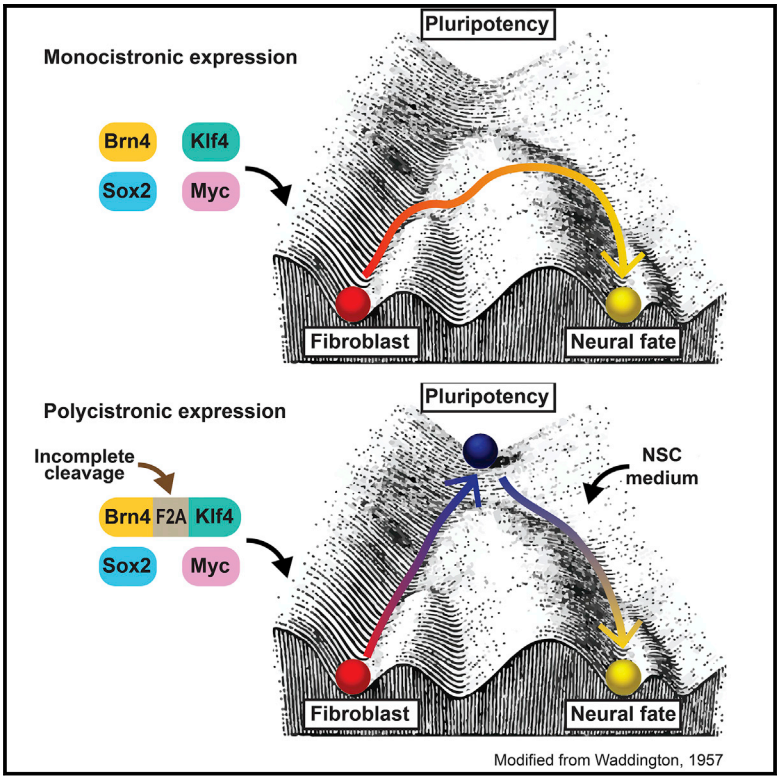
Cell Reports

Fusion of Reprogramming Factors Alters the Trajectory of Somatic Lineage Conversion

Graphical Abstract



Highlights

- Individually expressed BKSM generate iNSCs without a transient pluripotent state
- Polycistronic BKSM expression produces a Brn4-Klf4 fusion protein
- The Brn4-Klf4 fusion protein gains the ability to induce pluripotency
- A shift in Brn4 binding from *OctOct* to *SoxOct* motif enables induction of pluripotency

Authors

Sergiy Velychko, Kyuree Kang, Sung Min Kim, ..., Guangming Wu, Hans R. Schöler, Dong Wook Han

Correspondence

office@mpi-muenster.mpg.de (H.R.S.), dwhan@konkuk.ac.kr (D.W.H.)

In Brief

The pluripotency- and neuro-specific factors Brn4, Klf4, Sox2, and cMyc (BKSM), when expressed individually, directly transdifferentiate fibroblasts into neutral stem cells (iNSCs). Velychko et al. show that polycistronic expression produces a fusion protein that induces pluripotency rather than direct transdifferentiation.

Velychko et al., 2019, Cell Reports *27*, 30–39
April 2, 2019
https://doi.org/10.1016/j.celrep.2019.03.023

CellPress

OPEN ACCESS
CellPress

Cell Reports
Report

Fusion of Reprogramming Factors Alters the Trajectory of Somatic Lineage Conversion

Sergiy Velychko,[1,8] Kyuree Kang,[2,8] Sung Min Kim,[2] Tae Hwan Kwak,[2] Kee-Pyo Kim,[1] Chanhyeok Park,[3] Kwonho Hong,[3] ChiHye Chung,[4] Jung Keun Hyun,[5] Caitlin M. MacCarthy,[1] Guangming Wu,[1] Hans R. Schöler,[1,2,9,*] and Dong Wook Han[2,6,7,*]

[1]Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, 48149 Münster, Germany
[2]Department of Stem Cell Biology, School of Medicine, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea
[3]Department of Stem Cell and Regenerative Biotechnology, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea
[4]Department of Biological Sciences, Konkuk University, Seoul 05029, Republic of Korea
[5]Department of Nanobiomedical Science, Dankook University, Cheonan 330714, Republic of Korea
[6]KU Open-Innovation Center, Institute of Biomedical Science and Technology, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea
[7]Department of Advanced Translational Medicine, School of Medicine, Konkuk University, 120 Neungdong-ro, Gwangjin-gu, Seoul 05029, Republic of Korea
[8]These authors contributed equally
[9]Lead Contact
*Correspondence: office@mpi-muenster.mpg.de (H.R.S.), dwhan@konkuk.ac.kr (D.W.H.)
https://doi.org/10.1016/j.celrep.2019.03.023

SUMMARY

Simultaneous expression of Oct4, Klf4, Sox2, and cMyc induces pluripotency in somatic cells (iPSCs). Replacing Oct4 with the neuro-specific factor Brn4 leads to transdifferentiation of fibroblasts into induced neural stem cells (iNSCs). However, Brn4 was recently found to induce transient acquisition of pluripotency before establishing the neural fate. We employed genetic lineage tracing and found that induction of iNSCs with individual vectors leads to direct lineage conversion. In contrast, polycistronic expression produces a Brn4-Klf4 fusion protein that enables induction of pluripotency. Our study demonstrates that a combination of pluripotency and tissue-specific factors allows direct somatic cell transdifferentiation, bypassing the acquisition of a pluripotent state. This result has major implications for lineage conversion technologies, which hold potential for providing a safer alternative to iPSCs for clinical application both *in vitro* and *in vivo*.

INTRODUCTION

Forced expression of Oct4, Klf4, Sox2, and cMyc (OKSM) can convert fibroblasts into induced pluripotent stem cells (iPSCs) (Takahashi and Yamanaka, 2006). Although iPSCs are widely used for the derivation of different cells types, their differentiation often involves lengthy and inefficient protocols. Moreover, the tumorigenicity of the residual pluripotent cells presents a major risk for clinical application of these cells.

To overcome these limitations, numerous strategies for direct transdifferentiation of somatic cells into mature target cell types have recently been developed (Buganim et al., 2012; Huang et al., 2011; Ieda et al., 2010; Li et al., 2014; Margariti et al., 2012; Vierbuchen et al., 2010). Advances in direct lineage conversion have also led to the successful generation of self-renewing multipotent somatic stem cells; e.g., neural, hematopoietic, angioblast, hepatic, and cardiovascular progenitor or stem cells (Han et al., 2012; Kim et al., 2011a; Kurian et al., 2013; Lalit et al., 2016; Yu et al., 2013). Direct conversion of somatic cells into somatic stem cells has great advantages. First, the technology could potentially facilitate the derivation of specific mature cell types using faster and less expensive protocols compared with iPSC technology; second, direct lineage reprogramming of somatic cells into somatic stem cells could potentially be employed to repair aged or diseased tissues *in vivo* (Niu et al., 2013; Tapia et al., 2012).

We have demonstrated that the neuro-specific POU factor Brn4, in combination with Klf4, Sox2, and cMyc (BKSM), can convert fibroblasts into induced neural stem cells (iNSCs) but not iPSCs. Others have shown that use of the OKSM cocktail with time-restricted expression of Oct4, combined with specific culture conditions, can also generate iNSCs (Thier et al., 2012). However, two recent studies have proposed that conversion of somatic cells into iNSCs with either OKSM (Bar-Nur et al., 2015; Maza et al., 2015) or BKSM (Bar-Nur et al., 2015) is an indirect process that involves a transiently acquired pluripotent state that is subsequently converted into a specific cell type based on the culture conditions. These contradictory findings encouraged us to further dissect the process of directly converting fibroblasts into iNSCs.

RESULTS

A Lineage Tracing System Reveals Rare Direct Conversion of Fibroblasts into iNSCs

Several reasons might account for the discrepancies between the reports. First, in our previous studies, we did not use genetic lineage tracing and, therefore, could have missed a short-lived pluripotent state over the course of the conversion (Han et al.,

2012; Kim et al., 2014). Second, we used monocistronic retroviral vectors and not the polycistronic lentiviral vector (Bar-Nur et al., 2015; Sommer et al., 2009). We therefore decided to compare both gene delivery systems using genetic lineage tracing to check for endogenous Oct4 expression. In our first set of experiments, we used mouse embryonic fibroblasts (MEFs) carrying a tamoxifen-inducible Oct4-CreER allele in combination with a Rosa26-loxSTOPlox-diphtheria toxin fragment A allele (R26-lsl-diphtheria toxin A [DTA]) (Figure S1A). In this lineage tracing system, DTA selectively kills all cells that activate Oct4, eliminating cells that even transiently pass through a pluripotent state. We first used the polycistronic lentiviral vector encoding BKSM (pcBKSM) under the control of a doxycycline (dox)-inducible promoter (Figure S1A). The first NSC-like clusters appeared after 3–4 weeks of dox withdrawal (Figure S1B) (1–2 clusters from 5×10^4 transduced MEFs) whereas none appeared with control MEFs. The expanded clonal iNSCs (hereafter named DTA-iNSCs) expressed NSC markers (Figures S1C and S1D) and possessed pcBKSM, Oct4-CreER, and R26-lsl-DTA transgenes (Figure S1E), suggesting that the rare surviving DTA-iNSCs had not passed through a pluripotent state.

RNA sequencing (RNA-seq) showed that the gene expression pattern of DTA-iNSCs is similar to that of control NSCs (cNSC) isolated from embryonic day 16.5 fetal forebrain and distinct from MEFs (Figure S1F). DTA-iNSCs could differentiate into neurons, astrocytes, and oligodendrocytes *in vitro* (Figure S1G) and *in vivo* when transplanted into the cortical region of the rat brain (Figure S1H). We found no evidence of tumor formation in any of the transplanted rats (data not shown).

The Polycistronic BKSM Vector Induces a Transient iPSC-like State

To quantify the number of iNSCs generated by direct or indirect reprogramming, we used a lineage tracing system carrying R26-lsl-enhanced yellow fluorescent protein (EYFP) instead of DTA, which allowed us to discriminate between directly and indirectly generated iNSCs (Greder et al., 2012; Figure 1A). After 2 weeks of pcOKSM or pcBKSM induction, we observed the first cell clusters (6–8 clusters from 5×10^4 transduced MEFs) with typical NSC morphology (Figure 1B), which were expanded into stable lines (Figures 1C and 1D). As determined by fluorescence-activated cell sorting (FACS), the vast majority of both pcOKSM and pcBKSM iNSCs were $EYFP^+/Olig2^+$ (Figure 1E), supporting previous studies that showed that iNSCs are derived through an $Oct4^+$ intermediate stage (Bar-Nur et al., 2015; Maza et al., 2015). The identity of iNSCs was confirmed by RNA-seq (Figure 1F).

Notably, a small $EYFP^-$ iNSC population was consistently observed under both reprogramming conditions (Figure 1E). We sorted these $EYFP^-$ cells for further analysis. $EYFP^-$ iNSCs obtained with either pcOKSM or pcBKSM showed typical NSC morphology (Figure S2A) and expressed NSC marker genes (Figures 1G and 1H), similar to $EYFP^+$ iNSCs. Therefore, a small number of both pcOKSM and pcBKSM iNSCs might have directly transdifferentiated from MEFs without passing through a pluripotent state. However, we cannot exclude that the rare DTA and $EYFP^-$ iNSCs could represent false negative events resulting from incomplete excision by Cre.

Individually Transduced BKSM Vectors Cannot Induce Pluripotency

We next repeated the lineage-tracing experiment using monocistronic retroviral vectors encoding for Brn4, Klf4, Sox2, and cMyc (mcBKSM), as we described previously (Han et al., 2012; Kim et al., 2014).

The iNSC-like clusters appeared 2–3 weeks after infection (~10 clusters from 5×10^4 transduced MEFs) (Figures S2B and S2C) and could be expanded into stable iNSC lines (Figure S2D). The iNSCs had a morphology, gene expression, and promoter methylation status similar to cNSCs but distinct from starting MEFs (Figures 2A–2C). Retroviral transgenes have been shown previously to be silenced in fully reprogrammed iPSCs (Stadtfeld et al., 2008) and in iNSCs generated by pcOKSM- or pcBKSM-driven reprogramming (Bar-Nur et al., 2015). Notably, early-passage mcBKSM iNSCs had largely active retroviral transgenes (Figure 2D). This result confirms our previous report showing that retroviral transgenes are silenced stochastically in established iNSC lines only after 2 months of culture (Kim et al., 2014).

