Manipulation of KLF4 Expression Generates iPSCs Paused at Successive Stages of Reprogramming

Ken Nishimura,[1,*] Tetsuo Kato,[1] Chen Chen,[1] Lalhaba Oinam,[1] Emi Shiomitsu,[1] Daisuke Ayakawa,[1] Manami Ohtaka,[2] Aya Fukuda,[1] Mahito Nakanishi,[2] and Koji Hisatake[1,*]

[1]Laboratory of Gene Regulation, Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan
[2]Research Center for Stem Cell Engineering, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8562, Japan
*Correspondence: ken-nishimura@md.tsukuba.ac.jp (K.N.), kojihisa@md.tsukuba.ac.jp (K.H.)
 http://dx.doi.org/10.1016/j.stemcr.2014.08.014

SUMMARY

The detailed mechanism of reprogramming somatic cells into induced pluripotent stem cells (iPSCs) remains largely unknown. Partially reprogrammed iPSCs are informative and useful for understanding the mechanism of reprogramming but remain technically difficult to generate in a predictable and reproducible manner. Using replication-defective and persistent Sendai virus (SeVdp) vectors, we analyzed the effect of decreasing the expression levels of OCT4, SOX2, KLF4, and c-MYC and found that low KLF4 expression reproducibly gives rise to a homogeneous population of partially reprogrammed iPSCs. Upregulation of KLF4 allows these cells to resume reprogramming, indicating that they are paused iPSCs that remain on the path toward pluripotency. Paused iPSCs with different KLF4 expression levels remain at distinct intermediate stages of reprogramming. This SeVdp-based stage-specific reprogramming system (3S reprogramming system) is applicable for both mouse and human somatic cells and will facilitate the mechanistic analysis of reprogramming.

INTRODUCTION

Induced pluripotent stem cells (iPSCs) can be generated by introducing OCT4, SOX2, KLF4, and c-MYC (Takahashi et al., 2007; Takahashi and Yamanaka, 2006) or other combinations of reprogramming factors (Stadtfeld and Hochedlinger, 2010) into many types of mouse and human somatic cells. The understanding of its mechanism is crucial for generating high-quality iPSCs for cell therapy applications and also provides an insight into reprogramming and normal development. The reprogramming is an orderly process consisting of distinct stages, which can be distinguished by expression of specific markers, including stage-specific embryonic antigen-1 (SSEA-1), alkaline phosphatase (ALP), E-Cadherin (CDH1), and NANOG (Brambrink et al., 2008; Samavarchi-Tehrani et al., 2010; Stadtfeld et al., 2008). However, the precise molecular mechanisms behind iPSC generation remain to be elucidated.

Reprogramming mechanisms have been investigated by cell sorting a subpopulation of intermediate cells that express a specific combination of markers. These enriched cells are initially at a specific intermediate stage but continue reprogramming upon further cell culture, subsequently becoming a heterogeneous population of asynchronously reprogrammed cells (Hansson et al., 2012; O'Malley et al., 2013; Polo et al., 2012; Takahashi et al., 2014). Alternatively, more stable cell populations, isolated during iPSC generation, were also used for mechanistic analyses (Mikkelsen et al., 2008; Sridharan et al., 2009). These cells show phenotypes that are intermediate between somatic cells and iPSCs, thereby yielding mechanistic insights into reprogramming. Despite their usefulness, the stable cell populations are derived from rare by-products that appear sporadically during iPSC generation and difficult to obtain reproducibly in a predictable manner.

The efficiency of reprogramming and the characteristics of iPSCs are influenced by the expression levels and stoichiometry of reprogramming factors (Carey et al., 2011; Nagamatsu et al., 2012; Papapetrou et al., 2009; Sui et al., 2014; Tiemann et al., 2011). Detailed quantitative analyses show that merely increasing the expression levels of OCT4, SOX2, KLF4, and c-MYC does not necessarily lead to higher efficiency of iPSC generation; rather, the stoichiometry of the four factors plays a greater role in the efficiency of iPSC generation (Nagamatsu et al., 2012; Papapetrou et al., 2009; Tiemann et al., 2011). For example, high OCT4 and low SOX2 expression levels relative to other factors tend to generate iPSCs more efficiently than equal expression levels of the four factors. In addition, the reprogramming factor stoichiometry influences the gene expression pattern (Nagamatsu et al., 2012; Tiemann et al., 2011) and epigenetic status of iPSCs (Carey et al., 2011) as well as their ability to contribute to chimeric mice (Carey et al., 2011), suggesting that the stoichiometry of reprogramming factors may impact the reprogramming process as well.

We previously developed a unique gene transfer system, named SeVdp vectors, based on a mutant strain of Sendai virus, which remains persistently in the cytoplasm without integrating into the host genome (Nishimura et al., 2007). SeVdp vectors enable a long-term expression of multiple genes from a single vector with a constant stoichiometry and do not suffer from transcriptional silencing

CrossMark
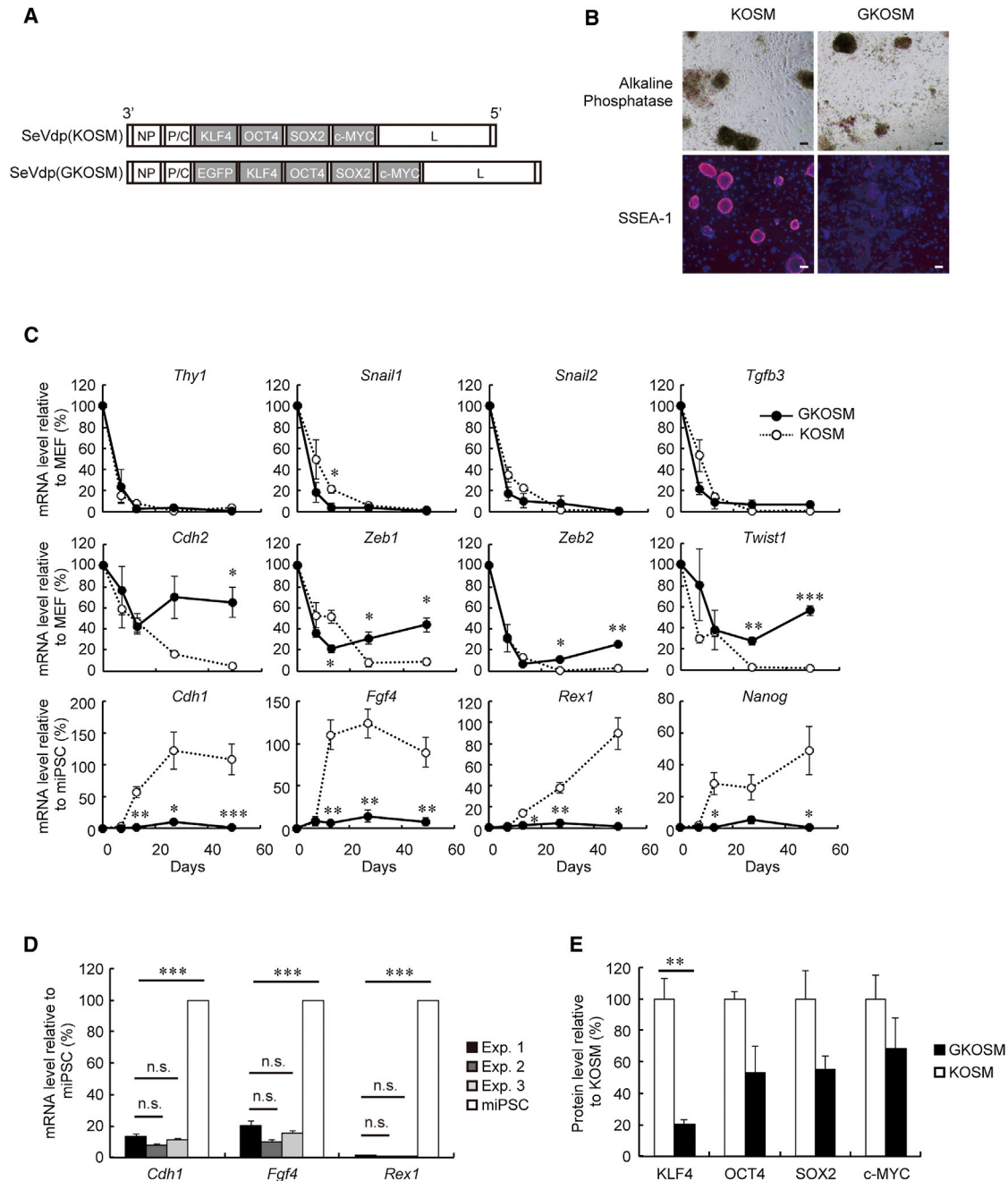
Figure 1. Generation of Partially Reprogrammed iPSCs by SeVdp Vectors

(A) Structures of SeVdp(KOSM) and SeVdp(GKOSM).

(B) Expression of ESC markers. Colonies induced by SeVdp(KOSM) or SeVdp(GKOSM) were stained for ALP or SSEA-1 at day 10 or 18, respectively. Scale bars, 100 μm.

(C) Time course of mRNA expression. Cells were collected at day 0, 6, 13, 25, and 49, and the mRNA levels of indicated genes were determined by quantitative RT-PCR and plotted relative to those in MEFs or mouse iPSCs (miPSC) (Nishimura et al., 2011). Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001 versus SeVdp(KOSM)-infected cells.

(D) Reproducibility of reprogramming by SeVdp(GKOSM). Cells were collected at day 25 (Exp. 1) or day 23 (Exp. 2 and 3), and the mRNA levels of Cdh1, Fgf4, and Rex1 were determined by quantitative RT-PCR. Data represent means ± SEM of three independent PCR reactions. n.s., not significant; ***p < 0.001 versus Exp. 1.

(legend continued on next page)

(Nishimura et al., 2011). These properties make SeVdp vectors an ideal tool for generating iPSCs. In fact, SeVdp vectors harboring four reprogramming factors (OCT4, SOX2, KLF4, and c-MYC) reprogram mouse and human somatic cells very efficiently (Nishimura et al., 2011; Nishimura et al., 2013; Takayama et al., 2010; Tateno et al., 2013; Wakao et al., 2013).

Here, we report the SeVdp-based stage-specific reprogramming system (3S reprogramming system) that generates paused iPSCs by controlling the KLF4 expression level with a destabilization domain (DD) and a small molecule, Shield1. We found that low KLF4 expression gives rise to partially reprogrammed cells, which are stably arrested at different stages of reprogramming depending on the level of KLF4 expression. Global gene expression analyses support that the partially reprogrammed iPSCs show intermediate gene expression profiles. Upregulation of KLF4 expression by Shield1 addition allows the partially reprogrammed cells to resume reprogramming toward pluripotency. The 3S reprogramming system establishes homogenous populations of paused iPSCs in a reproducible and predictable manner and is applicable for both mouse and human cells.

RESULTS

### Induction of Partially Reprogrammed Cells by SeVdp Vectors

In the course of developing SeVdp vectors for reprogramming, we created a vector named SeVdp(GKOSM) (Figure 1A), which is derived from SeVdp(KOSM) that induces fully reprogrammed iPSCs (Nishimura et al., 2011; Nishimura et al., 2013; Wakao et al., 2013). SeVdp(GKOSM) encodes EGFP in addition to KLF4, OCT4, SOX2, and c-MYC and enables monitoring of the SeVdp vector in living cells during reprogramming (Figure S1A available online). Mouse embryonic fibroblasts (MEFs) infected with SeVdp(GKOSM) or SeVdp(KOSM) gave rise to almost the same number of ALP-positive (ALP⁺) colonies (Table S1) consisting of small and rapidly growing cells (Figure 1B). Indeed, the colonies induced by SeVdp(GKOSM) were morphologically indistinguishable from SeVdp(KOSM)-induced ones, as exemplified by the cellular localization of β-CATENIN (Figures 1B and S1B). However, the colonies induced by SeVdp(GKOSM) were negative for SSEA-1 as opposed to the ones induced by SeVdp(KOSM), which were positive

for SSEA-1 (Figure 1B). Expression of *Thy1*, *Snail1*, *Snail2*, and *Tgfb3* was repressed 6 days after infection in the cells induced by SeVdp(GKOSM) and SeVdp(KOSM) (Figure 1C), consistent with the downregulation of somatic cell-specific genes at the early stage of reprogramming (Stadtfeld et al., 2008). On the other hand, at day 49, expression of mesenchymal-specific genes such as *Cdh2*, *Zeb1*, *Zeb2*, and *Twist1* remained higher in SeVdp(GKOSM)-induced cells than in SeVdp(KOSM)-induced ones (Figure 1C), indicating that the SeVdp(GKOSM)-induced cells have not completed the mesenchymal-to-epithelial transition (MET), which normally occurs during iPSC generation (Li et al., 2010; Samavarchi-Tehrani et al., 2010). Moreover, these cells expressed only low levels of embryonic stem cell (ESC) marker genes including *Cdh1*, *Fgf4*, *Rex1* (also called *Zfp42*), and *Nanog* even at day 49 (Figure 1C), showing that they persist as partially reprogrammed cells. Independent reprogramming by SeVdp(GKOSM) yielded cells with low expression levels of *Cdh1*, *Fgf4*, and *Rex1* reproducibly (Figure 1D). Quantitative western blots (Figure S1C) showed that expression of all the four reprogramming factors, especially KLF4, were lower in SeVdp(GKOSM)-infected cells than in those infected by SeVdp(KOSM) (Figure 1E), presumably because Sendai virus vectors express genes located in its 3′ region at a higher level than those in the 5′ region (Tokusumi et al., 2002) (Figure 1A), and the presence of the EGFP gene in the 3′ region of SeVdp(GKOSM) is detrimental for the expression of the four factors. These findings raise the possibility that lowered expression of reprogramming factors induces partially reprogrammed cells.

Low KLF4 Expression Induces Partially Reprogrammed Cells

To identify which factor was responsible for partial reprogramming, we created four different vectors, SeVdp(fK-OSM), SeVdp(K-fOSM), SeVdp(KO-fSM), and SeVdp(KOS-fM). In each vector, one of the four factors was tagged N-terminally with a destabilization domain (DD), derived from a mutant FKBP12 (Banaszynski et al., 2006) (Figure 2A). This domain facilitates degradation of the tagged protein and decreases its expression level (Figure 2B). In addition, the DD-tagged protein is stabilized by a low molecular weight ligand, Shield1, in a precisely controlled manner (Figure 2B). When each vector was infected to MEFs in the absence of Shield1, the expression levels of DD-tagged KLF4, OCT4, SOX2, and c-MYC were

---

(E) Protein levels of reprogramming factors expressed from SeVdp vectors. SeVdp(KOSM) or SeVdp(GKOSM)-infected MEFs were collected 1 day after infection, and the protein level of each reprogramming factor was determined by western blotting as described in the Experimental Procedures. The protein level of each factor is plotted relative to that in SeVdp(KOSM)-infected cells. Data represent means ± SEM of three independent experiments. **p < 0.01.
See also Figure S1 and Tables S1 and S2.
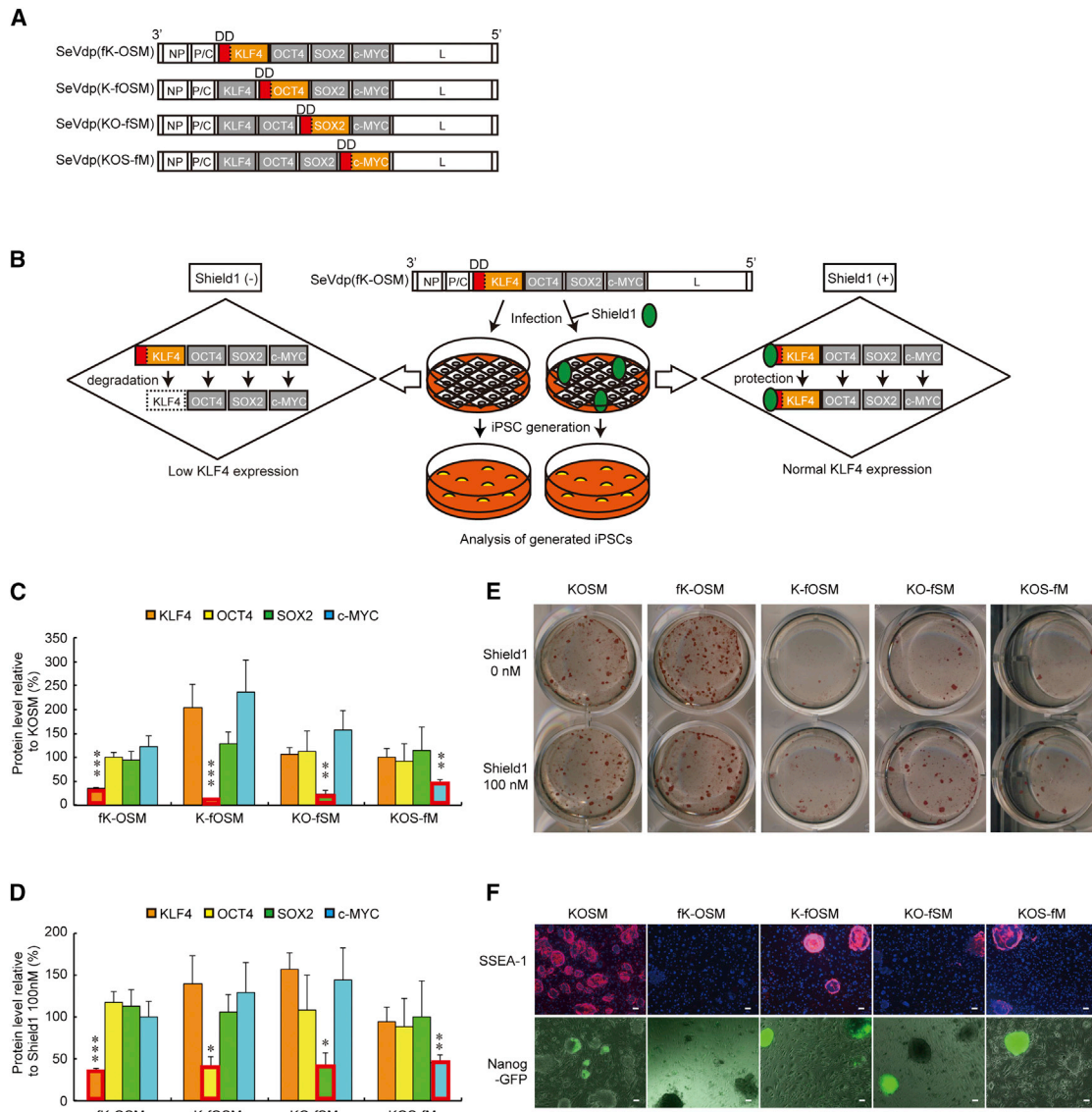
Figure 2. Effects of Downregulated KLF4, OCT4, SOX2, or c-MYC on Reprogramming

(A) Structures of SeVdp vectors in which one of the reprogramming factors is tagged N-terminally with DD.

(B) Stability of DD-tagged reprogramming factor during iPSC generation is regulated by the addition of Shield1.

(C and D) Protein levels of reprogramming factors expressed from SeVdp vectors. Cell extracts were prepared from MEFs infected with each indicated vector one day after infection in the presence or absence of 100 nM Shield1. The protein level of each reprogramming factor is determined by western blotting, and the level in the absence of 100 nM Shield1 is plotted relative to that in SeVdp(KOSM)-infected cells (C) or in the presence of 100 nM Shield1 (D). Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001 versus SeVdp(KOSM) (C) or versus in the presence of 100 nM Shield1 (D).

(E) ALP+ colonies induced by each vector. MEFs (1.0×10^4) infected with the indicated vector were cultured in the presence or absence of 100 nM Shield1 and stained for ALP at day 13.

(F) Expression of SSEA-1 and NANOG. SSEA-1 was detected by immunofluorescence staining at day 18, and NANOG was detected by GFP fluorescence from the *Nanog*-GFP reporter transgene at day 22. Scale bars, 100 μm.

See also Figure S2 and Table S1.

34.4% ± 6.0%, 11.5% ± 5.7%, 22.4% ± 15.8%, and 45.1% ± 14.5% of those in MEFs infected with SeVdp(KOSM), but the three other nontagged factors did not show similar decreases

in expression (Figure 2C). Moreover, the expression levels of DD-tagged KLF4, OCT4, SOX2, and c-MYC in the absence of Shield1 were 34.9% ± 6.2%, 40.3% ± 20.7%,

40.5% ± 28.6%, and 46.0% ± 14.8% of those in the presence of 100 nM Shield1, but the three other nontagged factors showed lesser fluctuations in their expression levels by the addition of Shield1 (Figure 2D).

SeVdp(K-fOSM), SeVdp(KO-fSM), and SeVdp(KOS-fM) gave rise to smaller numbers of ALP$^+$ colonies in the absence of Shield1 than in the presence of 100 nM Shield1 or by the infection with SeVdp(KOSM) (Figure 2E; Table S1). However, these colonies in the absence of Shield1 were positive for SSEA-1 and NANOG (Figure 2F) and far more advanced in reprogramming than the SeVdp(GKOSM)-induced colonies (Figure 1B), indicating that low levels of OCT4, SOX2, and c-MYC merely reduce the efficiency of generating fully reprogrammed iPSCs. By contrast, SeVdp(fK-OSM) gave rise to similar numbers of ALP$^+$ colonies as compared with SeVdp(KOSM) (Figure 2E; Table S1). Moreover, the number of ALP$^+$ colonies induced by SeVdp(fK-OSM) was essentially the same regardless of the presence of Shield1 (Figure 2E; Table S1), and the colonies induced in the absence of Shield1 were negative for SSEA-1 and NANOG (Figure 2F), displaying a phenotype common with the SeVdp(GKOSM)-induced ones (Figure 1B). Thus, low KLF4 expression in SeVdp(GKOSM)-infected cells is the main cause of generating partially reprogrammed iPSCs, and this phenomenon can be recapitulated by SeVdp(fK-OSM) in the absence of Shield1.

To further confirm that partially reprogrammed iPSCs were induced by low KLF4 expression, we prepared vectors that express KLF4 tagged C-terminally to the PEST or CL1 degradation signal (Gilon et al., 1998; Rogers et al., 1986) (Figure S2A). When KLF4 was tagged only with PEST [SeVdp(Kp-OSM)], KLF4 expression was decreased only moderately (48.0% ± 7.8%) (Figure S2B), and the induced cells were not only positive for ALP (Figure S2C) but also expressed ESC markers, *Cdh1*, *Fgf4*, and *Rex1* (Figure S2D). Tagging KLF4 with both PEST and CL1 [SeVdp(Kcp-OSM)] decreased KLF4 expression slightly lower than that by SeVdp(fK-OSM) in the absence of Shield1 (23.8% ± 13.6%) (Figure S2B). Indeed, SeVdp(Kcp-OSM) induced ALP$^+$ colonies (Figure S2C) that did not express ESC markers (Figure S2D), indicating that the induced cells are partially reprogrammed as in the case of SeVdp(fK-OSM) (Figures 2E and 2F). However, removal of KLF4 [SeVdp(OSM)] resulted in a drastically reduced number of colonies (Figure S2C). These results show that partially reprogrammed iPSCs are generated by low KLF4 expression and not by the untoward effect by the tagged DD.

Different Levels of KLF4 Expression Reprogram Cells to Different Degrees

We next examined if iPSCs generated by SeVdp(fK-OSM) acquire pluripotency when KLF4 expression is increased by Shield1. When SeVdp(fK-OSM) was infected to MEFs in the presence of 100 or 200 nM Shield1, the KLF4 expression level was restored almost to that by SeVdp(KOSM) (Figure 3A), which generates iPSCs that contribute to mouse chimeras (Nishimura et al., 2011). Expression analyses showed that cells induced by SeVdp(fK-OSM) in the presence of 100 nM Shield1 expressed the 12 marker genes in analogous time courses to those induced by SeVdp(KOSM) (Figure 3B). By contrast, the cells induced by SeVdp(fK-OSM) in the absence of Shield1 failed to completely downregulate MET-related genes and to upregulate ESC marker genes, displaying gene expression patterns similar to SeVdp(GKOSM)-induced cells (Figures 1C and 3B). As in the case of the colonies induced by SeVdp(GKOSM), the expression and localization of β-CATENIN in SeVdp(fK-OSM)-induced colonies were indistinguishable from those in SeVdp(KOSM)-induced ones (Figure S1B). Importantly, addition of Shield1 to the cells induced by SeVdp(Kcp-OSM) was not accompanied by the expression of the ESC-specific marker genes (Figure S2D), showing that Shield1 itself has no effect on induction of pluripotency.

These iPSCs, induced in the presence of 100 nM Shield1, were easily cleared of the SeVdp(fK-OSM) genome by an small interfering RNA (siRNA) against the *L* gene (Nishimura et al., 2011) that encodes the viral polymerase (Figures 2A and 3C), confirming that they have acquired a self-sustaining pluripotency state. When injected into mouse blastocysts, the vector-free iPSCs (Sh100-36) contributed to mouse chimeras, which subsequently gave rise to a live pup (236-18-3) carrying the *Nanog*-GFP transgene (Figure 3D). Thus, in the presence of 100 nM Shield1, DD-tagged KLF4 functions indistinguishably from nontagged KLF4 and permits SeVdp(fK-OSM) to reprogram MEFs to acquire pluripotency.

We then used SeVdp(fK-OSM) and different concentrations of Shield1 to analyze the dose-dependent effect of KLF4 on the reprogramming process. Infection of MEFs with SeVdp(fK-OSM) in the presence of 10, 30, or 100 nM Shield1 increased the KLF4 level in a dose-dependent manner while maintaining the expression of OCT4, SOX2, and c-MYC essentially unchanged (Figure 3E). In the presence of 10 or 30 nM Shield1, SeVdp(fK-OSM) gave rise to reprogrammed cells, most of which expressed SSEA-1, but not NANOG. The reprogramming of these cells is more advanced than ALP$^+$, SSEA-1-negative (SSEA-1$^-$), NANOG$^-$ cells induced by SeVdp(fK-OSM) in the absence of Shield1 but less advanced than ALP$^+$, SSEA-1$^+$, NANOG$^+$ cells induced by SeVdp(fK-OSM) in the presence of 100 nM Shield1 (Figure 3F) or by SeVdp(KOSM) (Figure 2F). The intermediate degree of reprogramming is also supported by the expression levels of *Cdh1*, *Fgf4*, and *Rex1* (Figure 3G).

