Lineage conversion induced by pluripotency factors involves transient passage through an iPSC stage

Ori Bar-Nur1–5, Cassandra Verheul1–5, Andrea G Sommer6,7, Justin Brumbaugh1–5, Benjamin A Schwarz1–5, Inna Lipchina1–5, Aaron J Huebner1–5, Gustavo Mostoslavsky6,7 & Konrad Hochedlinger1–5

Brief expression of pluripotency-associated factors such as Oct4, Klf4, Sox2 and c-Myc (OKSM), in combination with differentiation-inducing signals, has been reported to trigger transdifferentiation of fibroblasts into other cell types. Here we show that OKSM expression in mouse fibroblasts gives rise to both induced pluripotent stem cells (iPSCs) and induced neural stem cells (iNSCs) under conditions previously shown to induce only iNSCs. Fibroblast-derived iNSC colonies silenced retroviral transgenes and reactivated silenced X chromosomes, both hallmarks of pluripotent stem cells. Moreover, lineage tracing with an Oct4-CreER labeling system demonstrated that virtually all iNSC colonies originated from cells transiently expressing Oct4, whereas ablation of Oct4+ cells prevented iNSC formation. Lastly, an alternative transdifferentiation cocktail that lacks Oct4 and was reportedly unable to support induced pluripotency yielded iPSCs and iNSCs carrying the Oct4-CreER-derived lineage label. Together, these data suggest that iNSC generation from fibroblasts using OKSM and other pluripotency-related reprogramming factors requires passage through a transient iPSC state.

Transdifferentiation is defined as the experimentally induced direct conversion of one cell lineage into another. It is typically achieved by ectopic expression of small sets of cell type–specific transcription factors and has been used to convert fibroblasts into a variety of cell types, including myoblasts, cardiomyocytes, hepatocytes and neural cells1. Recently, a more general lineage conversion approach has been developed that uses transient expression of iPSC-inducing transcription factors. For example, transient expression of OKSM in mouse embryonic fibroblasts (MEFs) subjected to OKSM transdifferentiation protocols did not show activation of transgenic pluripotency reporters such as Nanog-GFP or Oct4-GFP2,4,7. Moreover, the kinetics of transdifferentiation was faster than that of induced pluripotency when using certain culture conditions, suggesting that iPSCs could not have formed2,4,7. Lastly, an alternative transdifferentiation cocktail, Brn4, Klf4, Sox2 and c-Myc (BKSM), has also generated iNSCs without apparent production of iPSCs3.

The use of OKSM in both induced pluripotency and transdifferentiation protocols raises the question of how these transcription factors induce such different cell fates. We therefore sought to compare reprogramming into iPSCs by OKSM and transdifferentiation to iNSCs by either OKSM or BKSM.

Reprogrammable MEFs (rep-MEFs), which carry the tetOP-OKSM allele in the Col1a1 locus, the M2rtTA locus in the Rosa26 locus and an EGFP reporter in the endogenous Oct4 locus (Oct4-GFP)10, exposed to doxycycline (dox), gave rise to iPSCs after 1–2 weeks in conventional embryonic stem cell (ESC) medium (serum and LIF; Fig. 1a,b). In contrast, treatment of rep-MEFs with dox in medium containing the NSC cytokines FGF and EGF, but lacking serum/LIF (NSC medium), yielded clusters of cells that morphologically and phenotypically resembled NSC colonies (Fig. 1a,b). For example, cellular projections emanating from iNSC-like colonies suggested formation of neurites, which are typically seen in brain-derived NSC cultures (Fig. 1b). We picked and expanded individual colonies in NSC medium, generating stable clonal cell lines. Immunofluorescence staining and microarray gene expression analysis of these cell lines revealed upregulation of key NSC markers, such as Sox1, Sox2, Nes (Nestin) and Glust (also known as Slc1a3), at levels similar to those of brain-derived NSCs, whereas MEF-related genes, including Col5a1 and Fbn5, were markedly downregulated (Fig. 1c,d and Supplementary Fig. 1a). Unsupervised clustering, correlation and functional annotation analyses of global mRNA expression data further showed that these iNSC-like cells are highly similar to brain-derived NSCs but different from MEFs or iPSCs/ESCs (Supplementary Fig. 1b–d). Cultures expanded from single iNSC-like colonies could differentiate into Tuj1+ neurons and GFAP+ astrocytes, demonstrating multilineage differentiation. For example, mouse embryonic fibroblasts (MEFs) subjected to OKSM transdifferentiation protocols did not show activation of transgenic pluripotency reporters such as Nanog-GFP or Oct4-GFP2,4,7. Moreover, the kinetics of transdifferentiation was faster than that of induced pluripotency when using certain culture conditions, suggesting that iPSCs could not have formed2,4,7. Lastly, an alternative transdifferentiation cocktail, Brn4, Klf4, Sox2 and c-Myc (BKSM), has also generated iNSCs without apparent production of iPSCs3.
differentiation potential (Fig. 1c). Together, these results show that OKSM expression with our reprogrammable system in the presence of NSC-inductive cytokines produces authentic iNSCs (OKSM-iNSCs) that exhibit multipotency, self-renewal potential and a high similarity to brain-derived NSCs.

We previously showed that the formation of stable iPSCs requires a minimum of 8–10 days of OKSM expression under conventional culture conditions \(^{11}\). To determine the minimal period of time needed to generate iNSCs, we induced rep-MEFs with dox in NSC medium for different lengths of time, followed by dox withdrawal before counting iNSC colonies at day 19. OKSM expression was required for at least 8 days to detect iNSC colonies, as determined by staining for Sox1 (Fig. 1f,g). The apparent similarity in the kinetics of iNSC and iPSC generation prompted us to ask whether iPSC colonies

---

**Figure 1** Generation of iPSCs and iNSCs from MEFs under neural culture conditions. (a) Experimental design to generate iPSCs and iNSCs in different culture conditions using reprogrammable MEFs (rep-MEFs) harboring the Col1α1-tetOP-OKSM and Rosa26-M2rtTA alleles and an Oct4-GFP reporter. (b) Representative image of Oct4-GFP+ iPSC colony generated in ESC medium (serum + LIF) (upper panel) and Oct4-GFP- iNSC-like colony generated in NSC medium (FGF/EGF/-LIF) (upper panel) showing typical spheroid morphology with neurite-like structures projecting from the center (bottom panel). Scale bars, 250 µM. (c) Representative Immunofluorescence images showing staining for indicated NSC markers in brain-derived NSCs and OKSM-iNSCs. MEFS served as a negative control. Scale bars, 100 µM. (d) Expression of NSC- or MEF-associated markers in the indicated cell lines based on microarray gene expression analysis. (e) Differentiation potential of OKSM-iNSCs into TuJ1+ neurons and Gfap+ astrocytes. Scale bars, 100 µM. (f) Representative Immunofluorescence image of a Sox1+ iNSC colony after 10 days of OKSM expression, followed by 9 days of dox-independent growth. Scale bars, 100 µM. (g) Contributions of NSC-medium-derived iPSCs to chimeras as indicated by agouti coat color (upper image). Germline offspring (white arrowheads) obtained from a male chimera (red arrowhead) (lower image). (k) Potential of sorted Ssea1+/Epcam+ and Ssea1+/Epcam- intermediates after 6 days of OKSM expression to produce iPSCs in ESC medium or iNSCs in conventional NSC medium, respectively. For each replicate, 10 x 10^5 cells were plated (n = 3 independent replicates; error bars, s.e.m.; *P < 0.05). (l) Representative images of an Oct4-GFP+ colony (top image) and a Sox1+ iNSC colony (bottom image) obtained from sorted Ssea1+/Epcam+ intermediates (day 6) in NSC medium. White arrowheads indicate an Oct4-GFP colony (top) or clusters of Sox1+ cells (bottom). Scale bars, 500 µM (Oct4-GFP) or 250 µM (Sox1). (m) Quantification of Oct4-GFP+ colonies, Oct4-GFP+/iNSC hybrid colonies and iNSC colonies at the indicated time points after dox withdrawal (w/d). For each replicate, 3 x 10^4 cells were used (n = 3 independent replicates for each time point; error bars, s.e.m.). (n) Representative image of iPSC/iNSC hybrid colonies detected during the transdifferentiation protocol. Note the dome-shaped iPSC-like colonies in the center and emanating neurites (indicated by black and red arrowheads, respectively). Bottom images show patches of Oct4-GFP reporter expression (white arrowheads) within the same colony as shown above. Scale bars, 250 µM. Smaller images above show magnification of neurites (top) and a representative Oct4-GFP colony (bottom).
could form in NSC medium, a condition that is incompatible with long-term culture of iPSCs. Notably, forced expression of OKSM in rep-MEFs consistently gave rise to rare Oct4-GFP+ iPSC-like cells (0.2% of total cell population) in this differentiation-stimulating condition (Fig. 1h.i). Unlike brain-derived NSCs or OKSM-iNSCs, these colonies co-expressed Oct4 and Nanog, ruling out the possibility that they were primitive neuroectodermal cells that continued to express Oct4 after exit from pluripotency12 (Supplementary Figs. 1e and 2a,b). The iPSC-like colonies could be stably propagated in ESC medium in the absence of dox (i.e., after discontinuation of exogenous transgene expression) and expressed the pluripotency-associated markers alkaline phosphatase, Pecam1, Epcam and Nanog at levels that were similar to those of ESCs, indicating self-renewal capacity and faithful molecular reprogramming to pluripotency (Supplementary Fig. 2c–e). Expanded iPSC-like clones contributed to coat-color chimeric mice and to their offspring after injection into blastocyst embryos—a rigorous measure of pluripotency (Fig. 1j). We conclude that forced OKSM expression overrides the requirement for serum and LIF/Stat3 signaling and counteracts differentiation-inducing signals (EGF, FGF), generating authentic, germline-competent iPSCs under adverse growth conditions.

These data raise the possibility that certain previously published transdifferentiation protocols using OKSM overexpression might have, in fact, generated transient iPSCs that then differentiated into other cell lineages in response to extracellular cues. To test this hypothesis, we first examined the order of events in transdifferentiation by tracking cell surface markers by flow cytometry. During reprogramming toward iPSCs, MEFs first downregulate the fibroblast-associated marker Thy1 and then upregulate the pluripotency-associated marker Ssea1 at day 3, followed by Epcam expression by day 6 (refs. 11,13). We observed a similar downregulation of Thy1 and an upregulation of Epcam in a fraction of cells (~5%) at day 6 of OKSM expression under INSC transdifferentiation conditions, although established INSCs and brain-derived NSCs do not express Epcam (Supplementary Fig. 2f,g). Analysis of Ssea1 expression was not informative because it is also expressed by NSCs14. These results show that OKSM-induced lineage conversion generates similar phenotypic intermediates as OKSM-induced pluripotency does, despite very different extracellular stimuli.

Sorted Epcam+ intermediates obtained after 6 days of OKSM expression in ESC medium give rise to rare dox-independent iPSC colonies, whereas Epcam– intermediates do not13. To prospectively determine whether these subpopulations can give rise to iNSCs, we sorted day-6 Ssea1+/Epcam–, Ssea1+/Epcam+ and Ssea1+/Epcam– intermediates generated in ESC medium and cultured them in either ESC or NSC medium without dox (Fig. 1k,l). Only Ssea1+/Epcam– intermediates could produce transgene-independent iPSCs or iNSCs (Fig. 1k,l and data not shown). These findings show that iPSCs and iNSCs originate from the same rare subpopulation of Ssea1+/Epcam– cells that emerge shortly after OKSM expression.

Considering that Epcam+ intermediates typically give rise to Oct4-GFP+ iPSCs under ESC culture conditions13, we next determined the fraction of nascent iNSC colonies that expressed the pluripotency gene Oct4 in situ over time using the same Oct4-GFP knock-in reporter. The majority of colonies forming at early time-points of OKSM expression (days 10–12) in NSC medium indeed contained small clusters of Oct4-GFP+ cells, whereas colonies at intermediate time points (days 17–22) consisted of Oct4-GFP+ clusters, iNSC-like clusters or a combination of both (Fig. 1m,n). Oct4-GFP expression was entirely silenced after picking and expansion of these iNSC-like colonies in NSC medium without dox (data not shown).

To accurately determine how many colonies expressed Oct4 before becoming iNSCs, we used cells carrying a 4-hydroxytamoxifen (4-OHT)-inducible Oct4-CreER allele in combination with a Rosa26-loxSTOPlox-EYFP (R26-Isl-EYFP) reporter cassette15 (Fig. 2a). We first determined the specificity and sensitivity of this lineage tracing system for labeling pluripotent cells. To demonstrate that primitive Oct4+ neuroectodermal cells12 are not labeled with this system in vivo, we treated pregnant females harboring Oct4-CreER; R26-Isl-EYFP bi-transgenic embryos with tamoxifen at E8.5 and examined resultant fetuses at E13.5. We did not detect EYFP expression in the brains of five E13.5 fetuses or in derivative NSC cultures (Fig. 2b and Supplementary Fig. 3a,b). However, we noticed strong EYFP signal in primordial germ cells in the genital ridge, which are known to express Oct4 at E8.5 (Fig. 2b). To ensure that culture of brain-derived NSCs did not spuriously activate the Oct4-CreER; R26-Isl-EYFP reporter, we derived NSCs from the brains of E13.5 fetuses with this genotype and exposed cells to 4-OHT. These NSC lines remained EYFP– regardless of whether 4-OHT was added to the culture (Supplementary Fig. 3c).

Next, to confirm that short-term expression of OKSM does not directly activate the endogenous Oct4 locus, we transduced Oct4-CreER; R26-Isl-EYFP MEFs with lentiviral vectors co-expressing tetOP-OKSM and M2-rtTA for 2 or 4 days before measuring EYFP expression by flow cytometry. No notable EYFP reporter signal was detected at these early time points (Supplementary Fig. 4a). We also failed to detect Oct4-GFP reporter expression after 2 and 4 days of OKSM expression (Supplementary Fig. 4b). In contrast, longer induction of OKSM in Oct4-CreER; R26-Isl-EYFP MEFs in ESC or NSC media generated EYFP+ iPSCs in a 4-OHT-dependent manner (Fig. 2c–e and Supplementary Fig. 5a and data not shown). EYFP+ iPSCs derived under NSC conditions expressed Nanog and Oct4, and they contributed to midgestation chimeras after injection into blastocysts, demonstrating pluripotency (Fig. 2c–e and Supplementary Fig. 5b,c). Based on these controls, we feel confident that this lineage tracing system and our culture conditions had high specificity and sensitivity, allowing us to determine whether nascent iNSCs went through a transient Oct4+ state. All recovered Sox1+ iNSC colonies derived from OKSM-expressing Oct4-CreER; R26-Isl-EYFP MEFs were EYFP+ in the presence of 4-OHT, whereas no EYFP+ iNSC colonies were detected in the absence of 4-OHT (Fig. 2f,g). EYFP+ iNSC clusters could be subdivided from single cells and propagated for over ten passages while maintaining NSC marker expression, indicating self-renewal capacity (Fig. 2h,i and Supplementary Fig. 6a–c). These results show that all iNSCs generated with this reprogramming system and transdifferentiation protocol pass through a transient Oct4+ state.

We next applied the Oct4-CreER lineage tracing system to an alternative transdifferentiation protocol that reportedly gives rise to iNSCs but not iPSCs after only 4–5 days of Oct4 expression, coupled with constitutive expression of Sox2, Klf4 and c-Myc2. We were able to derive Sox1+ iNSC lines after only 4–5 days of OKSM expression, followed by 14–21 days of OKSM-independent growth, similar to this previous report (Fig. 2j and Supplementary Fig. 7a). All examined Sox1+ iNSC colonies were also EYFP+ (Fig. 2j and Supplementary Fig. 7a). This alternative transdifferentiation protocol uses a modified NSC medium containing knockout serum replacement (KOSR) and small amounts of serum, which are not included in conventional NSC medium7. We surmise that the presence of ascorbic acid in KOSR stimulates the formation of iNSCs, whereas in ESC culture conditions, as was shown previously10, thus explaining the faster appearance of iPSCs and derivative NSCs. We used the modified transdifferentiation protocol in most subsequent experiments because of the higher recovery rate of iNSCs.
To rule out the possibility that our visual quantification of EYFP-marked cells missed single iNSCs or small iNSC colonies that underwent direct transdifferentiation, we repeated the experiment with MEFs harboring the Oct4-CreER allele in combination with a R26-IsI-DTA (diphtheria toxin fragment A) allele (Supplementary Fig. 7b). Expression of DTA in 4-OHT-induced cells undergoing transdifferentiation or reprogramming is expected to ablate nearly all Oct4-expressing cells (Supplementary Fig. 7b–d). Using this system, we did not detect any Sox1+ iNSC colonies or Sox1+ single cells in 4-OHT-treated cultures whereas we did in untreated cultures (Fig. 2k), confirming that virtually all iNSCs passed through an Oct4+ state.

To complement our Oct4 lineage labeling and ablation experiments with independent assays of pluripotency, we first examined retroviral transgene silencing in iNSC cultures. Unlike most somatic cells, including NSCs, iPSCs efficiently silence retroviral transgenes owing to the activity of de novo DNA methyltransferases11. We infected rep-MEFs with a retroviral vector constitutively expressing tdTomato (Fig. 3a), purified tdTomato+ rep-MEFs by fluorescence-activated cell sorting (FACS), and exposed them to dox and NSC medium (Fig. 3a,b). Approximately 90% of iNSC colonies were tdTomato−, and ~10% were tdTomato+. (Fig. 3c). We believe that the latter colonies represent nascent hybrid iPSC-iNSC colonies that have not yet undergone complete retroviral silencing (see also Fig. 1m). Notably, culture of sorted tdTomato+ MEFs in NSC medium without dox did not elicit silencing of retroviral transgenes (Supplementary Fig. 8).

We previously showed that reprogramming of female fibroblasts to iPSCs reactivates the silenced X chromosome, and subsequent
The silencing of retroviral transgenes and X chromosome reactivation during iNSC induction. (a) Experimental design to assess retroviral silencing during iNSC formation. (b) Representative images comparing tdTomato expression in MEFS, iPSCs, OKSM-iNSCs and brain-derived NSCs. Note that tdTomato nonreprogrammed MEFS surround the iPSC and iNSC colonies. Scale bars, 250 μM. (c) Quantification of iNSC colonies expressing tdTomato. For each replicate, 3 x 10^4 cells were used (n = 3 independent replicates; error bars, s.e.m.; **P < 0.005). (d) Experimental design to follow X chromosome inactivation status during iNSC formation. Tail fibroblasts from a female mouse heterozygous for an X-linked CMV-GFP reporter (mix of XiGFPXa and XaGFPXi cells) were sorted twice for cells that carried the GFP transgene on the inactive X chromosome (XiGFPXa). These cells were reprogrammed for 4–5 days in modified NSC medium, after which dox was withdrawn and GFP expression was assessed in emerging Sox1+ iNSC colonies 2–4 weeks later. (e) Representative Immunofluorescence images of Sox1+ iNSC colonies that show reactivation of the silent XiGFP reporter. Scale bars, 250 μM. (f) Quantification of Sox1+ iNSC colonies that reactivated the silenced X chromosome as determined by GFP reporter reactivation. All examined Sox1+ iNSC colonies showed patchy GFP expression. For each replicate, 3 x 10^4 cells were used (n = 3 independent replicates; error bars, s.e.m.; ***P < 0.0005). (g) Representative Immunofluorescence images showing Nanog knockout Sox1+ iNSC colony generated in the presence of ascorbate. Nanog-deficient iNSCs express the constitutive CAG-GFP marker, confirming their origin from Nanog-mutant MEFS. Scale bars, 100 μM. (h) Quantification of Sox1+ iNSC colonies generated from Nanog-deficient MEFS with and without ascorbate. For each replicate, 3 x 10^4 cells were used (n = 3 independent replicates; error bars, s.e.m.; **P < 0.005).

The percentage of GFP+ cells varied among established iNSC clones, confirming reactivation of the silenced X chromosome in seven clonally derived iNSC cultures by quantitative RT-PCR for an X-linked marker that distinguishes the maternal (Mus musculus musculus) and paternal (Mus musculus castaneous) X chromosome, respectively (Supplementary Fig. 9a). These data reinforce the interpretation that iNSCs transition through a pluripotent intermediate state with concomitant X reactivation before undergoing differentiation.

Next, we investigated whether interfering with genes involved in iPSC reprogramming affects the generation of iNSCs. Nanog is expressed in pluripotent cells but not in iNSCs (Supplementary Fig. 1e), and Nanog deficiency leads to a severe reduction in iNSC formation that can be rescued by exposure of cells to ascorbic acid19. Consistent with this finding, Nanog knockout MEFS gave rise to 6–8 times more Sox1+ iNSC colonies in the presence of ascorbic acid compared to control cultures lacking ascorbate (Fig. 3g.h). Ascorbate acts synergistically with the GSK3β inhibitor CHIR-99021 to increase iPSC formation efficiency20. Formation of iNSCs from wild-type MEFS was enhanced by exposure to both molecules compared to ascorbate alone, further demonstrating that iNSC generation and iPSC formation respond to the same small molecules (Supplementary Fig. 10a,b).

Finally, we studied a different iNSC transdifferentiation gene cocktail that reportedly does not yield iPSCs in ESC or NSC medium, namely Brn4, Klf4, Sox2 and c-Myc (BKS M)3. To ensure homogeneous expression of all four factors, we engineered a polycistronic
Figure 4 An alternative transdifferentiation cocktail generates iNSCs and iPSCs via an Oct4+ intermediate stage. (a) Replacement of the Oct4 gene within the Oct4-Klf4-Sox2-cMyc (OKSM cassette) by the neuronal transcription factor Brn4 (BKSM cassette). (b) Experimental design to generate iPSCs and iNSCs using the BKSM cassette from MEFs harboring a Sox2-GFP knock-in allele in ESC or NSC culture conditions. (c) Immunofluorescence of typical NSC markers in an expanded BKSM-iNSC clone. Scale bars, 100 µm. (d) Expression of selected NSC-associated genes in indicated cell lines based on microarray data. Microarray analysis was performed on expanded clonal iNSC cultures. (e) Immunofluorescence for indicated pluripotency markers in expanded BKSM-iPS clone. Scale bars, 250 µM. (f) Dendrogram based on microarray analysis of indicated samples. Note that BKSM-iNSCs and OKSM-iNSCs cluster with brain-derived NSCs whereas BKSM-iPSCs cluster with mouse ESCs. (g) Contribution of BKSM-iPSCs to an adult chimera as shown by agouti coat color. (h) Representative images of EYFP+ BKSM-iNSC colonies with emanating neurites generated from MEFs in conventional NSC medium. Scale bars, 500 µm for INSC colony (left) and 100 µm for neurites (right). Autofluor., autofluorescence control. (i) Quantification of EYFP+ BKSM-iNSC colonies, demonstrating that all of the examined BKSM-iNSC colonies transitioned through an Oct4+ state. For each replicate, 5 x 10^5 cells were used (n = 3 biological replicates; error bar, s.e.m.; ** P < 0.005). (j) Representative images showing EYFP+ cells in a cluster of beating cardiomyocytes (red arrowheads). Scale bars, 250 µm. (k) Quantification of EYFP+, EYFP− and mixed EYFP+/EYFP− contracting induced cardiomyocyte colonies derived from Oct4-CreER; R26-Isl-EYFP MEFs upon short OKSM expression in cardiac transdifferentiation conditions. For each replicate, 5 x 10^5 cells were used (n = 3 biological replicates; error bars, s.e.m.; ** P < 0.005). (l) Summary.

lentiviral vector encompassing the coding regions for Brn4, Sox2, Klf4 and c-Myc under the control of a dox-inducible promoter, akin to the OKSM/STEMCCA system8 (Fig. 4a,b and Supplementary Fig. 11a–c). Infection of Sox2-GFP MEFs with lentiviral constructs expressing tetOP-BKSM and M2rtTA gave rise to Sox2-GFP+ iNSC-like colonies in the presence of dox and neural-inducing cytokines, in agreement with previous work3 (Supplementary Fig. 12a). Of note, formation of iNSC clusters in conventional NSC medium required at least 12 days of BKSM expression, rather than 8 days with the OKSM cocktail, and we typically obtained ~10 fewer colonies (data not shown). BKSM-iNSCs could be propagated without dox, demonstrating independence from exogenous factor expression and self-renewal potential (Fig. 4c and Supplementary Fig. 12a and data not shown). Accordingly, BKSM-iNSCs activated NSC markers, suggesting faithful molecular reprogramming toward an NSC-like state (Fig. 4c,d and Supplementary Fig. 12b,c). Global gene expression analysis documented a high degree of similarity between OKSM-iNSCs, BKSM-iNSCs and brain-derived NSCs, regardless of their origin from either fibroblasts or fetal brains, whereas these samples were quite different from the starting MEF population (Fig. 4f and Supplementary Fig. 12b,c).

In contrast to previous observations1, the BKSM combination generated iPSC-like colonies in both ESC and NSC media (Supplementary Fig. 13a and data not shown). As with OKSM-iPSCs derived in NSC medium, BKSM-iPSC-like clones activated key endogenous pluripotency genes, including Nanog, Oct4 and Zfp42, when replated in ESC medium without dox, indicating transgene-independent self-renewal (Fig. 4e and Supplementary Fig. 13b). The global transcriptional profile of BKSM-iPS-like cells was highly similar to those of ESCs and OKSM-derived iPSCs (Fig. 4f and Supplementary Fig. 13b,c). Moreover, BKSM-iPSCs supported the formation of teratomas and the generation of fetal and adult chimeras after blastocyst injection, demonstrating pluripotency (Fig. 4g and Supplementary Fig. 13d,e). Like BKSM-iNSCs, BKSM-iPSCs were recovered at a lower frequency and with delayed kinetics compared to OKSM-iPSCs (data not shown), further supporting the notion of a linked process. Notably, infection of Oct4-CreER; R26-Isl-EYFP MEFs with lentiviral constructs
expressing M2rtTA and tetOP-BKSM yielded exclusively EYFP+ iNSC colonies in the presence of 4-OHT, dox, EGF and FGF (Fig. 4h,i). These BKSM-iNSC colonies continued to grow without dox and could be converted into iPSCs by readdiction of dox (Supplementary Fig. 14a–c). Together, these results suggest that MEFs undergoing transdifferentiation into iNSCs upon polycistronic BKSM expression inevitably pass through a transient Oct4+ state.

A key question is whether our observations extend to other types of lineage conversion using either pluripotency or lineage-specific factors. We examined activation of the Oct4-CreER, R26-Isl-EYFP reporter during OKSM-induced cardiomyocyte generation and 

Amel11 Brn2/Myt1l-driven induced neuron (iN) formation from MEFS. OKSM-induced cardiomyocyte clusters exhibited rhythmic contractions, expressed cardiac troponin and activated the Oct4-CreER; R26-Isl-EYFP reporter in nearly all cases (Fig. 4j,k and Supplementary Fig. 15a,b). By contrast, transdifferentiation of MEFS into iN-like cells did not activate the EYFP reporter, suggesting that passage through an Oct4+ state may be specific to the use of pluripotency-related transcription factors (Supplementary Fig. 15c–e).

Here we show that pulsed expression of OKSM or BKSM causes cells to pass through a transient Oct4+ state that silences retroviral transgenes and reactivates silenced X chromosomal genes in female somatic cells. We further demonstrate that forced OKSM and BKSM expression produces bona fide iPSCs under culture conditions that are incompatible with long-term culture of established pluripotent cell lines. These results imply that transdifferentiation protocols using pluripotency-associated factors often, if not always, involve reprogramming into short-lived iPSC that rapidly differentiate into various lineages as determined by the extracellular environment (Fig. 4l). Our results are consistent with previous findings that retroviral transgenes used to convert fibroblasts into iNSCs by brief OKSM expression are predominately silenced3,7, whereas induced neural progenitor cells generated by forced expression of lineage-specific factors22 remain transgene-dependent. Our data might also explain the recent observation that cells generated through transdifferentiation with pluripotency factors exhibit chromosomal aberrations typically associated with pluripotent stem cell lines23. Of note, an independent study, which appears in this issue, arrives at very similar conclusions as ours using a different transdifferentiation protocol and lineage tracing strategy24.

Transient iPSC formation was ruled out in previous transdifferentiation studies owing to (i) the lack of Nanog-GFP or Oct4-GFP signal, (ii) the inability to derive iPSCs in culture conditions lacking LIF and containing differentiation-associated cytokines and (iii) the different kinetics of direct lineage conversion and iPSC formation. We believe that the Nanog-GFP reporter may not be sufficiently sensitive to detect rare reprogramming events in heterogeneous bulk cultures. Consistent with this notion, previous studies recorded transient Nanog mRNA upregulation between days 6 and 8 of transdifferentiation by qPCR analysis even though no reporter signal was detected24–4. Moreover, transdifferentiation experiments have often been performed under optimized culture conditions compared to control iPSC experiments, which may inadvertently facilitate fast, yet transient, iPSC formation. Indeed, several transdifferentiation supplements (e.g., CHIR99021, BMP4, bFGF, KOSR/ascorbic acid) strongly enhance and accelerate iPSC formation16,20,25. Accordingly, we could reduce the minimum number of days of OKSM expression required to derive stable iNSCs from 8 to 4 days by modifying the culture medium.

Our results show that the absence of LIF and the presence of differentiation-inducing cytokines (e.g., FGF, EGF) are not sufficient to block iPSC formation. Thus, forced transcription factor expression can effectively compensate for the lack of growth factors important for iPSC self-renewal and counteract culture conditions that normally trigger ESC/iPSC differentiation. This finding is reminiscent of the observation that forced expression of Stat3 or Nanog bypasses the requirement for LIF signaling during ESC derivation and iPSC formation26,27. We believe that our single-copy transgenic reprogramming system allowed us to increase the efficiency of reprogramming through homogeneous expression of the four factors, facilitating mechanistic dissection of critical intermediate stages. Indeed, iNSC generation in some previous studies was extremely inefficient (1–3 iNSC colonies per infection with individual B, K, S and M vectors; 11 iNSC colonies per 130,000 infected cells or 0.008% using O, K, S and M vectors) and lengthy (4–6 weeks with BKSM expression; 3 weeks with OKSM)3,7, which may have been due to inefficient retroviral infection of fibroblasts or a suboptimal expression stoichiometry of all four transgenes. The low efficiency of iNSC formation may also explain why iPSCs were not detected upon BKSM overexpression, whereas our use of a polycistronic, dox-inducible expression vector gave rise to iPSCs in both ESC and NSC conditions, albeit at much lower efficiency than OKSM-mediated reprogramming. Considering the differences in transgenic and vector design between our study and previous reports, we cannot rule out that alternative reprogramming methods may induce direct lineage conversion at some frequency.

Pluripotency factor–driven transdifferentiation has been proposed as a safer and more efficient alternative to induced pluripotency for therapeutic applications as it does not require lengthy iPSC differentiation protocols and avoids potential teratoma formation from rare undifferentiated cells. Our study highlights the practical importance of dissecting the mechanisms underlying transdifferentiation to determine whether cells transit through a pluripotent state. Although transient passage through a pluripotent state is not expected to affect in vitro applications of this technology, it will be important to exclude the presence of residual iPSCs in potential clinical settings.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. NCBI Gene Expression Omnibus (GEO): microarray data have been deposited under accession number GSE65951.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank members of the Hochedlinger laboratory for critical evaluation and discussion of this manuscript. We thank E. Apostolou and S. Cheloufi for helpful comments and N. Maherali for suggesting the idea of X chromosome reactivation. We are also grateful to B. Payer and I. Lee for providing tail-tip fibroblasts carrying an X-linked GFP reporter and to A. Brack for sharing Rosa26-Isl-DTA mice. We are grateful to L. Prickett, M. Weglarz and K. Folling-Danahau at the MGH/HSCI flow cytometry core for their constant assistance and support. O.B.-N. is supported by a Gruss-Lipper postdoctoral fellowship from the EGL foundation. B.A.S. has been supported by a T32 (5-T32-CA-9216-33) grant and through a Postdoctoral Fellowship Award by the MGH Executive Committee on Research Fund for Medical Discovery. J.B. and I.L. are supported by Ruth L. Kirschstein F32 Post-doctoral Fellowships from the US National Institutes of Health (NIIH) (1F32HD078029-01A1 to J.B. and 1F32HD079225-01A1 to I.L.). Support to K.H. is from the NIH (R01HD058013) and Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

O.B.-N. and K.H. conceived the experiments, interpreted results and wrote the manuscript. O.B.-N. conducted all reprogramming experiments, performed statistical analysis, bioinformatics analysis of expression data and generated figures; C.V. assisted in most reprogramming experiments and interpretation of results.
A.G.S. and G.M. generated the BKSM lentiviral cassette; J.B. generated chimeric animals by blastocyst injections; B.A.S. conducted reprogramming experiments from FACS-sorted intermediates; L.I. performed allele-specific real-time PCR for an X-linked marker gene; A.J.H. confirmed specificity of the Oct4-CreER lineage tracing system in vivo.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

LETTERS

1. Vierbuchen, T. & Wernig, M. Molecular roadblocks for cellular reprogramming. Mol. Cell 47, 827–838 (2012).
2. Efe, J.A. et al. Conversion of mouse fibroblasts into cardiomyocytes using a direct reprogramming strategy. Nat. Cell Biol. 13, 215–222 (2011).
3. Han, D.W. et al. Direct reprogramming of fibroblasts into neural stem cells by defined factors. Cell Stem Cell 10, 465–472 (2012).
4. Kim, J. et al. Direct reprogramming of mouse fibroblasts to neural progenitors. Proc. Natl. Acad. Sci. USA 108, 7838–7843 (2011).
5. Li, K. et al. Small molecules facilitate the reprogramming of mouse fibroblasts into pancreatic lineages. Cell Stem Cell 14, 228–236 (2014).
6. Margariti, A. et al. Direct reprogramming of fibroblasts into endothelial cells capable of angiogenesis and reendothelialization in tissue-engineered vessels. Proc. Natl. Acad. Sci. USA 109, 13793–13798 (2012).
7. Thier, M. et al. Direct conversion of fibroblasts into stably expandable neural stem cells. Cell Stem Cell 10, 473–479 (2012).
8. Zhu, S. et al. Mouse liver repopulation with hepatocytes generated from human fibroblasts. Nature 508, 93–97 (2014).
9. Orkin, S.H. & Hochedlinger, K. Chromatin connections to pluripotency and cellular reprogramming. Cell 145, 835–850 (2011).
10. Stadtfeld, M., Maherali, N., Borkent, M. & Hochedlinger, K. A reprogrammable mouse strain from gene-targeted embryonic stem cells. Nat. Methods 7, 53–55 (2010).
11. Stadtfeld, M., Maherali, N., Breault, D.T. & Hochedlinger, K. Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell 2, 230–240 (2008).
12. DeVeale, B. et al. Oct4 is required ~E7.5 for proliferation in the primitive streak. PLoS Genet. 9, e1003957 (2013).
13. Polo, J.M. et al. A molecular roadmap of reprogramming somatic cells into iPS cells. Cell 151, 1617–1632 (2012).
14. Capela, A. & Temple, S. LeX/ssea-1 is expressed by adult mouse CNS stem cells, identifying them as nonependymal. Neuron 35, 865–875 (2002).
15. Greder, L.V. et al. Analysis of endogenous Oct4 activation during induced pluripotent stem cell reprogramming using an inducible Oct4 lineage label. Stem Cells 30, 2596–2601 (2012).
16. Esteban, M.A. et al. Vitamin C enhances the generation of mouse and human induced pluripotent stem cells. Cell Stem Cell 6, 71–79 (2010).
17. Maherali, N. et al. Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell 1, 55–70 (2007).
18. Sommer, C.A. et al. Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. Stem Cells 27, 543–549 (2009).
19. Schwarz, B.A., Bar-Nur, O., Silva, J.C. & Hochedlinger, K. Nanog is dispensable for the generation of induced pluripotent stem cells. Curr. Biol. 24, 347–350 (2014).
20. Bar-Nur, O. et al. Small molecules facilitate rapid and synchronous iPSC generation. Nat. Methods 11, 1170–1176 (2014).
21. Vierbuchen, T. et al. Direct conversion of fibroblasts to functional neurons by defined factors. Nature 463, 1035–1041 (2010).
22. Lujan, E., Chanda, S., Ahlenius, H., Sudhof, T.C. & Wernig, M. Direct conversion of mouse fibroblasts to self-renewing, tripotent neural precursor cells. Proc. Natl. Acad. Sci. USA 109, 2527–2532 (2012).
23. Weissbein, U., Ben-David, U. & Benvenisty, N. Virtual karyotyping reveals greater chromosomal stability in neural cells derived by transdifferentiation than those from stem cells. Cell Stem Cell 15, 687–691 (2014).
24. Maza, I. et al. Transient acquisition of pluripotency during somatic cell transdifferentiation with iPSC reprogramming factors. Nat. Biotechnol. doi:10.1038/nbt.3270 (22 June 2015).
25. Chen, J. et al. Rational optimization of reprogramming culture conditions for the generation of induced pluripotent stem cells with ultra-high efficiency and fast kinetics. Cell Res. 21, 884–894 (2011).
26. Chambers, I. et al. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. Cell 113, 643–655 (2003).
27. van Oosten, A.L., Costa, Y., Smith, A. & Silva, J.C. JAK/STAT3 signalling is sufficient identifying them as nonependymal. Neuron 35, 865–875 (2002).
**ONLINE METHODS**

**Cell culture.** Mouse embryonic fibroblasts (MEFs) were cultured in DMEM supplemented with 1% Glutamax, 1% nonessential amino acids (NEAA), 1% penicillin-streptomycin, 0.5% β-mercaptoethanol (all from Life Technologies) and 10% FBS (Gibco). Mouse embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs) were grown on irradiated MEFs (GlobalStem) in KO-DMEM medium (Life Technologies) supplemented with 15% FBS, 1% Glutamax, 1% NEAA, 1% penicillin-streptomycin, 1% β-mercaptoethanol and 1,000 unit/ml LIF (purified in-house). Normal stem cells (NSCs) were cultured in DMEM/F12 (Life Technologies) supplemented with 1X27 and 1X22 cell supplements (Life Technologies), 0.01% Fraction V BSA 7.5% solution (Life Technologies), 1% penicillin-streptomycin and 10 ng/ml BFGF (R&D) and 20 ng/ml mouse EGF (Life Technologies) (“conventional NSC medium”). Testing for Mycoplasma was routinely performed.

**Generation of iNSCs using conventional NSC medium.** Conventional NSC medium contained DMEM/F12 (Life Technologies) supplemented with 1X27 and 1X22 cell supplements (Life Technologies), 0.01% Fraction V BSA 7.5% solution (Life Technologies), 1% penicillin-streptomycin and 10 ng/ml BFGF (R&D) and 20 ng/ml mouse EGF (Life Technologies); this media formulation was previously used to generate iNSCs. We termed this medium “conventional NSC medium” because it does not contain LIF, serum or serum replacement components. For the generation of OKSM-iNSCs, reprogrammable MEFs (rep-MEFs) from mice carrying a tet-OPOP-OKSM allele in the ColIα1 locus and an M2rtTA transgene in the Rosa26 locus as well as an EGF transgene knocked into one allele of the Oct4 locus were used for some of the experiments. Alternatively, MEFs or tail tip fibroblasts (TTFs) were reprogrammed into iNSCs by infection with lentiviral vectors expressing M2rtTA in combination with either tetOP-OKSM or tetOP-BKSM cassettes. In conventional NSC medium, at least 8 days (OKSM) or 12 days (BKSM) of dox exposure, followed by 7–21 days (OKSM) and 12–24 days (BKSM) for dox-independent growth, were usually required to generate iNSCs. For most experiments, either 3 × 10^4 cells were plated per well of a 6-well plate or 5 × 10^5 cells per 10-cm plate in MEF medium a day before dox administration. Serum was added for the first 3 days of transdifferentiation to reduce cell death. Small molecules were routinely performed.

**Generation of iNSCs using modified NSC medium.** Modified NSC medium contained DMEM/F12, 2% FBS, 8% KOSR (Gibco), 1X22, 1% Glutamax, 1% NEAA and 0.5% β-mercaptoethanol (all purchased from Life Technologies) and 1,000 U/ml LIF (purified in-house). This medium was previously used to generate iNSCs. We termed this NSC medium “modified” because it contains low levels of serum, KOSR and LIF. To achieve iNSC induction with this medium, we transduced MEFs or TTFs with lentiviral vectors expressing M2rtTA and tetOP-OKSM. For each replicate, 5 × 10^3 cells were plated in a 10-cm dish a day before dox administration. OKSM expression was induced by adding dox for 4–5 days in modified NSC medium, followed by 2–4 weeks of dox-independent growth to select for transgene-independent iNSCs. The generation of iNSCs and iPSCs with modified NSC medium required a shorter duration of OKSM expression compared to conventional NSC medium (4 vs. 8 days). Moreover, low-passage MEFs (passage 3 and below) were reprogrammed more efficiently and faster than high-passage MEFs. A concentration of 2 µg/ml doxycycline (dox; Sigma) was used for all transdifferentiation and reprogramming experiments.

**Generation of iPSCs in NSC medium.** Reprogrammable MEFs or MEFs infected with lentiviral vectors expressing M2rtTA and tetOP-OKSM were reprogrammed in conventional NSC medium supplemented with dox for 8–12 days to initiate OKSM expression, before switching to ESC medium without dox. Oct4-GFP colonies were first observed in NSC medium at around days 10–12; however, these cells tended to differentiate into iNSCs after a few days without dox when kept in NSC medium. Switching to ESC medium (serum/LIF) prevented this rapid differentiation and allowed propagation of dox-independent iPSC colonies.

**Generation of Oct4-CreER; Rosa26-loxSTOPlox-EYFP MEFs.** Oct4-CreER mice, termed B6(SJL)-Pou5f1tm1.1(Cre/Esr1*)Yseg/I, were purchased from Jackson Laboratory (stock number 016829). Oct4-CreER MEFs were crossed with Rosa26-lox-STOP-lox-EYFP mice28 to produce transgenic reporter MEFs and NSCs30. Genotyping of the Oct4-CreER allele was performed as previously described. Cells with the correct genotype were infected with lentiviral vectors harboring M2rtTA and tetOP-OKSM alleles and reprogrammed in either modified or conventional NSC medium. MEFs carrying the same genotype were used to derive iPSCs in conventional and modified NSC medium (as shown in Fig. 2). 4-hydroxytamoxifen (4-OHT) was used at a concentration of 100 nM for all experiments and was continuously added to the culture medium during and after dox treatment to ensure labeling of Oct4+ cells at all stages of transdifferentiation. For in vivo induction, 1 mg tamoxifen (Sigma) was diluted in corn oil (10 mg/ml) and injected into the peritoneum of a pregnant mouse at embryonic day E8.5. The brains and gonads of recovered embryos were dissected and analyzed at E13.5.

**Derivation of NSCs from fetuses.** Brain-derived NSCs were derived from the heads of E13.5 fetal mice as previously described. We used conventional NSC medium for expansion and culture of NSCs.

**Scoring for X reactivation during iNSC formation.** Tail tip fibroblasts (TTFs) obtained from female mice carrying a CMV-GFP transgene on one X chromosome were infected with lentiviral vectors carrying the M2rtTA and tetOP-OKSM constructs without dox. Infected cells were sorted twice on a FACSariaI machine (BD Biosciences) to select for the GFP-negative fraction (sort purity 99.9–100% after two sorting rounds). For each replicate, 3 × 10^5 cells were sorted a day before dox administration and treated in a 6-well plate. iNSC induction was performed in modified NSC medium in the presence of dox for 4–5 days, followed by 2–4 weeks of dox-independent growth before scoring for the fraction of GFP+ Sox1+ iNSC colonies.

**Retroviral silencing during iNSC induction.** Rep-MEFs were infected with retroviral vectors carrying the Pmxs-tdTomato construct. Cells were then sorted on a FACSariaI machine (BD Biosciences) to isolate the tdTomato cell-fraction. For each replicate, 3 × 10^5 cells were sorted per well of a 6-well plate and cultured in ESC or NSC media supplemented with dox for 12 days. Scoring of iPSC or iNSC colonies was performed 7–10 days after dox withdrawal. iNSC induction was performed in the presence of serum for the first three days of OKSM expression to minimize cellular senescence. This modification is in accordance with other reports that used low serum during iNSC induction.31,32 Fetal brain–derived NSCs were infected with pMXs-tdTomato retroviral vectors, sorted for the tdTomato fraction and cultured for an additional 4 weeks in conventional NSC medium before measuring the fraction of tdTomato+ NSC spheres.

**Generation of Oct4-CreER; Rosa26-loxSTOPlox-DTA MEFs.** Oct4-CreER mice were crossed with Rosa26-loxSTOPlox-DTA animals. MEFs were isolated from bitransgenic fetuses at E13.5. Generation of iNSCs was performed in modified NSC medium. For each replicate, 5 × 10^5 cells infected with lentiviral vectors carrying the M2rtTA and tetOP-OKSM constructs were plated on a 10-cm plate a day before dox treatment. Dox was
administered for 4–5 days, followed by 2–3 weeks of dox-independent growth. 4-OHT was added throughout the transdifferentiation protocol at a concentration of 100 nM for all experiments.

Generation of induced neuronal (IN)-like cells from MEFs. Oct4-CreER; R26-lsl-EYFP MEFs were infected with tetoP-Brn2, tetoP-Ascl1, tetoP-Mylt1 and M2rtTA and reprogrammed as described previously[21]. Dox was added to the medium for 7 days, followed by 7 days of dox-independent growth before measuring Tuj1+ IN cells. 4-OHT was used at a concentration of 100 nM for all experiments.

Subcloning of iNSC clones. OKSM-iNSCs (EYFP+) derived from Oct4-CreER; R26-lsl-EYFP MEFs were plated at 50 per 6-well plate using a FACSariaII sorter (BD Biosciences). Colonial clones were detected between 7–10 days after plating. Subclones were picked onto poly-ornithine-coated plates and propagated until passage 25.

Generation of induced cardiomyocytes. Oct4-CreER; R26-lsl-EYFP MEFs were infected with lentiviral vectors expressing M2rtTA and tetoP-OKSM and treated precisely as reported previously to generate induced cardiomyocytes[22]. For each replicate, 5 × 10^5 cells were plated on a 10-cm plate the day before measuring Tuj1+ iN cells. 4-OHT was used at a concentration of 100 nM for all experiments.

Sequence data and immunostaining (see Supplementary Fig. 11). The primer sequences used were:

Brn4 5’ NotI/Brnt-F2A 3’ and F2A-Kf4 5’/Kf4 3’ BastZ171 (see below) under the following cycling conditions: initial denaturation at 95 °C for 2 min followed by 35 cycles of 15 s at 95 °C and 1 min at 68 °C. Aliquots of the two purified amplicons were then mixed in a 1:1 ratio and used in a third PCR with primers Brn4 5’ NotI and Kf4 3’ BastZ171 under the following conditions: initial denaturation at 95 °C for 2 min; 5 cycles of 15 s at 95 °C, 30 s at 58 °C, and 2 min at 68 °C, and 30 cycles of 15 s at 95 °C, 30 s at 63 °C, and 2 min at 68 °C. The resulting fragment (Brn4-F2A-Kf4) was gel-purified, digested and inserted by directional cloning into the NotI- and BastZ171-digested STEMCCA lentiviral vector. We confirmed the identity of the BKSM polycistronic vector by restriction analysis; DNA sequencing and immunostaining (see Supplementary Fig. 11). The primer sequences used were:

Brn4 5’ NotI: ATAAGAATTCGTGCCCCATGTCACGCTTCGTA ATGGCTAC
Brn4-F2A 3’: GTTGAAGAATTCGTTTTTGCTTTCCAGGCTCTCGGAGCTCTGC
Kf4 3’ BastZ171: TCTTTCGTATACTGGTGCAAATCCGCGCCGAGAGAC

Potential determination of iNSCs. To generate neurons, we grew OKSM-iNSCs in NSC medium after removal of the growth factors bFGF and EGF for 5–10 days; neuronal cells were identified by immunostaining with rabbit anti-Tuj1 antibody (Covance, IgG1, 1:200). Astrocytes were generated using NSC medium supplemented with 5% FBS for a duration of 10 days, after which they were stained with rabbit anti-GFAP antibody (Dako Z0334, IgG, 1:200). The secondary antibody was used donkey anti-rabbit IgG (A10040, Alexa Fluor 546) at a concentration of 1:250. See “Immunocytochemistry” section for a detailed protocol.

Immunocytochemistry. Cells were washed twice with 1 × PBS, cross-linked in 4% PFA solution for 5–10 min, washed with 1 × PBS and blocked for 1/2 hour at room temperature (RT) with blocking solution (2% BSA dissolved in PBS and 0.1% Triton-X-100). Primary antibody staining was performed at RT (1 h).

Appropriate secondary antibodies were added after two washes with 1 × PBS, followed by incubation for 1 h at RT. Primary and secondary antibodies were diluted in blocking solution. DAPI was used for nuclear counterstaining. The following antibodies were used in this study: rabbit anti-Nanog (AB60892, Abcam, IgG, 1:200), goat anti-Oct4 (Santa Cruz sc-6282, IgG 1:200), goat anti-Oct4 (R&D, AF3369, IgG1, 1:100), goat anti-Sox2 (Santa Cruz sc-71320, IgG, 1:200), mouse anti-Nestin (Abcam AB6142, IgG1, 1:200), goat anti-GFAP (EYFP)/EYFP antibodies to stain for EYFP when co-immunostaining colonies for Sox1 or cTnT and EYFP 1–2 weeks after dox withdrawal. 4-OHT was used at a concentration of 100 nM for all experiments.

Viral vector production. For retroviral generation, approximately 2 × 10^6 T-293 cells were transfected with 4.5 µg PCL–ECO (ectopic envelope) and 4.5 µg pMXs-tdTomato vectors, mixed in transfection solution consisting of 300 µl Opti-MEM (Gibco) and 27 µl of TransIT-T1 (Mirus). Transfection was done in regular medium without penicillin-streptomycin. 24 h after transfection, medium was replaced, and 48 and 2 h after transfection the supernatant was collected, filtered through a 0.45-µm filter (Westnet), supplemented with 4–8 µg/ml polybrene (Sigma) and added freshly to the cells. For lentiviral generation, ~6 × 10^6 T-293 cells were transfected in a 10-cm culture dish with a solution consisting of 770 µl Opti-MEM and 50 µl of TransIT-T1, Δ8.9 (8.5 µg), VSV-G (5.5 µg) and 11 µg of the target plasmid (either M2rtTA[11], tetoP-OKSM[11], tetoP-BKSM, tetoP-Ascl1 (ref. 21), tetoP-Brn2 (ref. 21) and tetoP-Mylt1[21]). Viral supernatant was harvested 48 and 72 h after transfection, filtered and added to the cells. Similar ratios of M2rtTA and tetoP-OKSM or tetoP-BKSM (1:1) were used for maximal infection efficiency.

Construction of the BKSM polycistronic expression vector. To achieve co-expression of Brn4, Kf4, Sox2 and cMyc, we generated a novel lentiviral construct (“BKSM”) by reengineering the STEMCCA reprogramming vector[18]. First, a DNA fragment consisting of the Oct4 (Pou5f1) gene followed by an F2A peptide and part of the Klf4 gene was excised from STEMCCA (OKSM) using NotI and BastZ171. Next, we used AccuPrimeTM Pfx DNA Polymerase (Life Technologies) to generate a fragment corresponding to Brn4-F2A-Kf4 by overlapping PCR. To accomplish this, two separate reactions were performed using the primer pairs Brn4 5’ NotI/Brnt-F2A 3’ and F2A-Kf4 5’/Kf4 3’ BastZ171 (see below) under the following cycling conditions: initial denaturation at 95 °C for 2 min followed by 35 cycles of 15 s at 95 °C and 1 min at 68 °C. Aliquots of the two purified amplicons were then mixed in a 1:1 ratio and used in a third PCR with primers Brn4 5’ NotI and Kf4 3’ BastZ171 under the following conditions: initial denaturation at 95 °C for 2 min; 5 cycles of 15 s at 95 °C, 30 s at 58 °C, and 2 min at 68 °C, and 30 cycles of 15 s at 95 °C, 30 s at 63 °C, and 2 min at 68 °C. The resulting fragment (Brn4-F2A-Kf4) was gel-purified, digested and inserted by directional cloning into the NotI- and BastZ171-digested STEMCCA lentiviral vector. We confirmed the identity of the BKSM polycistronic vector by restriction analysis; DNA sequencing and immunostaining (see Supplementary Fig. 11). The primer sequences used were:

Brn4 5’ NotI: ATAAGAATTCGTGCCCCATGTCACGCTTCGTA ATGGCTAC
Brn4-F2A 3’: GTTGAAGAATTCGTTTTTGCTTTCCAGGCTCTCGGAGCTCTGC
Kf4 3’ BastZ171: TCTTTCGTATACTGGTGCAAATCCGCGCCGAGAGAC

© 2015 Nature America, Inc. All rights reserved.
Cells were incubated on ice for 20 min, washed in 1 × PBS, resuspended in FACS buffer, filtered and then analyzed on a MACSQuant flow cytometer (Miltenyi Biotec). Analysis was performed using the software FlowJo.

**RNA extraction and microarrays.** RNA (DNase-treated) was extracted using RNAeasy Mini Kit (Qiagen) according to the manufacturer’s instructions and subjected to hybridization to the GeneChip Mouse 2.0 ST arrays (Affymetrix) at the Partners Center for Personalized Genetic Medicine. Expression Console (Affymetrix) was used for robust multi-array average (RMA). Hierarchical clustering and dendrogram analysis were performed using Expander software (EXPression Analyzer and DisplayER). Scatter plot, analysis of linear regression coefficients and TTEST analysis were performed using Excel. Classification and annotations of upregulated genes were performed using the DAVID online functional annotation tool (http://david.abcc.ncifcrf.gov/). RNA was extracted from expanded OKSM or BKSM-iNSC clusters that have been propagated for several passages. Expression data for MEFs and pluripotent stem cells were previously published. The microarray data have been deposited in NCBI Gene Expression Omnibus (GEO) under accession number GSE65951.

**Alkaline phosphatase staining.** To detect iPSC colonies, alkaline phosphatase kit (Vector Labs) was used according to the manufacturer’s recommendations.

**Teratoma formation.** Approximately 2 × 10⁶ BKSM-iPSCs were injected under the flanks of SCID-beige mice. 3–4 weeks after injection, mice were euthanized to harvest tumor masses. Paraffin section and hematoxylin and eosin (H&E) staining were performed at the Brigham and Women’s Hospital pathology core.

**Generation of chimeric mice.** Chimeras were generated by blastocyst injections as described previously. Briefly, superovulation was initiated by injecting five female B6D2F1/J mice (Jackson Laboratory) with pregnant mare serum (Sigma) by intraperitoneal injection. Two days later, human chorionic gonadotropin (Sigma) was administered by intraperitoneal injection, and female mice were mated with B6D2F1/J males. After 3 days, developing blastocysts were collected and injected with the indicated cell types. Injected blastocysts were transplanted into the uterus of pseudopregnant Swiss Webster female mice (Jackson Laboratory; ~2 months of age) and allowed to develop to midgestation or term. All procedures involving animals were performed according to a mouse protocol approved by the MGH Subcommittee on Research Animal Care.

**Allele-specific real-time PCR for X-linked marker gene.** DNase-treated RNA was extracted using the RNAeasy Mini Kit (Qiagen) according to the manufacturer’s instructions. Conversion of RNA to cDNA was performed using the Transcriptor First Strand cDNA Synthesis Kit (Roche) as per manufacturer’s instructions. mRNA levels were measured using Brilliant iii SybrGreen kit (Agilent) on a Stratagene Mx3005P qRT-PCR system (Agilent). mRNA levels were normalized to GAPDH. We used the following primer sequences for qRT-PCR: Uba1 129/Mus musculus musculus-specific primer—TGAGCTCCTCAAGACGGCAG. Uba1 Mus musculus castaneous-specific primer—TGAGCTCCTCAAGACGGCCA. Common Uba1 reverse primer—CCTTTTTGCGCGCCATTF.

28. Srinivas, S. et al. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* 1, 4 (2001).
29. