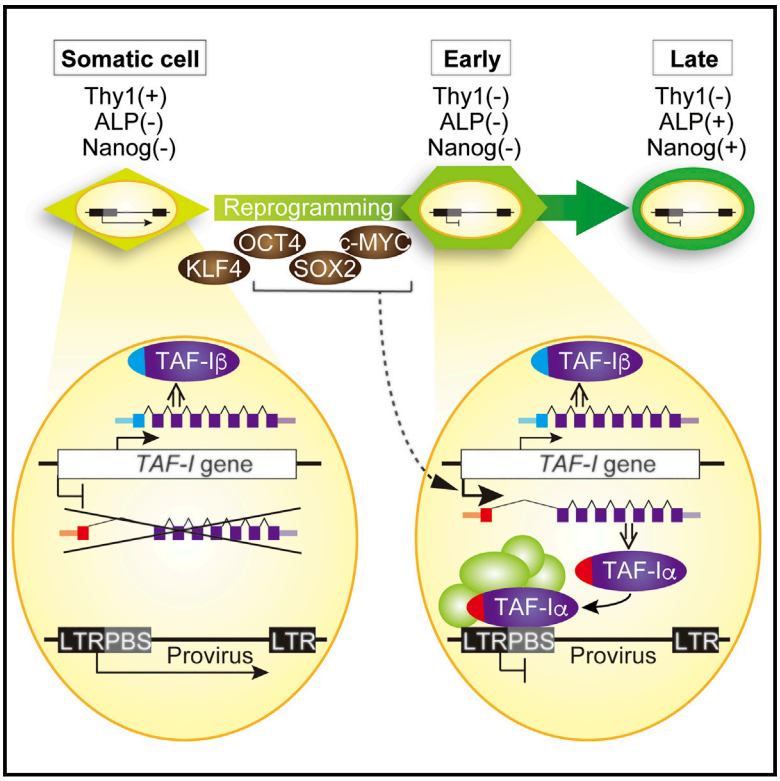
Article

Cell Reports

Template Activating Factor-I α Regulates Retroviral Silencing during Reprogramming

Graphical Abstract



Highlights

- Retroviral silencing precedes acquisition of pluripotency during reprogramming
- The primer-binding site in retrovirus is essential for retroviral silencing
- OCT4, SOX2, and c-MYC are required for retroviral silencing
- TAF-Iα is induced during reprogramming to facilitate retroviral silencing

Authors

Phuong Linh Bui, Ken Nishimura, Gonzalo Seminario Mondejar, ..., Yuzuru Ito, Mahito Nakanishi, Koji Hisatake

Correspondence

ken-nishimura@md.tsukuba.ac.jp (K.N.), kojihisa@md.tsukuba.ac.jp (K.H.)

In Brief

Bui et al. use replication-defective and persistent Sendai virus (SeVdp)-based vectors to monitor retroviral silencing during reprogramming and find that the silencing occurs earlier than the acquisition of pluripotency. Insertional chromatin immunoprecipitation (iChIP) identifies TAF-Iα, a SET/TAF-I isoform predominant in ESCs, as a factor that facilitates retroviral silencing.

Bui et al., 2019, Cell Reports *29*, 1909–1922
November 12, 2019
https://doi.org/10.1016/j.celrep.2019.10.010

CellPress

Cell Reports
Article

Template Activating Factor-I α Regulates Retroviral Silencing during Reprogramming

Phuong Linh Bui,[1,6] Ken Nishimura,[1,6,7,*] Gonzalo Seminario Mondejar,[1] Arun Kumar,[1] Shiho Aizawa,[1] Kensaku Murano,[2] Kyosuke Nagata,[3] Yohei Hayashi,[3,4] Aya Fukuda,[1] Yasuko Onuma,[5] Yuzuru Ito,[5] Mahito Nakanishi,[5] 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]Department of Molecular Biology, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan
[3]Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan
[4]iPS Cell Advanced Characterization and Development Team, Bioresource Research Center, RIKEN, 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
[5]Biotechnology Research Institute for Drug Discovery, National Institute of Advanced Industrial Science and Technology (AIST), Central 5, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8565, Japan
[6]These authors contributed equally
[7]Lead Contact
*Correspondence: ken-nishimura@md.tsukuba.ac.jp (K.N.), kojihisa@md.tsukuba.ac.jp (K.H.)
https://doi.org/10.1016/j.celrep.2019.10.010

SUMMARY

Reprogramming somatic cells to induced pluripotent stem cells (iPSCs) is accompanied by dramatic changes in epigenetic programs, including silencing of endogenous and exogenous retroviruses. Here, we utilized replication-defective and persistent Sendai virus (SeVdp)-based vectors to monitor retroviral silencing during reprogramming. We observed that retroviral silencing occurred at an early reprogramming stage without a requirement for KLF4 or the YY1-binding site in the retroviral genome. Insertional chromatin immunoprecipitation (iChIP) enabled us to isolate factors assembled on the silenced provirus, including components of inhibitor of histone acetyltransferase (INHAT), which includes the SET/TAF-I oncoprotein. Knockdown of SET/TAF-I in mouse embryonic fibroblasts (MEFs) diminished retroviral silencing during reprogramming, and overexpression of template activating factor-I α (TAF-Iα), a SET/TAF-I isoform predominant in embryonic stem cells (ESCs), reinforced retroviral silencing by an SeVdp-based vector that is otherwise defective in retroviral silencing. Our results indicate an important role for TAF-Iα in retroviral silencing during reprogramming.

INTRODUCTION

The development of induced pluripotent stem cells (iPSCs) holds great promise for regenerative medicine, disease modeling, and drug discovery (Takahashi and Yamanaka, 2013). Somatic cell reprogramming to generate iPSCs is elicited by ectopic expression of transcription factors such as OCT4, SOX2, KLF4, and c-MYC, which change somatic cells into a state essentially indistinguishable from embryonic stem cells (ESCs) (Okita et al., 2007; Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007). Generation of iPSCs progresses through distinct phases of reprogramming, which can be monitored by loss of Thy1 and appearance of alkaline phosphatase and SSEA-1, followed by expression of pluripotency-related genes such as NANOG and OCT4 (David and Polo, 2014). Underpinning this dramatic alteration of cell fate are massive changes in gene expression and epigenetic status (histone and DNA modifications) (Apostolou and Hochedlinger, 2013; Li et al., 2012; Papp and Plath, 2013).

One of the well-known epigenetic changes during reprogramming is silencing of retrovirus expression. In the course of retrovirus-mediated reprogramming, the integrated retroviruses expressing the four reprogramming factors become eventually silenced, especially in iPSCs with full pluripotency (Okita et al., 2007). Indeed, retroviral silencing is an indicator of high-quality iPSCs (Jaenisch and Young, 2008; Maherali and Hochedlinger, 2008) and can be used as a marker to select fully pluripotent iPSCs (Chan et al., 2009; Donai et al., 2013). Consistent with this, iPSCs that undergo retroviral silencing earlier tend to display characteristics of full pluripotency (Chan et al., 2009; Okada and Yoneda, 2011). Genome-wide analyses of reprogramming of mouse embryonic fibroblasts (MEFs) show that retroviral silencing occurs after 10–12 days when the reprogrammed cells become independent of the transgenes (Brambrink et al., 2008; Stadtfeld et al., 2008). In addition to exogenous retroviruses, endogenous retroviruses (ERVs) also undergo epigenetic changes. In the course of mouse and human somatic cell reprogramming, ERVs are initially reactivated upon loss of the repressive epigenetic marks and then become silenced again via a Trim28-mediated mechanism (Friedli et al., 2014). Failure to properly silence ERVs appears to impact the quality of derived iPSCs (Koyanagi-Aoi et al., 2013). Thus, silencing of both exogenous and endogenous retroviruses is fundamental to the epigenetic changes during reprogramming and may be a multifaceted process that requires numerous factors.

Retroviral silencing has been studied intensively using ESCs and embryonal carcinoma cells (ECCs) as model systems (Hotta and Ellis, 2008). An important role for DNA methylation for

retroviral silencing was noted initially (Jähner et al., 1982; Stewart et al., 1982). However, given that the silencing occurs even before DNA methylation (Linney et al., 1984; Niwa et al., 1983; Pannell et al., 2000), other mechanisms have also been explored, such as the absence of activators that act on the long terminal repeat (LTR) and the presence of repressive factors that may operate prior to DNA methylation (Schlesinger and Goff, 2015). The most important *cis* element for repressive factors in retroviral silencing is the primer-binding site (PBS), located immediately downstream of the 5′ LTR of retrovirus (Barklis et al., 1986; Feuer et al., 1989). The PBS is targeted by a repressive complex consisting of ZFP809, a zinc-finger DNA-binding protein (Wolf and Goff, 2009), and Trim28 (Wolf and Goff, 2007), a scaffold protein that presumably recruits and assembles epigenetic modifiers such as a histone H3-K9 methyltransferase, ESET (Matsui et al., 2010), the NURD complex (Schultz et al., 2001), heterochromatin 1 (HP1) (Wolf et al., 2008), Erb3-binding protein (EBP1) (Wang et al., 2014), and the polycomb group protein enhancer of zeste homolog 2 (EZH2) (Leeb et al., 2010) to generate repressive chromatin structure on the proviral DNA. In addition to the PBS, the YY1-binding site in the LTR is also known to play a role in retroviral silencing in ESCs (Schlesinger et al., 2013). A recent genome-wide screening revealed more than 300 factors potentially important for retroviral silencing, including histone chaperones and sumoylation factors (Yang et al., 2015). Thus, retroviral silencing may be a complex process involving a diverse array of cellular functions.

We previously developed a Sendai-virus-based reprogramming system based upon a temperature-sensitive strain (SeV Cl. 151) of Sendai virus, which possesses a negative-sense, single-stranded RNA genome. This mutant virus, termed replication-defective and persistent Sendai virus (SeVdp), remains stable in cytoplasm without cytopathic effect at a nonpermissive temperature (~38°C) (Nishimura et al., 2007, 2011). SeVdp-derived vectors allow for long-term expression of multiple factors at a relatively constant stoichiometry without integration into the host genome and reprogram MEFs into NANOG-positive iPSCs at a relatively high efficiency (~1%) (Nishimura et al., 2011). Because the expression of reprogramming factors is free from transcriptional repression throughout the progression of reprogramming, the SeVdp-based reprogramming system offers an ideal tool to analyze retroviral silencing during reprogramming, because the expression of reprogramming factors is free from transcriptional repression throughout the progression of reprogramming. Moreover, the SeVdp-based system allows for the generation of partially reprogrammed iPSCs, which permit mechanistic analyses of gene expression and epigenetic changes at various intermediate stages (Nishimura et al., 2014, 2017).

Here, we utilized the SeVdp-based vectors to set up a live-cell monitoring system that allows observation of retroviral silencing during reprogramming. We found that retroviral silencing occurs at a relatively early stage of reprogramming before the cells become positive for alkaline phosphatase. Retroviral silencing at this stage did not require KLF4 and was strictly dependent on the PBS, but not the YY1-binding site, in LTRs. We used insertional chromatin immunoprecipitation (iChIP) (Hoshino and Fujii, 2009) to isolate proteins assembled on the silenced provirus and identified components of inhibitor of histone acetyltransferase (INHAT), including SET/TAF-I, ANP32a, and HMG2. Knockdown of SET/TAF-I in MEFs diminished retroviral silencing during reprogramming, and overexpression of TAF-Iα, the SET/TAF-I isoform predominant in ESCs, enhanced retroviral silencing. These results indicate that TAF-Iα constitutes an integral component of the silenced retrovirus during reprogramming.

RESULTS

Monitoring of Retroviral Silencing during Reprogramming by a Sendai-Virus-Based Vector

To establish a live-cell monitoring system for retroviral silencing during reprogramming, we designed a Moloney murine leukemia virus (MLV)-based retrovirus, hereafter termed MLV(YY-PBS-hKO), that expresses humanized Kusabira orange (hKO) and the puromycin resistance gene (Figure 1A). MEFs, transduced by MLV(YY-PBS-hKO) and selected by puromycin, were infected by a Sendai virus vector, SeVdp(KOSM), that expresses KLF4, OCT4, SOX2, and c-MYC to generate fully reprogrammed iPSCs (Kyttälä et al., 2016; Matsumoto et al., 2016; Nishimura et al., 2011). At day 1, almost all MEFs expressed hKO from the integrated provirus. At day 5, some MEFs that were small, round, and positive for alkaline phosphatase (AP^+), did not appear to express hKO (Figure 1B). As the AP^+ cells formed colonies at day 7, the percentage of the colonies that repressed hKO expression increased from 4.5% to 98.6% after reprogramming (Figure 1B, day 7, and Figure 1C). Thus, the integrated provirus appeared to be silenced when MEFs were reprogrammed by KLF4, OCT4, SOX2, and c-MYC as early as day 5, and the silencing was largely complete by day 7, when AP^+ colonies appear. The retroviral silencing observed here is not due to the mere presence of a Sendai virus vector, because little silencing was observed in MEFs infected with SeVdp(BGC), which carries the blasticidin resistance gene, EGFP, and the *Cypridina* luciferase gene (CLuc) but lacks any reprogramming factors (Figure 1D).

To observe both silencing of the provirus and the presence of the SeVdp vector simultaneously in live cells, we used SeVdp(GKOSM) vector, which expresses EGFP in addition to the four reprogramming factors (Figure 1E). SeVdp(GKOSM) allows monitoring of the Sendai virus vector in live cells but generates only partially reprogrammed iPSCs, which are AP^+, SSEA-1^-, and $NANOG^-$, due to the lowered expression of the downstream reprogramming factors (Nishimura et al., 2014). However, these iPSCs are AP^+ and morphologically indistinguishable from fully reprogrammed iPSCs generated by SeVdp(KOSM) (Nishimura et al., 2014). Given that the retroviral silencing occurs by the time cells become AP^+ during reprogramming (Figure 1B), we tested whether hKO expression from the integrated provirus is also repressed in partially reprogrammed iPSCs generated by SeVdp(GKOSM). When MEFs expressing hKO were infected by SeVdp(GKOSM), AP^+ cells emerged by day 5 and became predominant by day 7 (Figure 1F), in a time course indistinguishable from that by SeVdp(KOSM) (Figure 1B). Importantly, $EGFP^+$ cells, which harbored the SeVdp vector, repressed hKO expression markedly by day 5, when the cells became AP^+; however, $EGFP^-$ cells, which were free of the

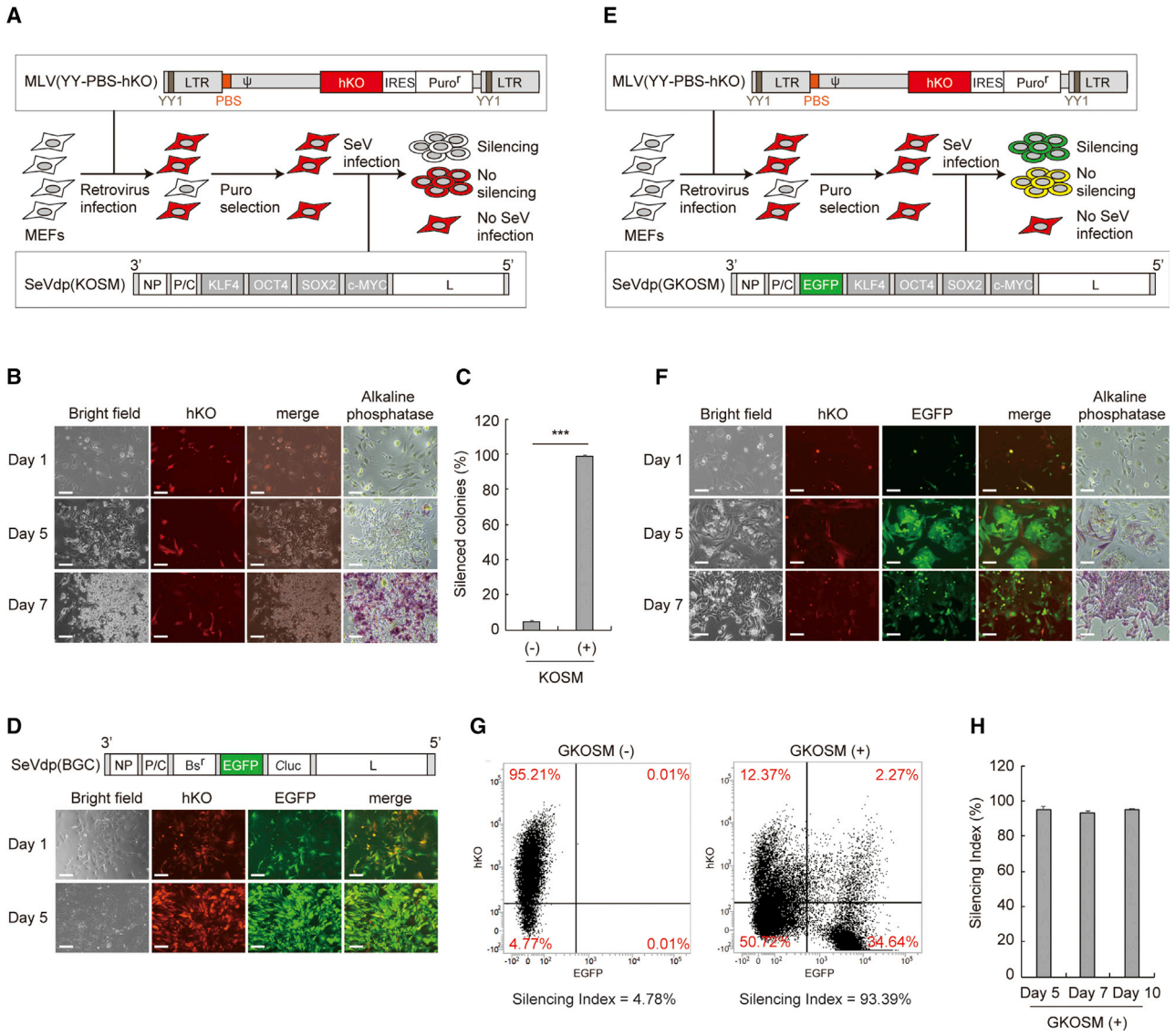


Figure 1. Development of a System for Monitoring Retroviral Silencing during Reprogramming

(A) Experimental procedure for monitoring hKO expression to assess retroviral gene silencing in cells reprogrammed by SeVdp(KOSM). Retrovirus is shown as a provirus integrated in the mouse genome. YY1, YY1-binding site; PBS, primer-binding site; ψ, packaging signal. The structure of SeVdp is also shown. NP, P/C, and L indicate genes encoding SeV NP protein, P/C proteins, and L protein, respectively.

(B) Induction of retroviral silencing in MEFs infected by SeVdp(KOSM). hKO expression and alkaline phosphatase activity were observed at the indicated days of reprogramming. Bright-field and hKO photos were overlaid to produce merged photos. Scale bars, 100 μm.

(C) The percentage of hKO^- colonies induced by SeVdp(KOSM) was calculated at day 7 of reprogramming. Data are represented as means ± SEM of three independent experiments. ***$p < 0.001$.

(D) hKO and EGFP expression in MEFs infected by SeVdp(BGC) at the indicated days of reprogramming. hKO and EGFP photos were overlaid to produce merged photos. Scale bars, 100 μm.

(E) Experimental procedure for monitoring hKO expression to assess retroviral gene silencing in MEFs infected by SeVdp(GKOSM). Structures of the provirus and SeVdp(GKOSM) are also shown. SeVdp(GKOSM), derived from SeVdp(KOSM), additionally carries the EGFP gene between the P/C gene and the *Klf4* gene.

(F) hKO and EGFP expression and alkaline phosphatase activity were observed in MEFs infected by SeVdp(GKOSM) at the indicated days of reprogramming. hKO and EGFP photos were overlaid to produce merged photos. Scale bars, 100 μm.

(G) FACS profiles of cells infected by SeVdp(GKOSM). hKO-expressing MEFs with or without SeVdp(GKOSM) infection were analyzed for hKO and EGFP expression at day 7 of reprogramming.

(H) Silencing indices were calculated from FACS analyses preformed at indicated days of reprogramming by SeVdp(GKOSM). Raw data from each FACS analysis are shown in Table S1. Data are represented as means ± SEM of three independent experiments.

See also Figure S1 and Table S1.

SeVdp vector, continued to express hKO (Figure 1F). When human cells were infected by SeVdp(GKOSM), retroviral silencing was also observed in a similar time course (Figure S1A). To quantify the degree of retroviral silencing, we performed fluorescence-activated cell sorting (FACS) analysis and defined the silencing index, which indicates the percentage of cells that have silenced retroviral expression (hKO^- cells) in the total population of SeVdp-infected cells ($EGFP^+$ cells) (Figures 1G and S1B). At day 7 of SeVdp(GKOSM)-based reprogramming, 93.84% of the $EGFP^+$ cells repressed hKO expression from the integrated provirus (Figure 1G, right panel), which remained silenced up to day 10 of reprogramming (Figure 1H; Table S1). By contrast, 95.21% of the MEFs without reprogramming expressed hKO (Figure 1G, left panel).

Previous studies showed that retroviral silencing occurs late during reprogramming, closely coupled with the acquisition of pluripotency (Brambrink et al., 2008; Chan et al., 2009; Donai et al., 2013; Stadtfeld et al., 2008). One possible reason that we observed retroviral silencing earlier during reprogramming could be that the hKO-expressing retroviral particles were transduced 6 days prior to reprogramming, which inadvertently allowed more time for the provirus to undergo silencing. To exclude this possibility, we transduced the hKO-expressing retroviral particles either 1 day or 6 days prior to reprogramming (Figure S1C). Regardless of the timing of retroviral transduction, the cells showed similar retroviral silencing (Figure S1D), and the degree of retroviral silencing was essentially indistinguishable (Figure S1E). Thus, SeVdp(GKOSM)-based reprogramming allows quantitative live-cell monitoring of retroviral silencing during reprogramming, and the silencing occurs earlier than previously reported.

The PBS Is Essential for Retroviral Silencing during Reprogramming

Retroviral DNA *cis* elements that are critical for retroviral silencing in pluripotent stem cells are the PBS located just downstream of the 5′ LTR and the YY1-binding sites located in the U3 region of the LTRs (Barklis et al., 1986; Feuer et al., 1989; Kempler et al., 1993; Petersen et al., 1991; Schlesinger et al., 2013). To test their requirement for retroviral silencing during reprogramming, we created three retroviruses carrying mutations in either or both of the *cis* elements (Figure 2A). MEFs were then transduced with the wild-type retroviral particles MLV(YY-PBS-hKO) or one of the mutant retroviral particles, MLV(ΔYY-PBS-hKO), MLV(YY-PBSQ-hKO), or MLV(ΔYY-PBSQ-hKO), and the untransduced cells were eliminated by puromycin selection. After 5 days of infection by SeVdp(GKOSM), the cells harboring the wild-type MLV(YY-PBS-hKO) or MLV(ΔYY-PBS-hKO) showed almost mutually exclusive expression of EGFP and hKO (Figure 2B), indicating that $EGFP^+$ cells reprogrammed by SeVdp(GKOSM) repressed hKO expression from these integrated provirus. By contrast, the cells harboring MLV(YY-PBSQ-hKO) or MLV(ΔYY-PBSQ-hKO) continued to express hKO despite reprogramming by SeVdp(GKOSM) (Figure 2B).

Quantitative FACS analyses showed that 90.80% of $EGFP^+$ cells harboring MLV(ΔYY-PBS-hKO) repressed hKO expression from the provirus, which was comparable to the silencing of the wild-type MLV(YY-PBS-hKO) (Figure 2C), indicating that retroviral silencing occurs during reprogramming even when the YY1-binding site is mutated. By contrast, only 53.52% of $EGFP^+$ cells harboring the PBS mutant MLV(YY-PBSQ-hKO) repressed hKO expression (Figure 2C), indicating that the mutation of the PBS severely compromises retroviral silencing at the early stage of reprogramming. As expected, mutation of YY in addition to PBSQ in MLV(ΔYY-PBSQ-hKO) resulted in silencing comparable to that observed for MLV(YY-PBSQ-hKO) (Figure 2C). In a control experiment, the mutated YY1-binding site of MLV(ΔYY-PBS-hKO) was confirmed to attenuate silencing in F9 mouse ECCs (Figure S2), as reported previously (Schlesinger et al., 2013). In conclusion, these results indicate that the PBS, but not the YY1-binding site, is critical for retroviral silencing during reprogramming.

KLF4 Is Dispensable for Retroviral Silencing during Reprogramming

As reprogramming is initiated solely by KLF4, SOX2, c-MYC, and OCT4, we addressed which factor is essential for triggering retroviral silencing. We designed a new set of SeVdp-based vectors that contain the EGFP gene but lack each reprogramming factor (Figure 3A). MEFs after 5 days of infection by each SeVdp vector were reprogrammed to minimal extents due to the absence of one factor (Figure 3B). The rapid proliferation and cell morphological changes, which are characteristic of an early stage of reprogramming, were observed only in the cells infected by SeVdp(GOSM) or SeVdp(GOKM) (Figure 3B). Because SeVdp(GSKM) caused massive cell death (Figure S3), the effect of Oct4 could not be investigated.

Fluorescence microscopy showed that the cells infected by SeVdp(GOSM), which lacks KLF4, silenced retrovirus to a similar degree by the SeVdp(GKOSM)-reprogrammed cells (Figure 3B). By contrast, the cells infected by SeVdp(GOKM) and SeVdp(GKOS) appeared to express hKO, indicating that the cells failed to silence retrovirus in the absence of either SOX2 or c-MYC. To determine if only SOX2 and c-MYC are sufficient for retroviral silencing, we constructed SeVdp(GSM), which expresses only SOX2, c-MYC, and EGFP (Figure 3A). The cells infected with SeVdp(GSM) largely failed to repress hKO expression (Figure 3B), indicating that SOX2 and c-MYC are necessary, but not sufficient, for retroviral silencing. Moreover, comparison of the cells infected with SeVdp(GSM) and those with SeVdp(GOSM) shows that addition of OCT4 increases the efficiency of retroviral silencing substantially, indicating that OCT4 also plays a role in retroviral silencing. We also performed FACS analysis to quantitatively assess the role for each factor in retroviral silencing during reprogramming. As shown in Figure 3C, retroviral silencing was 84.54% by SeVdp(GOSM), which is comparable to the silencing by SeVdp(GKOSM), confirming that KLF4 is dispensable for retroviral silencing. Omission of either SOX2 or c-MYC markedly diminished retroviral silencing to 43.72% or 13.85%, respectively. MEFs infected with SeVdp(GSM) repressed hKO expression in only 19.98% of cells. In summary, these results show that OCT4, SOX2, and c-MYC, but not KLF4, are involved in retroviral silencing.

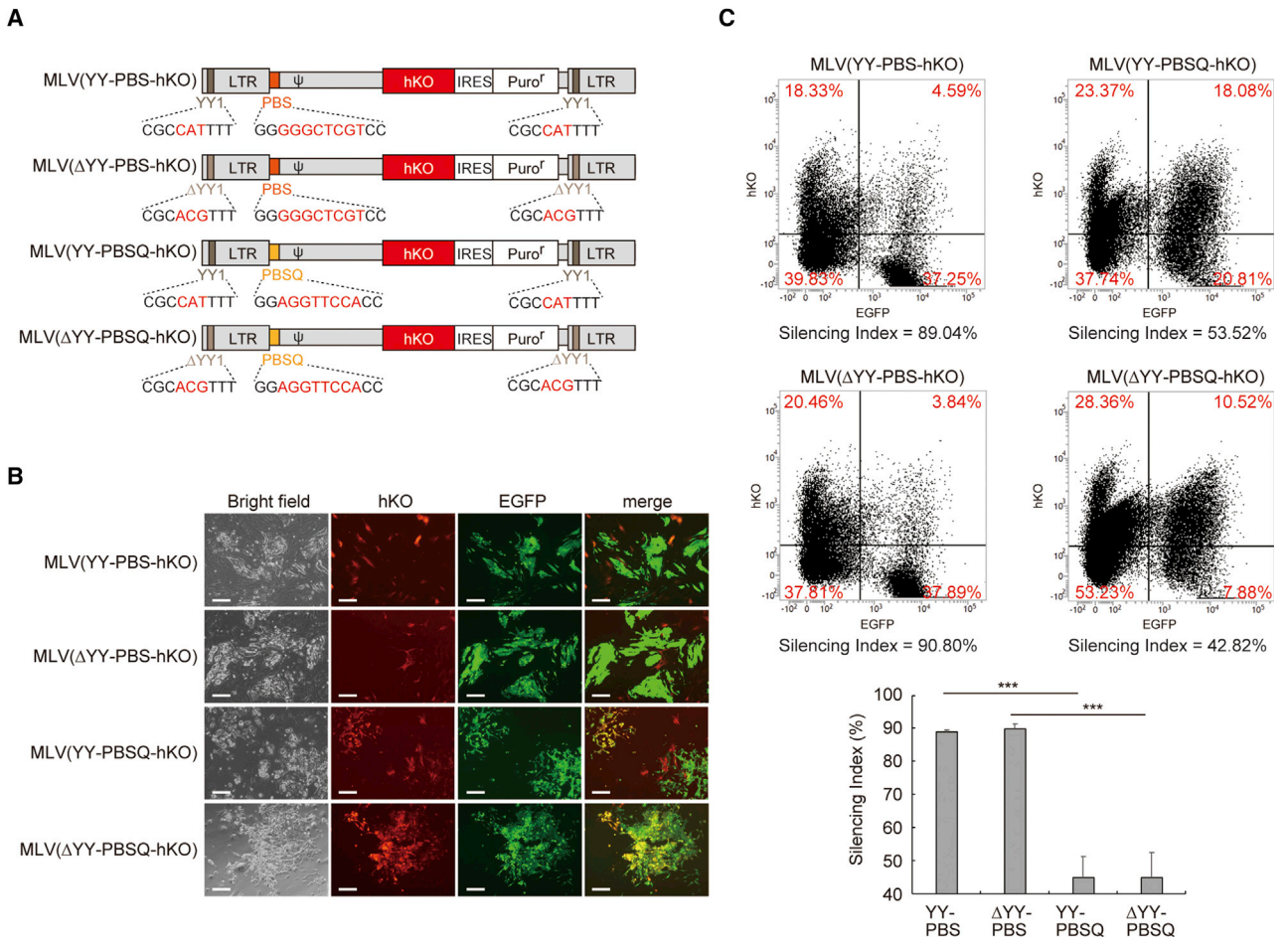


Figure 2. The PBS Is an Essential *cis* Element for Retroviral Silencing during Reprogramming

(A) Proviral structures of retroviruses possessing (a) mutation(s) in the YY1-binding site and PBS. The mutation of the YY1-binding site, introduced into the 3′ LTR of the retroviral plasmid vector, is copied into the 5′ LTR upon provirus production such that the YY1-binding sites in both the 5′ and 3′ LTRs become mutated in the provirus.

(B) Silencing of retroviruses with (a) mutation(s) in the YY1-binding site and PBS. MEFs transduced with the indicated retrovirus were infected by SeVdp(GKOSM). hKO and EGFP expression was observed at day 7 of reprogramming. Scale bars, 100 μm.

(C) Representative FACS profiles and silencing indices of cells transduced with mutated retroviral particles and reprogrammed for 7 days. The cells prepared in (B) were analyzed by FACS. Silencing indices are shown as bar graphs. Data are represented as means ± SEM of four independent experiments. ***$p < 0.001$. See also Figure S2 and Table S1.

Retroviral Silencing Precedes Acquisition of Pluripotency during Reprogramming

Pluripotent stem cells have a greater capacity to silence retrovirus expression than differentiated cells, indicating a close association between retroviral silencing and pluripotency (Schlesinger and Goff, 2015). However, our results here showed that retroviral silencing occurs even when the cells have not achieved the fully pluripotent state or remain partially reprogrammed (Figure 1). To further examine the relationship between the progression of reprogramming and retroviral silencing, we analyzed the gene expression changes and retroviral silencing throughout the process of reprogramming. qRT-PCR analysis showed that the hKO mRNA from the integrated MLV(YY-PBS-hKO) in MEFs infected by SeVdp(GKOSM) showed marked reduction of transcription at day 3, and silencing was almost complete at days 5–7 (Figure 4A). When the marked silencing was observed at day 5, the SeVdp(GKOSM)-infected MEFs repressed somatic genes (*Thy1* and *Snai1*), albeit incompletely, and became AP^+ at day 7 (Figure 1F). However, the cells were still $SSEA\text{-}1^-$ (Figure S4), and upregulation of the pluripotency-related genes *Cdh1*, *Oct4*, and *Nanog* was still negligible at day 5 (Figure 4B), confirming that retroviral silencing occurs before the acquisition of pluripotency.

We also used poorly reprograming cells induced by the SeVdp-based vectors carrying only a subset of reprograming factors (Figure 3A). The MEFs infected by SeVdp(GOSM) or SeVdp(GOKM) continued to proliferate, forming cell clusters reminiscent of colony formation, whereas those infected with SeVdp(GKOS) did not show rapid proliferation (Figure 4C). Among these MEFs, only those infected by SeVdp(GOKM)

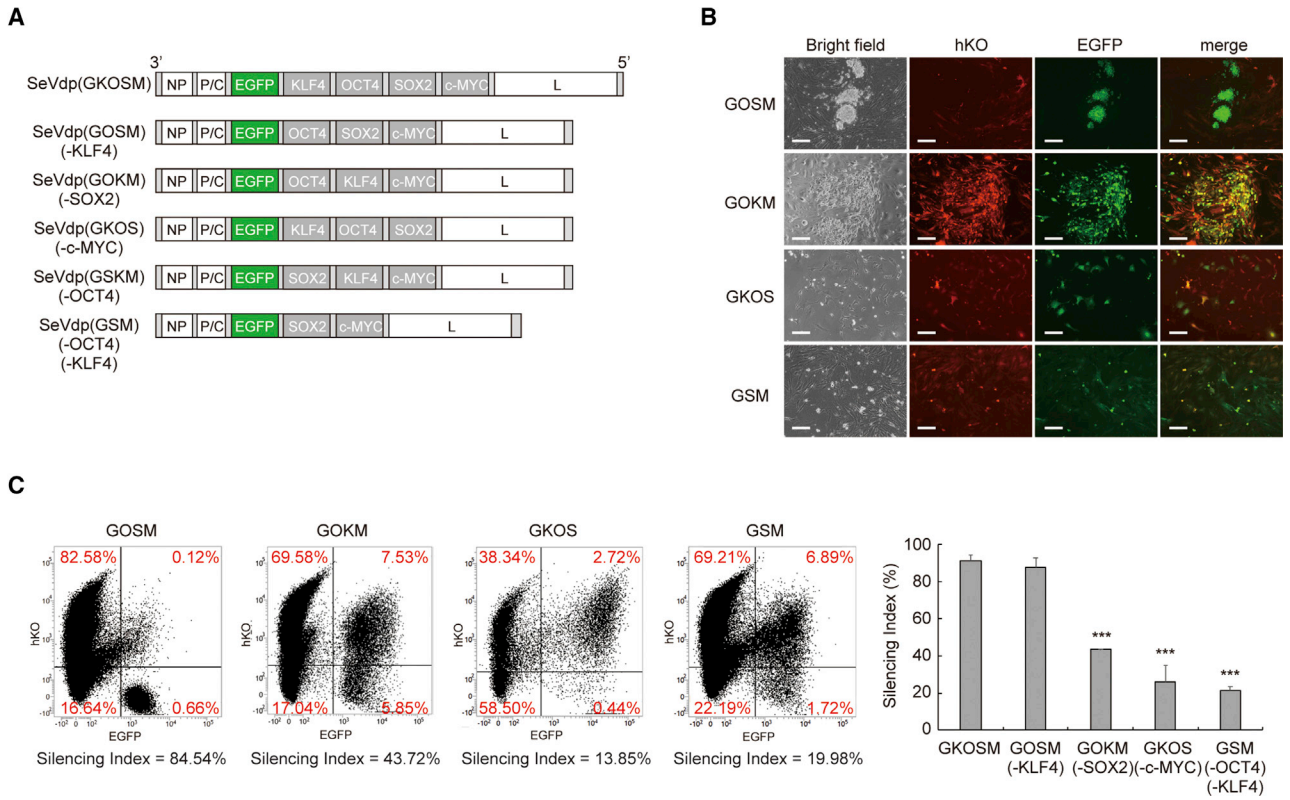


Figure 3. Requirement of Reprogramming Factors for Retroviral Silencing

(A) Structures of SeVdp vectors that express three or two reprogramming factors.

(B) MEFs expressing hKO from the MLV(YY-PBS-hKO) provirus were infected with the indicated SeVdp vector expressing three or two reprogramming factors. hKO and EGFP expression was observed at day 5 of SeVdp vector infection. Scale bars, 100 μm.

(C) Representative FACS profiles and silencing indices of cells infected with SeVdp vector expressing three or two reprogramming factors. Retroviral silencing was induced as in (B), and FACS analyses were performed at day 7 of SeVdp vector infection. Silencing indices are shown as bar graphs. Data are represented as means ± SEM of three independent experiments. ***$p < 0.001$.

See also Figure S3 and Table S1.

became AP^+ at day 10, suggesting that SeVdp(GOKM) reprograms MEFs to the highest extent (Figure 4C). Gene expression analyses at day 5 showed that all cells reprogrammed by these vectors failed to completely repress *Thy1* and *Snai1* expression and upregulated *Cdh1*, *Oct4*, and *Nanog* expression to a negligible extent (Figure 4D), indicating that these cells are at an early stage of reprogramming. Despite their comparable levels of gene expression changes, the cells infected with SeVdp(KOSM), SeVdp(GKOSM), or SeVdp(GOSM) showed strong retroviral silencing, while those infected with SeVdp(GOKM) or SeVdp(GKOS) showed only weak retroviral silencing (Figure 3C), indicating that retroviral silencing during reprogramming is correlated more closely with the presence of individual reprogramming factors rather than with the extent of reprogramming.

Identification of Factors that Mediate Retroviral Silencing during Reprogramming

To identify factors that function downstream of the reprogramming factors, we sought to identify proteins bound to the silenced proviral DNA during reprogramming. To this end, we performed iChIP followed by mass spectrometry, which allows nonbiased isolation of proteins assembled on a specific region of DNA (Hoshino and Fujii, 2009: Figure 5A). We designed a retrovirus in which the LexA-binding sites were inserted into the retrovirus genome at 11 or 451 bp downstream of the PBS site to create MLV(YY-PBS-Lex1-hKO) and MLV(YY-PBS-Lex2-hKO), respectively (Figure 5B). LexA protein was tagged with an N-terminal 3xFLAG epitope and expressed from a silencing-resistant retrovirus, MLV(3xFNLDD), (Figure 5B). When transduced into MEFs, MLV(YY-PBS-Lex1-hKO) showed a very low hKO expression (Figure 5C) presumably because multiple LexAs, bound only 11 bp downstream of the PBS, interfere with transcription from the LTR. However, MLV(YY-PBS-Lex2-hKO), in which LexA-binding sites are located 451 bp downstream of the PBS, displayed as high hKO expression as MLV(YY-PBS-hKO), in which there is no LexA-binding site (Figure 5C).