Surprisingly, FACS detected no $EYFP^+$ cells in either clonal iNSC lines (Figure 2E) or in the bulk cell population after 25 days of reprogramming (Figure 2F), suggesting that all mcBKSM-iNSCs were reprogrammed directly. To rule out the possibility that the contaminating neural cell population could be enriched by induction with reprogramming factors and subsequently contribute to iNSC generation, we repeated the experiment with MEFs pre-sorted for the fibroblast marker Thy1 (Figure S2E). We found that iNSCs could be generated from both $Thy1^+$ and $Thy1^-$ MEFs (Figure S2F), although reprogramming of $Thy1^+$ cells was less efficient (Figure S2G). mcBKSM-derived iNSCs established from both $Thy1^+$ and $Thy1^-$ MEFs expressed NSC markers (Figure S2H) and were $EYFP^-$ (Figure S2I).

Time Course Gene Expression Reveals Distinct Reprogramming Paths into iNSCs

To elucidate the molecular responses mediated by lentiviral pcBKSM versus retroviral mcBKSM reprogramming, we compared the time course expression of selected genes. We achieved comparable levels of induction for both systems by titrating different concentrations of dox or different volumes of concentrated retroviral soups and confirmed similar expression levels for both delivery systems using a qPCR titration kit (Figure S3A) and western blotting (Figure S3B).

The fibroblast markers were rapidly suppressed in MEFs after 1–2 weeks of induction in both systems (Figure S3C). mcBKSM did not activate endogenous pluripotency markers, whereas pcBKSM and mcOKSM or pcOKSM strongly upregulated those genes (Figure 3A). On the other hand, pcBKSM- and mcOKSM- or pcOKSM-reprogrammed cells exhibited delayed activation of NSC markers such as *Blbp* and *Olig2* compared with mcBKSM (Figures 3B and S3D). Similar trends were observed for lentiviral mcBKSM (Figure S3E), confirming that the differences arise from the polycistronic expression, not the virus type.

Structural Dissection of Oct4 and Brn4 Reveals an Inability of Brn4 to Induce Pluripotency

Oct4 has been shown to be the only reprogramming factor in the reprogramming cocktail that could not be replaced by other

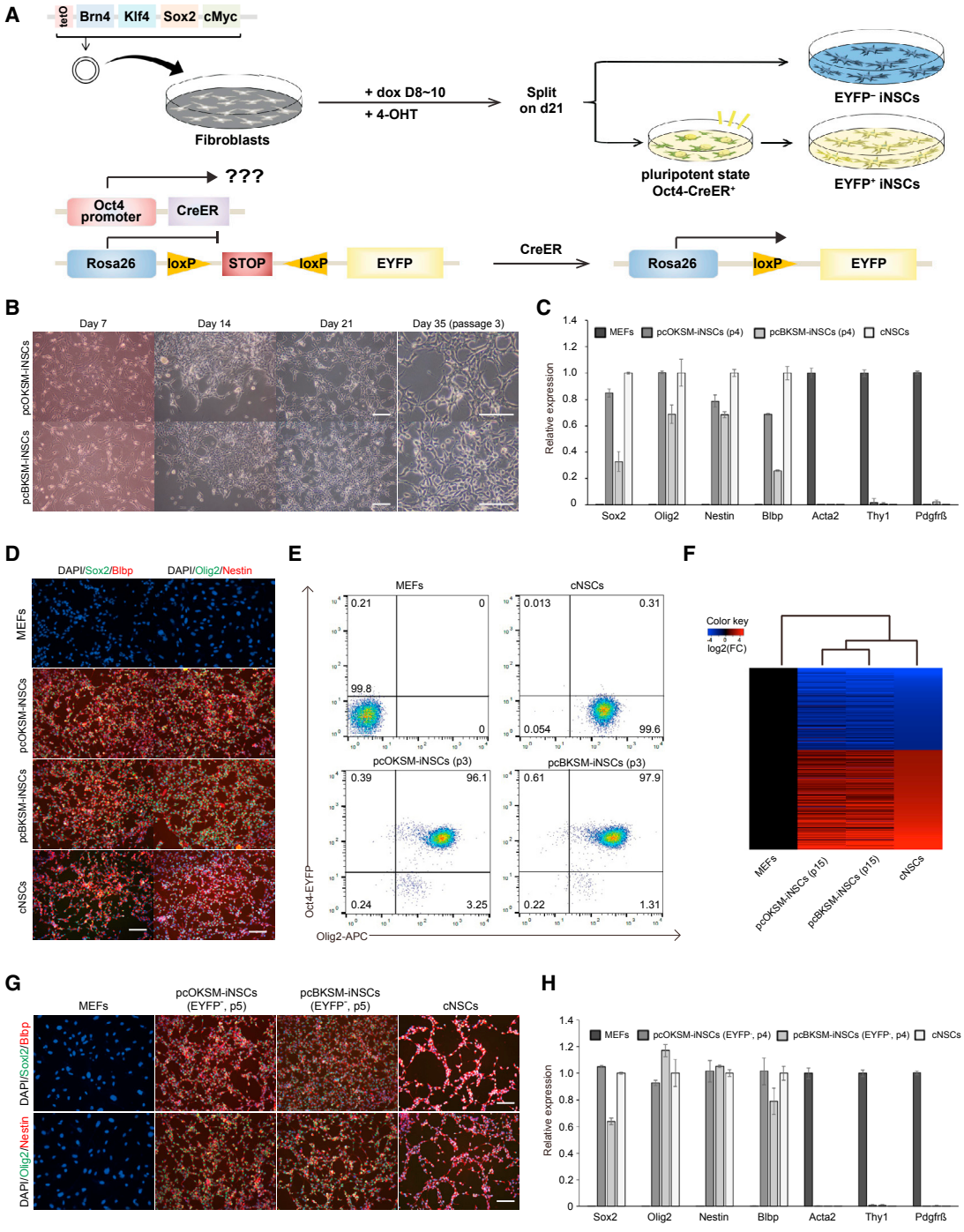


Figure 1. The Polycistronic BKSM Vector Induces Transient Activation of Oct4

(A) Schematic diagram of the Oct4-CreER; R26-lsl-EYFP lineage tracing system.

(B) Morphology of pcOKSM and pcBKSM iNSCs. Scale bar, 50 μm.

(C) qPCR analysis of NSC and fibroblast marker expression in pcOKSM and pcBKSM iNSCs (mean ± SD, n = 3).

(D) Immunofluorescence of pcOKSM and pcBKSM iNSC lines. Scale bar, 50 μm.

(E) FACS analysis of pcOKSM and pcBKSM iNSC lines after 3 passages.

(F) Heatmap of gene expression for pcOKSM and pcBKSM iNSC lines. Hierarchical clustering based on the gene expression profiles is shown at the top. Differentially expressed (DE) genes with fold change (FC) ≥ 4 between MEFs and cNSCs and fragments per kilobase of transcript per million (FPKM) ≥ 1 in at least one of the samples were plotted.

(G) Immunofluorescence of EYFP⁻ pcOKSM and pcBKSM iNSCs. Scale bar, 50 μm.

(H) qPCR analysis of the indicated marker gene expression in EYFP⁻ iNSCs (mean ± SD, n = 3).

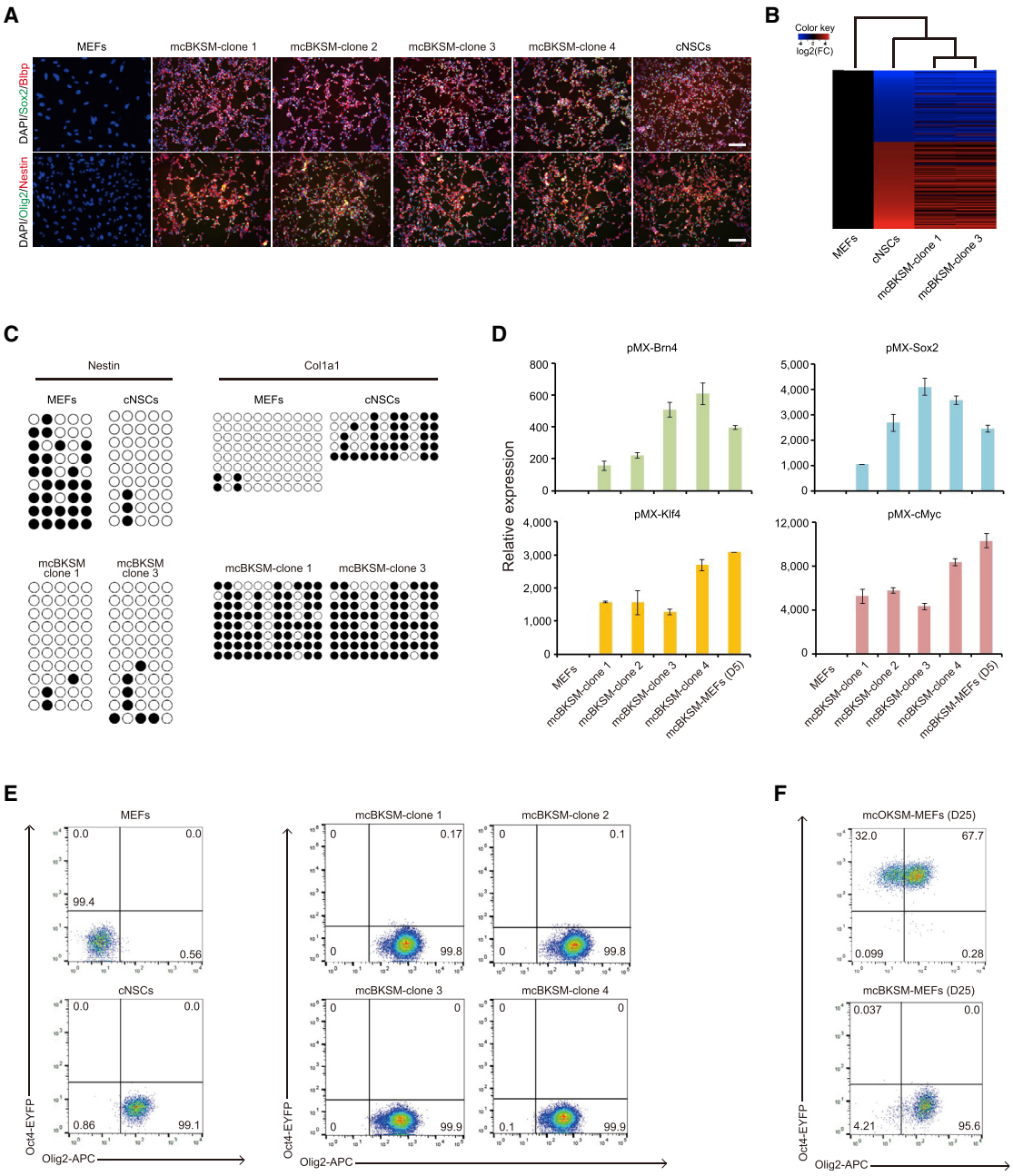


Figure 2. mcBKSM Does Not Activate Endogenous Oct4

(A) Immunofluorescence microscopy images of mcBKSM iNSCs. Scale bar, 50 μm.

