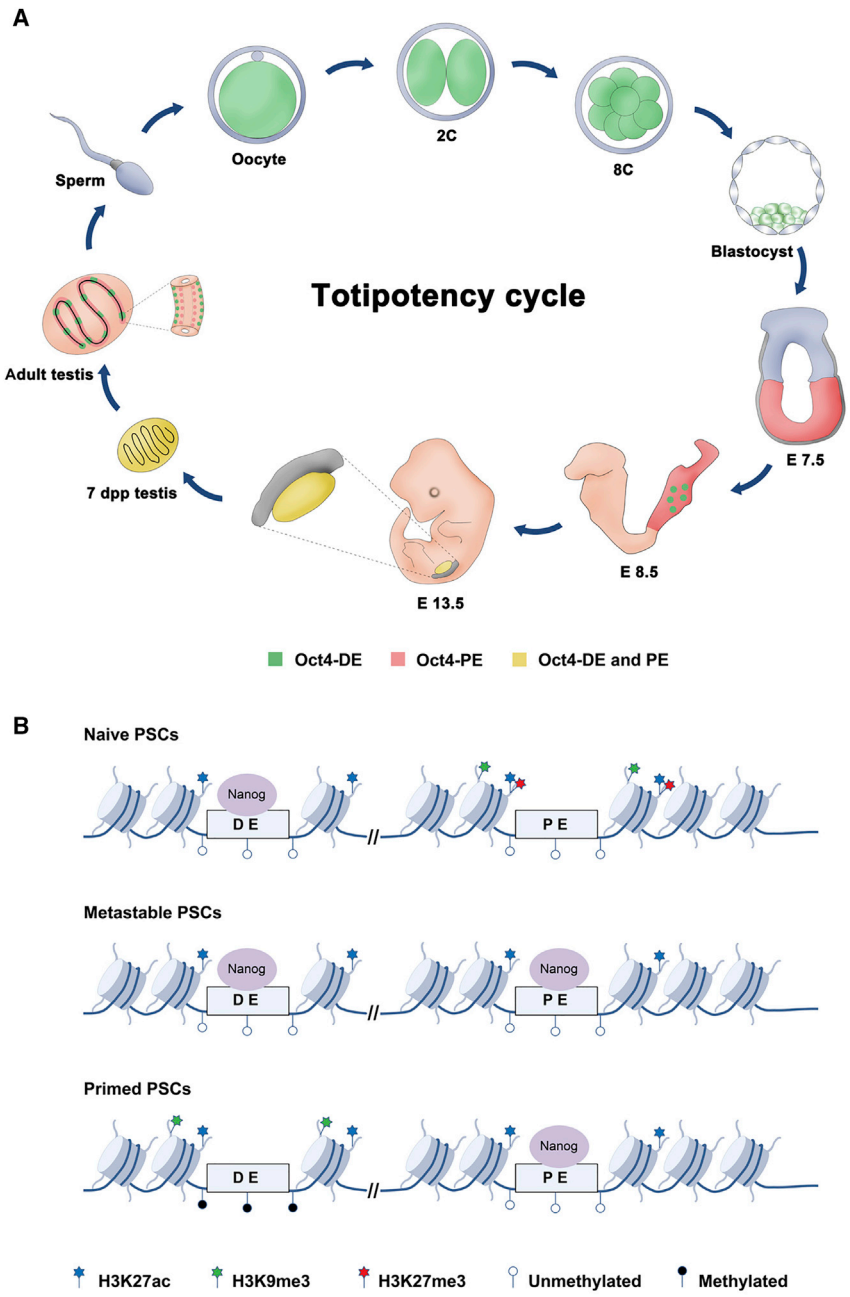

Figure 7. Schematic Representation of the *Oct4* Regulatory Elements

(A) The regulatory element of *Oct4* during the totipotent cycle.

(B) Epigenetic status of the *Oct4* enhancer in naive and primed pluripotency.

This study showed that ICM cells initially express only GFP but express both GFP and RFP during the derivation of ESCs (Figure 2B). This result supports the findings of recent studies indicating that the ability of ICM cells to self-renew as ESCs is acquired during epiblast specification (Boroviak et al., 2014; Tang et al., 2010). Boroviak et al. (2014) also suggested that early ICM cells are distinct from ESCs and that ESCs exhibit the greatest degree of identity to the embryonic day 4.5 (E4.5) epiblast. Our results confirm that the switch in enhancer activity occurs at approximately the E4.5 stage (Yeom et al., 1996), and enhancer activity of DE and PE overlapped in early post-implantation epiblast (5.5–6.5 dpc).

Self-renewing ESCs cultivated in conventional ESC culture medium (supplemented with LIF and serum) constitute a cell population at various stages of pluripotency (Hayashi et al., 2008; Martinez Arias and Brickman, 2011; Miyanari and Torres-Padilla, 2012). A defined ESC culture system, containing 2i (MEK inhibitor PD0325901 and glycogen synthase kinase 3 inhibitor CHIR99021) (Ying et al., 2008) together with LIF (2i + LIF) has been used to select naive pluripotent cells (Silva et al., 2008). Here we

show that ESCs from 2i + LIF medium still contain two cell populations when cultured in serum-containing medium and that the mixed population could be converted to homogeneous naive PSCs by using serum-free N2B27 medium. This result indicates that serum contains a factor(s) that induce differentiation or inhibit maintenance of naive pluripotency. Previous reports have suggested that serum contains ERK phosphorylation-inducing factors (Yamaji et al., 2013). Silva et al. (2008) also showed that the presence of MEK inhibitor in serum-containing medium accelerated the activation of *Nanog* and *Rex1* during reprogramming. This is also supported by the findings that GFP^+RFP^- ESCs cultured in serum-containing medium were molecularly distinguishable from GFP^+RFP^- ESCs cultured in serum-free medium. Collectively, the GFP^+RFP^- state is a necessary but not sufficient condition for naive PSCs, since special medium is also required for maintaining naive PSCs. Several reporter systems suggested that undifferentiated ESCs heterogeneously expressed *Nanog* (Chambers et al., 2007), *Rex1* (Marks et al., 2012; Toyooka et al., 2008), *Stella* (Hayashi et al., 2008), *Esrrb* (van den Berg et al., 2008), and *Tbx3* (Niwa et al., 2009). The heterogeneity of transcription factors in ESCs regulated self-renewal capacity, expression of developmental genes, and differentiation potential (Torres-Padilla and Chambers, 2014). However, the GFP^+RFP^- and GFP^+RFP^+ ESCs in serum-containing medium are similar to each other at the transcriptional, epigenetic, and functional levels. Our dual transgenic system allows monitoring of enhancer activity of DE and PE, but not *Oct4* expression levels. Although enhancer activity fluctuates between GFP^+RFP^- and GFP^+RFP^+ in serum-containing medium, these cells highly expressed *Oct4* gene.

Mixed populations of PSCs were also converted to naive PSCs in 2i + LIF medium without serum or to primed EpiLCs in bFGF + activin medium. EpiLCs induced from ESCs expressed only RFP, and the RFP^+ EpiLCs were indistinguishable from EpiSCs in terms of morphology, gene expression patterns, epigenetic status, and defective incorporation into ICM of blastocysts. Therefore, GFP^-RFP^+ state constitute a definitive marker of primed PSCs; GFP^-RFP^+ cells were not observed in any of the three different ESC culture media.

The conversion from GFP^+RFP^+ to GFP^+RFP^- cells and vice versa takes place within 2 days of changing the culture medium (Figure 2), indicating that the conversion of enhancer activity occurs within 2 days. It has been suggested that genome-wide demethylation and transcriptional changes occur in mouse ESCs within the first 24 hr of 2i addition (Ficz et al., 2013). Therefore, the regulatory machinery of *Oct4* enhancers seems to change rapidly in response to external cues accompanying epigenetic modification.