It is of note that independently performed reprogramming by SeVdp(fK-OSM) yielded cells with similar
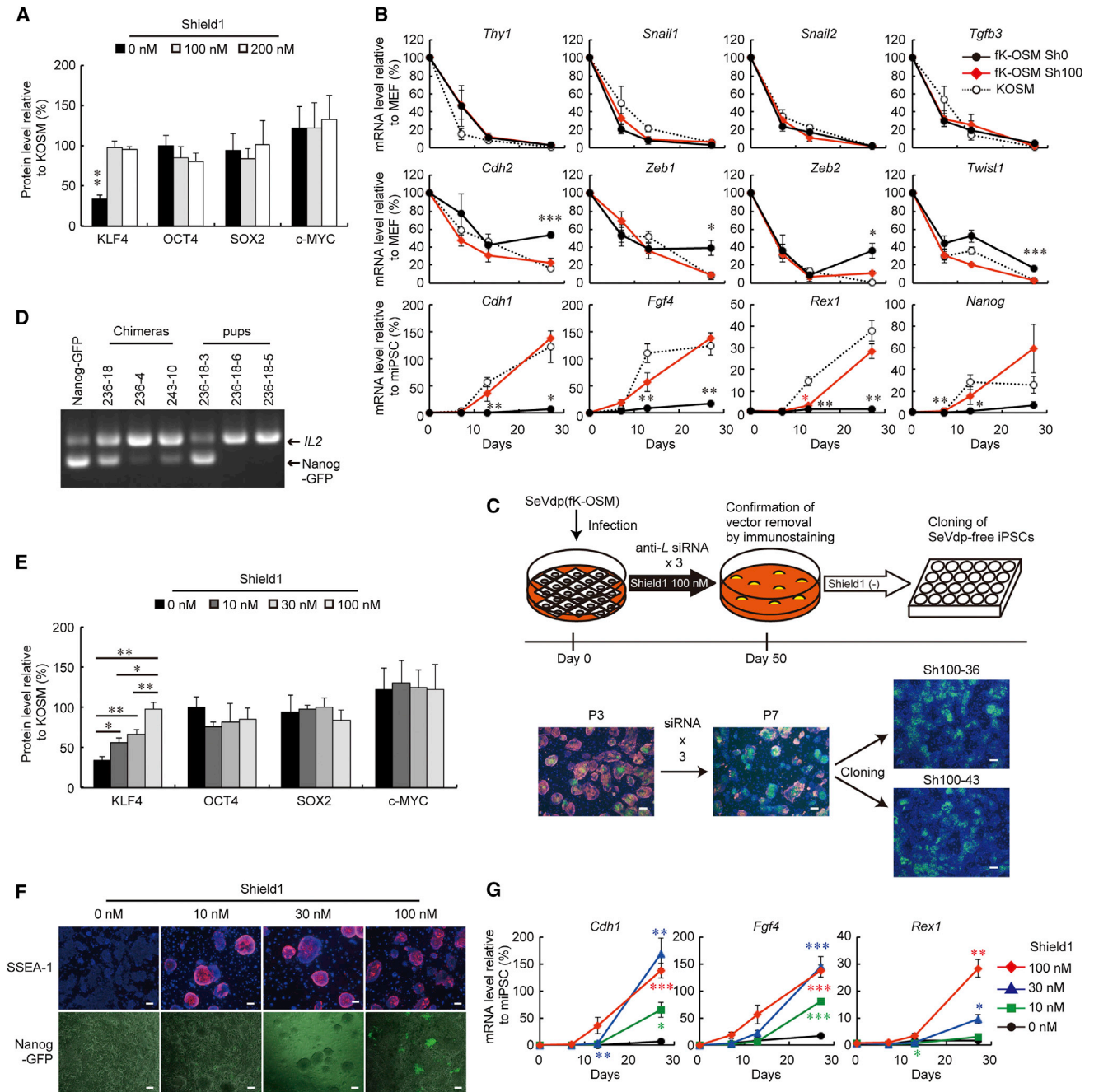
Figure 3. iPSC Generation by Manipulating the Protein Level of Exogenous KLF4

(A) Protein levels of reprogramming factors expressed from SeVdp(fK-OSM). Cell extracts were prepared from MEFs infected with SeVdp(fK-OSM) in the presence of the indicated concentration of Shield1 for one day. Protein levels are plotted relative to those in SeVdp(KOSM)-infected cells. Data represent means ± SEM of three independent experiments. **p < 0.01 versus SeVdp(KOSM).

(B) Time course of mRNA expression. SeVdp(fK-OSM)- or SeVdp(KOSM)-infected cells cultured in the presence (Sh100) or absence (Sh0) of 100 nM Shield1 were collected at days 7, 13, and 27 for quantitative RT-PCR. Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01 versus SeVdp(KOSM)-infected cells.

(C) Vector-free iPSCs generated by SeVdp(fK-OSM). The cells induced by SeVdp(fK-OSM) in the presence of 100 nM of Shield1 (P3) were treated three times with L527 siRNA (P7), and the vector-free iPSCs were isolated at day 50 (Sh100-36 and Sh100-43). Green, Nanog-GFP. Red, NP of SeV. Scale bars, 100 μm.

(legend continued on next page)

expression profiles of *Cdh1*, *Fgf4*, and *Rex1* (Figure S3A), showing that generation of partially reprogrammed iPSCs by SeVdp(fK-OSM) is, by and large, reproducible in independent experiments. Moreover, in the presence or absence of 10 nM Shield1, individual colonies isolated from the same plate showed similar gene expression profiles of *Cdh1* and *Fgf4* (Figure S3B), showing that SeVdp(fK-OSM)-induced colonies in the same plate are relatively homogeneous. Although *Rex1* was expressed at somewhat varied levels, between 2% and 30% of the miPSC level, in the independent clones induced in the presence of 10 nM Shield1, this may reflect the intrinsically heterogeneous expression of *Rex1* in mouse ESCs (Toyooka et al., 2008). Together, these results demonstrate that low KLF4 expression arrests the reprogramming process prematurely, and SeVdp(fK-OSM) yields partially reprogrammed iPSCs arrested at distinct stages depending on the amount of added Shield1. Given these unique properties, we named this as SeVdp-based stage-specific reprogramming system (3S reprogramming system).

Upregulation of KLF4 Expression Allows Partially Reprogrammed iPSCs to Resume Reprogramming

We addressed whether the partially reprogrammed cells have arrested but remained on a path toward pluripotency or have diverted from the path and become aberrantly reprogrammed. After MEFs were infected by SeVdp(fK-OSM) in the absence of Shield1 for 17 days, 100 nM Shield1 was added to the cell culture (Figure 4A). This increased the expression of DD-tagged KLF4 by more than 3-fold (Figure 4B), as revealed by anti-FKBP12 and anti-KLF4 antibodies. The upregulation of KLF4 expression at day 17 caused a gradual increase in the expression of ESC markers, NANOG (Figure 4C), *Cdh1*, *Fgf4*, and *Rex1* (Figure 4D), indicating that the cells resumed reprogramming toward pluripotency. Similarly, partially reprogrammed cells, initially induced by SeVdp(fK-OSM) in the presence of 10 or 30 nM Shield1, also showed a gradual increase in the expression of NANOG upon addition of 100 nM Shield1 (Figure S4). Thus, the partially reprogrammed iPSCs generated by the 3S reprogramming system are competent to resume reprogramming upon the increased KLF4 level and hence should be regarded as "paused iPSCs."

Next, we asked if the paused iPSCs are stable and retain the ability to resume reprogramming after a long period of cell culture. We induced partially reprogrammed cells using SeVdp(fK-OSM) in the absence of Shield1, and individual clones were isolated at day 20 and cultured for additional 30 days under the same condition (Figure 4E). Even after additional culture for 30 days (between day 20 and 50), the cloned cells, Sh0-2 and Sh0-5, remained negative for SSEA-1 (Figure 4F) as well as for *Cdh1*, *Fgf4*, and *Rex1* (Figure 4G). Moreover, addition of 100 nM Shield1 at day 50 allowed the cells to resume reprogramming and to express SSEA-1 (Figure 4F) as well as *Cdh1*, *Fgf4*, and *Rex1* (Figure 4G), in similar time courses to those observed for the cells to which 100 nM Shield1 was added at day 17 (Figure 4D).

To further confirm that increased expression of KLF4, but not other Shield1-induced genes, allowed the partially reprogrammed cells to resume reprogramming, we used retrovirus-mediated gene transfer to introduce KLF4, OCT4, SOX2, or c-MYC into the paused iPSCs generated by SeVdp(fK-OSM) in the absence of Shield1. As shown in Figure 4H, additional expression of KLF4, but not other reprogramming factors, caused the cells to express SSEA-1 to a similar extent to the iPSCs generated by SeVdp(fK-OSM) in the presence of 100 nM Shield1, suggesting that the cells resume reprogramming even when KLF4 is provided by retrovirus-mediated gene transfer. Together, these results show that the paused iPSCs generated by the 3S reprogramming system are relatively stable and retain the ability to resume reprogramming for over a month.

Global Gene Expression Analyses of Paused iPSCs

To further characterize the paused iPSCs, we performed the analysis of their global gene expression. Hierarchical clustering of cells using differently expressed genes (DEGs) between MEFs and mouse iPSCs (miPSCs) (Nishimura et al., 2011) showed that SeVdp(fK-OSM)-induced cells are more similar to miPSCs than MEFs, and the cells are clustered together according to the amount of added Shield1 (Figure 5A). Global gene expression profiles also

(D) Production of chimeras from iPSCs established by SeVdp(fK-OSM) to show germline transmission. Genomic DNAs of the chimeric mice generated by the microinjection of clone Sh100-36 or Sh100-43, and their pups were analyzed by PCR to detect the *Nanog*-GFP transgene. The IL-2 gene was amplified as an internal control.

(E) Gradual increase of KLF4 expression by Shield1. Cell extracts were prepared as in (A). Protein levels are plotted relative to those in SeVdp(KOSM)-infected cells. Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01.

(F) Expression of SSEA-1 (day 27) and NANOG (day 29) in the cells cultured at indicated concentrations of Shield1. Scale bars, 100 μm.

(G) The *Cdh1*, *Fgf4*, and *Rex1* mRNA levels were determined by quantitative RT-PCR at days 7, 13, and 27 for SeVdp(fK-OSM)-infected cells cultured at indicated concentrations of Shield1. Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001 versus SeVdp(fK-OSM) in the absence of Shield1.

See also Figures S1 and S3 and Table S2.
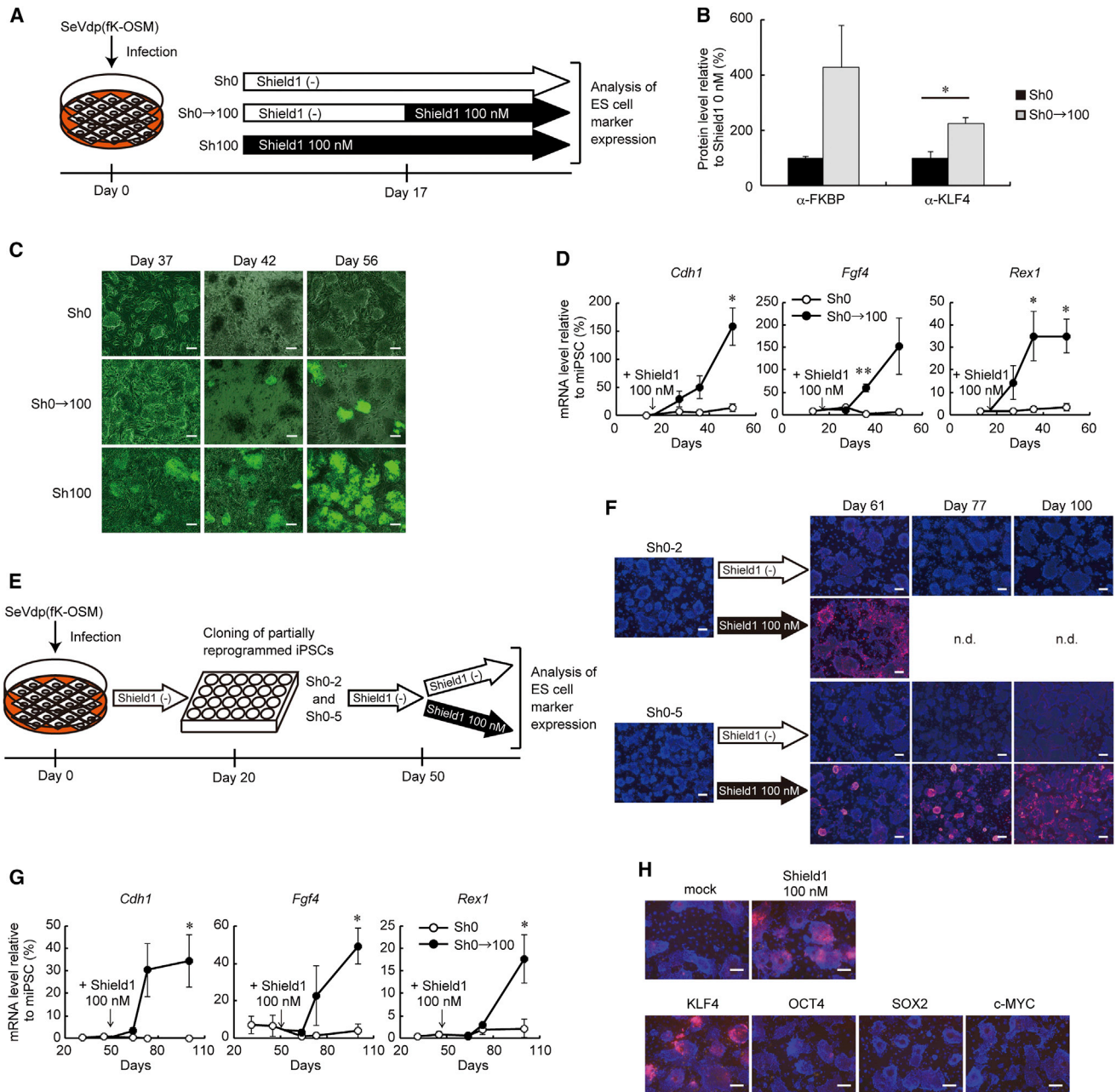
Figure 4. Upregulation of Exogenous KLF4 Resumes Reprogramming

(A) Outline of the experiments to restart reprogramming in partially reprogrammed iPSCs generated by SeVdp(fK-OSM) through the addition of Shield1 at day 17.