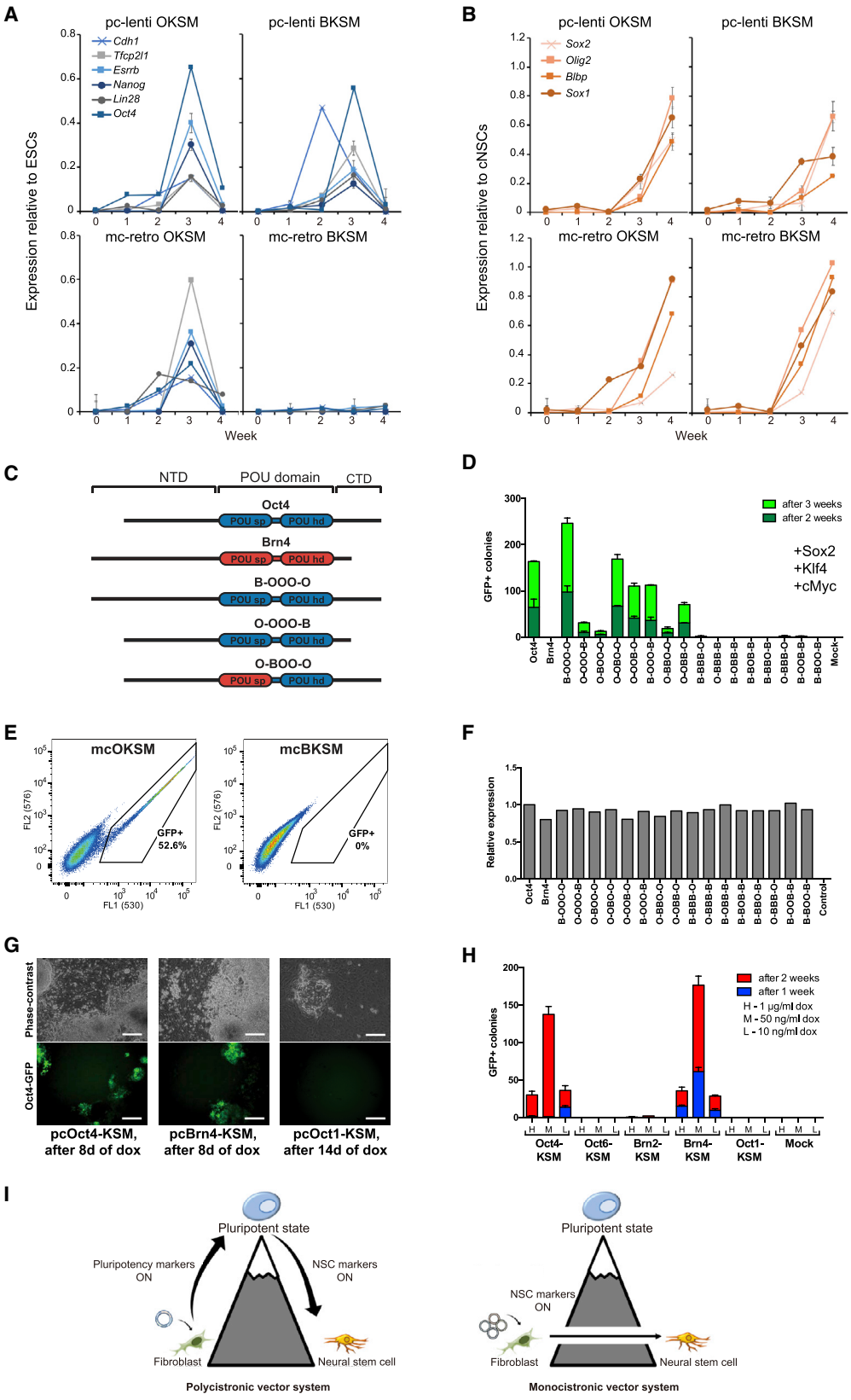
(B) Hierarchical clustering heatmap of global gene expression for two clonal mcBKSM iNSC lines. DE genes with FC ≥ 4 between MEFs and cNSCs and FPKM ≥ 1 in at least one of the samples were plotted.

(C) DNA methylation analysis of the second intron of *Nestin* and the promoter region of *Col1a1* in two mcBKSM iNSC lines, assessed by bisulfite sequencing PCR. Open and filled circles represent unmethylated and methylated CpGs, respectively.

(D) qPCR analysis of retroviral transgene expression in mcBKSM iNSC lines (passage 5) and mcBKSM-transduced MEFs (5 days post infection [dpi]). The levels were normalized to non-transduced MEFs (mean ± SD, n = 3).

(E and F) FACS analysis of mcBKSM iNSC lines (E) and the bulk population of mcOKSM- or mcBKSM-transduced MEFs 25 dpi (F).

members of the same transcription factor (TF) family (Nakagawa et al., 2008). However, Bar-Nur et al. (2015) showed that pcBKSM, carrying Brn4 instead of Oct4, could convert fibroblasts into chimera-grade iPSCs. Both Oct4 (*Pou5f1*) and Brn4 (*Pou3f4*) belong to the POU TFs (group IV and III, respectively) and share a conserved DNA-binding domain (POU domain)



(legend on next page)

flanked by N- and C-terminal transactivation domains (NTD and CTD, respectively). The POU domain itself consists of a POU-specific (POUsp) domain and a POU homeodomain (POUhd) connected by a non-conserved linker (Rosenfeld, 1991). To determine which parts of Oct4 and Brn4 are responsible for reprogramming to pluripotency, we generated a series of reciprocal Oct4-Brn4 chimeras by swapping their domains (Figure 3C). We tested for the ability of the chimeras to generate iPSCs by introducing them into Oct4-GFP MEFs together with Sox2, Klf4, and Myc (SKM) using individual retroviruses.

Unlike Oct4, wild-type Brn4 was unable to generate GFP$^+$ iPSC colonies (Figures 3D and 3E). The chimeras containing a single domain substitution from Brn4 (NTD, POUsp, linker, POUhd, or CTD) still maintained reprogramming function, albeit with decreased efficiency for POUsp, POUhd, and CTD. Notably, the POUsp domain appeared to be the most crucial for the reprogramming function of Oct4 because its replacement with the Brn4 counterpart (O-BOO-O chimera) decreased the TF's reprogramming ability by 30-fold. qPCR titration of the retroviral vectors confirmed equal transgene expression (Figure 3F). Taken together, our results demonstrate that Brn4 fails to induce pluripotency because of the functional inadequacy of its three domains, with the POUsp domain playing the most critical role.

Reprogramming Oct4-GFP MEFs to iPSCs using polycistronic vectors resulted in GFP$^+$ colonies after just 8 days of induction with either pcOKSM or pcBKSM (Figure 3G), with both displaying surprisingly similar reprogramming efficiencies (Figure 3H). Other POU factors cloned in the same polycistronic cassette, such as Oct1 and Oct6, were unable to generate iPSCs, whereas Brn2 gave rise to very few GFP$^+$ colonies (Figure 3H). Strikingly, the polycistronic vector can drive fibroblasts into a pluripotent state, whereas individually expressed factors mediate direct transdifferentiation of MEFs into iNSCs but cannot induce pluripotency (Figure 3I).

The Brn4-Klf4 Fusion Protein Gains the Function of Reprogramming to Pluripotency

The polycistronic reprogramming cassette (known as STEMCCA [single lentiviral stem cell cassette]) from Bar-Nur et al. (2015) and our current study carries TFs separated with two self-cleaving 2A peptides and an internal ribosome entry site (POU-F2A-Klf4-internal ribosome entry site [IRES]-Sox2-E2A-cMyc) (Figure 4A). F2A, which was used in both the original pcOKSM (Sommer et al., 2009) and in pcBKSM (Bar-Nur et al., 2015), displays the lowest self-cleaving efficiency among 2A peptides (Kim et al., 2011b). This suggests that Brn4-F2A-Klf4 can produce an unprocessed high-molecular-weight polyprotein that might alter the cellular response and change the reprogramming outcome. To test this idea, we generated a series of POU-KSM vectors where we replaced the F2A with the most efficient self-cleaving P2A or a non-cleavable mutant F2A (F2Am) obtained by replacing two residues of the self-cleavage site (GP → AA) (Figure 4A). Additionally, we used a synthetic polyglycine linker (GL) to exclude any effects specific to the F2A sequence. Western blot analysis confirmed full cleavage of POU and Klf4 by P2A, partial cleavage by F2A, or the absence of cleavage by F2Am (Figure 4B).

As expected, among the P2A constructs, only Oct4 but not other POU factors could generate iPSCs (Figure S4A). Surprisingly, both non-cleaving linkers, F2Am and GL, allowed reprogramming not only with Brn4 but also with Oct6 and Brn2 (Figure S4A), both of which belong to neuro-specific POU3 factors. Brn2-, Oct6-, and Brn4-F2Am-KSM iPSCs could all generate clonal cell lines (Figures 4C and 4D) that expressed the pluripotency markers Nanog and SSEA-1 (Figure S4B), differentiated into derivatives of all three germ layers in the teratoma assay (Figure S4C), and also effectively contributed to the germline of chimeric mice (Figure S4D).

Notably, Brn4-F2Am-KSM and Brn4-GL-KSM performed worse than the original F2A construct (Figure S4A), suggesting that a mixture of cleaved and uncleaved Brn4-Klf4 could be beneficial for efficient reprogramming. We therefore decided to examine the functionality of both POU and Klf4 within the fusion protein. We performed a pluripotency-rescuing assay using the ZHBTc4 embryonic stem cell (ESC) line (Niwa et al., 2000), where Oct4 can be conditionally ablated by dox treatment (Figure S4E). Unexpectedly, P2A-, F2A-, F2Am-, and GL-linked Oct4-Klf4 constructs rescued Oct4-depleted ESCs with similar efficiency (Figure S4F). The rescued cells had similar morphology, survival, and proliferation ability (Figure S4G), and they could be passaged at least 6 times while maintaining comparable expression of pluripotency genes (e.g., *Nanog*, *Sox2*, and *Esrrb*) (Figure S4H), suggesting that the POU TF function remains intact within the POU-Klf4 fusion protein. We next employed a primed-to-naive ESC conversion assay (Guo et al., 2009) to test the functionality of Klf4 within the fusion protein using Gof18 mouse epiblast stem cells (Han et al., 2010). Both F2Am and GL constructs showed significantly impaired conversion rates of Gof18$^-$ epiblast stem cells into the Gof18$^+$ ground state compared with vectors containing cleavable peptides, as

Figure 3. Monocistronically Expressed Brn4 Cannot Induce Pluripotency

(A and B) Time-course qPCR analysis of pluripotency (A) and NSC markers (B) in MEFs transduced with mcOKSM or pcOKSM or BKSM. The levels were normalized to ESCs (A) and cNSCs (B) (mean ± SD, n = 3).

(C) Generation of a series of reciprocal Oct4-Brn4 chimeras by domain swapping. B and O stands for Brn4 and Oct4, respectively, and the position corresponds to the domain starting from the N terminus: NTD, POUsp, linker, POUhd, and CTD.

(D) Reprogramming of Oct4-GFP MEFs into iPSCs using Oct4, Brn4, or chimeric proteins in combination with SKM (mean of GFP$^+$ colony count ± SD, n = 3).

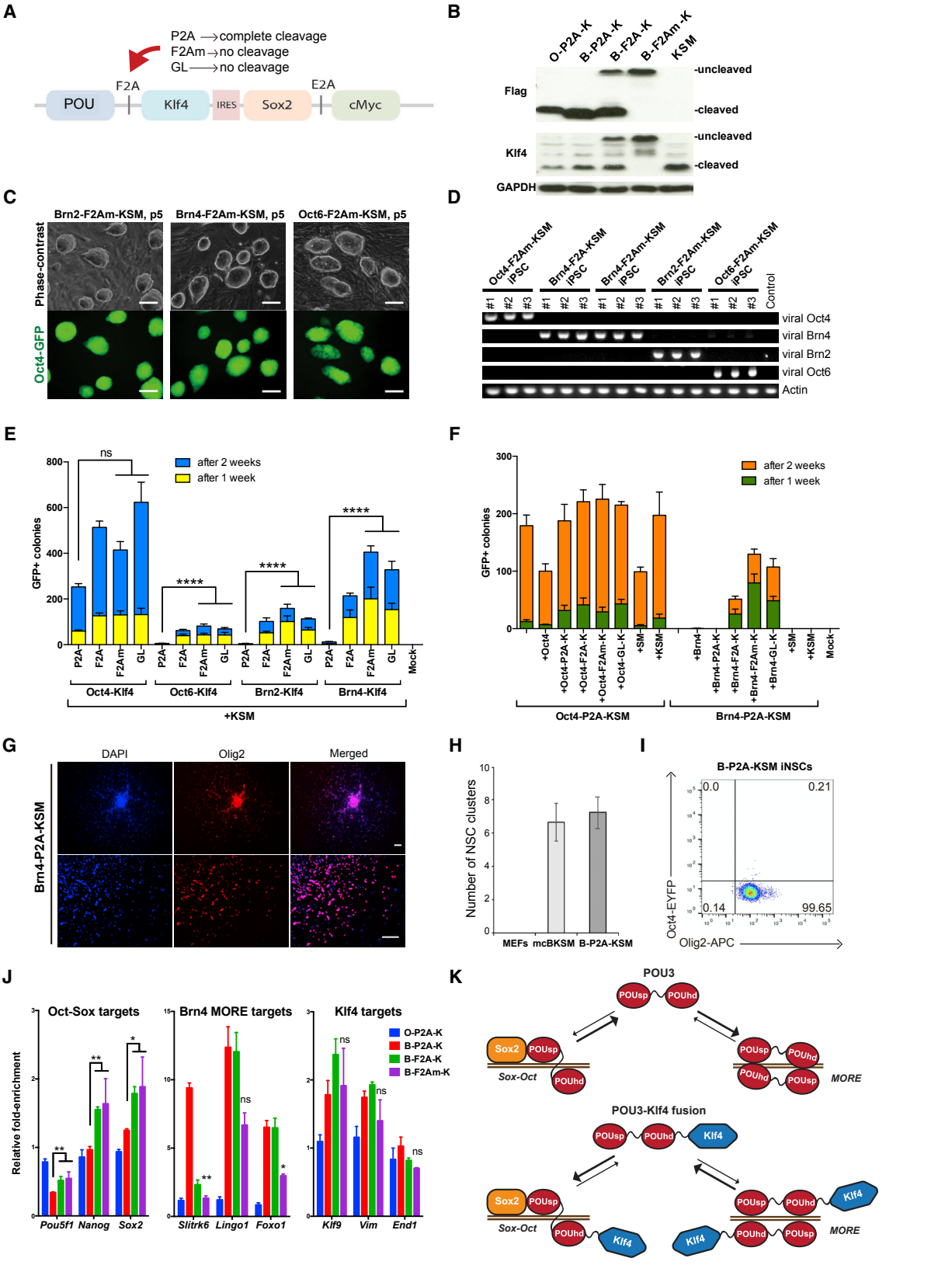
(E) Representative FACS analysis of the whole cell population after 3 weeks of reprogramming with mcOKSM or BKSM.