GFP^-RFP^+ EpiLCs derived from transgenic ESCs showed the molecular signature of EpiSCs, including expression of EpiSC markers and hypermethylation of germ cell markers (*Stella* and *Dppa5* region), *LINE1*, and *IAP*. Primed PSCs are derived from post-implantation stage epiblasts and express *Oct4* for self-renewal under the control of the PE (Brons et al., 2007; Tesar et al., 2007). RFP^+ cells were not incorporated into the ICM by aggregation with normal embryos. Although a single aggregated embryo with RFP^+ cells was observed, the GFP^-RFP^+ cells did not contribute to the ICM (Figure S6). This result could support the observation that GFP^-RFP^+ cells formed a pure population of primed PSCs.

Finally, we showed that the activity of the two *Oct4 cis*-regulatory elements was controlled by DNA methylation and histone modification concurring with the naive or primed pluripotency state. Consequently, the differential activity of DE and PE was closely related to binding affinity of Nanog on the two enhancers; Nanog was highly enriched on the DE in naive PSCs, but highly enriched on the PE in primed PSCs (Figures 6C and 7B). We also found that the activity of *Oct4* DE and PE is affected by the repressive histone marks, H3K9me3 and H3K27me3, and DNA methylation in a cell-type-specific manner. The enhancers of *Oct4* in primed PSCs is regulated by DNA methylation as well as H3K9me3, but those in naive PSCs is regulated by H3K9me3 and H3K27me3 but not by DNA methylation; DNA methylation was only correlated with inactive DE in primed PSCs and not with inactive PE in naive PSCs (Figure 7B). Thus, demethylation of the PE region seems to be the default state for PSCs and methylation of DE could constitute an epigenetic marker for primed PSCs. In contrast, H3K27ac was enriched in both the DE and PE of naive and primed PSCs, indicating that H3K27ac cannot be a histone mark for the active state of the *Oct4* DE and PE. Another explanation is that although one of the enhancers is not active based on fluorescent reporter expression, epigenetically these enhancer elements of *Oct4* are basically not completely silenced as H3K27ac remains enriched in both naive and primed PSCs. Similarly the PE is not methylated in 2i/LIF conditions, although the fluorescent mark is not on. These results suggest that overall both enhancers are active and never fully silenced in any of the states, although their level of activity is different under the tested conditions. Thus we should be more careful against using of terms "off" and "on" when describing enhancer activity of *Oct4*, the more accurate term rather being "dominant activity." This fact could also be the reason for the need of a dual reporter system for accurate separation of naive and primed PSCs.

In this study, we showed that enhancer-specific regulation of *Oct4* could constitute a determinant for distinguishing between naive and primed pluripotency. This is based on the ability to accurately separate naive and primed PSCs using the dual reporter system, thereby obtaining

pure populations of naive and primed pluripotent cells, which could be used for accurate analysis of distinct cell states. The dual reporter system could also be a useful tool for monitoring cellular reprogramming to naive or primed states as well as embryonic development under live-cell conditions.

EXPERIMENTAL PROCEDURES

Generation of O4-DE-GFP/O4-PE-RFP Mouse

To generate the *Oct4*-ΔDE-RFP transgene, we inserted the tdTomato (RFP) gene into genomic *Oct4* fragments of 6 kb in length (GOF6), which lack DE (Yeom et al., 1996). *Oct4*-ΔDE-RFP transgene expresses tdTomato transgene under the control of *Oct4* PE and PP. O4-PE-RFP embryos were generated by injecting Oct4-ΔDE-RFP (O4-PE-RFP) plasmid into normal zygote, which were developed to blastocysts. O4-PE-RFP blastocysts were transferred into the uterus of a pseudo-pregnant female mouse. We first generated five O4-PE-RFP founder transgenic mice, two male and three female (Figure S1). OG2 mice were purchased from the Jackson Laboratory and maintained in the mouse facility, and used for O4-DE-GFP founder mice (Szabo et al., 2002). The O4-PE-RFP mouse was crossed with homozygous O4-DE-GFP mouse. Finally, we generated O4-DE-GFP$^{+/-}$ O4-PE-RFP$^{+/-}$ mice carrying both *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-RFP.

ESC Derivation and Culture

ESC lines from blastocyst were derived essentially as described previously (Nichols et al., 2009). Embryos at the eight-cell stage were flushed from oviducts of C57/BL6 mice mated with male O4-DE-GFP/O4-PE-RFP mice at 2.5 dpc, cultured in G2 medium for 2 days, and transferred to conventional ESC medium with 15% fetal bovine serum (FBS), 1× penicillin/streptomycin/glutamine, 0.1 mM nonessential amino acids, 1 mM β-mercaptoethanol, and 10^3 units/mL LIF on feeder layers. The established ESCs in conventional ESC medium were transferred into conventional ESC medium supplemented with LIF and 2i inhibitors, CHIR99021(3 μM) and PD0325901 (1 μM), or serum-free N2B27 medium supplemented with LIF and 2i inhibitors. The serum-free N2B27 was prepared as described by Ying and Smith (2003).

EpiSCs-like Cell Differentiation from ESCs

EpiLCs from ESCs were derived essentially as described previously (Hayashi et al., 2011). The EpiLCs were induced by plating 1.0 × 10^5 ESCs on a well of a 12-well plate on a feeder layer or feeder free in N2B27 medium containing activin A (20 ng/mL), bFGF (12 ng/mL), and knockout serum replacement (1%). The medium was changed every day. The differentiated EpiLCs (*Oct4*-ΔDE-RFP-positive cells) were passaged every 2–3 days by dissociation with 1 mg/mL collagenase type IV (Invitrogen) or by pipetting with a glass pipette.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, six figures, and two tables and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2016.09.012.

AUTHOR CONTRIBUTIONS

The study was conceived and designed by H.W.C and J.T.D. H.W.C. carried out the majority of experiments, performed data analysis, and created the figures. J.Y.J., Y.J.H., G.W., H.S., and J.S.K. performed some data analysis. *Oct4*-ΔDE-RFP vector was constructed by J.W.L. H.R.S. performed data interpretation. The study was supervised by J.T.D. The manuscript was written by H.W.C. and J.T.D.

ACKNOWLEDGMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (grant nos. 2015R1A2A2A01003604 and 2015R1A5A1009701) and the Next-Generation BioGreen 21 Program (grant no. PJ01133802) funded by the Rural Development Administration, Republic of Korea.

Received: February 17, 2016
Revised: September 26, 2016
Accepted: September 28, 2016
Published: October 27, 2016

REFERENCES

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature *461*, 1292–1295.

Boiani, M., Eckardt, S., Schöler, H.R., and McLaughlin, K.J. (2002). Oct4 distribution and level in mouse clones: consequences for pluripotency. Genes Dev. *16*, 1209–1219.

Bonn, S., Zinzen, R.P., Girardot, C., Gustafson, E.H., Perez-Gonzalez, A., Delhomme, N., Ghavi-Helm, Y., Wilczyński, B., Riddell, A., and Furlong, E.E. (2012). Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. Nat. Genet. *44*, 148–156.

Boroviak, T., Loos, R., Bertone, P., Smith, A., and Nichols, J. (2014). The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification. Nat. Cell Biol. *16*, 516–528.

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

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. Cell *144*, 327–339.

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. Nat. Rev. Genet. *10*, 295–304.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. Nature *450*, 1230–1234.

Chenoweth, J.G., McKay, R.D., and Tesar, P.J. (2010). Epiblast stem cells contribute new insight into pluripotency and gastrulation. Dev. Growth Differ. *52*, 293–301.

Choi, H.W., Kim, J.S., Choi, S., Hong, Y.J., Kim, M.J., Seo, H.G., and Do, J.T. (2014). Neural stem cells differentiated from iPS cells spontaneously regain pluripotency. Stem Cells *32*, 2596–2604.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., and Sharp, P.A. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. USA *107*, 21931–21936.

Do, J.T., and Schöler, H.R. (2004). Nuclei of embryonic stem cells reprogram somatic cells. Stem Cells *22*, 941–949.