After transduction of MLV(YY-PBS-lex2-hKO) and MLV(3xFNLDD) into MEFs, SeVdp(GKOSM) was used to reprogram MEFs. After 5 days of reprogramming, hKO expression from the integrated MLV(YY-PBS-lex2-hKO), harboring the LexA sites, was repressed in a manner similar to

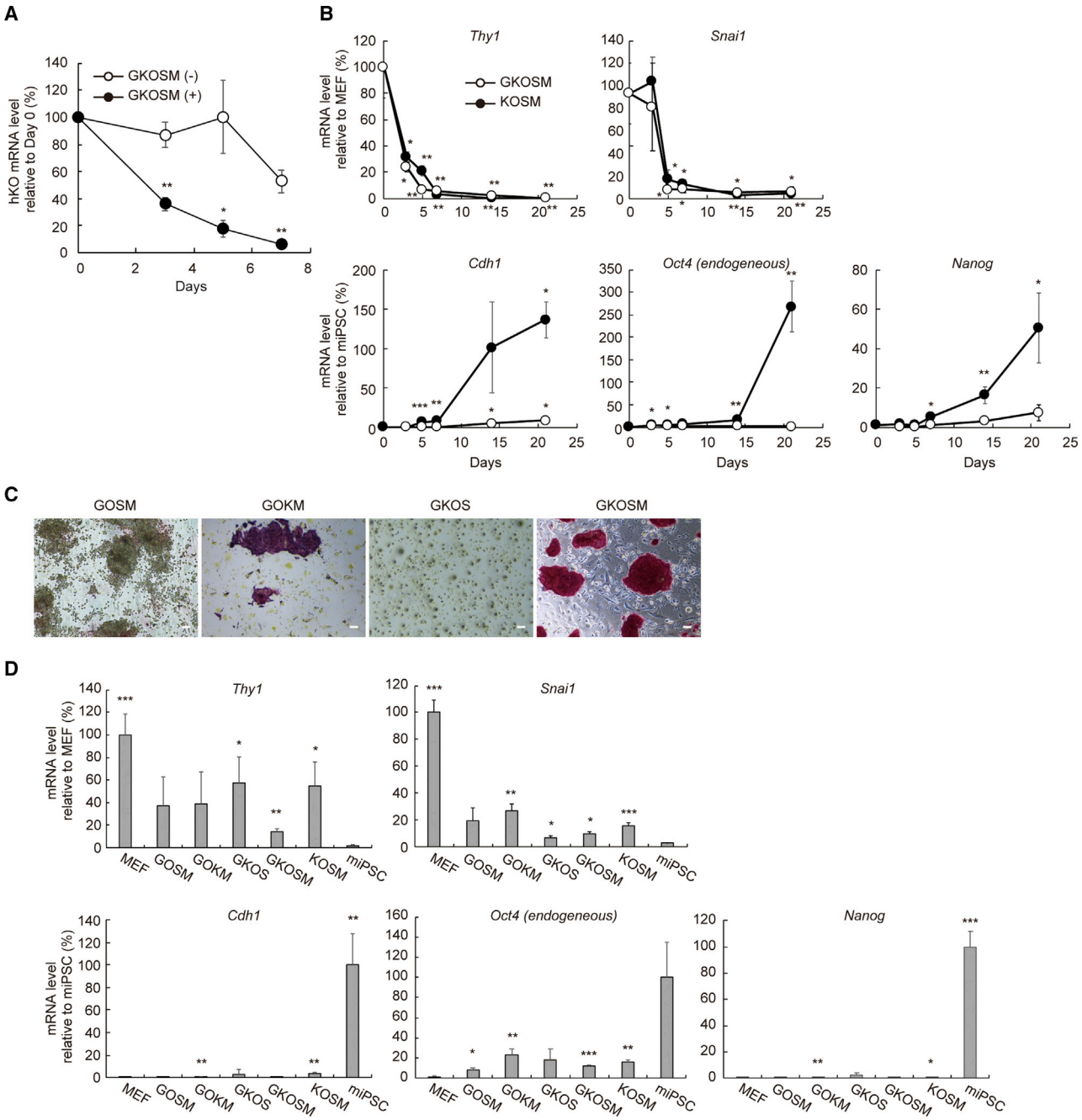


Figure 4. Retroviral Silencing Occurs before the Acquisition of Pluripotency

(A) MEFs expressing hKO from MLV(YY-PBS-hKO) were infected with or without SeVdp(GKOSM), and hKO mRNA levels were determined at the indicated days after SeVdp(GKOSM) infection. Data are represented as means ± SEM of three independent experiments. *p < 0.05, **p < 0.01 versus no SeVdp(GKOSM) infection.

(B) MEFs were reprogrammed by SeVdp(GKOSM) or SeVdp(KOSM), and *Thy1*, *Snai1*, *Cdh1*, *Oct4*, and *Nanog* mRNA levels were determined at days 0, 2, 5, 7, 14, and 21 of SeVdp vector infection. Data are represented as means ± SEM of three independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001 versus day 0.

(C) Alkaline phosphatase induction by three reprogramming factors. MEFs infected with each SeVdp vector were stained for alkaline phosphatase at day 10. Scale bars, 100 μm.

(D) Somatic or pluripotent marker expression after transduction with three reprogramming factors. *Thy1*, *Snai1*, *Cdh1*, *Oct4*, and *Nanog* mRNA levels in MEFs infected with each SeVdp vector were determined 5 days after SeVdp vector infection. Data are represented as means ± SEM of three independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001 versus mouse iPSCs (miPSCs) (*Thy1* and *Snai1*) or MEFs (*Cdh1*, *Oct4*, and *Nanog*).

See also Figure S4 and Table S4.

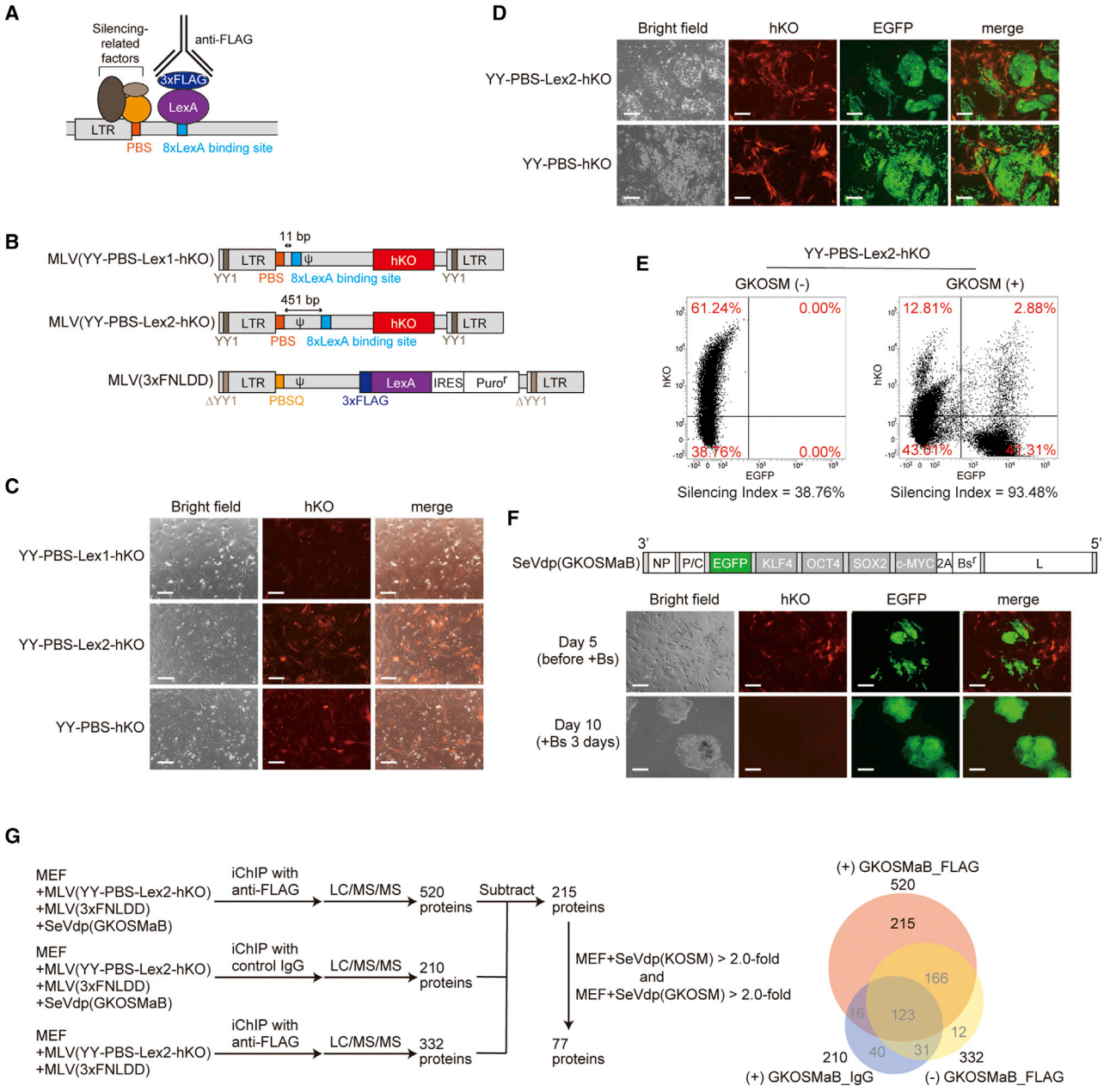


Figure 5. Identification of Factors Assembled on the Silenced Provirus during Reprogramming

(A) Schema for isolating factors associated with silenced proviral DNA using insertional chromatin immunoprecipitation (iChIP).

(B) Proviral structures of retroviruses harboring a LexA-binding site or expressing LexA protein for iChIP.

(C) hKO expression from the retrovirus with a LexA-binding site. hKO expression in MEFs infected with the indicated SeVdp vector was observed at day 7 of reprogramming. Scale bars, 100 μm.

(D) Silencing of the retrovirus with a LexA-binding site. MEFs transduced with the indicated retroviral particles as well as MLV(3xFNLDD) were infected by SeVdp(GKOSM). hKO and EGFP expression were observed at day 7 of reprogramming. Scale bars, 100 μm.

(E) Representative FACS profiles of cells transduced with the retroviral particles containing a LexA-binding site after reprogramming. MEFs transduced with MLV(YY-PBS-Lex2-hKO) and MLV(3xFNLDD) were cultured with or without infection of SeVdp(GKOSM). hKO and EGFP expression was analyzed by FACS at day 7 of SeVdp(GKOSM) infection.

(F) Selection of reprogrammed cells by blasticidin. MEFs transduced with MLV(YY-PBS-Lex2-hKO) and MLV(3xFNLDD) were reprogrammed by SeVdp(GKOSMaB). At day 7, blasticidin was added to select for reprogrammed cells. hKO and EGFP expression was observed at the indicated days. Scale bars, 100 μm.

(G) Schema for selecting the candidate factors assembled on the silenced provirus during reprogramming. The accession number for the microarray expression data used for selecting 77 proteins out of 215 is GSE134847.

See also Figure S5 and Tables S2 and S3.

MLV(YY-PBS-hKO) (Figure 5D). The FACS analysis also confirmed that hKO expression from MLV(YY-PBS-Lex2-hKO) and MLV(YY-PBS-hKO) become silenced to a similar extent in MEFs reprogrammed by SeVdp(GKOSM) (compare Figure 5E with Figure 1G). These results confirmed that the LexA-binding sites and the bound LexAs do not interfere with silencing of the integrated provirus during reprogramming. To ensure that 100% of MEFs are with SeVdp-based vectors, we further optimized the system by constructing another vector, SeVdp(GKOSMaB), which carries the blasticidin resistance gene (Figure 5F). Upon blasticidin selection, this vector eliminated cells not infected with the Sendai virus vector (Figure 5F), which would reduce precipitation of proteins bound to the proviral genome in MEFs not infected with the SeVdp vector. Almost all selected cells appeared to be reprogrammed and expressed 3xFLAG-tagged LexA from MLV(3xFNLDD) as well (Figure S5A).

The selected cells were then sonicated to fragment chromatin DNA into the size range of 500–1,000 bp. iChIP with an anti-FLAG antibody show that LexA bound near the 5′ LTR or 3′ LTR precipitated neighboring regions, but not the distant region encoding hKO or near the PBS, respectively (Figures S5B and S5C). Mass spectrometric analysis of the precipitated proteins identified 520 proteins, which were narrowed down to 215 proteins by excluding proteins precipitated similarly with control immunoglobulin G (IgG) or from the MEFs without reprogramming (Figure 5G). Among the 215 proteins, we selected 77 proteins (Table S2) that showed more than 2-fold mRNA induction during the first 2 days of infection by SeVdp(KOSM) or SeVdp(GKOSM) (Table S3). The selected proteins included EBP1 (*Pa2g4*), which was previously reported to be involved in retroviral silencing in ESCs and ECCs (Wang et al., 2014), demonstrating that the iChIP could correctly identify the proteins assembled on the provirus. The list included SET/TAF-I and acidic leucine-rich nuclear phosphoprotein (*Anp32a*), which are components of INHAT (Seo et al., 2001), a protein complex that inhibits histone acetyltransferases. Another protein, HMG2 (*Hmgb2*), which associates with SET/TAF-I (Fan et al., 2002), was also found among the list of specifically precipitated proteins.

TAF-Iα Facilitates Retroviral Silencing during Reprogramming

To test if SET/TAF-I, ANP32a, and HMG2 play any role in retroviral silencing during reprogramming, we first performed knockdown experiments using a pair of small interfering RNAs (siRNAs) for each protein. These siRNAs reduced the levels of the corresponding protein by 27.22% to 72.40% when introduced into NIH 3T3 cells (Figure S6A). The MEFs harboring MLV(YY-PBS-hKO) were treated with each siRNA and then infected by SeVdp(GKOSM) for 5 days (Figure 6A). Repression of hKO expression was observed under microscopy (Figure 6B) and quantified by FACS analysis (Figures 6C and S6B). Knockdown of SET/TAF-I with two different siRNAs clearly reduced retroviral silencing (Figures 6B and 6C). By contrast, knockdown of ANP32a and HMG2 did not show any significant effect (Figure 6C). Consistent with its known role in retroviral silencing (Wang et al., 2014), knockdown of EBP1 also reduced retroviral silencing (Figure 6C). These results show that an INHAT subunit, SET/TAF-I, plays an important role in retroviral silencing during reprogramming. A similar effect of SET/TAF-I knockdown was observed for silencing of the integrated lentivirus during reprogramming although the effect was smaller than that observed for silencing of MLV-based retrovirus (Figures 6D and 6E). No effect of SET/TAF-I knockdown was observed, however, for expression of ERVs in ESCs (Stocking and Kozak, 2008: Figure 6F). Thus, SET/TAF-I may play distinct roles in the silencing of different types of retroviruses.

The *Set/Taf-I* gene encodes two isoforms, TAF-Iα and TAF-Iβ, which are derived from two alternate promoters. TAF-Iα is expressed highly in ESCs, whereas TAF-Iβ is expressed ubiquitously in differentiated cells as a result of an isoform switch from TAF-Iα to TAF-Iβ during differentiation (Edupuganti et al., 2017). Consistent with the TAF-Iα/TAF-Iβ switch during ESC differentiation, analysis of the transcripts of the SET/TAF-I isoforms revealed that TAF-Iα expression increased while TAF-Iβ expression decreased after reprogramming, a reversal of the TAF-Iα/TAF-Iβ switch that occurs during ESC differentiation (Figure 7A). We therefore wished to know which SET/TAF-I isoform enhances retroviral silencing. Because the retroviral silencing by SeVdp(GKOSM) occurs rather rapidly, we chose SeVdp(GOKM), which is defective in retroviral silencing but reprograms MEFs to some extent by day 5 (Figures 3 and 4). MEFs that express hKO from the integrated MLV(YY-PBS-hKO) were transduced by retroviral particles expressing TAF-Iα or TAF-Iβ (Figure S7A) and then infected by SeVdp(GOKM). FACS analysis showed that the expression of TAF-Iα increased the hKO-negative cells from 34.87% ± 1.12% to 58.78% ± 2.02%, significantly enhancing the retroviral silencing during reprogramming by SeVdp(GOKM) (Figures 7B and S7B). By contrast, TAF-Iβ did not have a noticeable effect on retroviral silencing (Figures 7B and S7B). The effect of TAF-Iα expression on retroviral silencing was not observed in the absence of reprogramming (Figures 7B and S7B), indicating that TAF-Iα is effective for retroviral silencing only in conjunction with reprogramming. These results indicate that TAF-Iα, but not TAF-Iβ, plays an important role in retroviral silencing during reprogramming.

To confirm that TAF-Iα is recruited to the integrated provirus during reprogramming, MEFs that express hKO from the integrated MLV(YY-PBS-hKO) were transduced by retroviral particles expressing TAF-Iα and infected by SeVdp(GKOSM). Because TAF-Iα was FLAG tagged, we performed ChIP assays using an anti-FLAG antibody, and three sets of primers (Figure 7C; amplicons L, P, and O) were used to detect occupancy of TAF-Iα on the integrated provirus. As shown in Figure 7D, TAF-Iα occupancy was high on the 5′ LTR and PBS but lower on the hKO gene, which is remote from the 5′ LTR where transcription starts. Importantly, this high occupancy of TAF-Iα on the 5′ LTR and PBS was dependent on infection by SeVdp(GKOSM), indicating that TAF-Iα is recruited to the provirus in response to reprogramming. Regardless of reprogramming, however, no significant occupancy of TAF-Iα was observed for the *Nanog* gene (Figure 7D).

DISCUSSION

Pluripotent stem cells, such as ESCs and ECCs, potently silence exogenous and endogenous retroviruses, which otherwise may

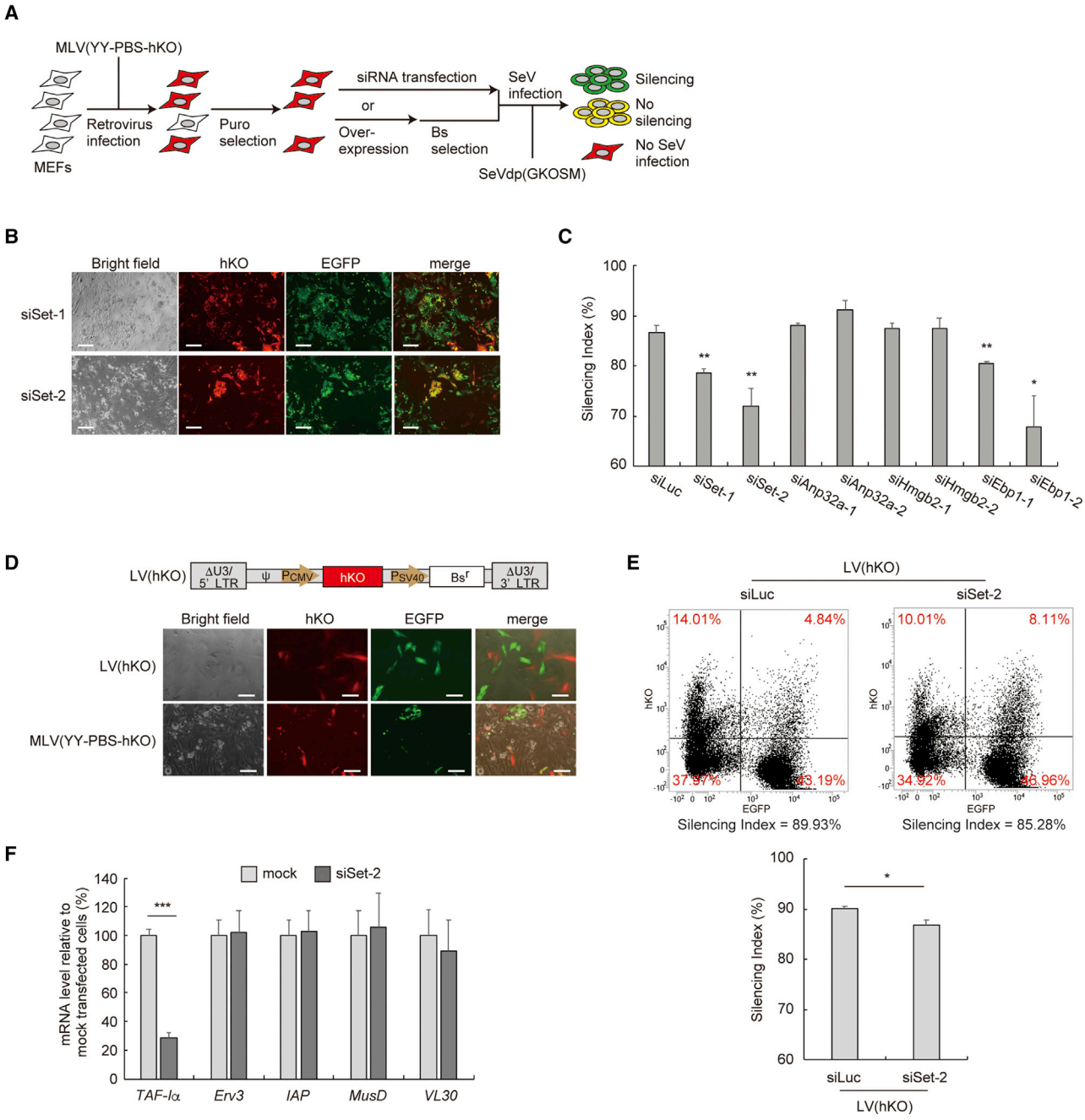


Figure 6. TAF-I Is Required for Retroviral Silencing during Reprogramming

(A) Knockdown and overexpression experiments of the candidate factors potentially involved in retroviral silencing.

(B) MEFs transduced with MLV(YY-PBS-hKO) were transfected with the indicated siRNA for 2 days and then infected by SeVdp(GKOSM). hKO and EGFP expression was observed at day 5 of reprogramming. Scale bars, 100 μm.

(C) Silencing was induced as in (B) with the indicated siRNA. Silencing indices were determined by FACS at day 7 of reprogramming by SeVdp(GKOSM). Data are represented as means ± SEM of three independent experiments. *$p < 0.05$, **$p < 0.01$ versus control cells treated with siRNA for the luciferase gene.

(D) Silencing of lentivirus during reprogramming of MEFs. Structure of a lentivirus expressing hKO, LV(hKO), is shown. MEFs were transduced with LV(hKO) or MLV(YY-PBS-hKO) and selected by blasticidin or puromycin, respectively. The hKO-expressing MEFs were infected by SeVdp(GKOSM), and hKO and EGFP expression were observed at day 5 of reprogramming. Scale bars, 100 μm.

(E) Representative FACS profiles and silencing indices of lentivirus-transduced cells infected by SeVdp(GKOSM) with prior knockdown of SET/TAF-I. MEFs expressing hKO from LV(hKO) were treated with or without siSet-2 for 2 days, and then the cells were infected by SeVdp(GKOSM). hKO and EGFP expression was

(legend continued on next page)

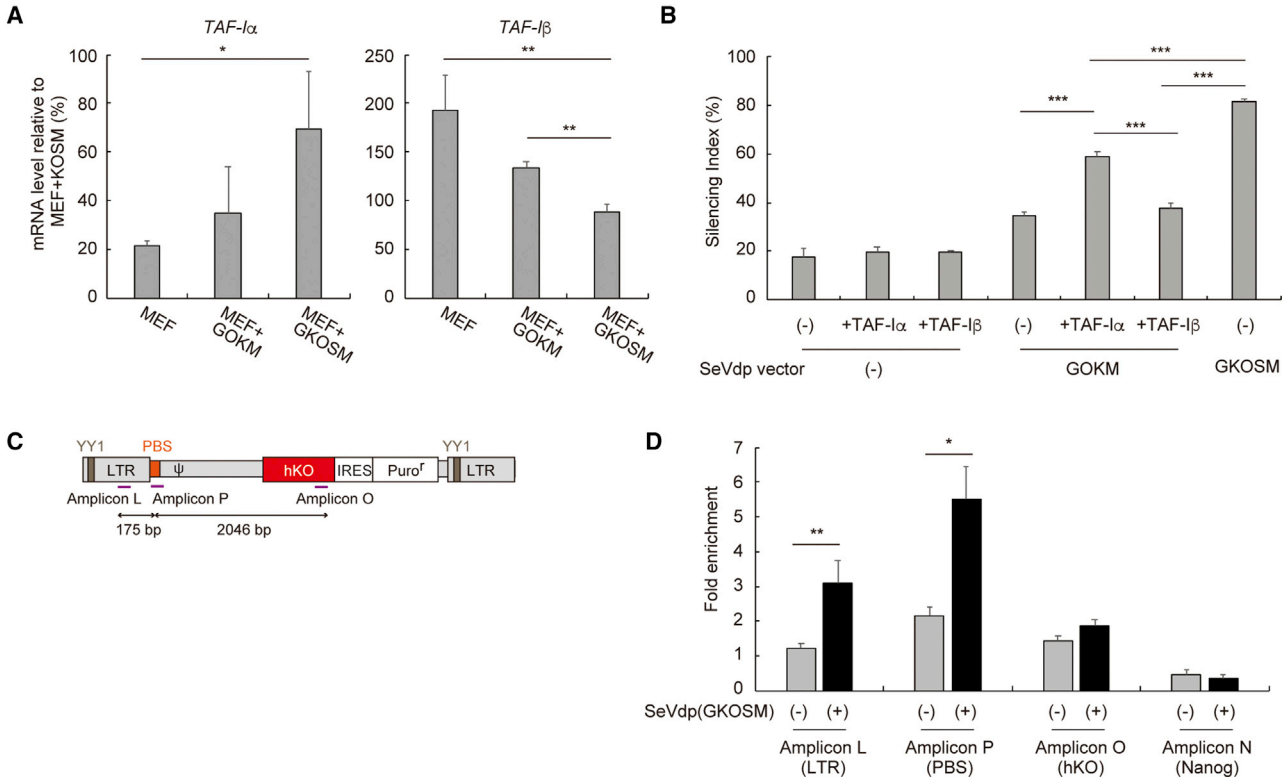


Figure 7. TAF-Iα, but Not TAF-Iβ, Facilitates Retroviral Silencing during Reprogramming

(A) *TAF-Iα* or *TAF-Iβ* mRNA levels in MEFs infected with SeVdp(GKOSM) or SeVdp(GOKM) were determined at day 5 of SeVdp vector infection. Data are represented as means ± SEM of three independent experiments. *$p < 0.05$, **$p < 0.01$.

(B) MEFs expressing hKO from MLV(YY-PBS-hKO) were transduced with or without MLV(TAF-Iα-Bs) or MLV(TAF-Iβ-Bs). After blasticidin selection for 2 days, the cells were infected by SeVdp(GKOSM) or SeVdp(GOKM). Silencing indices were determined by FACS analyses at day 7 of reprogramming. Data are represented as means ± SEM of three independent experiments. ***$p < 0.001$.

(C) Position of each amplicon in ChIP assay. The numbers indicate the distances between PBS and each amplicon.

(D) TAF-Iα binding to proviral DNA was analyzed by ChIP assays using MLV(YY-PBS-hKO)-harboring MEFs infected with or without SeVdp(GKOSM) for 10 days. Amplicon N is located at the promoter of the endogenous *Nanog* gene. Data are represented as means ± SEM of three independent experiments. *$p < 0.05$, **$p < 0.01$.

See also Figure S7 and Tables S1 and S4.

cause mutations by random integration in the host genome (Schlesinger and Goff, 2015). When somatic cells are reprogrammed into iPSCs, the retroviral vectors expressing reprogramming factors become transcriptionally repressed in iPSCs that have acquired full pluripotency (Okita et al., 2007), and inefficiently silenced retrovirus vectors indicate the low pluripotency of iPSCs (Jaenisch and Young, 2008; Maherali and Hochedlinger, 2008). Moreover, if reactivated, retrovirus vectors cause tumors after transplantation of the differentiated cells derived from iPSCs (Okita et al., 2007). Thus, silencing of the retroviral vectors is an epigenetic change that characterizes properly reprogrammed iPSCs (Jaenisch and Young, 2008).

Using SeVdp-based reprogramming, we have shown that retroviruses are strongly silenced during reprogramming as previously reported (Brambrink et al., 2008; Maherali and Hochedlinger, 2008; Okita et al., 2007; Stadtfeld et al., 2008). In contrast to earlier studies (Brambrink et al., 2008; Chan et al., 2009; Donai et al., 2013; Stadtfeld et al., 2008), however, we observed that retroviral silencing occurs at an unexpectedly early stage, even before the cells become positive for AP and upregulate pluripotency markers (Figures 1 and 4). Moreover, some of the SeVdp vectors with only three reprogramming factors showed retroviral silencing to some extent (Figure 4), indicating that pluripotency is separable from retroviral silencing. This discrepancy with previous studies may be due to a higher efficiency of reprogramming by SeVdp-derived vectors, which are capable of powerful expression of reprograming factors (Nishimura et al., 2011). More likely, however, is the use of retroviral reprogramming vectors, which

analyzed by FACS after 7 days of SeVdp(GKOSM) infection. Silencing indices were shown as bar graphs. Data are represented as means ± SEM of three independent experiments. **$p < 0.01$.

(F) Expression of ERVs was analyzed in EB5 mouse ESCs treated with or without siSet-2. The mRNA levels of *Taf-Iα* and the indicated ERVs were determined 3 days after the siRNA treatment. Data are represented as means ± SEM of three independent experiments. ***$p < 0.001$.

See also Figure S6 and Tables S1, S4, and S5.

resulted in overestimation of the time required for retroviral silencing. In the retrovirus-based reprogramming system, unless the retroviruses escape or delay silencing, the cells prematurely terminate reprogramming when they are still dependent upon exogenous reprogramming factors (Brambrink et al., 2008; Stadtfeld et al., 2008). Thus, the studies using a retrovirus-based reprogramming system may have inadvertently analyzed the cells that had successfully delayed retroviral silencing until a late reprogramming stage (Brambrink et al., 2008; Stadtfeld et al., 2008).

It is somewhat unexpected that out of the four reprogramming factors, KLF4 is not essential for silencing (Figure 3), because the role of KLF4 in the repression of somatic genes at an early stage of reprogramming is well established (Brambrink et al., 2008; Maherali and Hochedlinger, 2008; Stadtfeld et al., 2008), and an adequate level of KLF4 is important for the progression of reprogramming toward pluripotency (Kim et al., 2015; Nishimura et al., 2014). This may suggest that repression of somatic genes and silencing of the provirus employ distinct sets of factors or mechanisms. One possible relationship between reprogramming factors and the downstream effectors of silencing is that reprogramming factors induce the expression of effectors that generate a repressive chromatin structure. Consistent with this, the cells infected with the SeVdp vector that expresses OCT4, KLF4, and c-MYC but lacks SOX2 show low levels of TAF-Iα expression and reduced silencing activity, which suggests a role for SOX2 in the induction of TAF-Iα (Figure 7).

We found an important role for TAF-Iα, the SET/TAF-I isoform predominant in ESCs, in retroviral silencing (Figures 5, 6, and 7). As TAF-Iα is a subunit of the INHAT complex and possesses the INHAT activity by itself (Seo et al., 2001), one likely mechanism is that TAF-Iα protects histones from acetylation, thereby promoting the deposition of repressive marks and transcriptional repression of the provirus. Moreover, SET/TAF-I was shown to interact with KAP1 (also known as TRIM28) (Kalousi et al., 2015), which nucleates assemblage of the repressive chromatin structure by virtue of its binding to ZFP809 (Wolf and Goff, 2009). This repressive structure may repress transcription prior to DNA methylation, which further consolidates the repressive chromatin structure (Matsui et al., 2010). An alternative, but not mutually exclusive, possibility is that TAF-Iα functions as a histone chaperone (Hammond et al., 2017), depositing non-acetylated histones onto the integrated provirus to diminish transcription from the LTR. Consistent with this, a recent study demonstrated that Chaf1a and Chaf1b, two subunits of the replicative histone chaperone CAF1 (the chromatin assembly factor), are important factors for retroviral silencing in ESCs (Yang et al., 2015).

Structurally, SET/TAF-I is composed of three domains that are shared between TAF-Iα and TAF-Iβ: the dimerization domain, the earmuff domain for histone chaperone activity, and the acidic domain for INHAT activity (Muto et al., 2007). In addition, TAF-Iα and TAF-Iβ possess a unique N-terminal region, 36 and 24 amino acids in length, respectively, that extends N-terminally from the dimerization domain (Muto et al., 2007; Nagata et al., 1995). The N-terminal region of TAF-Iα is rich in proline and lysine but shows no significant homology to any known protein upon BLAST searching. Interestingly, however, the N-terminal region fails to yield the electron density upon crystallization, probably because of its flexible structure (Muto et al., 2007). Thus, the flexible N-terminal region may protrude from the dimerization domain to be readily available for interaction with other proteins, which could define the functional specificity of TAF-Iα and TAF-Iβ in retroviral silencing during reprogramming.

In conclusion, the SeVdp-based reprogramming system described here provide insights into the mechanism of retroviral silencing during reprogramming and nicely complements previous studies that used ESCs and ECCs. Because the defined factors initiate retroviral silencing during reprogramming, our system described here should facilitate identification and mechanistic analyses of the pathways that link the core pluripotency network to effectors in the repressive silencing complex.

STAR★METHODS

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

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Primary Cells
- METHOD DETAILS
  - Production of Retrovirus or Lentivirus
  - Production of SeVdp Vectors
  - Induction of Retroviral Silencing by Reprogramming
  - Fluorescence-activated Cell Sorting (FACS) Analysis
  - Characterization of Reprogrammed Cells
  - Collection of Whole Genomic Gene Expression Profile
  - Insertional Chromatin Immunoprecipitation (iChIP)
  - ChIP
  - Immunofluorescence Staining
  - Determination of Protein Expression by Immunoblotting
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.10.010.

ACKNOWLEDGMENTS

This work was funded by JSPS KAKENHI grants JP16K07244 (K. Nishimura), JP19H03203 (K. Nishimura), JP19K22945 (K. Nishimura), JP19K07343 (A.F.), and JP17H04036 (K.H.). This study was also supported by the NIMS Molecule & Material Synthesis Platform "Nanotechnology Platform Project" operated by the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan. We are grateful to K. Kato for technical advice regarding TAF-I and T. Nishimura for technical assistance.

AUTHOR CONTRIBUTIONS

K. Nishimura and K.H. designed the research. P.L.B., K. Nishimura, S.M.G., Y.H., Y.O., and Y.I. performed research and analyzed the data. A.K., S.A., K.M., K. Nagata, and M.N. prepared materials. K. Nishimura, A.F., and K.H. wrote the paper.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 26, 2019
Revised: August 2, 2019
Accepted: October 1, 2019
Published: November 12, 2019

REFERENCES

Apostolou, E., and Hochedlinger, K. (2013). Chromatin dynamics during cellular reprogramming. Nature *502*, 462–471.

Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. Cell *47*, 391–399.

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.

Chan, E.M., Ratanasirintrawoot, S., Park, I.H., Manos, P.D., Loh, Y.H., Huo, H., Miller, J.D., Hartung, O., Rho, J., Ince, T.A., et al. (2009). Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. Nat. Biotechnol. *27*, 1033–1037.

David, L., and Polo, J.M. (2014). Phases of reprogramming. Stem Cell Res. (Amst.) *12*, 754–761.

Donai, K., Kuroda, K., Guo, Y., So, K.H., Sone, H., Kobayashi, M., Nishimori, K., and Fukuda, T. (2013). Establishment of a reporter system to monitor silencing status in induced pluripotent stem cell lines. Anal. Biochem. *443*, 104–112.

Edupuganti, R.R., Harikumar, A., Aaronson, Y., Biran, A., Sailaja, B.S., Nissim-Rafinia, M., Azad, G.K., Cohen, M.A., Park, J.E., Shivalila, C.S., et al. (2017). Alternative SET/TAFI promoters regulate embryonic stem cell differentiation. Stem Cell Reports *9*, 1291–1303.

Fan, Z., Beresford, P.J., Zhang, D., and Lieberman, J. (2002). HMG2 interacts with the nucleosome assembly protein SET and is a target of the cytotoxic T-lymphocyte protease granzyme A. Mol. Cell. Biol. *22*, 2810–2820.

Feuer, G., Taketo, M., Hanecak, R.C., and Fan, H. (1989). Two blocks in Moloney murine leukemia virus expression in undifferentiated F9 embryonal carcinoma cells as determined by transient expression assays. J. Virol. *63*, 2317–2324.

Friedli, M., Turelli, P., Kapopoulou, A., Rauwel, B., Castro-Díaz, N., Rowe, H.M., Ecco, G., Unzu, C., Planet, E., Lombardo, A., et al. (2014). Loss of transcriptional control over endogenous retroelements during reprogramming to pluripotency. Genome Res. *24*, 1251–1259.

Fujita, T., Asano, Y., Ohtsuka, J., Takada, Y., Saito, K., Ohki, R., and Fujii, H. (2013). Identification of telomere-associated molecules by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP). Sci. Rep. *3*, 3171.

Hammond, C.M., Strømme, C.B., Huang, H., Patel, D.J., and Groth, A. (2017). Histone chaperone networks shaping chromatin function. Nat. Rev. Mol. Cell Biol. *18*, 141–158.

Hoshino, A., and Fujii, H. (2009). Insertional chromatin immunoprecipitation: a method for isolating specific genomic regions. J. Biosci. Bioeng. *108*, 446–449.

Hotta, A., and Ellis, J. (2008). Retroviral vector silencing during iPS cell induction: an epigenetic beacon that signals distinct pluripotent states. J. Cell. Biochem. *105*, 940–948.

Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell *132*, 567–582.

Jähner, D., Stuhlmann, H., Stewart, C.L., Harbers, K., Löhler, J., Simon, I., and Jaenisch, R. (1982). De novo methylation and expression of retroviral genomes during mouse embryogenesis. Nature *298*, 623–628.

Kalousi, A., Hoffbeck, A.S., Selemenakis, P.N., Pinder, J., Savage, K.I., Khanna, K.K., Brino, L., Dellaire, G., Gorgoulis, V.G., and Soutoglou, E. (2015). The nuclear oncogene SET controls DNA repair by KAP1 and HP1 retention to chromatin. Cell Rep. *11*, 149–163.

Kempler, G., Freitag, B., Berwin, B., Nanassy, O., and Barklis, E. (1993). Characterization of the Moloney murine leukemia virus stem cell-specific repressor binding site. Virology *193*, 690–699.

Kim, S.I., Oceguera-Yanez, F., Hirohata, R., Linker, S., Okita, K., Yamada, Y., Yamamoto, T., Yamanaka, S., and Woltjen, K. (2015). KLF4 N-terminal variance modulates induced reprogramming to pluripotency. Stem Cell Reports *4*, 727–743.

Koyanagi-Aoi, M., Ohnuki, M., Takahashi, K., Okita, K., Noma, H., Sawamura, Y., Teramoto, I., Narita, M., Sato, Y., Ichisaka, T., et al. (2013). Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. Proc. Natl. Acad. Sci. USA *110*, 20569–20574.

Kyttälä, A., Moraghebi, R., Valensisi, C., Kettunen, J., Andrus, C., Pasumarthy, K.K., Nakanishi, M., Nishimura, K., Ohtaka, M., Weltner, J., et al. (2016). Genetic variability overrides the impact of parental cell type and determines iPSC differentiation potential. Stem Cell Reports *6*, 200–212.

Leeb, M., Pasini, D., Novatchkova, M., Jaritz, M., Helin, K., and Wutz, A. (2010). Polycomb complexes act redundantly to repress genomic repeats and genes. Genes Dev. *24*, 265–276.

Li, M., Liu, G.H., and Izpisua Belmonte, J.C. (2012). Navigating the epigenetic landscape of pluripotent stem cells. Nat. Rev. Mol. Cell Biol. *13*, 524–535.

Linney, E., Davis, B., Overhauser, J., Chao, E., and Fan, H. (1984). Non-function of a Moloney murine leukaemia virus regulatory sequence in F9 embryonal carcinoma cells. Nature *308*, 470–472.

Maherali, N., and Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. Cell Stem Cell *3*, 595–605.

Matsui, T., Leung, D., Miyashita, H., Maksakova, I.A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M.C., and Shinkai, Y. (2010). Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature *464*, 927–931.

Matsumoto, T., Fujimori, K., Andoh-Noda, T., Ando, T., Kuzumaki, N., Toyoshima, M., Tada, H., Imaizumi, K., Ishikawa, M., Yamaguchi, R., et al. (2016). Functional neurons generated from T cell-derived induced pluripotent stem cells for neurological disease modeling. Stem Cell Reports *6*, 422–435.

Muto, S., Senda, M., Akai, Y., Sato, L., Suzuki, T., Nagai, R., Senda, T., and Horikoshi, M. (2007). Relationship between the structure of SET/TAF-Ibeta/INHAT and its histone chaperone activity. Proc. Natl. Acad. Sci. USA *104*, 4285–4290.

Nagata, K., Kawase, H., Handa, H., Yano, K., Yamasaki, M., Ishimi, Y., Okuda, A., Kikuchi, A., and Matsumoto, K. (1995). Replication factor encoded by a putative oncogene, set, associated with myeloid leukemogenesis. Proc. Natl. Acad. Sci. USA *92*, 4279–4283.

Nagata, K., Saito, S., Okuwaki, M., Kawase, H., Furuya, A., Kusano, A., Hanai, N., Okuda, A., and Kikuchi, A. (1998). Cellular localization and expression of template-activating factor I in different cell types. Exp. Cell Res. *240*, 274–281.

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, K., Kato, T., Chen, C., Oinam, L., Shiomitsu, E., Ayakawa, D., Ohtaka, M., Fukuda, A., Nakanishi, M., and Hisatake, K. (2014). Manipulation of KLF4 expression generates iPSCs paused at successive stages of reprogramming. Stem Cell Reports *3*, 915–929.