(F) qPCR titration of the retroviral vectors from (D).

(G) Representative images of iPSCs generated from Oct4-GFP MEFs. Scale bar, 250 μm.

(H) Reprogramming of Oct4-GFP MEFs with the pPOU-KSM construct using high (H), medium (M), and low (L) levels of dox induction (mean of GFP$^+$ colony count ± SD, n = 3).

(I) A scheme describing the divergent routes of somatic cells into iNSCs.



(legend on next page)

determined by FACS analysis (Figure S4I), indicating that Klf4 loses most but not all of its function when fused to the C terminus of a POU factor.

To compensate for the impaired Klf4, we repeated the reprogramming experiment using POU-Klf4 bicistronic constructs along with KSM. Again, although P2A abolished reprogramming of fibroblasts into iPSCs with POU factors other than Oct4, non-cleavable peptides rescued the reprogramming efficiency and performed better than the original F2A (Figure 4E). These results support the idea that Brn4 gains its reprogramming function in the polycistronic cassette by fusion with Klf4. In agreement with this observation, addition of non-cleavable Brn4-Klf4 constructs rescued the reprogramming ability of the Brn4-P2A-KSM cassette to a similar level as the Oct4-carrying constructs (Figures 4F and S4J). Moreover, the Brn4-P2A-Klf4 construct could generate EYFP^{-} iNSCs (Figures 4G and 4H), corroborating the ability of Brn4 to drive direct transdifferentiation into the neural fate when not fused with Klf4 (Figure 4I).

To understand how the Brn4-Klf4 fusion protein gained its reprogramming activity, we performed chromatin immunoprecipitation (ChIP)-qPCR for POU and Klf4 targets. We tagged the N terminus of Brn4 or Oct4 with 3xFLAG and confirmed the functionality of FLAG-tagged proteins in reprogramming experiments (Figure S4K). We next performed ChIP for the bulk of the reprogrammed cells after 4 days of dox induction. The Brn4-Klf4 fusion showed a modest but significant increase in occupancy of pluripotency-related Sox-Oct motif-containing loci (*Pou5f1*, *Nanog*, and *Sox2*) while reducing the occupancy of POU3-specific more palindromic Oct-1 responsive element (MORE)-containing targets (*Slitrk6*, *Lingo1*, and *Foxo1*) (Figure 4J). Consistent with our epiblast-to-naive ESC conversion results, prominent Klf4 targets (Chronis et al., 2017) were not affected by the impaired cleavage.

DISCUSSION

We show a striking discrepancy between iNSCs reprogrammed by mono- versus polycistronically transduced BKSM. Although polycistronic induction results mostly in acquisition of a transient Oct4^{+} state (Figure 1E), individually delivered vectors convert fibroblasts directly into fully functional iNSCs without passing through an intermediate pluripotent stage (Figure 2E).

We determined that the conflicting results of the expression vectors used arises from the incomplete cleavage of the POU-F2A-Klf4 polyprotein. When F2A was replaced with P2A, the ability of pcBKSM to reprogram MEFs into pluripotent cells was completely lost as cleavage was more complete (Figures 4E, 4F, and S4A), and iNSCs could be generated without an intermediate Oct4^{+} state (Figures 4G–4I). On the other hand, when F2A was replaced with cleavage-deficient linkers, not only Brn4 but also other tissue-specific POU3 factors, such as Brn2 and Oct6, were capable of reprogramming MEFs into iPSCs (Figures 4C–4F and S4A). Based on our ChIP-qPCR data, we propose that a POU-Klf4 fusion protein could enable activation of the pluripotency network by redistributing the occupancy of POU3 factors from their normally preferred Oct-Oct MORE motif (Mistri et al., 2015) onto the Sox-Oct motif (Figures 4J and 4K). Previously, we showed that an Oct6 point mutation that diminished homodimerization of Oct6 on the MORE motif converted the TF into a pluripotency inducer (Jerabek et al., 2017). The residue responsible for homodimerization of POU3 factors on the MORE motif, Met 151, is located at the very end of the POUhd. Fusion of the large Klf4 to the short CTD of the POU3 TF could disrupt homodimerization while leaving heterodimerization with Sox2 intact, the latter of which depends on a region within the POUsp domain (Reményi et al., 2003). It is possible that Klf4 plays additional roles in the fusion protein, such as facilitating protein-protein interactions otherwise inaccessible by POU3 factors.

A fusion of Oct4 and Klf4 has already been reported by others but deemed non-functional, although its properties were never tested (Carey et al., 2011). The stoichiometry of the reprogramming factors was thought to negatively affect the developmental potential of the OKSM iPSCs because none of them could generate all-iPSC mice (Carey et al., 2011; Stadtfeld et al., 2012). The OKSM cassette was also used for multiple mechanistic studies of reprogramming (Cheloufi et al., 2015; Chronis et al., 2017; Polo et al., 2012) and might even be used to generate human iPSCs for clinical applications. The results presented in this study demonstrate that the unprocessed POU-Klf4 fusion protein not only fully maintains POU functionality but also leads to unpredictable additional properties (Figures 4 and S4).

Figure 4. Gaining Pluripotency-Inducing Function by POU3-Klf4 Polyproteins

(A) A schematic of replacement of the F2A peptide linking the POU and Klf4 TFs in the STEMCCA construct.

(B) Western blot analysis, using an anti-FLAG antibody, of MEFs transduced with FLAG-tagged Oct4- and Brn4-Klf4 cassettes.

(C) Representative images of the clonal iPSCs generated with POU3-F2Am-KSM. Scale bar, 250 μm.

(D) PCR genotyping of established iPSC lines.

(E) Reprogramming of Oct4-GFP MEFs with tet-inducible POU-Klf4 constructs linked by different peptides in combination with KSM (mean GFP^{+} colony count ± SD, n = 3).

(F) Reprogramming of Oct4-GFP MEFs with fully cleavable Oct4- and Brn4-P2A-KSM polycistronic cassettes in combination with additional constructs (mean of GFP^{+} colony count ± SD, n = 3).

(G) Immunofluorescence of Brn4-P2A-Klf4 iNSCs generated from Thy1^{+} MEFs. Scale bar, 50 μm.

(H) Reprogramming of Thy1^{+} MEFs with pcBrn4-P2A-Klf4 or mcBKSM. The efficiencies are shown as number of iNSC clusters expressing Olig2 (mean ± SD, n = 3).

(I) FACS analysis of Brn4-P2A-Klf4-generated iNSCs.

(J) ChIP-qPCR of reprogrammed MEFs on day 4 of induction with 3xFLAG-POU-Klf4 in combination with KSM. Data are represented as fold enrichment relative to the Oct4 sample (mean ± SD of biological replicates).

(K) A schematic of the putative mechanism for the gain of pluripotency-inducing function by POU3-Klf4 polyproteins.

Statistical significance in (I) and (J) was measured using a Student's t test.

Therefore, caution must be taken when employing polycistronic reprogramming cassettes, and more research is needed to evaluate the effects of fusion proteins resulting from incomplete cleavage by 2A peptides. On the other hand, this surprising finding might help us to better understand how reprogramming factor cooperation can drive cell fate conversion.

STAR★METHODS

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

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mice
- METHOD DETAILS
  - Generation of iNSCs
  - Differentiation of iNSCs
  - Vector Construction
  - Reprogramming of MEFs into iPSCs
  - Characterization of iPSCs
  - Immunocytochemistry
  - Flow Cytometry Analysis
  - Gene Expression Analysis by Quantitative PCR
  - DNA Methylation Analysis
  - Chromatin Immunoprecipitation (ChIP)
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - RNA Sequencing Analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.celrep.2019.03.023.

ACKNOWLEDGMENTS

We thank Konrad Hochedlinger for providing the pcBKSM and OKSM constructs and Oct4-CreER x Rosa26-loxSTOPlox-DTA MEFs. We are grateful to Martin Stehling for FACS, Martina Sinn for teratoma processing, Ingrid Gelker for 293T cell culture, and Areti Malapetsas for editing the manuscript. S.V. is a member of the International Max Planck Research School and the Graduate School of the "Cells in Motion" Cluster of Excellence at the University of Münster (CiM-IMPRS). This work was supported by the Max Planck Society, the European Research Council (ERC; grant agreement 669168), and National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIP) (NRF-2016R1A2B3011860, NRF-2016K1A3A1A61006005, and NRF-2017M3C7A1047640).

AUTHOR CONTRIBUTIONS

K.K., S.M.K., S.V., H.R.S., and D.W.H. conceived the study. K.K. and S.M.K. performed iNSC generation and *in vitro* characterization. S.V. preformed Oct4-Brn4 chimera reprogramming, analyzed the effect of the fusion protein on the reprogramming process, and wrote the manuscript. T.H.K. helped with time-course qPCRs. C.P. and K.H. performed RNA-seq and analyzed the results. J.K.H. performed *in vivo* differentiation of iNSCs. C.M.M. edited the manuscript. G.W. generated chimeric mice. C.C. and K.-P.K. analyzed and interpreted the data. H.R.S. and D.W.H. supervised the project and wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 17, 2018
Revised: January 23, 2019
Accepted: March 7, 2019
Published: April 2, 2019

REFERENCES

Bar-Nur, O., Verheul, C., Sommer, A.G., Brumbaugh, J., Schwarz, B.A., Lipchina, I., Huebner, A.J., Mostoslavsky, G., and Hochedlinger, K. (2015). Lineage conversion induced by pluripotency factors involves transient passage through an iPSC stage. Nat. Biotechnol. *33*, 761–768.

Buganim, Y., Itskovich, E., Hu, Y.C., Cheng, A.W., Ganz, K., Sarkar, S., Fu, D., Welstead, G.G., Page, D.C., and Jaenisch, R. (2012). Direct reprogramming of fibroblasts into embryonic Sertoli-like cells by defined factors. Cell Stem Cell *11*, 373–386.

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.

Cheloufi, S., Elling, U., Hopfgartner, B., Jung, Y.L., Murn, J., Ninova, M., Hubmann, M., Badeaux, A.I., Euong Ang, C., Tenen, D., et al. (2015). The histone chaperone CAF-1 safeguards somatic cell identity. Nature *528*, 218–224.

Chronis, C., Fiziev, P., Papp, B., Butz, S., Bonora, G., Sabri, S., Ernst, J., and Plath, K. (2017). Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell *168*, 442–459.e20.

Greder, L.V., Gupta, S., Li, S., Abedin, M.J., Sajini, A., Segal, Y., Slack, J.M., and Dutton, J.R. (2012). Analysis of endogenous Oct4 activation during induced pluripotent stem cell reprogramming using an inducible Oct4 lineage label. Stem Cells *30*, 2596–2601.

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

Han, D.W., Do, J.T., Araúzo-Bravo, M.J., Lee, S.H., Meissner, A., Lee, H.T., Jaenisch, R., and Schöler, H.R. (2009). Epigenetic hierarchy governing Nestin expression. Stem Cells *27*, 1088–1097.

Han, D.W., Tapia, N., Joo, J.Y., Greber, B., Araúzo-Bravo, M.J., Bernemann, C., Ko, K., Wu, G., Stehling, M., Do, J.T., and Schöler, H.R. (2010). Epiblast stem cell subpopulations represent mouse embryos of distinct pregastrulation stages. Cell *143*, 617–627.

Han, D.W., Tapia, N., Hermann, A., Hemmer, K., Höing, S., Araúzo-Bravo, M.J., Zaehres, H., Wu, G., Frank, S., Moritz, S., et al. (2012). Direct reprogramming of fibroblasts into neural stem cells by defined factors. Cell Stem Cell *10*, 465–472.

Hogan, B. (1994). Manipulating The Mouse Embryo: A Laboratory Manual,, Second Edition (CSHL Press).

Hong, J.Y., Lee, S.H., Lee, S.C., Kim, J.W., Kim, K.P., Kim, S.M., Tapia, N., Lim, K.T., Kim, J., Ahn, H.S., et al. (2014). Therapeutic potential of induced neural stem cells for spinal cord injury. J. Biol. Chem. *289*, 32512–32525.

Huang, P., He, Z., Ji, S., Sun, H., Xiang, D., Liu, C., Hu, Y., Wang, X., and Hui, L. (2011). Induction of functional hepatocyte-like cells from mouse fibroblasts by defined factors. Nature *475*, 386–389.

Ieda, M., Fu, J.D., Delgado-Olguin, P., Vedantham, V., Hayashi, Y., Bruneau, B.G., and Srivastava, D. (2010). Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. Cell *142*, 375–386.