Favaedi, R., Shahhoseini, M., and Akhoond, M.R. (2012). Comparative epigenetic analysis of Oct4 regulatory region in RA-induced differentiated NT2 cells under adherent and non-adherent culture conditions. Mol. Cell. Biochem. *363*, 129–134.

Ficz, G., Hore, T.A., Santos, F., Lee, H.J., Dean, W., Arand, J., Krueger, F., Oxley, D., Paul, Y.L., Walter, J., et al. (2013). FGF signaling inhibition in ESCs drives rapid genome-wide demethylation to the epigenetic ground state of pluripotency. Cell Stem Cell *13*, 351–359.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., and Zviran, A. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282–286.

Galonska, C., Ziller, M.J., Karnik, R., and Meissner, A. (2015). Ground state conditions induce rapid reorganization of core pluripotency factor binding before global epigenetic reprogramming. Cell Stem Cell *17*, 462–470.

Ginsburg, M., Snow, M.H., and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. Development *110*, 521–528.

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

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

Hanna, J., Markoulaki, S., Mitalipova, M., Cheng, A.W., Cassady, J.P., Staerk, J., Carey, B.W., Lengner, C.J., Foreman, R., Love, J., et al. (2009). Metastable pluripotent states in NOD-mouse-derived ESCs. Cell Stem Cell *4*, 513–524.

Hanna, J.H., Saha, K., and Jaenisch, R. (2010). Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell *143*, 508–525.

Hayashi, K., Lopes, S.M., Tang, F., and Surani, M.A. (2008). Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell *3*, 391–401.

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

Kim, J.B., Sebastiano, V., Wu, G., Arauzo-Bravo, M.J., Sasse, P., Gentile, L., Ko, K., Ruau, D., Ehrich, M., van den Boom, D., et al. (2009). Oct4-induced pluripotency in adult neural stem cells. Cell *136*, 411–419.

Kunath, T., Saba-El-Leil, M.K., Almousailleakh, M., Wray, J., Meloche, S., and Smith, A. (2007). FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. Development *134*, 2895–2902.

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

Martinez Arias, A., and Brickman, J.M. (2011). Gene expression heterogeneities in embryonic stem cell populations: origin and function. Curr. Opin. Cell Biol. *23*, 650–656.

Miyanari, Y., and Torres-Padilla, M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470–473.

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

Nichols, J., and Smith, A. (2011). The origin and identity of embryonic stem cells. Development *138*, 3–8.

Nichols, J., Silva, J., Roode, M., and Smith, A. (2009). Suppression of Erk signalling promotes ground state pluripotency in the mouse embryo. Development *136*, 3215–3222.

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

Ong, C.T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. *12*, 283–293.

Ong, C.T., and Corces, V.G. (2012). Enhancers: emerging roles in cell fate specification. EMBO Rep. *13*, 423–430.

Pesce, M., Wang, X., Wolgemuth, D.J., and Scholer, H. (1998). Differential expression of the Oct-4 transcription factor during mouse germ cell differentiation. Mech. Dev. *71*, 89–98.

Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. Nature *418*, 293–300.

Sasaki, H., and Matsui, Y. (2008). Epigenetic events in mammalian germ-cell development: reprogramming and beyond. Nat. Rev. Genet. *9*, 129–140.

Scholer, H.R. (1991). Octamania: the POU factors in murine development. Trends Genet. *7*, 323–329.

Scholer, H.R., Dressler, G.R., Balling, R., Rohdewohld, H., and Gruss, P. (1990). Oct-4: a germline-specific transcription factor mapping to the mouse t-complex. EMBO J. *9*, 2185–2195.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. PLoS Biol. *6*, e253.

Szabo, P.E., Hubner, K., Scholer, H., and Mann, J.R. (2002). Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. Mech. Dev. *115*, 157–160.

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

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

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., Fan, Z.P., Maetzel, D., Ganz, K., Shi, L., et al. (2014). Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. Cell Stem Cell *15*, 471–487.

Torres-Padilla, M.E., and Chambers, I. (2014). Transcription factor heterogeneity in pluripotent stem cells: a stochastic advantage. Development *141*, 2173–2181.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development *135*, 909–918.

Vallier, L., Touboul, T., Chng, Z., Brimpari, M., Hannan, N., Millan, E., Smithers, L.E., Trotter, M., Rugg-Gunn, P., Weber, A., et al. (2009). Early cell fate decisions of human embryonic stem cells and mouse epiblast stem cells are controlled by the same signalling pathways. PLoS One *4*, e6082.

van den Berg, D.L., Zhang, W., Yates, A., Engelen, E., Takacs, K., Bezstarosti, K., Demmers, J., Chambers, I., and Poot, R.A. (2008). Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. Mol. Cell. Biol. *28*, 5986–5995.

Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., and Smith, A. (2011). Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat. Cell Biol. *13*, 838–845.

Wu, G., and Scholer, H.R. (2014). Role of Oct4 in the early embryo development. Cell Regen. *3*, 7.

Yamaji, M., Ueda, J., Hayashi, K., Ohta, H., Yabuta, Y., Kurimoto, K., Nakato, R., Yamada, Y., Shirahige, K., and Saitou, M. (2013). PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. Cell Stem Cell *12*, 368–382.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881–894.

Ying, Q.-L., and Smith, A.G. (2003). Defined conditions for neural commitment and differentiation. Methods Enzymol. *365*, 327–341.

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

Zentner, G.E., Tesar, P.J., and Scacheri, P.C. (2011). Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. *21*, 1273–1283.

Stem Cell Reports, Volume *7*

Supplemental Information

Distinct Enhancer Activity of *Oct4* in Naive and Primed Mouse Pluripotency

Hyun Woo Choi, Jin Young Joo, Yean Ju Hong, Jong Soo Kim, Hyuk Song, Jeong Woong Lee, Guangming Wu, Hans R. Schöler, and Jeong Tae Do

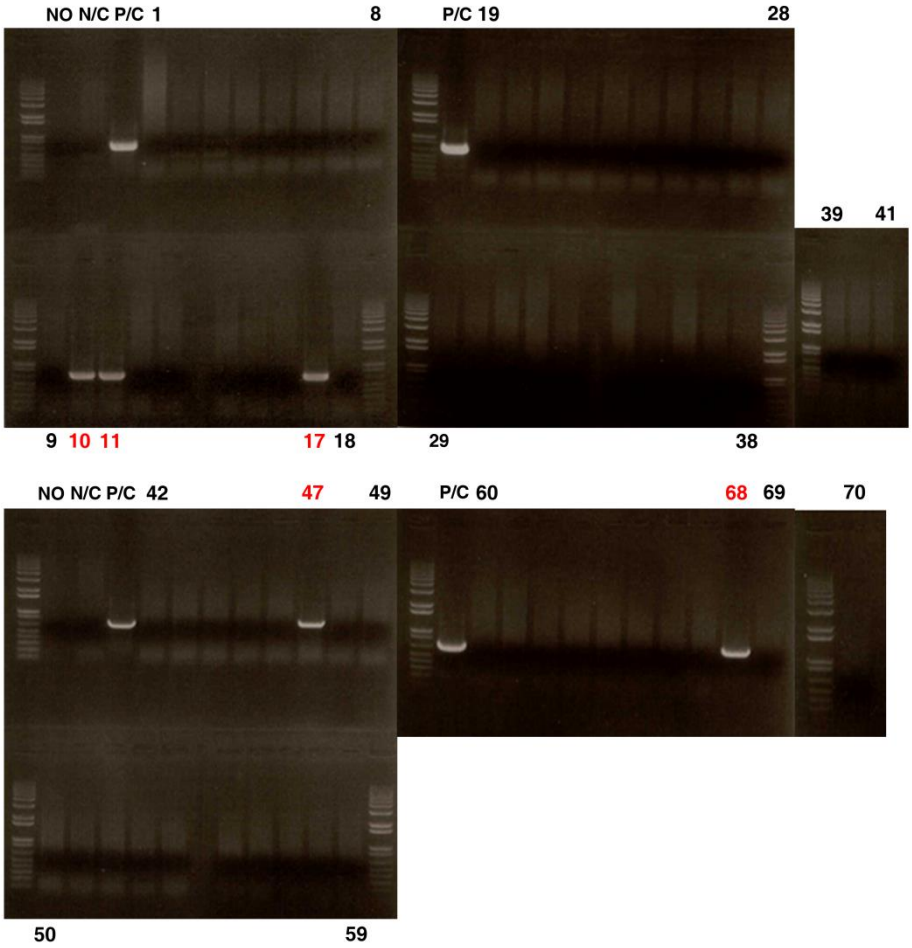

Figure S1. Genotyping of *Oct4*-ΔDE-tdTomato mice by using tdTomato specific primer. (related to Figure 1). We obtained five *Oct4*-ΔDE-tdTomato founder transgenic mice (5/70). NO: no template, N/C: normal mouse genomic DNA, P/C: positive (plasmid DNA) control.