Nishimura, K., Aizawa, S., Nugroho, F.L., Shiomitsu, E., Tran, Y.T.H., Bui, P.L., Borisova, E., Sakuragi, Y., Takada, H., Kurisaki, A., et al. (2017). A role for KLF4 in promoting the metabolic shift via TCL1 during induced pluripotent stem cell generation. Stem Cell Reports *8*, 787–801.

Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983). Independent mechanisms involved in suppression of the Moloney leukemia virus genome during differentiation of murine teratocarcinoma cells. Cell *32*, 1105–1113.

Okada, M., and Yoneda, Y. (2011). The timing of retroviral silencing correlates with the quality of induced pluripotent stem cell lines. Biochim. Biophys. Acta *1810*, 226–235.

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

Pannell, D., Osborne, C.S., Yao, S., Sukonnik, T., Pasceri, P., Karaiskakis, A., Okano, M., Li, E., Lipshitz, H.D., and Ellis, J. (2000). Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. EMBO J. *19*, 5884–5894.

Papp, B., and Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. Cell *152*, 1324–1343.

Petersen, R., Kempler, G., and Barklis, E. (1991). A stem cell-specific silencer in the primer-binding site of a retrovirus. Mol. Cell. Biol. *11*, 1214–1221.

Schlesinger, S., and Goff, S.P. (2015). Retroviral transcriptional regulation and embryonic stem cells: war and peace. Mol. Cell. Biol. *35*, 770–777.

Schlesinger, S., Lee, A.H., Wang, G.Z., Green, L., and Goff, S.P. (2013). Proviral silencing in embryonic cells is regulated by Yin Yang 1. Cell Rep. *4*, 50–58.

Schultz, D.C., Friedman, J.R., and Rauscher, F.J., 3rd. (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. Genes Dev. *15*, 428–443.

Seo, S.B., McNamara, P., Heo, S., Turner, A., Lane, W.S., and Chakravarti, D. (2001). Regulation of histone acetylation and transcription by INHAT, a human cellular complex containing the set oncoprotein. Cell *104*, 119–130.

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.

Stewart, C.L., Stuhlmann, H., Jähner, D., and Jaenisch, R. (1982). De novo methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA *79*, 4098–4102.

Stocking, C., and Kozak, C.A. (2008). Murine endogenous retroviruses. Cell. Mol. Life Sci. *65*, 3383–3398.

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., and Yamanaka, S. (2013). Induced pluripotent stem cells in medicine and biology. Development *140*, 2457–2461.

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.

Wang, G.Z., Wolf, D., and Goff, S.P. (2014). EBP1, a novel host factor involved in primer binding site-dependent restriction of moloney murine leukemia virus in embryonic cells. J. Virol. *88*, 1825–1829.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318–324.

Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. Cell *131*, 46–57.

Wolf, D., and Goff, S.P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. Nature *458*, 1201–1204.

Wolf, D., Cammas, F., Losson, R., and Goff, S.P. (2008). Primer binding site-dependent restriction of murine leukemia virus requires HP1 binding by TRIM28. J. Virol. *82*, 4675–4679.

Yang, B.X., El Farran, C.A., Guo, H.C., Yu, T., Fang, H.T., Wang, H.F., Schlesinger, S., Seah, Y.F., Goh, G.Y., Neo, S.P., et al. (2015). Systematic identification of factors for provirus silencing in embryonic stem cells. Cell *163*, 230–245.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-SSEA-1	Santa Cruz	Cat# sc-21702; RRID: AB_626918
Mouse monoclonal anti-SeV NP	Nishimura et al. 2007	N/A
Mouse monoclonal anti-DYKDDDDK tag	Wako	Cat# 018-22381; RRID: AB_10659453
Rabbit polyclonal anti-KLF4	Nishimura et al. 2014	N/A
Mouse monoclonal anti-TAF-I	Nagata et al. 1998	N/A
Mouse monoclonal anti-ANP32A	Santa Cruz	Cat# sc-374552; RRID: AB_10989733
Mouse monoclonal anti-HMGB2	Santa Cruz	Cat# sc-271689; RRID: AB_10709027
Mouse monoclonal anti-EBP1	Santa Cruz	Cat# sc-393114
Mouse monoclonal anti-α-TUBULIN	Abcam	Cat# ab7291; RRID: AB_2241126
Normal mouse IgG	Santa Cruz	Cat# sc-2025; RRID: AB_737182
AlexaFlour 555 goat anti-mouse IgG	Thermo Fisher	Cat# A32727; RRID: AB_2633276
AlexaFlour 647 goat anti-mouse IgG	Thermo Fisher	Cat# A28181; RRID: AB_2536165
AlexaFlour 555 goat anti-mouse IgM	Thermo Fisher	Cat# A21426; RRID: AB_1500929
Bacterial and Virus Strains		
MLV(ΔYY-PBSQ-hKO)	Nishimura et al. 2017	N/A
MLV(YY-PBS-hKO)	This paper	N/A
MLV(ΔYY-PBS-hKO)	This paper	N/A
MLV(YY-PBSQ-hKO)	This paper	N/A
MLV(YY-PBS-Lex1-hKO)	This paper	N/A
MLV(YY-PBS-Lex2-hKO)	This paper	N/A
MLV(YY-PBS-Lex3-hKO)	This paper	N/A
MLV(3xFNLDD)	This paper	N/A
MLV(TAF-Iα-Bs)	This paper	N/A
MLV(TAF-Iβ-Bs)	This paper	N/A
LV(hKO)	This paper	N/A
SeVdp(KOSM)	Nishimura et al. 2014	N/A
SeVdp(GKOSM)	Nishimura et al. 2014	N/A
SeVdp(GOSM)	This paper	N/A
SeVdp(GOKM)	This paper	N/A
SeVdp(GKOS)	This paper	N/A
SeVdp(GSKM)	This paper	N/A
SeVdp(GSM)	This paper	N/A
SeVdp(GKOSMaB)	This paper	N/A
Chemicals, Peptides, and Recombinant Proteins		
FBS	GIBCO	Cat# 10270
Hexadimethrine bromide (Polybrene)	Sigma	Cat# 107689; CAS: 28728-55-4
Puromycin	Nacalai	Cat# 29455-12; CAS: 58-58-2
Blasticidin	Wako	Cat# 029-18701; CAS: 3513-03-9
GlutaMAX	Invitrogen	Cat# 35050061
Non-essential amino acids	Nacalai	Cat# 06344-56
2-mercaptoethanol	Invitrogen	Cat# 21985023
Penicillin-streptomycin	Nacalai	Cat# 26252-94
Leukemia Inhibitory Factor (LIF)	Wako	Cat# 125-05603

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Stemsure DMEM	Wako	Cat# 197-16275
StemSure Serum Replacement	Wako	Cat# 191-18375
iMatrix-511	TaKaRa	Cat# 892011
StemFit AK02N	TaKaRa	Cat# AK02N
Trypsin	Promega	Cat# V5113
Critical Commercial Assays		
In-fusion HD cloning kit	TaKaRa	Cat# Z9633N
ViraPower Lentiviral Packaging Mix	Invitrogen	Cat# K4975-00
Lipofectamine 2000 Transfection Reagent	Invitrogen	Cat# 11668019
Lipofectamine LTX Reagent with PLUS Reagent	Invitrogen	Cat# 15338100
Lipofectamine RNAiMAX Transfection Reagent	Invitrogen	Cat# 13778075
ISOGEN	Nippon Gene	Cat# 319-90211
Superscript III First-Strand Synthesis System	Invitrogen	Cat# 18080-051
GoTaq qPCR Master Mix	Promega	Cat# A6001
Alkaline Phosphatase Staining Kit II	Stemgent	Cat# 00-0055
Protease Inhibitor Cocktail	Sigma	Cat# P2714
Dynabeads Protein G	Thermo Fisher	Cat# 1004D
Quick Amp Labeling Kit	Agilent	Cat# 5190-0442
Whole Mouse Genome DNA microarray 4x44K	Agilent	Cat# G4846A
Gene Expression Hybridization Kit	Agilent	Cat# 5188-5242
Deposited Data		
DNA microarray data	This paper	GSE: 134847
Experimental Models: Cell Lines		
C57BL/6 mouse fibroblasts	This paper	N/A
Mouse iPSCs	Nishimura et al. 2011	N/A
F9	JCRB Cell Bank	Cat# JCRB0721
TIG-3	JCRB Cell Bank	Cat# JCRB0510
Plat-E	Cell Biolabs.	Cat# VPK-300
293FT	Invitrogen	Cat# R70007
BHK/T7/151M(SE)	Nishimura et al. 2011	N/A
Oligonucleotides		
Nucleotide sequences of siRNA – See Table S5	This paper	N/A
Primers for RT-qPCR and ChIP-qPCR – See Table S4	This paper	N/A
Recombinant DNA		
pLenti6/UbC/mSlc7a1	Takahashi et al. 2007	Addgene #17224
pMCsΔYY-hKO-IRES-Puro	Nishimura et al. 2017	N/A
3xFNLDD/pMXs-puro	Fujita et al. 2013	Addgene #49536
pEGFP-C1	Clontech	Cat# 6084-1
pMXs	Cell Biolabs.	Cat# RTV-010
pLenti6.3/V5-TOPO	Invitrogen	Cat# K531520
pCMV-NP	Nishimura et al. 2011	N/A
pCMV-P	Nishimura et al. 2011	N/A
pCMV-L	Nishimura et al. 2011	N/A
pMKIT-151M	Nishimura et al. 2011	N/A
pSRD-HN-Fmut	Nishimura et al. 2011	N/A
Software and Algorithms		
BD FACSuite software	BD Bioscience	N/A
Feature Extraction software	Agilent	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Mascot server	MATRIX Science	N/A
FUSION Capt Software	Vilber-Lourmat	N/A

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ken Nishimura (ken-nishimura@md.tsukuba.ac.jp). SeVdp vectors are available from the Lead Contact with a completed Materials Transfer Agreement. Other materials generated in this study are available from the Lead Contact without restriction.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Primary Cells

Mouse embryonic fibroblasts (MEFs) were isolated from embryos of C57BL/6 male mouse. The cells and F9 cells (JCRB0721, JCRB Cell Bank) were cultured in DMEM medium (DMEM (Nacalai) supplemented with 10% fetal bovine serum (FBS) (GIBCO), and, penicillin-streptomycin (Nacalai)). TIG-3 human fibroblasts (JCRB0510, JCRB Cell Bank) were also cultured in DMEM medium.

METHOD DETAILS

Production of Retrovirus or Lentivirus

pMXsYY-hKO-IRES-Puro plasmid was constructed by inserting a BamHI-SalI fragment of pMCsΔYY-hKO-IRES-Puro (Nishimura et al., 2017) (a DNA fragment containing humanized Kusabira Orange (hKO)-IRES-Puro) into pMXs (Cell Biolabs.) plasmid. pMXsΔYY-hKO-IRES-Puro plasmid was constructed by replacing the 5′ LTR and PBS region of pMCsΔYY-hKO-IRES-Puro with that of pMXsYY-hKO-IRES-Puro by using SspI and SpeI restriction sites. pMCsYY-hKO-IRES-Puro plasmid was constructed by similar replacement of the 5′ LTR and PBS region of pMXsYY-hKO-IRES-Puro with that of pMCsΔYY-hKO-IRES-Puro. A sequence of 3xFLAG-NLS-LexA in 3xFNLDD/pMXs-puro (Fujita et al., 2013) (Addgene) was re-cloned into pMCsΔYY-IRES-Puro to construct pMCsΔYY-3xFNLDD-IRES-Puro. The DNA fragment containing a 158 bp LexA-binding sites was synthesized by annealing two oligonucleotides and inserted into pMXsYY-hKO plasmid using In-fusion HD cloning kit (TaKaRa) at three locations 11 bp, 451 bp or 1945 bp from the PBS to create pMXsYY-Lex1-hKO, pMXsYY-Lex2-hKO, or pMXsYY-Lex3-hKO, respectively. pMCsΔYY-IRES-Bs containing the blasticidin-resistant gene was constructed by replacing the puromycin-resistant gene using In-fusion HD cloning kit. Mouse TAF-Iα or TAF-Iβ gene fused with a FLAG-tag was cloned by PCR into pEGFP-C1 plasmid (Clontech). The FLAG-TAF-Iα or FLAG-TAF-Iβ DNA fragment from this plasmid was re-cloned into pMCsΔYY-IRES-Bs to construct pMCsΔYY-TAF-Iα-IRES-Bs or pMCsΔYY-TAF-Iβ-IRES-Bs, respectively. For lentivirus construction, we inserted hKO cDNA in pLenti6.3/V5-TOPO plasmid (Invitrogen) by TA cloning to construct pLenti-hKO-Bs.

Each plasmid for retrovirus was transfected into PLAT-E cells by using Lipofectamine 2000 Transfection Reagent (Invitrogen), and after overnight incubation, the medium was replaced afresh. The culture supernatant was collected 3 days after transfection and filtered with a 0.45 μm cellulose acetate filter. The virus stock was frozen in liquid nitrogen and stored at −80°C until use. Titers of the retroviruses were determined by counting the number of hKO-expressing cells or puromycin or blasticidin-resistant cells. MLV(YY-PBS-hKO), MLV(ΔYY-PBS-hKO), MLV(YY-PBSQ-hKO), MLV(ΔYY-PBSQ-hKO), MLV(YY-PBS-Lex1-hKO), MLV(YY-PBS-Lex2-hKO), MLV(YY-PBS-Lex3-hKO), MLV(3xFNLDD), MLV(TAF-Iα-Bs), and MLV(TAF-Iβ-Bs) were constructed from pMXsYY-hKO-IRES-Puro, pMXsΔYY-hKO-IRES-Puro, pMCsYY-hKO-IRES-Puro, pMCsΔYY-hKO-IRES-Puro, pMXsYY-Lex1-hKO, pMXsYY-Lex2-hKO, pMXsYY-Lex3-hKO, pMCsΔYY-3xFNLDD-IRES-Puro, pMCsΔYY-TAF-Iα-IRES-Bs, and pMCsΔYY-TAF-Iβ-IRES-Bs, respectively. For LV(hKO) production, pLenti-hKO-Bs and plasmids for lentivirus packaging (pLP1, pLP2, and pLP/VSVG [ViraPower Lentiviral Packaging Mix]; Invitrogen) were transfected to 293FT (Invitrogen), and the supernatant was collected same as retroviruses.

Production of SeVdp Vectors

Based on the SeVdp vector cDNA encoding OCT4, SOX2, KLF4, and c-MYC, the EGFP gene was used to replace one of the four factors to construct cDNAs for SeVdp(GOSM), SeVdp(GOKM), SeVdp(GKOS), and SeVdp(GSKM). The cDNA for SeVdp(GSM) was constructed by removing the Klf4 gene from SeVdp(GSKM). The cDNA for SeVdp(GKOSMaB) was constructed by insertion of the blasticidin-resistant gene together with the T2A peptide sequence after the c-Myc gene of SeVdp(GKOSM) (Nishimura et al., 2014).

To produce SeVdp vectors, each SeVdp cDNA plasmid and the plasmids encoding SeV genes (NP, P, M, F, HN, and L) were transfected into BHK/T7/151M(SE) cells (Nishimura et al., 2011) by using Lipofectamine LTX Reagent with PLUS Reagent (Invitrogen), followed by cell culture at 32°C for 5 days. Then, the cells were transfected again with plasmids encoding SeV M, F, and HN proteins and

cultured at 32°C for more than 6 days to produce SeVdp vector. The supernatant containing the SeVdp vector was collected and filtered with a 0.45 μm cellulose acetate filter, frozen in liquid nitrogen, and stored at −80°C until use. Titers of the SeVdp vectors were determined by the number of EGFP positive cells (in case of EGFP-expressing vector) or by immunostaining using an anti-SeV NP antibody (Nishimura et al., 2007).

Induction of Retroviral Silencing by Reprogramming

MEFs were transduced with retrovirus or lentivirus particle expressing hKO in the presence of 8 μg/mL hexadimethrine bromide (polybrene; Sigma) followed by a selection with 2 μg/mL puromycin (Nacalai) or 10 μg/mL blasticidin (Wako) for 4 days. The selected MEFs were infected with SeVdp vector at 32°C for 1 day to induce a somatic cell reprogramming. One day after SeVdp vector infection, the MEFs were seeded onto mitomycin C-treated SNL 76/7 feeder cells and maintained at 37°C in mES2 medium (Stemsure DMEM (Wako) supplemented with 15% StemSure Serum Replacement (Wako), 2 mM GlutaMAX (Invitrogen), 0.1 mM non-essential amino acids (Nacalai), 0.055 mM 2-mercaptoethanol (Invitrogen), 100 U/mL penicillin-streptomycin, and 1,000 U/mL Leukemia Inhibitory Factor (LIF) (Wako)) for 6 days, and then in mES1 medium (DMEM (Nacalai) supplemented with 15% fetal bovine serum (FBS) (GIBCO), 0.055 mM 2-mercaptoethanol, 0.1 mM non-essential amino acids, 100 U/mL penicillin-streptomycin and 1,000 U/mL LIF). In case of a silencing induction with short-term culture after retroviral transduction, SeVdp vector was infected one day after retroviral transduction, and, after the passage onto feeder cells, the reprogramming cells were cultivated with puromycin in first two days.

For silencing induction in human cells, TIG-3 human fibroblast was first transduced with lentivirus particle expressing mouse Slc7a1 protein produced from pLenti6/UbC/mSlc7a1 (Takahashi et al., 2007) (Addgene) to support MLV infection to human cells. The TIG-3/Slc7a1 cells were then transduced with MLV(YY-PBS-hKO) with Puromycin selection for 4 days. SeVdp(GKOSM) vector was infected to the cells at room temperature for two hours then 37°C for a day followed by passage on to iMatrix-511 (TaKaRa) coated plate and cultivation in StemFit AK02N (TaKaRa) without solution C.

When candidate genes were knocked down before reprogramming, cells transduced with MLV(YY-PBS-hKO) were transfected with siRNA against each factor by Lipofectamine RNAiMAX Transfection Reagent (Invitrogen). Two days after the transfection, cells were infected with SeVdp vector to induce retroviral silencing. The cells treated again with the same siRNA 4 days after SeVdp infection. The efficiency of siRNA-mediated knockdown was analyzed in NIH 3T3 cells 3 days after the siRNA transfection by RT-qPCR with primers specific for each gene. Target sequences of siRNAs for each gene are listed in Table S5. To test the effect of overexpression of TAF-Iα or TAF-Iβ retroviral in the silencing, hKO-expressing MEFs were transduced with MLV(TAF-Iα-Bs) or MLV(TAF-Iβ-Bs), respectively.

To observe retroviral silencing in ECCs, F9 cells were transduced with hKO-expressing retrovirus particle in the presence of 8 μg/mL polybrene overnight. The selection and reprogramming were performed as described above.

Fluorescence-activated Cell Sorting (FACS) Analysis

Cells were collected and re-suspended in 500 μl of phosphate buffered saline (PBS) supplemented with 2% FBS. At least 1.0×10^5 viable cells were used for each analysis by BD FACSVerse cytometer (BD Bioscience) using BD FACSuit software (BD Bioscience).

Characterization of Reprogrammed Cells

Total RNA was extracted using ISOGEN (Nippon Gene), and reverse transcription was performed using Superscript III First-Strand Synthesis System (Thermo Fisher). qPCR analyses were performed using 7500 Fast Real-time PCR System (Applied Biosystems) with GoTaq qPCR Master Mix (Promega). As a control of RT-qPCR, RNA from MEF or mouse iPSCs (miPSCs) generated by the SeVdp(KOSM) (Nishimura et al., 2011) was used. The expression levels were normalized against that of TATA-box binding protein (TBP). The DNA sequences of the primers used for the RT-qPCR are listed in Table S4. Detection of alkaline phosphatase was carried out using Alkaline Phosphatase Staining Kit II (Stemgent) according to manufacturer's instructions.

Collection of Whole Genomic Gene Expression Profile

Total RNAs were extracted from MEF, miPSCs and MEF infected with SeVdp(KOSM) or SeVdp(GKOSM) two days after the infection. The RNAs were labeled using Quick Amp Labeling Kit (Agilent) followed by hybridization to Whole Mouse Genome DNA microarray 4x44K (Agilent) using Gene Expression Hybridization Kit (Agilent) according to manufacturer's instructions. The array was scanned by DNA microarray scan system (Agilent) and the scanned images were analyzed with Feature Extraction software (Agilent). The data of normalized signal intensity was analyzed by Excel software to obtain whole genomic gene expression profile.

Insertional Chromatin Immunoprecipitation (iChIP)

MEFs were first transduced with MLV(3xFNLDD) and selected with puromycin for 2 days. The selected cells were then subjected to another transduction with MLV(YY-PBS-Lex2-hKO). Because MLV(YY-PBS-Lex2-hKO) lacks a drug resistant gene, high infectivity (> 50%) of the virus as confirmed by observation of hKO expression. Two days after MLV(YY-PBS-Lex2-hKO) transduction, the cells were reprogrammed by SeVdp(GKOSMaB) and cultured onto feeder cells. At day 7 of reprogramming, blasticidin selection was performed to select only reprogrammed cells, thereby reducing the proteins precipitated from unreprogrammed cells in the samples for Liquid Chromatography/Mass Spectrometry (LC/MS/MS).

The reprogrammed cells were cross-linked for 10 min at room temperature with 0.75% formaldehyde, followed by 5-min quenching with 125 mM glycine, and then washed three times with ice-cold PBS. Collected cells were lysed in FA lysis buffer (50 mM HEPES-KOH, pH7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS, and Protease Inhibitor Cocktail (Sigma)) and sonicated to shear the chromatin using Sonifier S-250 (Branson). The sheared chromatin was cleared by centrifugation at 15,000 rpm for 5 min at 4°C, and the supernatant was used as a chromatin suspension for subsequent immunoprecipitation. The chromatin was diluted 10-fold with Dilution buffer (16.7 mM Tris-HCl (pH8.0), 167 mM NaCl, 1.2 mM EDTA-NaOH (pH8.0), 1.1% Triton X-100, 0.01% SDS). The diluted chromatin was incubated with anti-DYKDDDDK tag antibody (FLAG-tag) (018-22381; Wako) or mouse normal IgG (sc-2025; Santa Cruz) overnight at 4°C, followed by incubation with Dynabeads Protein G (Invitrogen) for 6 h at 4°C. The antibody-chromatin mixture was collected by a magnetic stand and washed twice with 500 μL of ice-cold Wash Buffer (20 mM Tris-HCl (pH8.0), 150 mM NaCl, 2 mM EDTA-NaOH (pH8.0), 1% Triton X-100, 0.1% SDS), once with 500 μL of ice-cold Final Wash Buffer (20 mM Tris, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), and then with 500 μL of ice-cold LiCl Wash Buffer (250 mM LiCl, 10 mM Tris-HCl, 1 mM EDTA, 1% sodium deoxycholate and 1% NP-40). The beads were subjected to additional wash with 1,000 μL of ice-cold TBS (50 mM Tris-HCl (pH 7.5), 150 mM NaCl) with 0.1% NP-40 twice. After addition of 500 μl of FLAG elution buffer (0.5 mg/mL 3xFLAG peptide (Protein Ark) in TBS with 0.1% NP-40), the precipitated beads were incubated at 37°C for 20 min. The elution step was repeated twice, and the eluted proteins were precipitated overnight by 2-propanol at −30°C. The precipitated protein sample was digested overnight by trypsin (Promega) at 37°C, followed by purification using SPE C-TIP-T300 (Wako). Peptide analysis was performed with Liquid Chromatography by Nano-Advance (BRUKER), followed by Mass Spectrometry by Q Exactive Plus (Thermo Fisher). The MS/MS spectra of the peptides were interpreted using Mascot server (MATRIX Science).

For DNA precipitation, the washed beads were subjected to elution with Elution buffer (10 mM Tris, 1 mM EDTA, 1% SDS) for 30 min at 30°C. The eluted chromatin was reverse cross-linked by overnight incubation at 65°C, followed by Proteinase K treatment at 45°C for 2 h. The extracted DNA was analyzed by qPCR. The DNA sequences of the primers for the ChIP-qPCR are listed in Table S4.

ChIP

For confirmation of FLAG-tagged TAF-Iα binding to retroviral sequence, MEFs transduced with MLV(YY-PBS-Lex2-hKO) and MLV(TAF-Iα-Bs) were cultivated for 10 days with or without reprogramming by SeVdp(GKOSM). The cells were cross-linked and sonicated to collect chromatin suspension as described above. The chromatin suspensions were incubated with anti-DYKDDDDK tag antibody or mouse normal IgG overnight at 4°C, followed by incubation with Dynabeads Protein G for 6 h at 4°C. The antibody-chromatin mixture was collected and washed as described above. The washed beads were treated with Elution buffer as described in DNA precipitation in iChIP. The extracted DNA was analyzed by qPCR using primers listed in Table S4.

Immunofluorescence Staining

Cells were fixed with 3.7% formaldehyde in PBS and permeabilized by 0.1% Triton X-100 in PBS. The cells were stained with anti-DYKDDDDK tag, anti-SSEA-1 (sc-21702; Santa Cruz) and/or SeV NP antibody followed by anti-mouse IgG conjugated with AlexaFlour 555 (1:500, A32727; Thermo Fisher), or AlexaFlour 647 (1:500, A28181; Thermo Fisher), or anti-mouse IgM conjugated with AlexaFlour 555 (1:500, A21426; Thermo Fisher). Nuclei were counterstained with DAPI using Fluoro-KEEPER Antifade Reagent (Nacalai).

Determination of Protein Expression by Immunoblotting

NIH 3T3 cells treat with siRNA were collected two days after the treatment and lysed on ice in Cell lysis buffer (10 mM Tris-HCl (pH8.0), 0.3 M NaCl, 1 mM EDTA-NaOH (pF8.0), 0.5% NP-40), followed by incubation for 30 min at 4°C and centrifugation by 15,000 rpm for 5 min, and the supernatant were subjected to western blot analysis as a soluble protein fraction. Protein expression levels were quantified using FUSION FX7.EDGE and FUSION Capt Software (Vilber-Lourmat). We employed the following primary antibodies; anti-KLF4 purified from immunized rabbit serum (1:5,000) (Nishimura et al., 2014), anti-TAF-I (1:250, KM1725) (Nagata et al., 1998), anti-ANP32A (1:200, sc-374552; Santa Cruz), anti-HMG2 (1:200, sc-271689; Santa Cruz), anti-EBP-1 (1:200, sc-393114; Santa Cruz), and, anti-α-TUBULIN (1:10,000, ab7291; Abcam).

QUANTIFICATION AND STATISTICAL ANALYSIS

Experiments were repeated at least three times, and unpaired 2-tailed Student's t tests by Excel software were used to determine statistically significant difference between two groups. Data are represented as means ± SEM. A value of $p < 0.05$ was considered statistically significant.

DATA AND CODE AVAILABILITY

The accession number for the microarray datasets reported in this paper is NCBI GEO: GSE134847.