Jerabek, S., Ng, C.K., Wu, G., Arauzo-Bravo, M.J., Kim, K.P., Esch, D., Malik, V., Chen, Y., Velychko, S., MacCarthy, C.M., et al. (2017). Changing POU dimerization preferences converts Oct6 into a pluripotency inducer. EMBO Rep. *18*, 319–333.

Kim, J., Efe, J.A., Zhu, S., Talantova, M., Yuan, X., Wang, S., Lipton, S.A., Zhang, K., and Ding, S. (2011a). Direct reprogramming of mouse fibroblasts to neural progenitors. Proc. Natl. Acad. Sci. USA *108*, 7838–7843.

Kim, J.H., Lee, S.R., Li, L.H., Park, H.J., Park, J.H., Lee, K.Y., Kim, M.K., Shin, B.A., and Choi, S.Y. (2011b). High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PLoS ONE *6*, e18556.

Kim, S.M., Flaßkamp, H., Hermann, A., Araúzo-Bravo, M.J., Lee, S.C., Lee, S.H., Seo, E.H., Lee, S.H., Storch, A., Lee, H.T., et al. (2014). Direct conversion of mouse fibroblasts into induced neural stem cells. Nat. Protoc. *9*, 871–881.

Kim, S.M., Kim, J.W., Kwak, T.H., Park, S.W., Kim, K.P., Park, H., Lim, K.T., Kang, K., Kim, J., Yang, J.H., et al. (2016). Generation of Integration-free Induced Neural Stem Cells from Mouse Fibroblasts. J. Biol. Chem. *291*, 14199–14212.

Kurian, L., Sancho-Martinez, I., Nivet, E., Aguirre, A., Moon, K., Pendaries, C., Volle-Challier, C., Bono, F., Herbert, J.M., Pulecio, J., et al. (2013). Conversion of human fibroblasts to angioblast-like progenitor cells. Nat. Methods *10*, 77–83.

Lalit, P.A., Salick, M.R., Nelson, D.O., Squirrell, J.M., Shafer, C.M., Patel, N.G., Saeed, I., Schmuck, E.G., Markandeya, Y.S., Wong, R., et al. (2016). Lineage Reprogramming of Fibroblasts into Proliferative Induced Cardiac Progenitor Cells by Defined Factors. Cell Stem Cell *18*, 354–367.

Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. Nat. Protoc. *1*, 729–748.

Li, K., Zhu, S., Russ, H.A., Xu, S., Xu, T., Zhang, Y., Ma, T., Hebrok, M., and Ding, S. (2014). Small molecules facilitate the reprogramming of mouse fibroblasts into pancreatic lineages. Cell Stem Cell *14*, 228–236.

Margariti, A., Winkler, B., Karamariti, E., Zampetaki, A., Tsai, T.N., Baban, D., Ragoussis, J., Huang, Y., Han, J.D., Zeng, L., et al. (2012). Direct reprogramming of fibroblasts into endothelial cells capable of angiogenesis and reendothelialization in tissue-engineered vessels. Proc. Natl. Acad. Sci. USA *109*, 13793–13798.

Maza, I., Caspi, I., Zviran, A., Chomsky, E., Rais, Y., Viukov, S., Geula, S., Buenrostro, J.D., Weinberger, L., Krupalnik, V., et al. (2015). Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. Nat. Biotechnol. *33*, 769–774.

Mistri, T.K., Devasia, A.G., Chu, L.T., Ng, W.P., Halbritter, F., Colby, D., Martynoga, B., Tomlinson, S.R., Chambers, I., Robson, P., and Wohland, T. (2015). Selective influence of Sox2 on POU transcription factor binding in embryonic and neural stem cells. EMBO Rep. *16*, 1177–1191.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat. Biotechnol. *26*, 101–106.

Niu, W., Zang, T., Zou, Y., Fang, S., Smith, D.K., Bachoo, R., and Zhang, C.L. (2013). In vivo reprogramming of astrocytes to neuroblasts in the adult brain. Nat. Cell Biol. *15*, 1164–1175.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. Nat. Genet. *24*, 372–376.

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.

Reményi, A., Lins, K., Nissen, L.J., Reinbold, R., Schöler, H.R., and Wilmanns, M. (2003). Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. Genes Dev. *17*, 2048–2059.

Rosenfeld, M.G. (1991). POU-domain transcription factors: pou-er-ful developmental regulators. Genes Dev. *5*, 897–907.

Sommer, C.A., Stadtfeld, M., Murphy, G.J., Hochedlinger, K., Kotton, D.N., and Mostoslavsky, G. (2009). Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells *27*, 543–549.

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.

Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R.M., Chen, T., Ooi, S.S., Kim, S.Y., Bestor, T.H., Shioda, T., et al. (2012). Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all-iPS cell mice from terminally differentiated B cells. Nat. Genet. *44*, 398–405.

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.

Tapia, N., Han, D.W., and Schöler, H.R. (2012). Restoring stem cell function in aged tissues by direct reprogramming? Cell Stem Cell *10*, 653–656.

Thier, M., Wörsdörfer, P., Lakes, Y.B., Gorris, R., Herms, S., Opitz, T., Seiferling, D., Quandel, T., Hoffmann, P., Nöthen, M.M., et al. (2012). Direct conversion of fibroblasts into stably expandable neural stem cells. Cell Stem Cell *10*, 473–479.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C., and Wernig, M. (2010). Direct conversion of fibroblasts to functional neurons by defined factors. Nature *463*, 1035–1041.

Yoshimizu, T., Sugiyama, N., De Felice, M., Yeom, Y.I., Ohbo, K., Masuko, K., Obinata, M., Abe, K., Schöler, H.R., and Matsui, Y. (1999). Germline-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. Dev. Growth Differ. *41*, 675–684.

Yu, B., He, Z.Y., You, P., Han, Q.W., Xiang, D., Chen, F., Wang, M.J., Liu, C.C., Lin, X.W., Borjigin, U., et al. (2013). Reprogramming fibroblasts into bipotential hepatic stem cells by defined factors. Cell Stem Cell *13*, 328–340.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse anti-Nestin	Millipore	MAB353; RRID: AB_94911
Goat anti-Sox2	Santa Cruz Biotechnology	CSB-PA16539A0Rb
Mouse anti-SSEA1	Santa Cruz Biotechnology	SC-21702 AF488; RRID: AB_626918
Rabbit anti-Olig2	Millipore	AB9610; RRID: AB_570666
Mouse anti-Tuj1	Covance	MMS-435P; RRID: AB_2313773
Rabbit anti-GFAP	DAKO	M0761; RRID: AB_2109952
Rat anti-MBP	Abcam	ab7349; RRID: AB_305869
Rat anti-Nanog	eBioscience	eBioMLC-51; RRID: AB_763613
Mouse anti-FLAG	Sigma	F1804; RRID: AB_262044
Chemicals, Peptides, and Recombinant Proteins		
4-hydroxytamoxifen	Sigma	H7904-5MG
Recombinant Human FGF-basic (154 a.a.)	Peprotech	100-18B
Epidermal growth factor EGF	Peprotech	Q-AF-100-15
Ascorbic acid	Sigma	A7506
rhPDGF-AA, CF	Sigma	221-AA-010
T3	Sigma	T6397
Paraformaldehyde	Sigma	P6148-500G
Triton X-100	Sigma	X100
Hoechst 33342	Sigma	14533
Critical Commercial Assays		
Hybrid-R^{TM} kit	GeneAll	305-101
GoTag green master mix	Promega	M3005
SYBR Green PCR Master Mix	Applied Biosystems	ABS-4367659
EpiTect Bisulfite Kit	QIAGEN	59104
SuperTaq polymerase	Ambion	AM2052
PCR 2.1-TOPO vector	Invitrogen	K2500-20
QIAprep Spin Miniprep Kit	QIAGEN	27104
Experimental Models: Organisms/Strains		
B6(SJL)-Pou5f1$^{tm1.1(cre/Esr1*)Yseg}$/J	Jackson laboratory	016829
B6.129X1-Gt(ROSA)26Sor$^{tm1(EYFP)Cos}$/J	Jackson laboratory	006148
Oct4-CreER x Rosa26-loxSTOPlox-DTA mouse	Harvard Stem Cell Institute	
Software and Algorithms		
CLC Genomics Workbench 11	Qiagen	
R	R Core Team	http://www.R-project.org/

CONTACT FOR REAGENT AND RESOURCE SHARING

Requests for further information should be directed to and will be fulfilled by Sergiy Velychko (sergii.velychko@mpi-muenster.mpg.de) and Kyuree Kang (krkang@konkuk.ac.kr)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mice

All mice used were bred and housed at the mouse facility of Konkuk University (KU) or Max Planck Institute (MPI) in Muenster, and animal handling was in accordance with the KU and MPI animal protection guidelines. A protocol for animal handling and maintenance for this study was approved by the Landesamt für Natur, Umwelt und Verbraucherschutz Nordrhein-Westfalen under the supervision of a certified veterinarian in charge of the MPI animal facility. Oct4-CreER mice (stock number: 016829, Jackson Laboratory) were crossed with Rosa26-lox-STOP-lox-EYFP mice (stock number: 006148). MEFs were derived from Oct4-CreER x Rosa26-loxSTOPlox-EYFP mouse embryos at embryonic day 13.5 after carefully removing the head and all the internal organs including the spinal cord. Cells with the correct genotype were maintained in DMEM (Biowest) containing 10% FBS (Biowest), 5 mL of penicillin/streptomycin/glutamine (Invitrogen), and 5 mL of MEM NEAA solution (Invitrogen) in 500 mL of MEF medium. MEFs from an Oct4-CreER x Rosa26-loxSTOPlox-DTA mouse were kindly given by the lab of Prof. Konrad Hochedlinger (Harvard Stem Cell Institute). The MEFs derived from embryos of both sexes were pooled together, thus, both sexes were included in our analysis.

METHOD DETAILS

Generation of iNSCs

To generate iNSCs by using polycistronic vectors, 5×10^4 MEFs were infected with lentivirus encoding for either pcOKSM or pcBKSM. After 24 hours of incubation, the medium was replaced with NSC medium containing 3 μg/ml of dox. After 7–8 days of induction with pcOKSM or pcBKSM, the cells were cultured for 23–24 days in NSC medium without dox treatment to allow for the expansion of initial iNSC clusters. 1 μM of 4-hydroxytamoxifen (4-OHT) was continuously added to the culture medium. To generate individual retroviral vector–mediated iNSCs, the MEFs were transduced with individual retroviral particles encoding for either mcOKSM or mcBKSM and cultured as previously described (Kim et al., 2014). Briefly, 5×10^4 fibroblasts were plated onto gelatin-coated 35-mm dishes and incubated with ecotropic retroviruses. After 48 hours of incubation, the medium containing retroviral particles was replaced with NSC medium. To enrich for iNSCs, non-reprogrammed fibroblasts or unwarranted cells were removed with a cell scraper as previously described. To establish the clonal iNSC lines, the iNSC bulk culture was stained for SSEA1, and $SSEA1^+$ single cells were sorted using BD FACSAria™ (BD Biosciences) and plated onto laminin/poly-D-lysine-coated 96-well plates. The cNSCs and established iNSCs were maintained in NSC culture medium: DMEM/F-12 supplemented with 2% B27 (GIBCO), 10 ng/ml of EGF (Peprotech), 10 ng/ml of bFGF (Peprotech), and 1% penicillin/streptomycin/glutamine (Invitrogen).

Differentiation of iNSCs

For differentiation into neurons, iNSCs were plated onto laminin/polylysine-coated dishes at 2.5×10^4 cells/cm^2 in NSC medium. The next day, the medium was replaced by neural differentiation medium: DMEM/F-12 supplemented with 2% B27 (GIBCO), 1% penicillin/streptomycin/glutamine (Invitrogen), and 10 ng/ml of bFGF (Peprotech). On day 4 of differentiation, the medium was changed to neural differentiation medium containing 200 mM ascorbic acid (Sigma) without growth factors for 8–10 more days. For differentiation into astrocytes, iNSCs were cultured in DMEM/F-12 supplemented with 10% FBS and 1% penicillin/streptomycin/glutamine on gelatin-coated dishes for 5 days. Finally, for differentiation into oligodendrocytes, iNSCs were plated onto laminin/polylysine-coated dishes at 2.5×10^4 cells/cm^2 in NSC medium. The next day, the medium was replaced with oligodendrocyte differentiation medium: DMEM/F-12 supplemented with 2% B27, 1% penicillin/streptomycin/glutamine, 10 ng/ml of bFGF, and 10 ng/ml of PDGF (Sigma). On day 4 of differentiation, the medium was changed to oligodendrocyte differentiation medium containing 30 ng/ml of T3 (Sigma) and 200 mM ascorbic acid for 4 more days. The differentiation medium was replaced every other day.