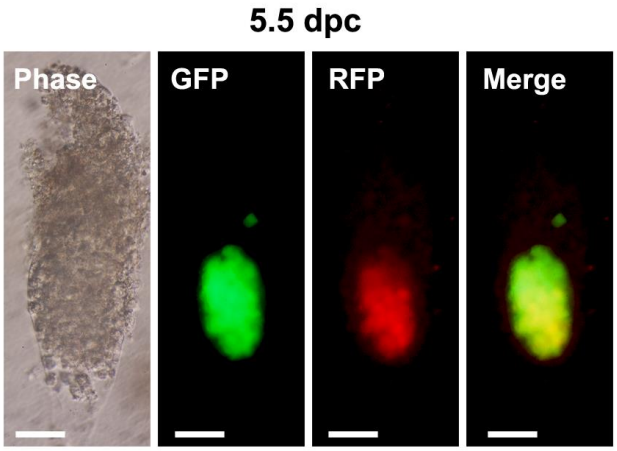

Figure S2. Oct4 enhancer activity in early post-implantation embryo (5.5 dpc) Phase and flurescence images of 5.5 dpc embryo, Scale bar = 100 μm, The epiblast of 5.5 dpc embryo expressed both O4-DE-GFP and O4-PE-RFP. However, proximal epiblast cells weakly expressed O4-PE-RFP than distal epiblast cells.

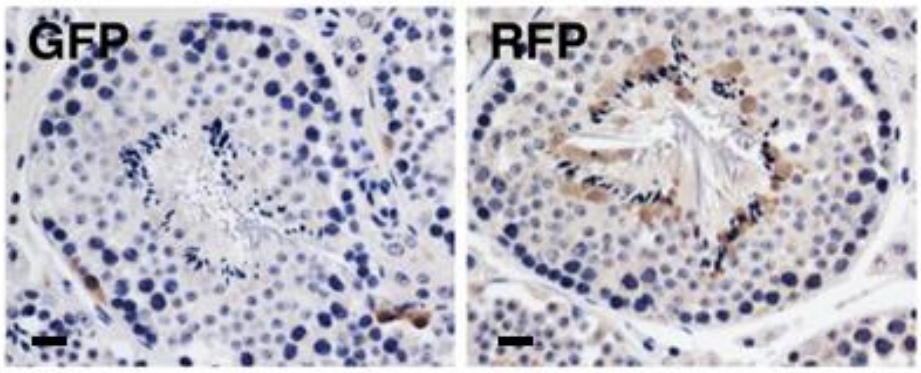

Figure S3. Immunohistochemistry analysis in testis of adult double transgenic mouse (4 weeks). (related Figure 1). GFP (brown) were positive in spermatogonia at periphery (nearby basement membrane) of seminiferous tubules in testis of adult RFP (brown) were positive from the middle to the center of seminiferous tubules. scale bar = 50 μm.

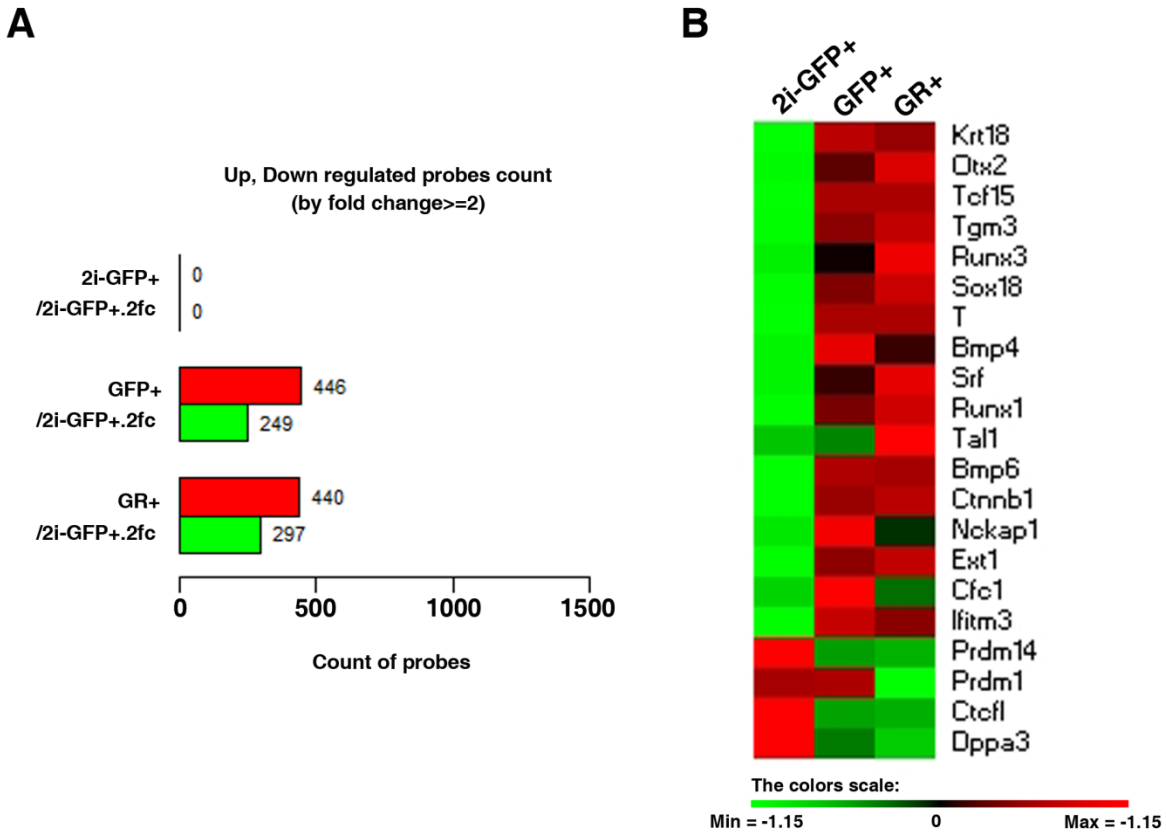

Figure S4. Different gene expression 2i-GFP$^+$ compared with GFP$^+$ or GR$^+$ cells. (related Figure 3). (A) Counting of up and down regulated probes in 2i-GFP$^+$ compared with GFP$^+$ or GR$^+$ cells. (B) Development related genes were up-regulated in GFP$^+$ and GR$^+$ cells in serum+LIF medium as ectoderm (*Krk18*, *Otx2*, *Tcf15*, and *Tgm3*), mesoderm (*Bmp4*, *Srf*, *Runx1*, *Tal1*, and *Bmp6*), and endoderm (*Ctnnb1*, *Ext1*,and *Cfc1*).

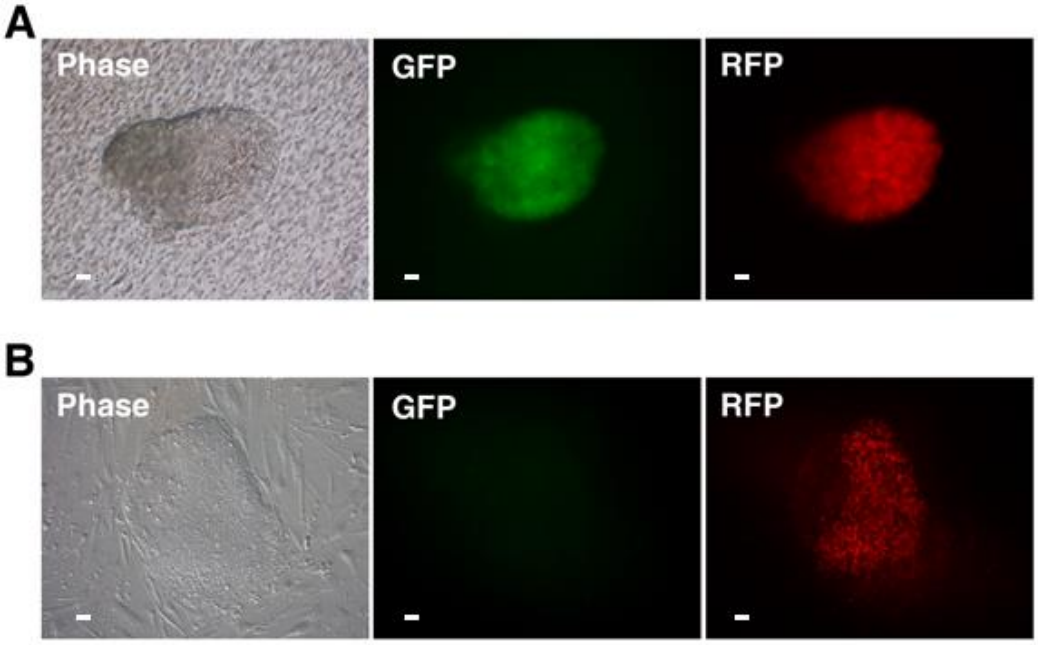

Figure S5. Derivation of EpiSCs from double transgenic embryo. (related Figure 4). **(**A) Isolated epiblast from 6.5 dpc double transgenic embryo expressed *Oct4*-ΔPE-GFP and *Oct4*-ΔDE-tdTomato (See also Figure 1C). scale bar = 50 μm. (B) Only *Oct4*-ΔDE-tdTomato positive cells were expanded in EpiSCs medium on feeder at day 1. scale bar = 50 μm.

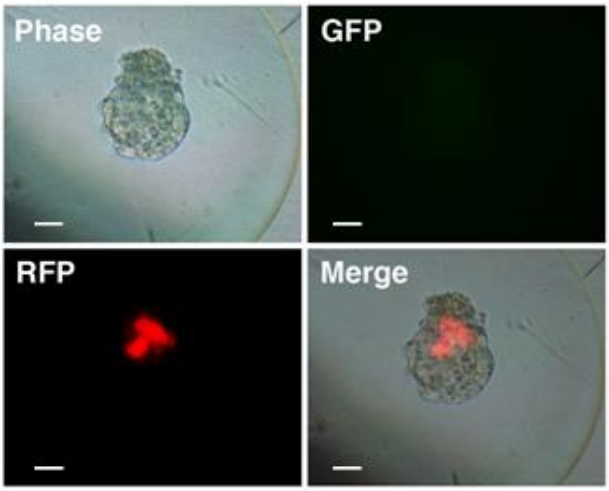

Figure S6. Only *Oct4*-ΔDE-RFP positive cells did not contribute ICM. (related Figure 5). One embryo (1/27) was aggregated with *Oct4*-ΔDE-RFP positive cells. However, *Oct4*-ΔDE-RFP positive cells did not incorporate into ICM (See also Figure 5A and B). scale bar = 50 μm.

Table S1. Gene Ontology (GO) analysis of up-regulated gene in 2i-GFP+cells

Annotation Cluster 1	Enrichment Score: 3.29			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
sterol biosynthetic process	6	MVD, DHCR7, INSIG1, FDPS, SC4MOL, NSDHL	1.30E-05	1.70E-02
sterol metabolic process	7	MVD, DHCR7, INSIG1, FDPS, PCSK9, SC4MOL, NSDHL	1.50E-04	6.20E-02
cholesterol metabolic process	6	MVD, DHCR7, INSIG1, FDPS, PCSK9, NSDHL	8.10E-04	2.30E-01
cholesterol biosynthetic process	5	MVD, DHCR7, INSIG1, FDPS, NSDHL	8.70E-05	5.30E-02
lipid biosynthetic process	9	SCD1, SCD2, A4GALT, MVD, DHCR7, INSIG1, FDPS, SC4MOL, NSDHL	1.00E-02	6.90E-01

Table S2. Gene Ontology (GO) analysis of up-regulated gene in GFP or GR positive cells

Annotation Cluster 1	Enrichment Score: 3.47			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
gland morphogenesis	12	BMP4, EGFR, WNT5A, DDR1, CAV1, FGF7, RARG, CD44, CSF1, TNC, AREG, ETV4	8.90E-07	1.50E-03
epithelium development	20	COL18A1, EGFR, WNT5A, BMP4, CEBPB, RARG, CSF1, LMO4, TNC, GJA1, GREM1, SPRR2K, DDR1, LOC100048295, CD44, ID1, POU3F1, AREG, CAR2, ETV4	2.40E-06	2.00E-03
mammary gland morphogenesis	7	BMP4, WNT5A, DDR1, CAV1, CSF1, AREG, ETV4	4.40E-05	2.40E-02
gland development	15	EGFR, WNT5A, BMP4, CAV1, CEBPB, FGF7, RARG, SOCS2, CSF1, TNC, BCL2L11, DDR1, CD44, AREG, ETV4	4.50E-05	1.90E-02
mammary gland development	10	BMP4, WNT5A, DDR1, CAV1, CEBPB, SOCS2, CSF1, AREG, BCL2L11, ETV4	5.60E-05	1.80E-02
morphogenesis of an epithelium	13	WNT5A, BMP4, EGFR, RARG, CSF1, LMO4, TNC, GREM1, DDR1, CD44, AREG, ETV4, CAR2	2.00E-04	4.00E-02
tissue morphogenesis	15	WNT5A, BMP4, EGFR, FGF7, RARG, CSF1, LMO4, TNC, GREM1, TPM1, DDR1, CD44, AREG, CAR2, ETV4	3.30E-04	5.90E-02
branching involved in mammary gland duct morphogenesis	5	WNT5A, DDR1, CSF1, AREG, ETV4	4.90E-04	7.80E-02
tube morphogenesis	12	BMP4, WNT5A, DDR1, CD44, LMO4, CSF1, GJA1, AREG, GREM1, ZIC3, BCL2L11, ETV4	6.80E-04	8.40E-02
tube development	15	WNT5A, BMP4, FGF7, CSF1, LMO4, GJA1, GREM1, FOXP1, BCL2L11, ZIC3, DDR1, CD44, ID1, AREG, ETV4	9.30E-04	1.00E-01
mammary gland duct morphogenesis	5	WNT5A, DDR1, CSF1, AREG, ETV4	1.70E-03	1.40E-01
epithelial tube morphogenesis	9	BMP4, WNT5A, DDR1, CD44, LMO4, CSF1, AREG, GREM1, ETV4	1.80E-03	1.40E-01
branching morphogenesis of a tube	8	BMP4, WNT5A, DDR1, CD44, CSF1, AREG, GREM1, ETV4	2.70E-03	1.80E-01
developmental growth involved in morphogenesis	4	BMP4, WNT5A, CSF1, AREG	3.70E-03	2.20E-01
morphogenesis of a branching structure	9	BMP4, WNT5A, DDR1, FGF7, CD44, CSF1, AREG, GREM1, ETV4	3.70E-03	2.10E-01
developmental growth	6	BMP4, WNT5A, CSF1, GJA1, HBEGF, AREG	5.10E-02	6.40E-01

Annotation Cluster 2	Enrichment Score: 2.89			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
vasculature development	16	BMP4, COL18A1, CAV1, SPHK1, EFNB2, GJA1, ELK3, THY1, ANXA2, LOC100048295, ID1, ROBO4, HBEGF, PLCD1, PPAP2B, TNFAIP2	1.60E-04	3.80E-02
blood vessel development	14	BMP4, COL18A1, CAV1, SPHK1, GJA1, ELK3, THY1, ANXA2, LOC100048295, ROBO4, HBEGF, PLCD1, TNFAIP2, PPAP2B	1.30E-03	1.40E-01
blood vessel morphogenesis	12	BMP4, COL18A1, CAV1, LOC100048295, ROBO4, GJA1, HBEGF, PLCD1, ELK3, TNFAIP2, ANXA2, THY1	2.20E-03	1.60E-01
angiogenesis	9	BMP4, COL18A1, ROBO4, HBEGF, PLCD1, ELK3, TNFAIP2, ANXA2, THY1	5.40E-03	2.60E-01

Annotation Cluster 3	Enrichment Score: 2.67			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
mammary gland morphogenesis	7	BMP4, WNT5A, DDR1, CAV1, CSF1, AREG, ETV4	4.40E-05	2.40E-02
regulation of morphogenesis of a branching structure	4	BMP4, WNT5A, FGF7, ETV4	1.30E-02	4.00E-01
negative regulation of epithelial cell proliferation	4	BMP4, WNT5A, SERPINF1, ETV4	1.60E-02	4.30E-01

Annotation Cluster 4	Enrichment Score: 2.31			
GO Term	**Count**		**P_Value**	**Benjamini**
negative regulation of protein kinase activity	7	CAV1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, THY1	5.30E-04	7.60E-02
negative regulation of kinase activity	7	CAV1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, THY1	5.30E-04	7.60E-02
negative regulation of transferase activity	7	CAV1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, THY1	6.50E-04	8.60E-02
negative regulation of molecular function	10	CLN3, CAV1, TRIB3, SPRED1, GADD45A, MYC, PDCD4, SPRY4, DDIT3, THY1	1.40E-03	1.30E-01
negative regulation of catalytic activity	8	CLN3, CAV1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, THY1	4.20E-03	2.20E-01
regulation of phosphorylation	14	BMP4, EGFR, CAV1, CSF1, SPHK1, TRIB3, PDCD4, SPRY4, IL11, THY1, CDKN1C, SPRED1, GADD45A, VLDLR	5.90E-03	2.60E-01
negative regulation of MAP kinase activity	4	CAV1, SPRED1, PDCD4, SPRY4	7.10E-03	3.00E-01
regulation of phosphate metabolic process	14	BMP4, EGFR, CAV1, CSF1, SPHK1, TRIB3, PDCD4, SPRY4, IL11, THY1, CDKN1C, SPRED1, GADD45A, VLDLR	8.00E-03	3.20E-01
regulation of phosphorus metabolic process	14	MP4, EGFR, CAV1, CSF1, SPHK1, TRIB3, PDCD4, SPRY4, IL11, THY1, CDKN1C, SPRED1, GADD45A, VLDLR	8.00E-03	3.20E-01
regulation of protein kinase activity	10	CAV1, CSF1, SPHK1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, VLDLR, THY1	1.30E-02	4.10E-01
regulation of kinase activity	10	CAV1, CSF1, SPHK1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, VLDLR, THY1	1.60E-02	4.30E-01
regulation of transferase activity	10	CAV1, CSF1, SPHK1, TRIB3, SPRED1, GADD45A, PDCD4, SPRY4, VLDLR, THY1	1.90E-02	4.80E-01
regulation of MAP kinase activity	5	CAV1, TRIB3, SPRED1, PDCD4, SPRY4	8.90E-02	7.60E-01

Annotation Cluster 5	Enrichment Score: 1.89			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
mammary gland morphogenesis	7	BMP4, WNT5A, DDR1, CAV1, CSF1, AREG, ETV4	4.40E-05	2.40E-02
developmental growth involved in morphogenesis	4	BMP4, WNT5A, CSF1, AREG	3.70E-03	2.20E-01
mammary gland epithelial cell proliferation	3	WNT5A, CEBPB, AREG	1.30E-02	4.10E-01
epithelial cell proliferation	4	BMP4, WNT5A, CEBPB, AREG	1.50E-02	4.20E-01
developmental growth	6	BMP4, WNT5A, CSF1, GJA1, HBEGF, AREG	5.10E-02	6.40E-01
growth	8	BMP4, WNT5A, INHBA, CSF1, GJA1, HBEGF, AREG, EMP1	9.60E-02	7.60E-01
cell proliferation	7	BMP4, WNT5A, SATB1, KLK8, CEBPB, AREG, PES1	3.80E-01	9.60E-01

Annotation Cluster 6	Enrichment Score: 1.83			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
negative regulation of multicellular organismal process	8	INHBA, BBS2, KLK8, THBD, SOCS2, ADORA2B, ANXA5, ANXA2	4.00E-03	2.20E-01
negative regulation of coagulation	3	THBD, ANXA5, ANXA2	1.30E-02	4.10E-01
regulation of coagulation	3	THBD, ANXA5, ANXA2	6.00E-02	6.80E-01

Annotation Cluster 7	Enrichment Score: 1.68			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
collagen metabolic process	4	MMP10, ID1, MMP3, MMP13	1.10E-02	3.70E-01
multicellular organismal macromolecule metabolic process	4	MMP10, ID1, MMP3, MMP13	1.20E-02	4.00E-01
multicellular organismal metabolic process	4	MMP10, ID1, MMP3, MMP13	1.50E-02	4.20E-01
collagen catabolic process	3	MMP10, MMP3, MMP13	4.50E-02	6.40E-01
multicellular organismal catabolic process	3	MMP10, MMP3, MMP13	5.00E-02	6.50E-01

Annotation Cluster 8	Enrichment Score: 1.6			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
membrane organization	13	CLN3, CAV1, ITGA3, PMAIP1, ELMO1, SYP, DAB2, SH3KBP1, CAP1, BIN1, STEAP2, UPK2, VLDLR	9.00E-03	3.40E-01
membrane invagination	10	SYP, CLN3, DAB2, CAV1, SH3KBP1, CAP1, STEAP2, BIN1, VLDLR, ELMO1	1.40E-02	4.10E-01
endocytosis	10	SYP, CLN3, DAB2, CAV1, SH3KBP1, CAP1, STEAP2, BIN1, VLDLR, ELMO1	1.40E-02	4.10E-01
vesicle-mediated transport	13	CLN3, CAV1, AP1M2, SYT4, ELMO1, SYP, DAB2, TRIM36, SH3KBP1, CAP1, BIN1, STEAP2, VLDLR	2.30E-01	8.90E-01

Annotation Cluster 9	Enrichment Score: 1.57			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
prostate gland epithelium morphogenesis	4	BMP4, RARG, CD44, TNC	1.50E-02	4.20E-01
prostate gland morphogenesis	4	BMP4, RARG, CD44, TNC	1.60E-02	4.30E-01
urogenital system development	8	MP4, WNT5A, LOC100048295, RARG, CD44, TNC, GREM1, BCL2L11	2.80E-02	5.50E-01
prostate gland development	4	BMP4, RARG, CD44, TNC	4.30E-02	6.30E-01
reproductive structure development	7	BMP4, WNT5A, LOC100048295, RARG, CD44, TNC, BCL2L11	4.70E-02	6.40E-01

Annotation Cluster 10	Enrichment Score: 1.46			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
in utero embryonic development	12	EGFR, VCAM1, DAB2, WNT7B, CEBPB, DNMT3L, GJA1, HBEGF, PLCD1, CAPN2, TPM1, BCL2L11	2.00E-02	4.80E-01
chordate embryonic development	15	BMP4, EGFR, CEBPB, LMO4, GJA1, CAPN2, TPM1, BCL2L11, VCAM1, DAB2, WNT7B, DNMT3L, HBEGF, PLCD1, TCF15	4.50E-02	6.40E-01
embryonic development ending in birth or egg hatching	15	BMP4, EGFR, CEBPB, LMO4, GJA1, CAPN2, TPM1, BCL2L11, VCAM1, DAB2, WNT7B, DNMT3L, HBEGF, PLCD1, TCF15	4.80E-02	6.50E-01

Annotation Cluster 11	Enrichment Score: 1.43			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
reproductive developmental process	14	WNT5A, BMP4, RARG, TDRD7, CSF1, TNC, BCL2L11, ZFP37, TAF7L, BBS2, LOC100048295, CD44, DDX25, AREG	2.70E-03	1.70E-01
reproductive cellular process	10	BMP4, BBS2, LOC100048295, TRIM36, TDRD7, DDX25, TNC, PIWIL2, TAF7L, ZFP37	8.40E-03	3.30E-01
germ cell development	7	BMP4, BBS2, LOC100048295, TDRD7, DDX25, TAF7L, ZFP37	1.60E-02	4.30E-01
sexual reproduction	14	BMP4, NANOS3, TDRD7, ITGA3, BCL2L11, ZFP37, TAF7L, BBS2, TRIM36, LOC100048295, TCFL5, DDX25, OVOL1, PIWIL2	4.90E-02	6.50E-01
multicellular organism reproduction	14	BMP4, NANOS3, CAV1, SOCS2, TDRD7, BCL2L11, ZFP37, TAF7L, BBS2, LOC100048295, TCFL5, DDX25, OVOL1, PIWIL2	6.90E-02	7.10E-01
reproductive process in a multicellular organism	14	BMP4, NANOS3, CAV1, SOCS2, TDRD7, BCL2L11, ZFP37, TAF7L, BBS2, LOC100048295, TCFL5, DDX25, OVOL1, PIWIL2	6.90E-02	7.10E-01
gamete generation	12	BMP4, NANOS3, BBS2, LOC100048295, TCFL5, TDRD7, DDX25, OVOL1, PIWIL2, BCL2L11, TAF7L, ZFP37	7.10E-02	7.20E-01
male gamete generation	9	NANOS3, BBS2, TCFL5, DDX25, OVOL1, PIWIL2, BCL2L11, TAF7L, ZFP37	1.40E-01	8.00E-01
spermatogenesis	9	NANOS3, BBS2, TCFL5, DDX25, OVOL1, PIWIL2, BCL2L11, TAF7L, ZFP37	1.40E-01	8.00E-01

Annotation Cluster 12	Enrichment Score: 1.42			
GO Term	**Count**	**Genes**	**P_Value**	**Benjamini**
regulation of epithelial cell proliferation	7	BMP4, EGFR, WNT5A, FGF7, SERPINF1, FOXP1, ETV4	1.60E-03	1.50E-01
regulation of morphogenesis of a branching structure	4	BMP4, WNT5A, FGF7, ETV4	1.30E-02	4.00E-01
respiratory system development	6	BMP4, WNT5A, FGF7, RARG, ID1, FOXP1	1.00E-01	7.60E-01
lung development	5	BMP4, WNT5A, FGF7, ID1, FOXP1	1.80E-01	8.50E-01
respiratory tube development	5	BMP4, WNT5A, FGF7, ID1, FOXP1	1.90E-01	8.60E-01

EXTENDED EXPERIMENTAL PROCEDURES

RNA isolation and qRT-PCR analysis

Total RNA was isolated using the RNeasy Mini Kit (QIAGEN) and was treated with DNase to remove genomic DNA contamination. One microgram of total RNA was reverse-transcribed with SuperScript III Reverse Transcriptase Kit (Invitrogen) and oligo(dT) primer (Invitrogen) according to the manufacturer's instructions. Quantitative polymerase chain reaction (PCR) reactions were set up in duplicate with the Power SYBR Green Master Mix (Takara) and analyzed with the Roche LightCycler 5480 (Roche). The primers for qRT-PCR used were as follows: *Oct4* sense 5'-GATGCTGTGAGCCAAGGCAAG-3', *Oct4* antisense 5'-GGCTCCTGATCAACAGCATCAC-3'; *Nanog* sense 5'-CTTTCACCTATTAAGGTGCTTGC-3', *Nanog* antisense 5'-TGGCATCGGTTCATCATGGTAC-3'; *Sox2* sense 5'-CATGAGAGCAAGTACTGGCAAG-3', *Sox2* antisense 5'- CCAACGATATCAACCTGCATGG-3'; *Rex1* sense 5'-TCCATGGCATAGTTCCAACAG-3', *Rex1* antisense 5'-TAACTGATTTTCTGCCGTATGC-3'; *Esrrb* sense 5'-CAGGCAAGGATGACAGACG-3', *Esrrb* antisense 5'-GAGACAGCACGAAGGACTGC-3'; *Klf2* sense 5'-TCGAGGCTAGATGCCTTGTGA-3', *Klf2* antisense 5'-AAACGAAGCAGGCGGCAGA-3'; *Klf4* sense 5'-AGGAGCCCAAGCCAAAGAGG-3', *Klf4* antisense 5'-CGCAGGTGTGCCTTGAGATG-3'; *Tcl1* sense 5'-TGGCCTCACTAGAACAAGAGG-3', *Tcl1* antisense 5'-CTCGGTCAAGGATGGAAGC-3'; *Tbx3* sense 5'-TTATTTCCAGGTCAGGAGATGGC-3', *Tbx3* antisense 5'-GGTCGTTTGAACCAAGTCCCTC-3'; *Prdm14* sense 5'-ACAGCCAAGCAATTTGCACTAC-3', *Prdm14* antisense 5'-TTACCTGGCATTTTCATTGCTC-3'; *Dnmt3a* sense 5'-GACTCGCGTGCAATAACCTTAG -3', *Dnmt3a* antisense 5'-GGTCACTTTCCCTCACTCTGG -3', *Dnmt3b* sense 5'-CTCGCAAGGTGTGGGCTTTTGTAAC-3', and *Dnmt3b* antisense 5'-CTGGGCATCTGTCATCTTTGCACC-3', *Dnmt3l* sense 5'-CCAGGGCAGATTTCTTCCTAAGGTC-3', and *Dnmt3l* antisense 5'-TGAGCTGCACAGAGGCATCC-3', *T*/*Brachyury* sense 5'-ATCAGAGTCCTTTGCTAGGTAG-3', and *T*/*Brachyury* antisense 5'-GTTACAATCTTCTGGCTATGC-3', *Fgf5* sense 5'-AAAACCTGGTGCACCCTAGAAG-3', and *Fgf5* antisense 5'-GCTAAACCGTCTGTGGTTTCTG-3', *Fgfr1* sense 5'-CTACCAACCCTGTCCCCAGT-3', and *Fgfr1* antisense 5'-CACAGGAAGGCCTCAGTCAG-3', *Fgfr2* sense 5'-CAAGGAGCTCTTGTTCTTCAGG-3', and *Fgfr2* antisense 5'-TAACACTGCCGTTTATGTGTGG-3'.

Bisulfite genomic sequencing

To differentiate between methylated and unmethylated CG dinucleotides, genomic DNA was treated with sodium bisulfite to convert all unmethylated cytosine residues into uracil residues using the EpiTect Bisulfite Kit (QIAGEN) according to the manufacturer's protocol. Briefly, purified genomic DNA (0.5–1 μg) was denatured at 99℃ and then incubated at 60℃. After desulfonation, neutralization, and desalting, the modified DNA was diluted in 20 μl of distilled water. Subsequently, bisulfite PCR (BS-PCR) amplification was carried out using 1- to 2-μl aliquots of modified DNA for each PCR reaction. The primers used for BS-PCR were as follows: *Oct4*-DE sense 5'- TTTAGGTTTTAGAGGTTGGTTTTG-3', *Oct4*-DE antisense 5'- CCAATTTCTATACATTCATTATAAAACAAT-3'; *Oct4*-PE first sense 5'- GGTTTTTTGAGGTTGTGTGATTTAT-3', *Oct4*-PE first antisense 5'- CTCCCCTAAAAACAACTTCCTACTC-3'; *Oct4*-PE second sense 5'- GGGATTTTTAGATTGGGTTTAGAAAA-3', *Oct4*-PE second antisense 5'- CTCCTCAAAAACAAAACCTCAAATA-3', *Oct4*-PP first sense 5'- TTTGTTTTTTTATTTATTTAGGGGG-3', *Oct4*-PP first antisense 5'- ATCCCCAATACCTCTAAACCTAATC-3'; *Oct4*-PP second sense 5'- GGGTTAGAGGTTAAGGTTAGAGGG-3', *Oct4*-PP second antisense 5'- CCCCCACCTAATAAAAATAAAAAAA-3'; *Nanog* first sense 5'- TTTGTAGGTGGGATTAATTGTGAA-3', *Nanog* first antisense 5'- AAAAAATTTTAAACAACAACCAAAAA-3', *Nanog* second sense 5'- TTTGTAGGTTGGGATTAATTGTGAA-3', *Nanog* second antisense 5'- AAAAAAACAAAACACCAACCAAAT-3'; *Stella* first sense 5'- TTTTTTTATTTTGTGATTAGGGTTG-3', *Stella* first antisense 5'- CTTCACCTAAACTACACCTTTAAAC-3'; *Stella* second sense 5'- TTTGTTTTAGTTTTTTTGGAATTGG-3', *Stella* second antisense 5'- CTTCACCTAAACTACACCTTTAAAC-3', *Dppa5* first sense 5'- GGTTTGTTTTAGTTTTTTTAGGGGTATA-3', *Dppa5* first antisense 5'- CCACAACTCCAAATTCAAAAAAT-3'; *Dppa5* second sense 5'- TTTAGTTTTTTTAGGGGTATAGTTTG-3', *Dppa5* second antisense 5'- CACAACTCCAAATTCAAAAAATTTTA-3', *LINE* sense 5'- TCAAACACTATATTACTTTAACAATTCCCA-3', *LINE* antisense 5'-CCCCCACCTAATAAAAATAAAAAAA-3'; *IAP* first sense 5'- TTGATAGTTGTGTTTTAAGTGGTAAATAAA-3', *IAP* first antisense 5'- AAAACACCACAAACCAAAATCTTCTAC-3', *IAP* second sense 5'- TTGTGTTTTAAGTGGTAAATAAATAATTTG-3', and *IAP* second antisense 5'- CAAAAAAAACACACAAACCAAAAT -3'.

Briefly, the amplified products were verified by electrophoresis on a 1% agarose gel. The desired PCR products

were used for subcloning using the TA cloning vector (pGEM-T Easy Vector, Promega). The reconstructed plasmids were purified, and individual clones were sequenced (Solgent Corporation).

Luciferase Assay

For quantifying the relative Oct4 enhancer activity, an *Oct4* upstream sequence ~2 kb (Oct4-2kb) containing distal enhancer (DE) and proximal enhancer (PE), and either ΔDE, or ΔPE was cloned into pGL3 basic vector (Promega, USA). The Oct4 upstream sequence (~2 kb) containing distal enhancer (DE) and proximal enhancer (PE) was derived from pOct4-GFP plasmid which was digested and ligated to the *Kpn*I/*Bgl*II sites of the pGL3 basic vector. The pGL3-Oct4/ΔDE or pGL3-Oct4/ΔPE reporter constructs were prepared in two steps. First, a fragment of DE 5′ or PE 5′ was PCR-amplified from pOct4-GFP plasmid using specific primer pairs, digested with *Kpn*I and *Mlu*I restriction enzymes and cloned into pGL3 basic vector to obtain pGL3-DE 5′ or pGL3-PE 5′ plasmids, respectively. Subsequently, a fragment of DE 3′ or PE 3′ was PCR amplified from pOct4-GFP plasmid using primer pairs carrying *Mlu*I and *Bgl*II restriction sites, respectively. The amplified fragment was digested and ligated to *Mlu*I/*Bgl*II sites of either pGL3-DE 5′ or pGL3-PE 5′ plasmids. Luciferase assays were performed by using the Dual-Luciferase Reporter Assay System (Promega, USA). For the reporter assay of Oct4 enhancer activity, the pGL3-Oct4-2 kb vectors, pGL3-Oct4-ΔDE, or pGL3-Oct4/ΔPE vectors (for firefly luciferase activity) and pRL-TK vector (for *Renilla* luciferase activity) were transfected individually into respectively cells. After 48 h of transfection, growth medium was removed and cells were rinsed in 1× PBS. Subsequently, the cells were lysed using 1× passive lysis buffer (PLB) and incubated for 10 min at room temperature with shaking. The cell lysate was then transferred to a 1.5 ml new tube and centrifuged at 10, 000 rpm for 5 min at 4℃. Ten microliters of the supernatant was transferred to a 96-well plate and then analyzed for luciferase expression by luminometry. Each experiment was performed in triplicate and the values obtained were recorded as relative light units (RLU).

Chromatin Immunoprecipitation (ChIP)

Cultured cells were cross-linked with 1% formaldehyde and then washed with PBS containing protease inhibitors. Genomic DNA extraction and chromatin immunoprecipitation (ChIP) were performed using SimpleChIP Plus Enzymatic Chromatin IP kit (Cell Signaling) according to the manufacturer's instructions. The antibodies used were Nanog (Bethyl, A300-397A), H3K27ac (Abcam, ab729), H3K27me3 (Cell signaling,

#9733), and H3K9me3 (Abcam, ab8898). The primers used for ChIP-qPCR were as follows: *Oct4*-DE sense 5'-GGCTGCAGGCATACTTGAAC-3', *Oct4*-DE antisense 5'-AGGGCAGAGCTATCATGCAC-3'; *Oct4*-PE sense 5'-TCCTCCTAATCCCGTCTCCT-3', and *Oct4*-PE antisense 5'-GGACTCCGGTGTTCATCCT-3'.

Microarray-based analysis

Total RNA was isolated with the RNeasy Mini Kit (Qiagen) and digested with DNase I (RNase-free DNase, Qiagen) according to the manufacturer's instructions. Total RNA was amplified, biotinylated, and purified using the Ambion Illumina RNA amplification kit (Ambion) according to the manufacturer's instructions. Labeled cRNA samples (750 ng) were hybridized to each MouseRef-8 v2 Expression BeadChip. Signal detection was performed with Amersham Fluorolink Streptavidin-Cy3 (GE Healthcare Bio-Science) according to the bead array manual. Arrays were scanned with an Illumina Bead Array Reader according to the manufacturer's instructions.

Raw data were extracted using the software provided by the manufacturer (Illumina GenomeStudio v2011.1, Gene Expression Module v1.9.0). Array data were filtered by detection p-value < 0.05 in at least 50% samples. Selected probe signal was log-transformed and normalized by the quantile method. Comparative analysis was performed using LPE test and fold-change. False discovery rate (FDR) was controlled by adjusting the p-value with the Benjamini-Hochberg algorithm. Hierarchical clustering was performed using complete linkage and Pearson distance as a measure of similarity.

Aggregation with normal embryo

The ESCs or EpiLCs were aggregated with denuded post-compacted eight-cell-stage embryos to obtain an aggregate chimera. Eight-cell embryos flushed from 2.5-dpc B6D2F1 female mice were cultured in microdrops of embryo culture medium under mineral oil. The clumps of ESCs or EpiLCs (4–10 cells) were selected and transferred into microdrops containing zona-free eight-cell embryos. Morula-stage embryos aggregated with ESCs were cultured overnight at 37 ℃ under 5% CO_2.

Accession Numbers

Microarray data for each gene are available at the Gene Expression Omibus under accession number GSE67031.