(B) Upregulation of the DD-tagged KLF4 protein level in partially reprogrammed iPSCs after 5 days of Shield1 addition. The protein levels of DD-tagged KLF4 or total KLF4 was determined by western blotting using an anti-FKBP or anti-KLF4 antibody, respectively, and were plotted relative to those in the cells cultured in the absence of Shield1. Data represent means ± SEM of three independent experiments. *p < 0.05.

(C) GFP fluorescence of the *Nanog*-GFP transgene was observed at indicated time points. Scale bars, 100 μm.

(D) The *Cdh1*, *Fgf4*, and *Rex1* mRNA levels were determined by quantitative RT-PCR at day 16, 27, 36, and 49. 100 nM Shield1 was added at day 17. Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01 versus no addition of Shield1.

(E) Partially reprogrammed iPSC clones were isolated at day 20 (Sh0-2 and Sh0-5) and cultured in the absence of Shield1 until day 50, when Shield1 was added.

(F) SSEA-1 expression was analyzed at indicated time points. Scale bars, 100 μm.

(legend continued on next page)

corroborate that the paused iPSCs generated in the absence of Shield1 are stable (Sh0 22d and Sh0 79d, Figure 5A) and resume reprogramming upon the addition of 100 nM Shield1 (Sh0→100 79d, Figure 5A). Principal component analysis indicates that SeVdp(fK-OSM)-induced cells have paused at intermediate stages between MEFs and miPSCs, and that the addition of a higher amount of Shield1 reprograms the cells closer to miPSCs (Figure S5A). The analysis also showed that the paused iPSCs generated by SeVdp(fK-OSM) in the absence of Shield1 resume reprogramming and advance toward pluripotency upon the addition of 100 nM Shield1 (Figure S5B). A recent study reported that mesendodermal genes are transiently activated during reprogramming and subsequently repressed again in iPSCs (Takahashi et al., 2014). Consistent with this, paused iPSCs express some of the mesendodermal genes (*Eomes*, *Mixl1*, and *T*), which are nonetheless repressed in fully reprogrammed iPSC produced by SeVdp(KOSM) or SeVdp(fK-OSM) in the presence of 100 nM Shield1 (Figures S5C–S5E). These gene expression profiles again confirm that paused iPSCs are intermediates during reprogramming.

Hierarchical clustering of DEGs also revealed the presence of five distinct gene groups (Figure 5B). Group D includes a set of genes whose expression is already downregulated in paused iPSCs (Sh0 22d). The genes in groups A and B increase expression in a stepwise manner from MEFs to paused iPSCs (Sh0 22d) and from paused iPSCs (Sh0 22d) to iPSCs (Sh100 22d), respectively. Pathway analysis shows that Podocyte-related genes, which are involved in the epithelial-to-mesenchymal transition (EMT) of podocytes, are enriched in groups A and D (Figure 5C), indicating that the mesenchymal-to-epithelial transition (MET), a reverse process of EMT, separates paused iPSCs (Sh0 22d) from MEFs. This result is consistent with a previous report that showed the important role for KLF4 in MET during reprogramming (Li et al., 2010). Furthermore, PluriNetWork genes, which are associated with pluripotency, are enriched in groups A and B (Figure 5C), and heatmaps show that all the PluriNetWork genes in groups A and B increase their expression in a Shield1-dependent manner (Figure 5D), as if increasing the amount of KLF4 recapitulates the chronological progression of reprogramming. Independent quantitative RT-PCR analysis of selected pluripotency-related genes (Kim et al., 2008; O'Malley et al., 2013; Xu et al., 2010) identified two categories of genes that show distinct expression profiles in response to the Shield1 amount; genes that increase expression gradually as the amount of Shield1 is increased (upper panel, Figures 5E and S5F–S5M) and genes that increase expression markedly at 30 nM or higher concentrations of Shield1 (lower panel, Figures 5E and S5N–S5U). The former and the latter categories of genes correspond more or less to the genes in group A and group B, respectively (Figures 5D and 5E), demonstrating that pluripotency-related genes respond to the KLF4 level by at least two distinguishable modes. Together, the global gene expression analyses indicate that the iPSCs generated by the 3S reprogramming system are paused at critical stages of reprogramming and suggest that the KLF4-dependent transitions from one stage to the next is accompanied by changes in the expression of genes that are crucial for reprogramming.
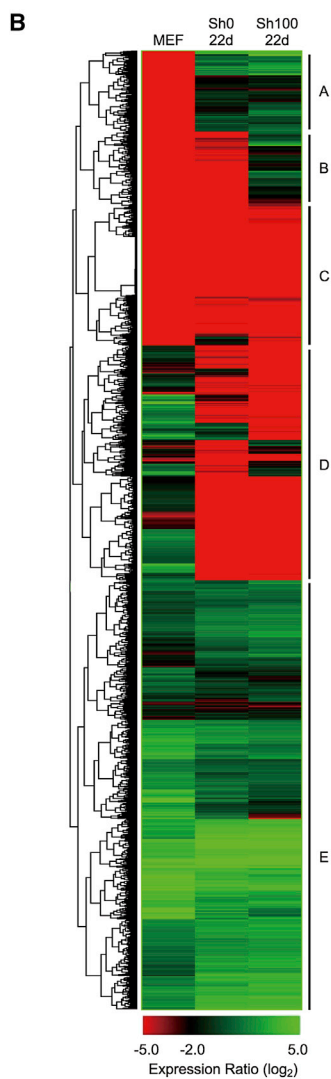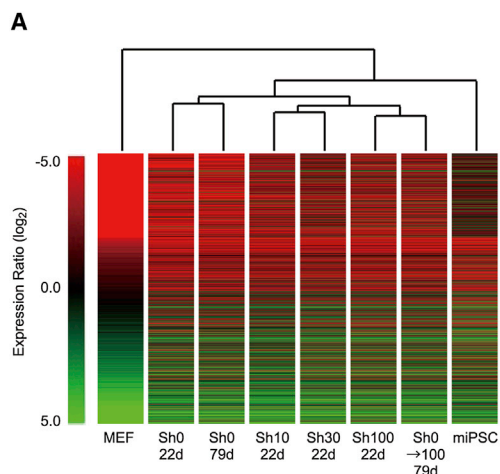
Generation of Partially Reprogrammed Human iPSCs

Although molecular mechanisms of mouse iPSC generation have been investigated substantially, there have been relatively few mechanistic analyses of human iPSCs. Because our SeVdp vectors have been used successfully to generate human iPSCs from various somatic cells (Nishimura et al., 2013; Takayama et al., 2010; Tateno et al., 2013; Wakao et al., 2013), we tested if the 3S reprogramming system generates human iPSCs that have paused during reprogramming. When SeVdp(KOSM) and SeVdp(GKOSM) were infected to human fibroblasts, both vectors produced ALP$^+$ colonies of small, rapidly growing cells in a similar manner to MEFs, although SeVdp(GKOSM) induced slightly more colonies than SeVdp(KOSM) (Figure 6A; Table S1). SeVdp(GKOSM)-induced human cells are similar to their mouse counterparts (Figure 1B) in that they are positive for ALP but negative for ESC markers, *CDH1*, *FGF4*, and *REX1* (Figure 6B).

When human fibroblasts were infected with SeVdp(fK-OSM) in the presence or absence of 100 nM Shield1, similar numbers of ALP$^+$ colonies were obtained regardless of the presence of Shield1 (Figure 6A; Table S1). However, induction in the absence of Shield1 generated cells that were negative for *CDH1*, *FGF4*, and *REX1*, whereas induction in the presence of 100 nM Shield1 generated cells that expressed these ESC marker genes (Figure 6B). Moreover, as in the case of mouse iPSCs (Figure 3C), hiPSCs generated by SeVdp(fK-OSM) in the presence of 100 nM Shield1 were cleared of the vector by the siRNA against the *L* gene to establish vector-free hiPSCs (Figure S6). These results
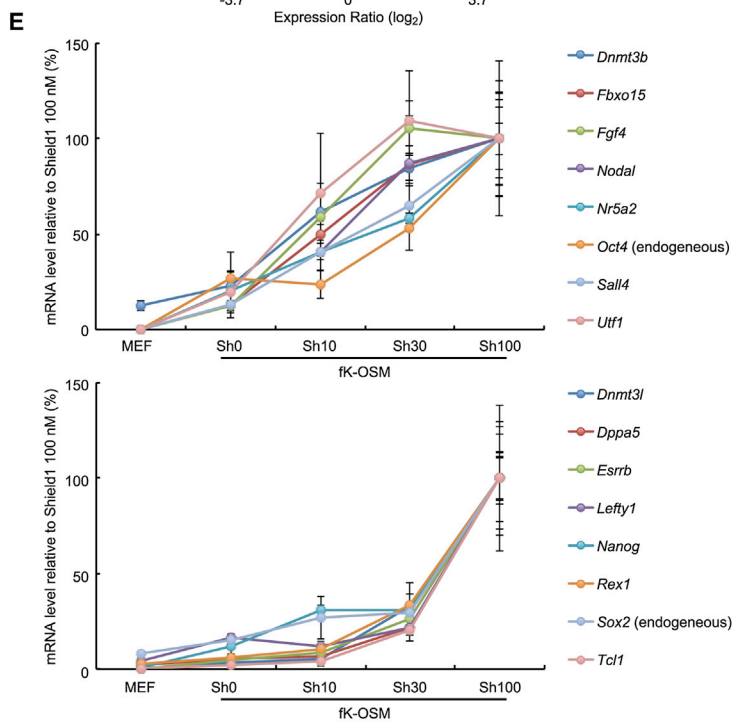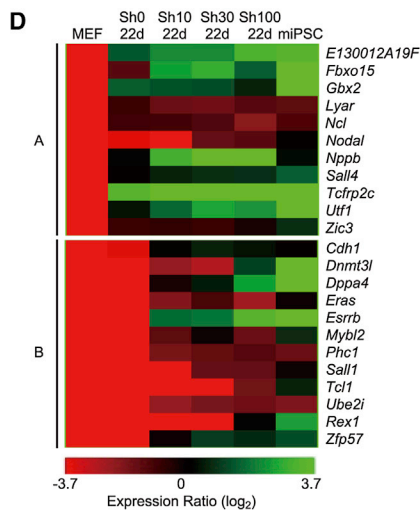
(G) Clone Sh0-5 was cultured in the absence of Shield1 until day 50, when Shield1 was added. The mRNA levels of *Cdh1*, *Fgf4*, and *Rex1* were determined at day 31, 45, 64, 73, and 100. Data represent means ± SEM of three independent experiments. *p < 0.05 versus no addition of Shield1.

(H) SSEA-1 expression was analyzed in paused iPSCs (Sh0-5) 29 days after the addition of 100 nM Shield1 or infection with a retroviral vector expressing KLF4, OCT4, SOX2, or c-MYC. Scale bars, 100 μm.

See also Figure S4 and Table S2.

(legend on next page)

indicate that SeVdp(fK-OSM) generates partially reprogrammed human iPSCs that are similar to the paused mouse iPSCs as well as fully reprogrammed iPSCs that have acquired a self-sustaining pluripotency state. Thus, the 3S reprogramming system can be utilized to generate partially reprogrammed human iPSCs, and these cells will be useful for analyzing the mechanism of human iPSC generation.

DISCUSSION

Understanding the mechanisms of iPSC generation is important for application for cell therapy and also provides insights into the molecular mechanisms of normal developmental processes. The analysis, however, suffers from technical limits because the intermediate cells in the process of reprogramming are transient, asynchronous and heterogeneous. Cell sorting enriches intermediate cells, but they are transient and limited in amount (Hansson et al., 2012; O'Malley et al., 2013; Polo et al., 2012). For example, SSEA-1$^+$ Oct4$^-$ reprogramming intermediates, initially isolated by fractionation using SSEA-1 antigen, diversify into a heterogeneous population of cells upon continued cell culture; some of them revert to lose SSEA-1 expression, whereas others advance to initiate Oct4 expression (Polo et al., 2012). Partially reprogrammed cell populations obtained during iPSC generation are relatively stable and expandable, but they are usually sporadic and unpredictable byproducts of reprogramming (Mikkelsen et al., 2008; Sridharan et al., 2009), sometimes generating aberrantly reprogrammed cells that do not transit into the pluripotent state (Mikkelsen et al., 2008).

Unlike these previous methods, the 3S reprogramming system generates partially reprogrammed iPSCs in a reproducible and predictable manner by regulating the amount of KLF4. These cells resume reprogramming upon increased KLF4 expression and therefore are not aberrantly reprogrammed but have paused during iPSC generation. Different KLF4 levels yield paused iPSCs at a series of stages during iPSC generation, enabling the molecular analysis of orderly events that proceed during reprogramming. Moreover, the paused iPSCs are relatively homogeneous and phenotypically stable, allowing expansion of cells for further analyses. Last, this system is based upon a defective Sendai virus, which replicates exclusively in the cytoplasm (Nishimura et al., 2011), and thus is free from both the positional effects and the silencing of transgenes that may arise in the systems based upon the retrovirus or lentivirus (Jaenisch et al., 1981). Therefore, our system to generate paused iPSCs is a significant advance over the sorted cells and partially reprogrammed cell populations previously used for mechanistic analyses of iPSC generation.

It remains to be determined how the low expression of KLF4, but not the other three factors, generates paused iPSCs. It is known that KLF4 represses somatic gene expression together with c-MYC at an early stage of iPSC generation and enhances pluripotency gene expression in cooperation with OCT4 and SOX2 at its late stage (Polo et al., 2012). Therefore, reduced levels of KLF4 may be sufficient for repressing somatic cell-specific genes, but not for activating pluripotency-related genes. A recent study also showed that KLF4 organizes long-range chromosomal interactions with the *Oct4* locus by recruiting cohesion (Wei et al., 2013). These chromosomal interactions are observed not in SSEA-1$^-$ cells but only in SSEA-1$^+$ cells during reprogramming, suggesting that KLF4 may mediate the chromosomal reorganization during the transition from the SSEA-1$^-$ to SSEA-1$^+$ cells. Interestingly, this corresponds to a stage where SeVdp(fK-OSM) generates paused iPSCs in the absence of Shield1. It is thus possible that such dynamic chromosomal reorganization requires a high level of KLF4 and presents a roadblock for reprogramming with a low expression level of KLF4.

Recently, Sui et al. reported production of partially reprogrammed iPSCs using trimethoprim (TMP)-regulated destabilizing domain to control the expression of reprogramming factors from a *piggyBac* transposon vector (Sui et al., 2014). In contrast to our 3S reprogramming system, SOX2 rather

Figure 5. Global Gene Expression Profiles of Paused iPSCs
(A) Hierarchical clustering of paused iPSCs based on gene expression profiles of the differentially expressed genes (DEGs) between MEFs and miPSCs (p < 0.05). Analyzed samples include cells induced by SeVdp(fK-OSM) at the indicated concentrations (0, 10, 30, or 100 nM) of Shield1 for 22 or 79 days. Cells that resumed reprogramming (Sh0→100 79d) were first induced by SeVdp(fK-OSM) in the absence of Shield1 for 25 days and then cultured in the presence of 100 nM Shield1 before collection at day 79.
(B) Hierarchical clustering of DEGs. DEGs between MEFs and miPSCs (p < 0.05) were clustered based on gene expression profiles of the indicated cells and classified into five groups (group A–E).
(C) Enriched pathways in the genes of groups A, B, or D.
(D) Heatmap of the expression profiles of PluriNetWork genes clustered in group A or B.
(E) Two distinct expression profiles of pluripotency-related genes at increasing concentrations of Shield1. SeVdp(fK-OSM)-infected cells were cultured in the presence of indicated concentrations of Shield1, and the cells were collected at day 27. The mRNA levels of pluripotency-related genes were determined by quantitative RT-PCR. Data represent mean ± SEM of three independent experiments. See also Figure S5 and Table S2.

A

B
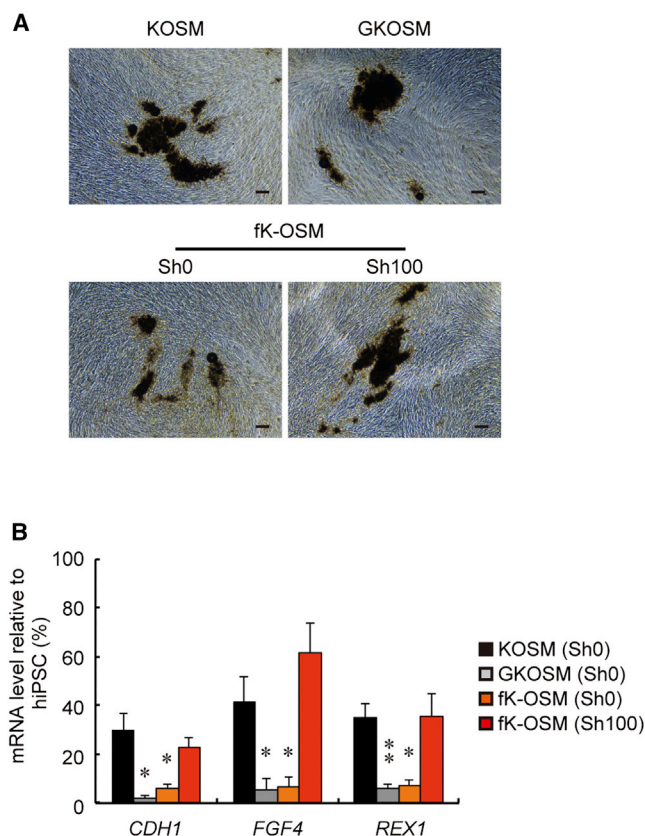
Figure 6. Generation of Partially Reprogrammed iPSCs from Human Somatic Cells

(A) Generation of ALP[+] human iPSCs by SeVdp(fK-OSM). The indicated SeVdp vectors were used to generate human iPSCs, and the cells were stained for ALP at day 16. Scale bars, 100 μm.

(B) Expression of ESC markers in partially reprogrammed human iPSCs. The mRNA levels of *CDH1*, *FGF4*, and *REX1* in the cells infected by the indicated vector in the presence or absence of 100 nM Shield1 were determined by quantitative RT-PCR at day 18. Data represent means ± SEM of three independent experiments. *p < 0.05, **p < 0.01 versus SeVdp(KOSM)-infected cells.
See also Figure S6 and Tables S1 and S2.

than KLF4 plays a predominant role in regulating the progression of reprogramming. Several reasons may account for this apparent discrepancy. First, each system employed a different combination of the vector and reprogramming factors for iPSC generation. Sui et al. used only OCT4, SOX2, and KLF4 in *piggyBac* transposon vectors whereas our system used OCT4, SOX2, KLF4, and c-MYC in SeVdp vectors. Second, unlike the FKBP-based domain used in our system, the destabilizing domain used by Sui et al. seems to promote more complete protein degradation. Third, reprogramming may proceed via distinct pathways in each system and therefore the role for each reprogramming factor may not be identical between different systems. In fact,

recent studies indicate the presence of multiple pathways by which reprogramming proceeds toward pluripotency (O'Malley et al., 2013; Parchem et al., 2014; Rais et al., 2013). Given this possibility, the availability of different systems that may generate iPSCs stalled at alternative intermediate states should be invaluable for more comprehensive understanding of the reprogramming process.

Regardless of the effect of the KLF4 expression level on the progression of reprogramming, the 3S reprogramming system described here provides a powerful tool for analyzing the mechanism of iPSC generation, particularly with respect to its ill-defined intermediate stages. One such example is MET, a reverse process of EMT that occurs during normal development. Consistent with the role for KLF4 in MET (Li et al., 2010), the paused iPSCs generated in the absence of Shield1 have stalled around MET and show gene expression profiles reminiscent of the ongoing MET. Another is the primitive streak-like mesendodermal (PSMN) state, where cells transiently display gene expression profiles resembling mesendoderm (Takahashi et al., 2014), which forms part of the primitive streak during early embryonic development. Paused iPSCs generated in the presence of 10–30 nM Shield1 express mesendoderm-related genes but have not reached the fully reprogrammed state. Because the addition of Shield1 permits the paused iPSCs to advance toward the fully reprogrammed state in a controlled manner, the 3S reprogramming system may be especially suitable for analyzing the MET stage, the mesendodermal stage as well as the late stage toward pluripotency. Finally, in addition to MEF-based iPSC generation, the current system also works for human somatic cells. Thus, the 3S reprogramming system will offer a unique opportunity to analyze transient and unstable stages of iPSC generation from mouse and human somatic cells.

EXPERIMENTAL PROCEDURES

Cell Culture and Generation of iPSCs

Mouse iPSCs were generated from mouse embryonic fibroblasts (MEFs) isolated from a transgenic mouse carrying the *Nanog*-GFP-IRES-Puro[r] reporter construct (Okita et al., 2007) (provided by the RIKEN BioResource Center) (*Nanog*-GFP MEFs). After *Nanog*-GFP MEFs were infected with each SeVdp vector at 32°C for 14 hr at an MOI of about 5, 2.0×10^4 of infected cells were seeded onto mitomycin C-treated MEFs in 6-well plates and cultured in KSR medium (Knockout-DMEM (Life Technologies) supplemented with 15% Knockout Serum Replacement (Life Technologies), 2 mM GlutaMAX (Life Technologies), 0.1 mM nonessential amino acids (Life Technologies), 55 μM 2-mercaptoethanol (Life Technologies), 100 U/ml penicillin-streptomycin (Wako), and 1,000 U/ml LIF [Wako]) for 7 days, and then in mES medium (DMEM; nacalai tesque) supplemented with 15% fetal bovine serum (Hyclone), 0.1 mM nonessential amino acids, 55 μM 2-mercaptoethanol, 100 U/ml penicillin-streptomycin and 1,000 U/ml LIF), in the

presence or absence of Shield1 (Takara Bio). The presence of the vector in MEFs was confirmed 1 day after infection by immuno-staining using an anti-SeV NP monoclonal antibody (Nishimura et al., 2011) as described in Supplemental Experimental Procedures. Reprogrammed cells were maintained on mitomycin C-treated SNL76/7 feeder cells (DS Pharma Biomedical) in the same mES medium, which was changed every 2 days. To remove SeVdp vectors from the reprogrammed cells, the cells were treated with 40 nM of L527 siRNA (Nishimura et al., 2011) mixed with Lipofectamine RNAi MAX Transfection Reagent (Life Technologies) in every passage after about day 20.

Human iPSCs were generated from human skin fibroblast NB1RGB cells (RIKEN BioResource Center). The cells were infected with SeVdp vectors at room temperature for 2 hr followed by 37°C for 14 hr at an MOI of about 5, and 2.0×10^4 of infected cells were seeded onto mitomycin C-treated MEFs in 6-well plates and cultured in Primate ESC culture medium (ReproCell) with 5 ng/ml basic FGF (Wako).

Retroviral vectors were infected to partially reprogrammed Sh0-5 cells cultured under a feeder-free condition at an MOI of about 10 for 2 days with 8 μg/ml of Hexadimethrine bromide (polybrene; Sigma). The cells were then passaged onto feeder cells and cultured in mES medium.

Determination of Protein Expression by Western Blotting

After SeVdp vectors were infected to MEFs, the cells were collected one day after infection and lysed on ice in SDS-PAGE gel loading buffer (50 mM Tris-HCl [pH 6.8], 2% SDS, 10% glycerol, 0.1 mg/ml Bromophenol blue, 5% 2-mercaptoethanol), followed by boiling for 5 min, and the whole-cell extracts were subjected to western blot analysis. Antibodies against SOX2, KLF4 or c-MYC were prepared by immunizing rabbits with each purified recombinant protein, as described previously (Fukuda et al., 2013). Each full-length recombinant protein was expressed in BL21(DE3)pLysS, separated by SDS-PAGE, and electroeluted. The purified recombinant proteins were used to immunize rabbits by Japan Bio Serum, and the obtained sera were affinity-purified by a corresponding antigen column, in which glutathione S-transferase (GST)-tagged proteins was crosslinked to glutathione Sepharose 4B (GE Healthcare). Anti-OCT4 (1:5,000, ab19857, Abcam) and anti-α-TUBULIN (1:10,000, ab7291, Abcam) were purchased commercially. Blots were probed with anti-rabbit or anti-mouse IgG-HRP secondary antibody (1:5,000, GE Healthcare) and visualized using Luminol, p-Coumaric acids, and H_2O_2 (Mruk and Cheng, 2011). Intensity of the chemiluminescence was quantified by LAS4000 (GE Healthcare) using ImageQuant TL software (GE Healthcare). We calculated the relative protein expression level of each reprogramming factor by normalizing on the basis of α-TUBULIN protein expression as described in Figure S1C. Determination of protein expression was carried out from biologically triplicate experiments.

Characterization of Mouse and Human iPSCs

Detailed methods for immunofluorescence, ALP staining, quantitative RT-PCR, microarray analysis, and generation of chimeric mice are described in Supplemental Experimental Procedures.

All animal experiments were performed according to the Guide for the Care and Use of Laboratory Animals at the University of Tsukuba.

REFERENCES

Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G., and Wandless, T.J. (2006). A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. Cell 126, 995–1004.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2, 151–159.

Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creyghton, M.P., et al. (2011). Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. Cell Stem Cell 9, 588–598.

Fukuda, A., Shimada, M., Nakadai, T., Nishimura, K., and Hisatake, K. (2013). Heterogeneous nuclear ribonucleoprotein R cooperates with mediator to facilitate transcription reinitiation on the c-Fos gene. PLoS ONE 8, e72496.

Gilon, T., Chomsky, O., and Kulka, R.G. (1998). Degradation signals for ubiquitin system proteolysis in Saccharomyces cerevisiae. EMBO J. 17, 2759–2766.

Hansson, J., Rafiee, M.R., Reiland, S., Polo, J.M., Gehring, J., Okawa, S., Huber, W., Hochedlinger, K., and Krijgsveld, J. (2012). Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. Cell Reports 2, 1579–1592.

Jaenisch, R., Jähner, D., Nobis, P., Simon, I., Löhler, J., Harbers, K., and Grotkopp, D. (1981). Chromosomal position and activation of

retroviral genomes inserted into the germ line of mice. Cell *24*, 519–529.

Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. Cell *132*, 1049–1061.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., He, W., Chen, J., Li, F., Zhuang, Q., et al. (2010). A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. Cell Stem Cell *7*, 51–63.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature *454*, 49–55.

Mruk, D.D., and Cheng, C.Y. (2011). Enhanced chemiluminescence (ECL) for routine immunoblotting: An inexpensive alternative to commercially available kits. Spermatogenesis *1*, 121–122.

Nagamatsu, G., Saito, S., Kosaka, T., Takubo, K., Kinoshita, T., Oya, M., Horimoto, K., and Suda, T. (2012). Optimal ratio of transcription factors for somatic cell reprogramming. J. Biol. Chem. *287*, 36273–36282.

Nishimura, K., Segawa, H., Goto, T., Morishita, M., Masago, A., Takahashi, H., Ohmiya, Y., Sakaguchi, T., Asada, M., Imamura, T., et al. (2007). Persistent and stable gene expression by a cytoplasmic RNA replicon based on a noncytopathic variant Sendai virus. J. Biol. Chem. *282*, 27383–27391.

Nishimura, K., Sano, M., Ohtaka, M., Furuta, B., Umemura, Y., Nakajima, Y., Ikehara, Y., Kobayashi, T., Segawa, H., Takayasu, S., et al. (2011). Development of defective and persistent Sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. J. Biol. Chem. *286*, 4760–4771.

Nishimura, T., Kaneko, S., Kawana-Tachikawa, A., Tajima, Y., Goto, H., Zhu, D., Nakayama-Hosoya, K., Iriguchi, S., Uemura, Y., Shimizu, T., et al. (2013). Generation of rejuvenated antigen-specific T cells by reprogramming to pluripotency and redifferentiation. Cell Stem Cell *12*, 114–126.

O'Malley, J., Skylaki, S., Iwabuchi, K.A., Chantzoura, E., Ruetz, T., Johnsson, A., Tomlinson, S.R., Linnarsson, S., and Kaji, K. (2013). High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. Nature *499*, 88–91.

Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. Nature *448*, 313–317.

Papapetrou, E.P., Tomishima, M.J., Chambers, S.M., Mica, Y., Reed, E., Menon, J., Tabar, V., Mo, Q., Studer, L., and Sadelain, M. (2009). Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. Proc. Natl. Acad. Sci. USA *106*, 12759–12764.

Parchem, R.J., Ye, J., Judson, R.L., LaRussa, M.F., Krishnakumar, R., Blelloch, A., Oldham, M.C., and Blelloch, R. (2014). Two miRNA clusters reveal alternative paths in late-stage reprogramming. Cell Stem Cell *14*, 617–631.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J.,

et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell *151*, 1617–1632.

Rais, Y., Zviran, A., Geula, S., Gafni, O., Chomsky, E., Viukov, S., Mansour, A.A., Caspi, I., Krupalnik, V., Zerbib, M., et al. (2013). Deterministic direct reprogramming of somatic cells to pluripotency. Nature *502*, 65–70.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. Science *234*, 364–368.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell *7*, 64–77.

Sridharan, R., Tchieu, J., Mason, M.J., Yachechko, R., Kuoy, E., Horvath, S., Zhou, Q., and Plath, K. (2009). Role of the murine reprogramming factors in the induction of pluripotency. Cell *136*, 364–377.

Stadtfeld, M., and Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. Genes Dev. *24*, 2239–2263.

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell *2*, 230–240.

Sui, D., Sun, Z., Xu, C., Wu, Y., Capecchi, M.R., Wu, S., and Li, N. (2014). Fine-tuning of iPSC derivation by an inducible reprogramming system at the protein level. Stem Cell Rep. *2*, 721–733.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., Nakamura, M., Sutou, K., Osafune, K., and Yamanaka, S. (2014). Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. Nat. Commun. *5*, 3678.

Takayama, N., Nishimura, S., Nakamura, S., Shimizu, T., Ohnishi, R., Endo, H., Yamaguchi, T., Otsu, M., Nishimura, K., Nakanishi, M., et al. (2010). Transient activation of c-MYC expression is critical for efficient platelet generation from human induced pluripotent stem cells. J. Exp. Med. *207*, 2817–2830.

Tateno, H., Matsushima, A., Hiemori, K., Onuma, Y., Ito, Y., Hasehira, K., Nishimura, K., Ohtaka, M., Takayasu, S., Nakanishi, M., et al. (2013). Podocalyxin is a glycoprotein ligand of the human pluripotent stem cell-specific probe rBC2LCN. Stem Cells Transl. Med. *2*, 265–273.

Tiemann, U., Sgodda, M., Warlich, E., Ballmaier, M., Schöler, H.R., Schambach, A., and Cantz, T. (2011). Optimal reprogramming factor stoichiometry increases colony numbers and affects molecular characteristics of murine induced pluripotent stem cells. Cytometry A *79*, 426–435.

Tokusumi, T., Iida, A., Hirata, T., Kato, A., Nagai, Y., and Hasegawa, M. (2002). Recombinant Sendai viruses expressing different levels of a foreign reporter gene. Virus Res. 86, 33–38.

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.

Wakao, H., Yoshikiyo, K., Koshimizu, U., Furukawa, T., Enomoto, K., Matsunaga, T., Tanaka, T., Yasutomi, Y., Yamada, T., Minakami, H., et al. (2013). Expansion of functional human mucosal-associated invariant T cells via reprogramming to pluripotency and redifferentiation. Cell Stem Cell 12, 546–558.

Wei, Z., Gao, F., Kim, S., Yang, H., Lyu, J., An, W., Wang, K., and Lu, W. (2013). Klf4 organizes long-range chromosomal interactions with the oct4 locus in reprogramming and pluripotency. Cell Stem Cell 13, 36–47.

Xu, H., Lemischka, I.R., and Ma'ayan, A. (2010). SVM classifier to predict genes important for self-renewal and pluripotency of mouse embryonic stem cells. BMC Syst. Biol. 4, 173.

Manipulation of KLF4 Expression Generates iPSCs Paused at Successive Stages of Reprogramming

Ken Nishimura, Tetsuo Kato, Chen Chen, Lalhaba Oinam, Emi Shiomitsu, Daisuke Ayakawa, Manami Ohtaka, Aya Fukuda, Mahito Nakanishi, and Koji Hisatake

Figure S1

A

| EGFP | KLF4 | EGFP + KLF4 | DAPI |

B

fK-OSM

| KOSM | GKOSM | Sh0 | Sh100 |

C

| | MEF | GKOSM | | | KOSM | | | applied volume (µL) |
| KLF4 | 4 | 1 | 2 | 4 | 1 | 2 | 4 | |

KLF4
OCT4
SOX2
c-MYC
α-TUBULIN

Quantification of signal intensities

< KOSM >

KLF4

signal intensity

6,000,000

3,753,590

4,000,000

2,000,000

371,913

applied volume (µL)

3,753,590 - 371,913 = 3,381,677

Protein expression level of KLF4

3,381,677 ÷ 5,259,461 = 0.643

Normalized
KLF4 expression level
from KOSM

α-TUBULIN

signal intensity

12,000,000

8,031,145

8,000,000

4,000,000

2,771,684

applied volume (µL)

8,031,145 - 2,771,684 = 5,259,461

Protein expression level of α-TUBULIN

Figure S1. (Related to Figures 1 and 3)

(A) Infection of MEFs with SeVdp(GKOSM) and monitoring of the gene expression from the vector. MEFs were infected with SeVdp(GKOSM), and expression of EGFP and one of the reprogramming factors (KLF4) was detected at day 2 by the green fluorescence and immunostaining, respectively. Scale bars, 100 μm. (B) Cell morphology of mouse iPSCs produced by several SeVdp vectors. Mouse iPSCs colonies induced by SeVdp(KOSM), SeVdp(GKOSM) or SeVdp(fK-OSM) (in the presence or absence of 100 nM of Shied1) were stained for β-CATENIN at day 30. Scale bars, 100 μm. (C) Quantitative western blots to determine the relative protein levels of reprogramming factors. MEFs were infected with SeVdp(GKOSM) or SeVdp(KOSM) and collected one day after infection to prepare cell extract. The cell extracts were applied onto the SDS-PAGE gel and subjected to western blotting, and the signals were analyzed by LAS4000 and quantified using ImageQuant TL software. The applied volumes and the signal intensities from each western blot were plotted to obtain the linear regression line. Pilot experiments were performed to determine the appropriate amounts of cell extracts that showed a linear correlation with the obtained signals. Then the relative expression level of each factor was calculated from the intensity of the signal that falls within the linear range. The protein expression level of each reprogramming factor was normalized against that of α-TUBULIN, which was determined in the same quantitative manner. The quantifications were biological triplicates.

Figure S2

A

B

C

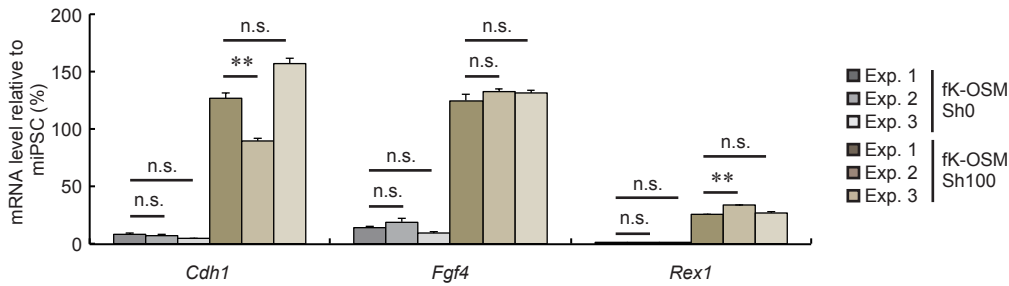KOSM   GKOSM   Kp-OSM   Kcp-OSM   OSM

D
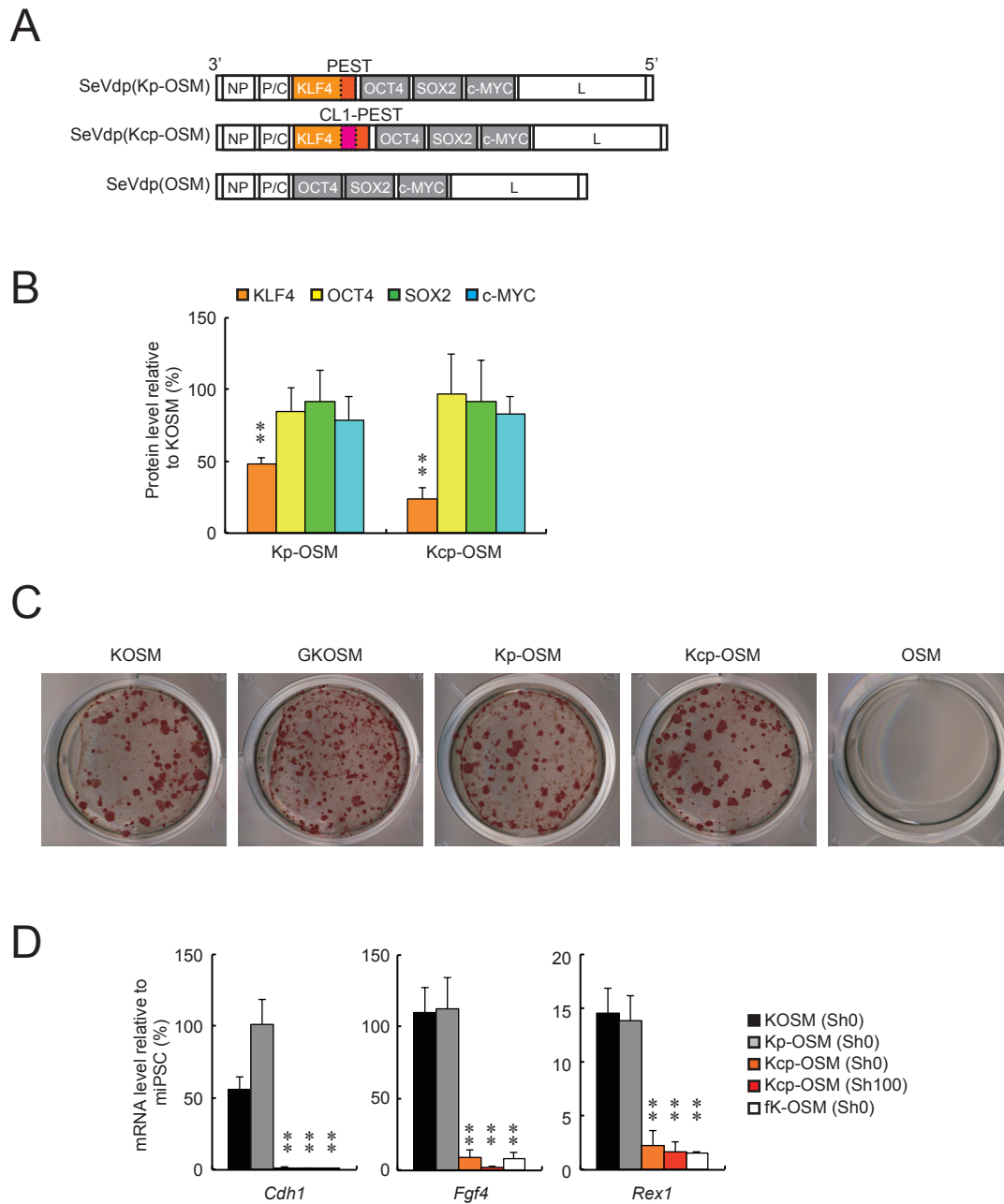
Figure S2. (Related to Figure 2)

(A) Structures of the SeVdp-derived vectors. In SeVdp(Kp-OSM) and SeVdp(Kcp-OSM), KLF4 is tagged C-terminally with PEST or CL1-PEST. In SeVdp(OSM), there are only three reprogramming factors; namely, OCT4, SOX2 and c-MYC. (B) Protein levels of reprogramming factors in MEFs infected with SeVdp(Kp-OSM) or SeVdp(Kcp-OSM). Cell extracts were prepared a day after infection with the indicated vector. Protein expression levels are indicated relative to those in SeVdp(KOSM)-infected cells. Data represent means ± SEM of three independent experiments. **$P < 0.01$ versus SeVdp(KOSM). (C) Alkaline phosphatase (+) colonies induced by each vector. MEFs infected with the indicated vector were stained for alkaline phosphatase at day 13. (D) Expression of ES cell markers in reprogrammed iPSCs. The mRNA levels of *Cdh1*, *Fgf4* and *Rex1* in the cells infected by the indicated vector in the presence or absence of 100 nM Shield1 were determined by quantitative RT-PCR at day 16. Data represent means ± SEM of three independent experiments. **$P < 0.01$ versus SeVdp(KOSM).

Figure S3

A

mRNA level relative to miPSC (%)

200 — 150 — 100 — 50 — 0

n.s.
**
n.s.
n.s.
n.s.
n.s.
n.s.
n.s.
**
n.s.

Cdh1   Fgf4   Rex1

Legend:
- Exp. 1 | Exp. 2 | Exp. 3 — fK-OSM Sh0
- Exp. 1 | Exp. 2 | Exp. 3 — fK-OSM Sh100

B

mRNA level relative to miPSC (%)

1000 — 100 — 10 — 1 — 0.1 — 0.01

Cdh1   Fgf4   Rex1

Legend:
- Sh0-11 | Sh0-12 | Sh0-13 | Sh0-14 | Sh0-15 — fK-OSM Sh0
- Sh10-1 | Sh10-2 | Sh10-3 | Sh10-4 | Sh10-5 — fK-OSM Sh10

Cdh1

Sh0-12	Sh0-13	Sh0-14	Sh0-15	
n.s.	n.s.	n.s.	n.s.	Sh0-11
	n.s.	n.s.	n.s.	Sh0-12
		n.s.	n.s.	Sh0-13
			n.s.	Sh0-14

Sh10-2	Sh10-3	Sh10-4	Sh10-5	
n.s.	n.s.	n.s.	n.s.	Sh10-1
	n.s.	n.s.	< 0.01	Sh10-2
		n.s.	n.s.	Sh10-3
			n.s.	Sh10-4

Fgf4

Sh0-12	Sh0-13	Sh0-14	Sh0-15	
n.s.	n.s.	n.s.	n.s.	Sh0-11
	n.s.	n.s.	n.s.	Sh0-12
		n.s.	n.s.	Sh0-13
			n.s.	Sh0-14

Sh10-2	Sh10-3	Sh10-4	Sh10-5	
n.s.	n.s.	n.s.	n.s.	Sh10-1
	n.s.	n.s.	n.s.	Sh10-2
		n.s.	n.s.	Sh10-3
			n.s.	Sh10-4

Rex1

Sh0-12	Sh0-13	Sh0-14	Sh0-15	
n.s.	n.s.	n.s.	n.s.	Sh0-11
	n.s.	n.s.	n.s.	Sh0-12
		n.s.	n.s.	Sh0-13
			n.s.	Sh0-14

Sh10-2	Sh10-3	Sh10-4	Sh10-5	
< 0.001	n.s.	< 0.001	n.s.	Sh10-1
	< 0.001	n.s.	< 0.001	Sh10-2
		< 0.001	n.s.	Sh10-3
			< 0.001	Sh10-4

fK-OSM Sh0 (top tables), fK-OSM Sh10 (bottom tables)

Figure S3. (Related to Figure 3)

(A) Reproducibility of reprogramming by SeVdp(fK-OSM). SeVdp(fK-OSM)-infected cells were cultured in the presence or absence of Shield1. The cells were collected at day 27 (Exp.1), day 24 (Exp.2) or day 26 (Exp.3), and the mRNA levels of *Cdh1*, *Fgf4* and *Rex1* were determined by quantitative RT-PCR. Data represent means ± SEM of three independent PCR reactions. n.s., not significant; **$P < 0.01$ versus Exp. 1. (B) Homogeneity of partially reprogrammed clones generated by SeVdp(fK-OSM). Partially reprogrammed cells were induced by SeVdp(fK-OSM) in the presence (Sh10-1 to 5) or absence (Sh0-11 to 15) of 10 nM Shield1, and individual colonies were isolated and expanded for analyses. The mRNA levels of *Cdh1*, *Fgf4* and *Rex1* in each clone were determined by quantitative RT-PCR at day 56. Data represent means ± SEM of three independent PCR reactions. *P*-values of individual data sets are listed in the tables. n.s., not significant.
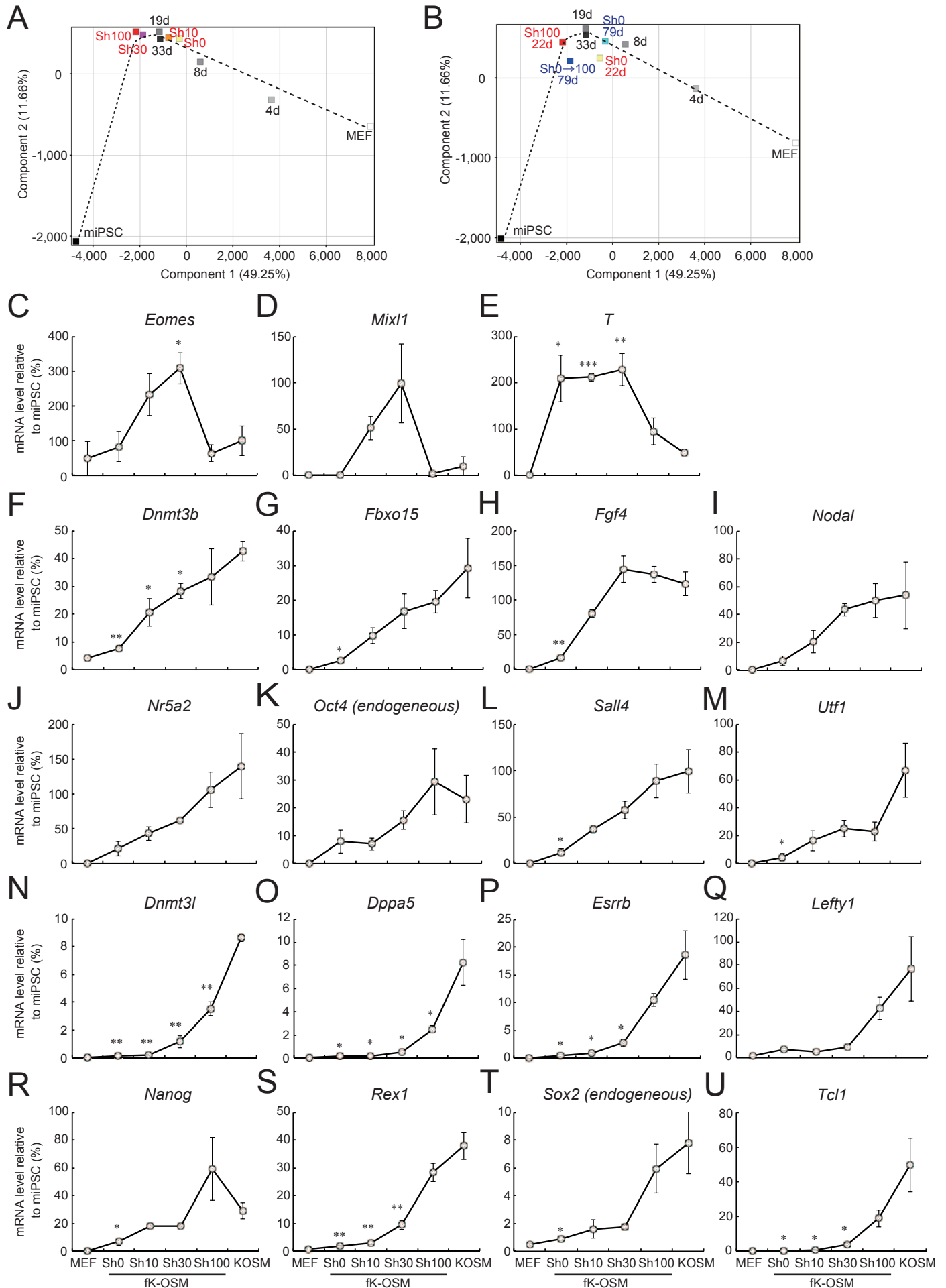
Figure S4
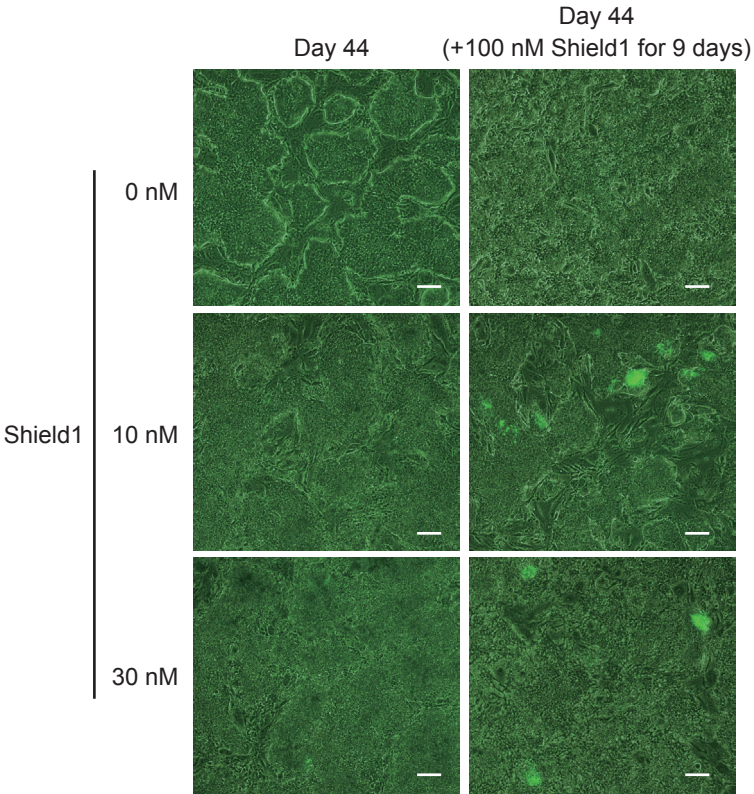
Figure S4. (Related to Figure 4)
Paused iPSCs were generated in the presence of the indicated concentration of Shield1, and 100 nM Shield1 was added to the cells at day 35. The NANOG expression was observed by GFP fluorescence from the *Nanog*-GFP reporter transgene at day 44. Scale bars, 100 μm.
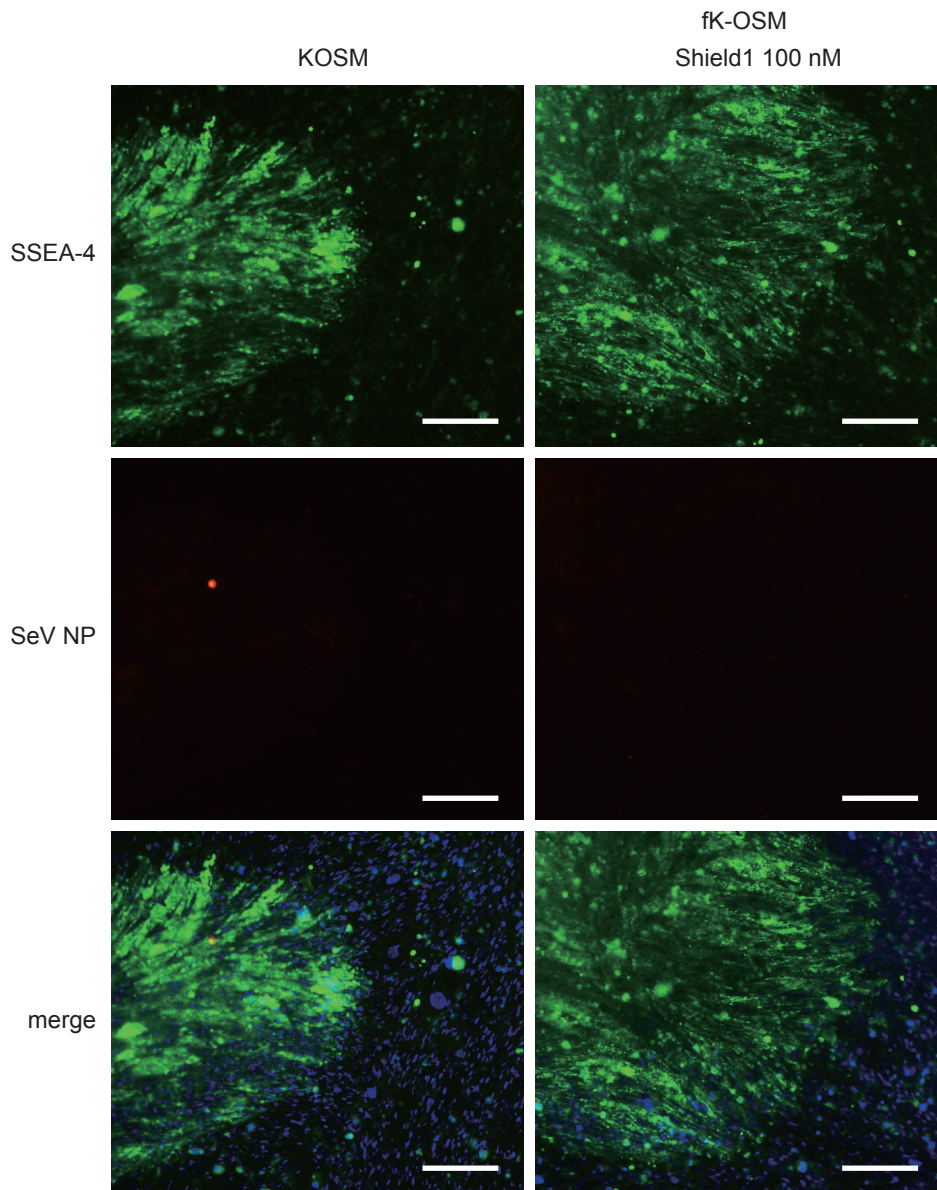
Figure S5

Figure S5. (Related to Figure 5)

(A, B) Principal component analysis (PCA) of the paused iPSCs. The expression profiles of the differentially expressed genes (DEGs) between MEFs and miPSCs ($P < 0.05$) were used for the PCA. PCA of SeVdp(KOSM)-reprogrammed MEFs collected at day 4, 8, 19 and 33 is compared to paused iPSCs generated by SeVdp(fK-OSM) in the presence of 0, 10, 30 or 100 nM of Shield1 collected at day 22 (A) or iPSCs that had paused and then resumed reprogramming (B). Sh0→100 79d; MEFs were reprogrammed by SeVdp(fK-OSM) in the absence of Shield1, and 100 nM Shield1 was added to the cells at day 25, which were further cultured until day 79. (C-U) Gene expression profiles of mesendodermal genes and pluripotency-related genes in paused iPSCs. SeVdp(fK-OSM)-infected cells were reprogrammed in the presence of indicated concentrations of Shield1, and the cells were collected at day 27. The mRNA levels of mesendodermal genes (C-E) and pluripotency-related genes (F-U) were determined by quantitative RT-PCR. Data represent means ± SEM of three independent experiments. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$ versus SeVdp(KOSM).

Figure S6

Figure S6. (Related to Figure 6)
Human iPSCs generated by SeVdp(KOSM) or SeVdp(fK-OSM) in the presence of 100 nM Shield1 were treated three times with L527 siRNA, and the cells were stained for SSEA-4 and SeV NP at day 44. Scale bars, 100 μm.

Table S1. Number of alkaline phosphatase (+) or *Nanog*-GFP (+) colonies in mouse or human iPSC production
Related to Figures 1, 2 and 6.

< mouse iPSCs >

Alkaline Phosphatase

		KOSM	GKOSM	fK-OSM	K-fOSM	KO-fSM	KOS-fM
Shield1	0 nM	705±221	712±170	986±331	9±2*	40±9*	2±1*
	100 nM	n.d.	n.d.	693±232	16±3*	339±119	85±38

Nanog-GFP

		KOSM	fK-OSM	K-fOSM	KO-fSM	KOS-fM
Shield1	0 nM	63±14	0±0**	1±1*	10±5*	0±1*
	100 nM	n.d.	30±12	3±1*	37±11	9±5*

< human iPSCs >

Alkaline Phosphatase

		KOSM	GKOSM	fK-OSM
Shield1	0 nM	36±5	63±5*	19±4
	100 nM	n.d.	n.d.	25±5

Colony numbers were counted 16 days after infection. Data represent means ± SEM of three independent experiments.

*$P < 0.05$, **$P < 0.01$ versus KOSM in the absence of Shield1.

n.d., not done

Table S2. Primer sequence for quantitative RT-PCR analyses
Related to Figures 1, 3, 4, 5 and 6.

• For mouse genes

Gene	Sequence
Cdh1	ACGTCCCCCTTTACTGCTG
	TATCCGCGAGCTTGAGATG
Cdh2	ATCAACCCCATCTCAGGACA
	CAATGTCAATGGGGTTCTCC
Dnmt3b	CCACATTTGCTGGAGGATG
	TACCAAAGCAAGGGGAAGG
Dnmt3l	GAACCGACGGAGCATTGAAG
	TCCCTCAAACAAGGGGTGCCG
Dppa5	CCAGTCGCTGGTGCTGAAATATCTG
	GCCCGAATCTTGTTGTTTTGAGAGC
Eomes	TCCAAGACTCAGACCTTCACC
	AGGGACAATCTGATGGGATG
Esrrb	TGGCAGGCAAGGATGACAGA
	TTTACATGAGGGCCGTGGGA
Fbxo15	CTCATCTGTCACGAAGCAGC
	AGGTCACCGCATCCAAGTAA
Fgf4	GCAGCGAGGCGTGGTGAGCATCTT
	CCCCTTCTTGGTCCGCCCGTTCTT
Lefty1	ACAGCTGGAGCTGCACA
	GTTCTCGGCCCACTTCATC
Mixl1	TGGAGCTCGTCTTCCGACA
	AATGACTTCCCACTCTGGCG
Nanog	ACCTGAGCTATAAGCAGGTTAAGAC
	GTGCTGAGCCCTTCTGAATCAGAC
Nodal	ACATGTTGAGCCTCTACCGAG
	GTGAAAGTCCAGTTCTGTCCG
Nr5a2	CAAAGTGGAGACGGAAGCC
	ATCGCCACACACAGGACATA
Oct4 (endogeneous)	CCAACGAGAAGAGTATGAGGC
	GTGCTTTTAATCCCTCCTCAG
Rex1	TTGATGGCTGCGAGAAGAG
	ACCCAGCCTGAGGACAATC
Sall4	CACGAAAGGCAACCTGAAG
	GGAGCTCCTTGGAAAACACC
Snail1	TGTGTCTGCACGACCTGTG
	AGTGGGAGCAGGAGAATGG
Snail2	ACACATTGCCTTGTGTCTGC
	GCCCTCAGGTTTGATCTGTC
Sox2 (endogeneous)	AGAGAAGTTTGGAGCCCGAG
	ATCTGGCGGAGAATAGTTGG
T	AAGGCGCCTGTGTCTTTCAG
	CTCACGATGTGAATCCGAGG
Tbp	GGGGAGCTGTGATGTGAAGT
	CCAGGAAATAATTCTGGCTCA
Tcl1	GGAGAAGCACGTGTACTTGGA
	TGACTGGGGGACATAGCTTC
Tgfb3	GAAGGCTGCACTCAGGAGAC
	TGAGGACACATTGAAACGAA
Thy1	TCGCTCTCCTGCTCTCAGTC
	TTATTCTCATGGCGGCAGTC
Twist1	CCCACACCTCTGCATTCTG
	CTGTCAGTGGCTGATTGGC
Utf1	GTCCCTCTCCGCGTTAGC
	GGGGCAGGTTCGTCATTT
Zeb1	TGCCAGCAGACCAGACAGTA
	TTCGGATCATGGTTTTGCTC
Zeb2	GCTAACCCAAGGAGCAGGTAAC
	TGAACTGTAGGACCCAGAATGA

• For human genes

Gene	Sequence
CDH1	GAGACAGTTTCGCTCCATCG
	AGCTTGGGCAACATAGCAAG
FGF4	ACCTTGGTGCACTTTCTTCG
	AAAAAACACACCCGCAGAAC
REX1	GCTCCTACACTTTGTGATACCG
	TCACTGACCTTCCTGTTGGA
TBP	TCTCATGTACCCTTGCCTCC
	GGCACTTACAGAAGGGCATC

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Production of SeVdp vectors

The SeVdp vector genomic cDNA was constructed as described (Nishimura et al., 2011). First, we changed the order of reprograming factors in SeVdp(*c-Myc/Klf4/Oct4/Sox2*) cDNA to the order of KLF4, OCT4, SOX2 and c-MYC to construct SeVdp(KOSM) cDNA. Using the SeVdp(KOSM) cDNA, we constructed the following vector cDNAs. The DNA fragment encoding the EGFP gene was amplified by polymerase chain reaction (PCR) using pEGFP-1 (Takara Bio) as a template. The amplified EGFP expression cassette containing transcription initiation and termination signals for Sendai virus (SeV) was inserted between the P/C and KLF4 genes of SeVdp(KOSM) vector to create SeVdp(GKOSM) cDNA. The DNA fragment of the destabilization domain (DD) derived from an FKBP12 mutant was amplified by PCR from pPTuner vector (Takara Bio) and inserted into the N-terminus of the coding region of each reprogramming factor (KLF4, OCT4, SOX2, or c-MYC) to create a fusion protein using In-Fusion HD Cloning Kit (Takara Bio). These cDNAs were used to construct vector cDNAs of SeVdp(fK-OSM), SeVdp(K-fOSM), SeVdp(KO-fSM) or SeVdp(KOS-fM). The DNA fragment coding for CL1 or PEST was amplified from pST-blue-T7-GFP-CL/PEST and inserted into the C-terminus of the KLF4 gene to construct SeVdp(Kcp-OSM) or SeVdp(Kp-OSM) cDNA. We removed the KLF4 gene expression cassette from SeVdp(KOSM) cDNA to create SeVdp(OSM) cDNA.

To prepare vector-packaging cells, 2.0×10^5 BHK/T7/151M(SE) cells in a 6-well plate were transfected with each vector cDNA (2 μg) and the expression vector plasmids for SeV proteins (NP, P/C, M, F, HN, L) (1 μg each) using Lipofectamine LTX with Plus Reagent (Life Technologies) and cultured at 32 ˚C for 6 days. Then, the packaging cells were further transfected with the expression vector plasmids for SeV M, F and HN proteins and cultured at 32 ˚C for additional 4 days to rescue the SeVdp vector, which was released into the culture supernatant. The supernatant was filtered through a 0.45 μm cellulose acetate filter and stored at -80 ˚C until use. Titers of the SeVdp vectors were determined by examining NIH3T3 cells infected with a diluted vector suspension by immunostaining using an anti-SeV NP antibody.

Production of retroviral vector

The cDNAs encoding the OCT4, SOX2, KLF4 and c-MYC proteins were amplified from mouse embryonic stem cell cDNA and inserted into pMXs (Addgene) to construct the retroviral vector expressing each protein. PLAT-E cells were transfected with each vector using

Lipofectamine LTX with Plus Reagent. Viral supernatant was collected 2 days after transfection and filtered through a 0.45 μm cellulose acetate filter and stored at -80 ˚C until use. Titers of the retroviral vectors were determined by examining NIH3T3 cells infected with a diluted vector suspension by immunostaining using an antibody against OCT4 (1:1,000, ab19857, Abcam), SOX2 (1:250, sc-17320, Santa Cruz), KLF4 (1:250, sc-20691, Santa Cruz) or c-MYC (1:250, sc-764, Santa Cruz).

Generation of chimeric mice

Chimeric mice were generated by a conventional method in which SeVdp(fK-OSM)-free iPSCs (Sh100-36 and Sh100-43) were microinjected into blastocysts (3.5 days post-coitum) of C57BL/6J mice, followed by transferring to the uterus of foster mothers (Joyner, 1993). Genomic DNAs were extracted from the resultant chimeric mouse tails by HotSHOT method (Truett et al., 2000), and used for genotyping the mice by PCR using primers; 5'-TGG GAT CCC TAT GCT ACT CCG TCG AAG TTC-3' and 5'-CTA GGC AAA CTG TGG GGA CCA GGA AGA C-3' to detect the *Nanog*-GFP construct derived from the iPSCs as described in http://www2.brc.riken.jp/animal/pdf/02290_PCR.pdf.

Immunofluorescence and alkaline phosphatase staining

For immunostaining, cells were fixed with 3.7% formaldehyde in phosphate buffered saline (PBS). After permeablization in 0.1% Triton X-100/PBS, cells were stained with primary antibodies followed by secondary antibodies (AlexaFlour, Life Technologies). Nuclei were counterstained with DAPI using VECTASHIELD Mounting Medium with DAPI (Vector). We employed the following primary antibodies; anti-SeV NP (1:1,000) (Nishimura et al., 2011), anti-SSEA-1 (1:250, sc-21702, Santa Cruz), anti-β-CATENIN (1:250, 610153, BD Biosciences) and anti-SSEA-4 (1:1,000, MAB4304, Millipore). Detection of alkaline phosphatase was carried out using VECTOR Red Alkaline Phosphatase Substrate Kit (Vector) according to manufacturer's instructions.

Quantitative RT-PCR

To avoid any contamination of feeder cells, reprogrammed cells were grown initially on feeder cells and then passaged under a feeder-free condition before RNA extraction. Total RNA was extracted using ISOGEN (Nippon Gene), and reverse transcription was performed using Superscript III First-Strand Synthesis System (Life Technologies). Quantitative PCR (qPCR)

analyses were performed in technical triplicates using 7500 Fast Real-time PCR System (Applied Biosystems) with GoTaq qPCR Master Mix (Promega). The expression levels were normalized against that of TATA-box binding protein (TBP). In case of human iPSCs, we collected the reprogrammed cells together with feeder cells but used primers specific for human genes for qPCR analyses. All primer sequences are listed in Table S2.

Microarray analysis

Total RNA was purified by ReliaPrep RNA Cell Miniprep System (Promega), and then reverse-transcribed and labeled using Illumina TotalPrep RNA Amplification Kit (Life Technologies). Arrays (MouseRef-8 v2.0 Expression BeadChip Kit), processed according to manufacture's instructions, were scanned with BeadArray Reader (Illumina). Each scanned image was analyzed with GenomeStudio software (Illumina). All of the data analyses were done by GeneSpring 12.5 software (Agilent). The data was filtered to remove low-confidence measurements and normalized globally per array such that 75 percentile of the signal intensity was set at 1.0. Differentially expressed genes were identified with a student's t-test ($P < 0.05$) for multiple comparisons. Hierarchical clustering was performed using the Ward's method. Pathway analysis was performed by referring to WikiPathway.

Statistical analysis

The statistical analyses were performed using StatView-J 5.0 software. Student's t-tests were used to test for statistically significant difference between data sets.

SUPPLEMENTAL REFERENCES

Joyner, A.L. (1993). Gene targeting : a practical approach (Oxford, IRL Press at Oxford University Press).

Truett, G.E., Heeger, P., Mynatt, R.L., Truett, A.A., Walker, J.A., and Warman, M.L. (2000). Preparation of PCR-quality mouse genomic DNA with hot sodium hydroxide and tris (HotSHOT). Biotechniques *29*, 52, 54.