For *in vivo* differentiation, 1 × 106 GFP-labeled iNSCs were transplanted into the cortical region of rat brain, as we had described previously (Hong et al., 2014; Kim et al., 2016).

Vector Construction

The Oct4-Brn4 chimera library was generated using overlap extension assembly of two or three PCR fragments with approximately 20 bp–long overlapping ends. The final chimeric PCR products were inserted into the pMX backbone (Takahashi and Yamanaka, 2006) using EcoRI and XhoI restriction sites and sequenced.

The pHAGE2-tetO-OKSM (STEMCCA) and pHAGE-tetO-BKSM lentiviral vectors were kindly provided by Konrad Hochedlinger. All the other lentiviral expression cassettes used in this study were cloned into the same backbone. To completely disable the self-cleaving ability of the F2A peptide, the last two residues, which constitute the self-cleaving Pro-Gly site, were mutated into Ala-Ala. The following sequences of the P2A self-cleaving peptide and poly-glycine linker were used: GSGATNFSLLKQAGDVEENPGP and GSGGGSGGGSGGGGSGGGGSG, respectively. The detailed cloning information and sequence maps can be provided upon request.

Reprogramming of MEFs into iPSCs

For iPSC generation, we used Oct4-GFP (OG2 MEFs or heterozygous ROSA26-rtTA/GOF18) MEFs after 3 or 4 passages (Yoshimizu et al., 1999). The MEFs were plated onto 12-well plates at the density of $2.5–3 \times 10^4$ per well in 1 mL of fresh MEF medium. On the

same day, the cells were transduced with 60 μl of non-concentrated tetO-OKSM and 30 μl of rtTA (in the case of OG2 MEFs) viral supernatants supplemented with 6 μg/ml protamine sulfate (Sigma-Aldrich). The volume of the rest of viral sups was adjusted according to qPCR titration. After 48 hours, the medium was replaced with embryonic stem cell medium (high-glucose DMEM, 15% KSR, LIF) supplemented with dox (1 μg/ml unless otherwise mentioned). The number of GFP^+ colonies was counted after 1 and 2 weeks of induction. Dox was removed after 12 days of induction (2 days before the second count). The lentivirus stocks were titrated using Q-PCR for *WPRE* with cDNA samples from MEFs after 24 hours of dox induction. The primers used are listed in Table S1.

Characterization of iPSCs

For PCR genotyping, we isolated genomic DNA from established iPSC lines. The primers used for genotyping are listed in Table S3.

For the *in vivo* differentiation assay (teratoma formation), approximately 5×10^6 iPSCs were harvested and injected subcutaneously into the flank of immunodeficient (scid) mice. After 4–5 weeks, the teratomas were removed, fixed in 4% paraformaldehyde, and subjected to histological examination with hematoxylin and eosin staining.

The chimera contribution assay was performed using a previously published protocol (Hogan, 1994). Briefly, 8–10 trypsinized iPSCs were aggregated with 8-cell embryos derived from (C57BL/6 x C3H) F1 female mice crossed with CD1 mice at 2.5 days post-coitum (dpc), and, after 24 hours of culture, transferred into 2.5-dpc pseudopregnant recipients.

Immunocytochemistry

The cells were fixed with 4% paraformaldehyde (Sigma) for 20 min at room temperature and then blocked with Dulbecco's PBS (Biowest) containing 0.3% Triton X-100 (Sigma) and 5% FBS (Biowest) for 2 hours at room temperature. The cells were then incubated with primary antibodies at 4°C for 16 hours, washed three times with Dulbecco's PBS, and then incubated with the appropriate fluorescence-conjugated secondary antibody for 2 hours at room temperature in the dark. Nuclei were stained with Hoechst 33342 (Sigma). Primary antibodies used for immunofluorescence are as follows: mouse anti-Nestin (Millipore, 1:200), goat anti-Sox2 (Santa Cruz Biotechnology, 1:200), rabbit anti-Blbp (Santa Cruz Biotechnology, 1:200), mouse anti-SSEA1 (Santa Cruz Biotechnology, 1:100), rabbit anti-Olig2 (Millipore, 1:200), mouse anti-Tuj1 (Covance, 1:500), rabbit anti-GFAP (DAKO, 1:500), rat anti-MBP (Abcam, 1:100), and rat anti-Nanog (eBioscience, eBioMLC-51, 1:1000).

Flow Cytometry Analysis

To generate iNSC clonal lines, cells were dissociated with trypsin, washed once with PBS, and resuspended in FACS buffer (PBS containing 5% FBS). 1×10^6 cells were incubated with FITC-conjugated anti-SSEA1 antibody (Santa Cruz Biotechnology, 1:10) for 15 min at 4°C. Cells were washed once with FACS buffer and resuspended in PBS for analysis using a BD FACSAria™ (BD Biosciences).

Gene Expression Analysis by Quantitative PCR

Total RNA was isolated using the Hybrid-R™ kit (GeneAll), and 1 μg of total RNA was reverse transcribed into cDNA using a high capacity cDNA reverse transcription kit (Applied Biosystems) according to the manufacturer's instructions. RT-PCR was performed using the GoTag green master mix (Promega). qPCR was performed using SYBR Green PCR Master Mix (Applied Biosystems) on the ABI 7500 real-time PCR system (Applied Biosystems). ΔC_t values were calculated by subtracting the *Gapdh* or *Rpl37a* C_t value from that of target genes. Relative expression levels were calculated using the $2^{-\Delta\Delta}Ct$ method. The sequence of primer sets used are listed in Table S1.

DNA Methylation Analysis

To determine the DNA methylation status of iNSCs, genomic DNA was treated with sodium bisulfite to convert all unmethylated cytosine residues into uracil residues using EpiTect Bisulfite Kit (QIAGEN) according to the manufacturer's instructions. Briefly, PCR amplifications were performed using SuperTaq polymerase (Ambion) in a total volume of 25 μL and a protocol of a total of 40 cycles of denaturation at 94°C for 30 s, annealing at the appropriate temperature for each target region for 30 s, extension at 72°C for 30 s with a 1st denaturation at 94°C for 5 min, and a final extension at 72°C for 10 min. Primer sequences and annealing temperatures were described in our previous study and listed in Table S2 (Han et al., 2009; Hong et al., 2014). For each primer set, 3 μL of product from the first round of PCR was used in the second round of PCR as template. The amplified products were verified by electrophoresis on 1% agarose gel. PCR products were subcloned using the PCR 2.1-TOPO vector (Invitrogen) according to the manufacturer's instructions. Reconstructed plasmids were purified using the QIAprep Spin Miniprep Kit (QIAGEN) and individual clones were sequenced (Macrogen, Republic of Korea). Clones were analyzed using QUMA software.

Chromatin Immunoprecipitation (ChIP)

Chromatin immunoprecipitation was performed as described before (Lee et al., 2006). Briefly, 2×10^7 cells were cross-linked with formaldehyde to a final concentration of 1% for 20 min at 4°C and quenched with glycine (final concentration 0.1375M). Cross-linked cells were sonicated with the Bioruptor (Diagenode) for 45 rounds of 30 s pulse and 30 s rest cycles. 25 μg of sonicated chromatin from each sample was used for immunoprecipitation with anti-FLAG antibody (Sigma, F1804). The primers used for qPCR are listed in Table S4.

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA Sequencing Analysis

Amount and quality of RNAs extracted from the cells were measured using Bioanalyzer RNA Chip (Agilent Technologies). RNA samples with RNA integrity number (RIN) $\geq$ 8.0 were used for library preparation. The RNA-seq libraries were produced as instructed by the manufacturer (TruSeq Stranded mRNA Sample Preparation Kit, Illumina) and sequenced on a Nextseq500 platform (Illumina). The 75-bp paired end read data were mapped to UCSC mm10 mouse genome build and counted using CLC Genomics Workbench 11. The data were visualized with heatmap.2 function (gplots) in R package (http://www.r-project.org/, v3.3.2).

DATA AND SOFTWARE AVAILABILITY

RNA-seq data are available from Gene Expression Omnibus under accession number GEO: GSE125740.

Cell Reports, Volume *27*

Supplemental Information

Fusion of Reprogramming Factors Alters the Trajectory of Somatic Lineage Conversion

Sergiy Velychko, Kyuree Kang, Sung Min Kim, Tae Hwan Kwak, Kee-Pyo Kim, Chanhyeok Park, Kwonho Hong, ChiHye Chung, Jung Keun Hyun, Caitlin M. MacCarthy, Guangming Wu, Hans R. Schöler, and Dong Wook Han

Figure S1

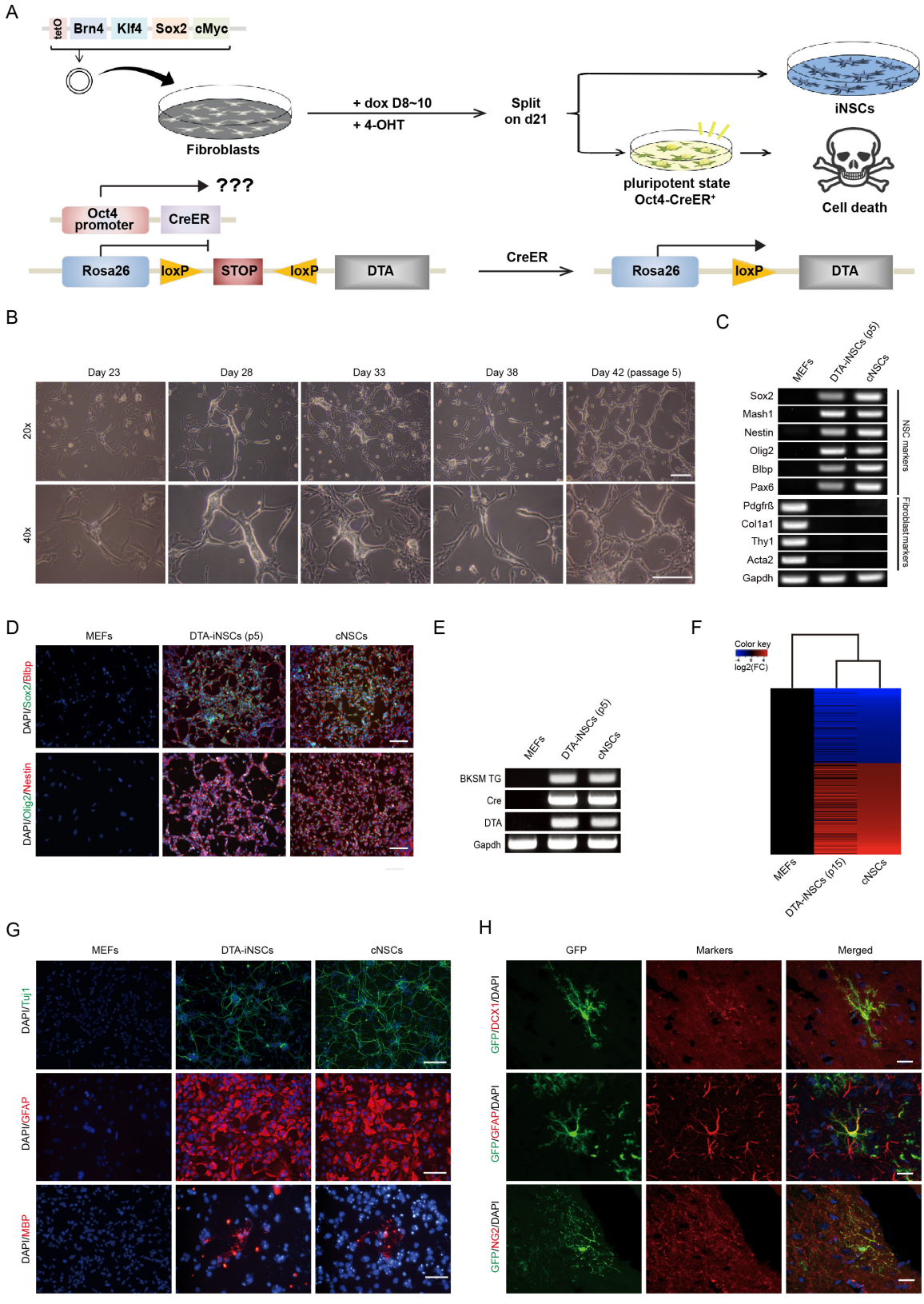


Figure S1 Lineage tracing during iNSC generation. Related to Figure 1

(a) Schematic diagram depicting the lineage-tracing experiment for the direct conversion of fibroblasts into iNSCs using the Oct4-CreER; R26-lsl-DTA reporter.

(b) Time-course images showing the morphological changes occurring during the induction of iNSCs from Oct4-CreER; R26-lsl-DTA MEFs after dox withdrawal. Scale bar, 50 μm.

(c) RT-PCR analysis of NSC- and fibroblast-specific marker gene expression in DTA iNSCs.

(d) Immunofluorescence microscopy of DTA-iNSCs. The nuclei were stained with DAPI. Scale bar, 50 μm.

(e) PCR genotyping of DTA iNSCs for integration of pcBKSM, Oct4-CreER, and DTA transgenes.

(f) Heatmap of differentially expressed genes in DTA-iNSCs. Hierarchical clustering analysis based on the gene expression profiles is shown at the top. The differentially expressed genes with FC ≥4 between MEFs and cNSC controls, and FPKM ≥1 in at least one of the samples were plotted.

(g) *In vitro* differentiation of DTA-iNSCs into neurons, astrocytes, and oligodendrocytes after immunostaining with antibodies against Tuj1, GFAP, and MBP, respectively. Scale bar, 50 μm.

(h) Immunocytochemistry of *in vivo*–engrafted GFP-labelled DTA-iNSCs that differentiated into neurons (GFP+/DCX1+), astrocytes (GFP+/GFAP+), and oligodendrocytes (GFP+/NG2+). Scale bar, 25 μm.

Figure S2

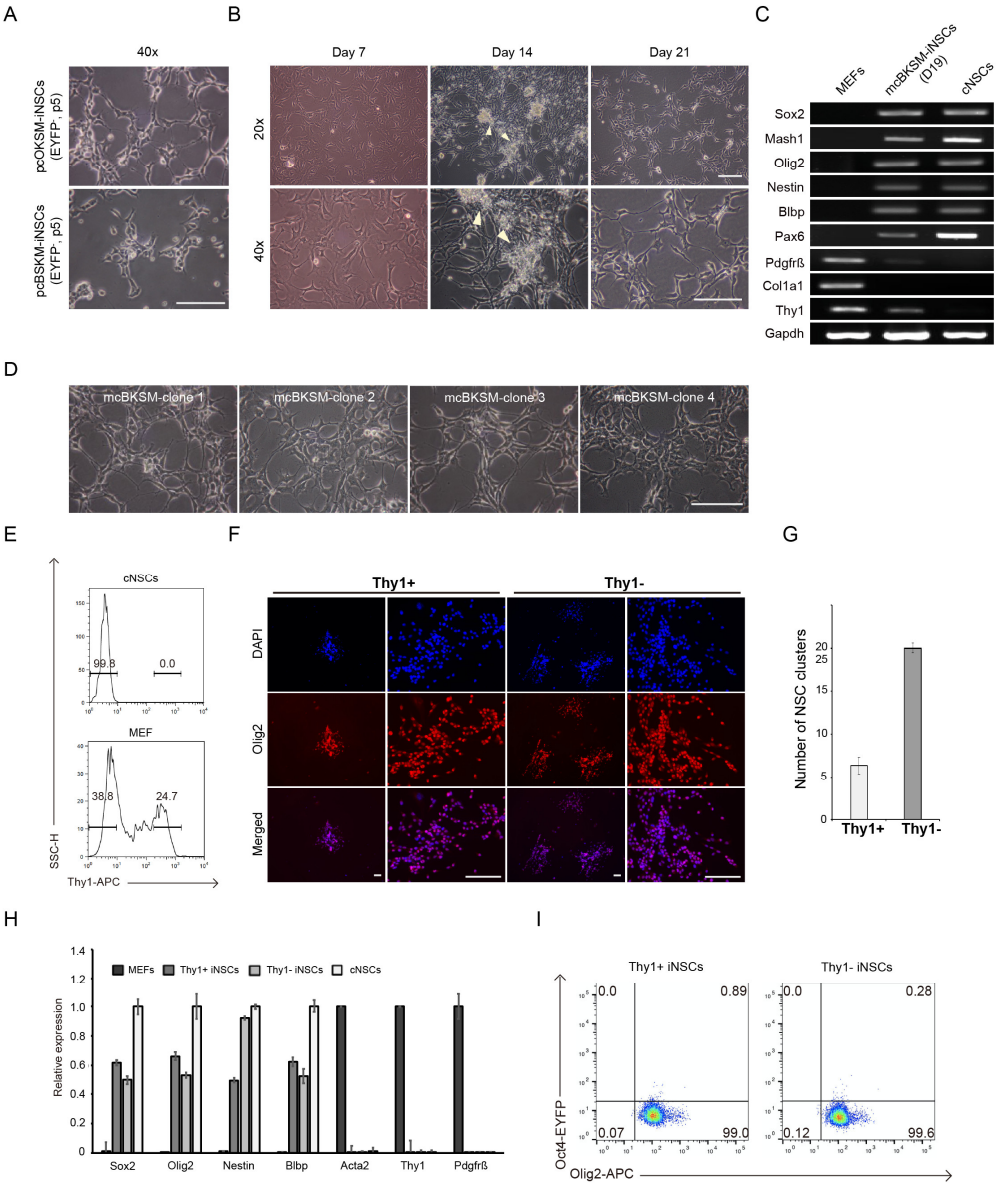


Figure S2 Characterization of iNSCs generated with poly- and monocistronic cassettes. Related to Figure 2

(a) Phase-contrast images of EYFP^{-} pcOKSM and pcBKSM iNSCs. Scale bar, 50 µm.

(b) The morphology of mcBKSM-transduced MEFs and initial NSC-like clusters (arrow heads). Scale bar, 50 µm.

(c) RT-PCR analysis of NSC and fibroblast marker gene expression of mcBKSM-transduced MEFs after 19 days of infection.

(d) The morphology of four established clonal mcBKSM-derived iNSC lines. Scale bar, 50 µm.

(e) FACS analysis of MEFs to pre-sort Thy1^{+} and Thy1^{-} populations.

(f) Immunofluorescence of iNSC generated from Thy1^{+} and Thy1^{-} MEFs transduced with mcBKSM. Scale bar, 50 µm.

(g) The number of iNSC cluster expressing Olig2 from Thy1^{+} and Thy1^{-} MEFs transduced with mcBKSM. Data are represented as mean ± SD, n=3.

(h) qPCR analysis of NSC- and fibroblast-specific marker expression in mcBKSM-derived iNSC lines from Thy1^{+} and Thy1^{-} MEFs.

(i) FACS analysis to determine the number of directly (Olig2^{+}/EYFP^{-}) and indirectly (Olig2^{+}/EYFP^{+}) generated iNSCs for mcBKSM-induced iNSC lines from Thy1^{+} and Thy1^{-} MEFs.

Figure S3

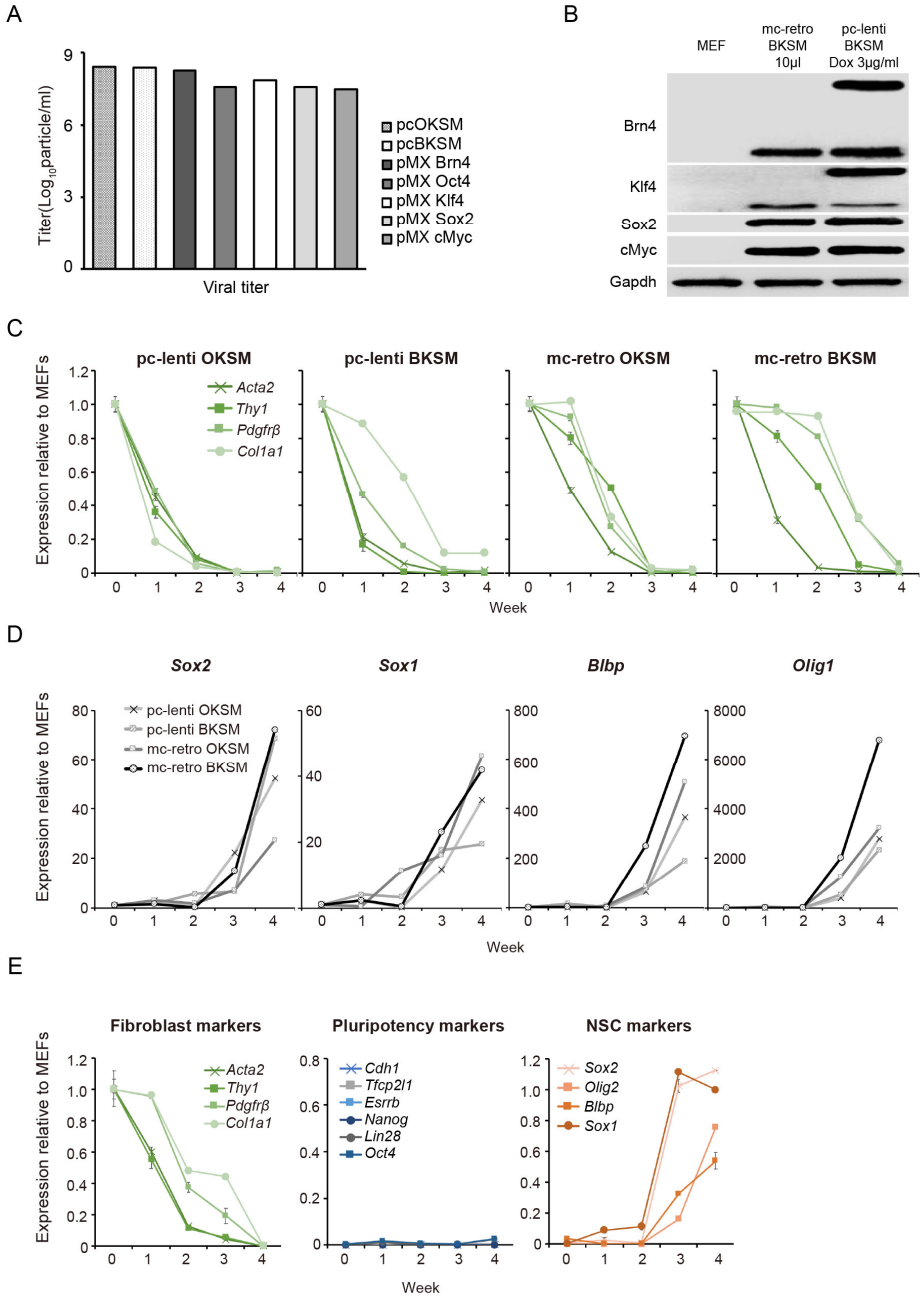


Figure S3 Monocistronic lentiviral induction also leads to direct transdifferentiation into iNSCs. Related to Figure 3

(a) Measured number of monocistronic retrovirus and polycistronic lentivirus particles used in experiments.

(b) Western blot analysis of mono- and polycistronically introduced Brn4, Sox2 and cMyc.

(c) Time-course qPCR analysis of fibroblast marker gene expression in MEFs transduced with lentiviral pcOKSM, pcBKSM, and retroviral mcOKSM, or mcBKSM. The levels were normalized to non-transduced MEF control and are represented as mean ± SD, n=3.

(d) Time-course qPCR analysis of NSC marker gene expression in MEFs transduced with pcOKSM, pcBKSM, mcOKSM, or mcBKSM. The levels were normalized to non-transduced MEF control and are represented as mean ± SD, n=3.

(e) Time-course qPCR analysis of fibroblast, pluripotency, and NSC marker gene expression in MEFs transduced with lentiviral mcBKSM. The levels were normalized to non-transduced MEF control (fibroblast markers), ESC control (pluripotency markers), and or cNSCs (NSC markers) and are represented as mean ± SD, n=3.

Figure S4

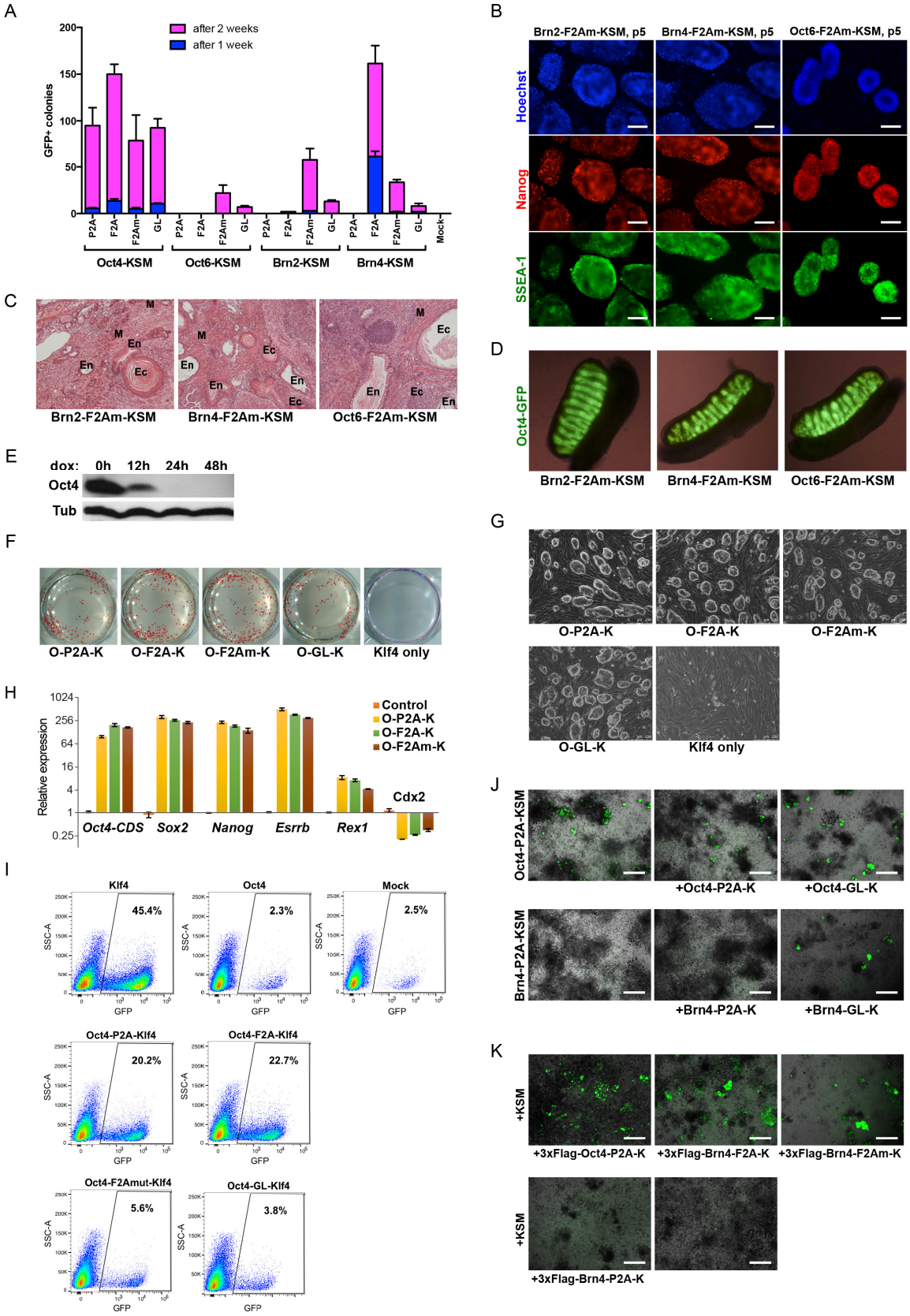


Figure S4. POU factors retain functionality within the POU-Klf4 fusion protein. Related to Figure 4

(a) Reprogramming of Oct4-GFP MEFs into iPSCs with POU-KSM polycistronic cassettes, where POU and Klf4 were separated by P2A, F2A, cleavage-deficient mutated F2A (F2Am), or a glycine linker (GL). The GFP^{+} colonies were counted after 1 and 2 weeks of dox induction. Data are represented as mean ± SD, n=3.

(b) Immunostaining for Nanog and SSEA1 of clonal iPSCs generated with Brn2-, Brn4-, and Oct6-F2Am-KSM constructs. Scale bar, 250 μm.

(c) Teratoma formation assay with Brn2-, Brn4-, and Oct6-F2Am-KSM–generated iPSCs. The iPSCs could contribute to all three germ layers: ectoderm (keratinized epithelium, neural tissue), mesoderm (connective tissue, smooth muscle), and endoderm (intestinal, ciliated, and cuboidal epithelia).

(d) Brightfield/Oct4-GFP merged images of germline contribution of the iPSCs.

(e) Western blotting of ZHBTc4 ESCs treated with dox for indicated number of hours.

(f) Alkaline phosphatase (AP) staining of Oct4-knock-out ZHBTc4 ESCs rescued with Oct4-Klf4 constructs with different linker peptides after 7 days of dox induction.

(g) Representative images of conditionally Oct4-knock-out ZHBTc4 ESCs rescued with Oct4-Klf4 constructs with different linker peptides on the second passage.

(h) qPCR gene expression analysis of p6 ZHBTc4 ESCs rescued with Oct4-P2A-Klf4, Oct4-F2A-Klf4, Oct4-F2Am-Klf4, or control (ZHBTc4 ESCs treated with dox for 1 week). *Rpl37a* was used as a house-keeping gene. Data are represented as mean ± SD, n=3 (technical replicates).

(i) FACS analysis of Gof18 epiblast stem cells converted into naïve ESCs with Oct4, Klf4, and Oct4-Klf4 separated by self-cleavable peptides, non-cleavable peptides, or mock constructs after 1 week of induction.

(j) Representative wide-field images of Oct4-GFP MEFs reprogramed with fully cleavable Oct4- and Brn4-P2A-KSM polycistronic cassettes, in combination with fully cleaved or uncleaved POU-Klf4 constructs (related to Figure 4F). Scale bar, 1 mm.

(k) Representative wide-field images of Oct4-GFP MEFs reprogramed 3xFLAGed constructs used for Chip (related to Figure 4G). Scale bar, 1 mm.

Table S1. Primers for qPCR. Related to STAR Methods.

Gene Name	Accession number	Sequence	Annealing temperature
Sox2	NM_ 011443	5'- ACGGCCATTAACGGCACACT -3'	60°C
		5'- TTTTGCACCCCTCCCAATTC -3'	
Pax6	NM_ 001244198	5'- CAAGTTCCCGGGAGTGAACC -3'	60°C
		5'- TCCACATAGTCATTGGCAGA -3'	
Nestin	NM_ 016701	5'- TCCTGGTCCTCAGGGGAAGA -3'	60°C
		5'- TCCACGAGAGATACCACAGG -3'	
Olig2	NM_ 016967	5'- ACCACCACGTGTCGGCTATG -3'	60°C
		5'- TGGTCCAGCTCCCCTTCTTG -3'	
Mash1	NM_ 008553	5'- CAGAGGAACAAGAGCTGCTG -3'	60°C
		5'- GATCTGCTGCCATCCTGCTT -3'	
Blbp	NM_ 021272	5'- GGATGGCAAGATGGTCGTGA -3'	60°C
		5'- TGGGACTCCAGGAAACCAAG -3'	
Sox1	NM_009233	5'- GGTGGAGGCCACAACAACAA -3'	60°C
		5'- GGCCTCTTTGGCAAGTGGTT -3'	
Cdh1	NM_009864	5'- AACAACTGCATGAAGGCGGGAATC -3'	60°C
		5'- CCTGTGCAGCTGGCTCAAATCAAA -3'	
Tfcp2l1	NM_023755	5'- AACCCGCCCAGGTAGAGCT -3'	60°C
		5'- AGGGCAGCCACGTGGGAAGA -3'	
Esrrb	NM_001159500	5'- GCCTTTACTATCTG GCCTGGT -3'	60°C
		5'- TAGTGCTTCTCTTTGGTGCTGT -3'	
Nanog	NM_001289828	5'- CTTTCACCTATTAAGGTGCTTGC -3'	60°C
		5'- TGGCATCGGTTCATCATGGTAC -3'	
Lin28	NM_145833	5'- CCGCAGTTGTAGCACCTGTCT -3'	60°C
		5'- GAAGAACATGCAGAAGCGAAGA -3'	
Oct4	NM_013633	5'- CTTGGGCTAGAGAAGGATGTGGTT -3'	60°C
		5'- TCTTGTCTACCTCCCTTGCCTTG -3'	
Thy1	NM_009382	5'- TTCCCTCTCCCTCCTCCAAGC -3'	60°C
		5'- TCGAGGGCTCCTGTTTCTCCTT -3'	
Col1a1	NM_007742	5'- CCCTGCCTGCTTCGTGTAAA -3'	60°C
		5'- TCGTCTGTTTCCAGGGTTGG -3'	
Acta2	NM_007392	5'-ATCGTCCACCGCAAATGCTT -3'	60°C
		5'-AACTGGAGGCGCTGATCCAC-3'	
Pdgfrβ	NM_001146268	5'-TGGGTGGAGATTCGCAGGAGG -3'	60°C
		5'-CCACTAAGGCCAGGATGGCTGA -3'	
WPRE		5'-TGTTGCCACCTGGATTCTGC-3'	60°C
		5'- AGGAAGGTCCGCTGGATTGA-3'	
Rpl37A	NC_000067	5'- GTGGTTCCTGCATGAAGACAGTG-3'	60°C
		5'- TTCTGATGGCGGACTTTACCG-3'	

Hprt1	NC_000086	5'-CTGGTGAAAAGGACCTCTCGA-3'	60°C
		5'-CTGAAGTACTCATTATAGTCAAGGGCAT-3'	
Gapdh	NM_008084	5'- CCAATGTGTCCGTCGTGGAT -3'	60°C
		5'- TGCCTGCTTCACCACCTTCT -3'	
Oct4-CDS	NM_013633	5'- GGCTAGAGAAGGATGTGGTTCGAG -3'	60°C
		5'- CCTGGGAAAGGTGTCCCTGTAG -3'	
Rex1	NM_009556	5'- GGCTGCGAGAAGAGCTTTATTCA -3'	60°C
		5'- AGCATTTCTTCCCGGCCTTT -3'	
Cdx2	NM_007673	5'- ACCGGAATTGTTTGCTGCTGT -3'	60°C
		5'- TCCCGACTTCCCTTCACCAT -3'	

Table S2. Primers for DNA methylation analysis. Related to STAR Methods.

Gene Name	Sequence	Annealing temperature
Nestin enhancer1st	5'- GATCCCAGTGTGGTGGTACG -3'	45°C
	5'- GGCTTCAGCTCCGTCTCCAT -3'	
Nestin enhancer 2nd	5'- GTGTGGTGGTACGGGAAATC -3'	60°C
	5'- GAGAAGGACGGGAGCAGAG -3'	
*Col1a1*1st	5'- GTTAGGTAGTTTTGATTGGTTGG -3'	55°C
	5'- ACAATAACCCCTAAAAAAAACAAAAA -3'	
*Col1a1*2nd	5'- TGGTATAAAAGGGGTTTAGGTTAGT -3'	60°C
	5'- ACAATAACCCCTAAAAAAAACAAAAA -3'	

Table S3. Primers for PCR genotyping. Related to STAR Methods.

Gene Name	Sequence	Annealing temperature
Viral Oct4	5'- GAGACGCCATCCACGCTGT-3'	55°C
	5'- GGTGAGAAGGCGAAGTCTGAAG-3'	
Viral Brn4	5'- GAGACGCCATCCACGCTGT-3'	55°C
	5'- ATGGACAAGGGAGCTGGAAC-3'	
Viral Oct6	5'- CACCACCACACACTGCCCGGCTCTG-3'	55°C
	5'- CCCTTTTTCTGGAGACTAAATAAAATC-3'	
Viral Brn2	5'- AGTAGGGACACGCCACCACACCACG-3'	55°C
	5'- GTGGAGAAGGACGGGAGCAG-3'	
ActB	5'- ACTGCCGCATCCTCTTCCTC-3'	55°C
	5'- CCGCTCGTTGCCAATAGTGA-3'	

Table S4. Primers for ChIP-qPCR. Related to STAR Methods.

Gene Name	Sequence	Annealing temperature
grDNA-IGS1	5'- GGCCAGTTCCTCCTGCCTTCTGTT -3'	60°C
	5'- ACTGTGGATGGAGCGTGCATGTGT -3'	
grDNA-IGS2	5'- TGACTTTCTGGTTTGCAAGTTAAG -3'	60°C
	5'- CCCATTAATTGGGGTACTCTACTG -3'	
Sox2	5'- AGTCCAAGCTAGGCAGGTTCCCCT -3'	60°C
	5'- TGCCCGAGCCCGGGAAATTCTTTT -3'	
Nanog (DE)	5'- CCTGTCCCTAGTCCCCGCTCCTTT -3'	60°C
	5'- TGGAACTAGCTGTGTGGGTGGGGA -3'	
Pou5f1(CR4)	5'- TGGGCAGACGGCAGATGCATAACA -3'	60°C
	5'- GGGACCCCTCCCCAACCATCTTCT -3'	
Slitrk6	5'- TGAAAGGCTTTAGTTTCATTTGC -3'	60°C
	5'- TTTTCACAAACTGAAAAGCAGA -3'	
Foxo	5'- TAAGCATTTCTATACATGTGTGCG -3'	60°C
	5'- AAAACTGACAGGCTTGTTCCAC -3'	
Lingo1	5'- AGTGTCAGGAGAAGTTGCGG -3'	60°C
	5'- GATGAGCTTGGGACCACTCC -3'	
Vim	5'- AGGCAGTGGCTCCTAATCCT -3'	60°C
	5'- CCCCTGATCTAACGGACTCA -3'	
End1	5'- CAGGAGATTCCACAGCAGGG -3'	60°C
	5'- GCCTGAGTCAGACACGAACA -3'	
Klf9	5'- TTGCTGGGACTGTAGCGTTT -3'	60°C
	5'- CTAGGCAGTTACGCCCATCC -3'	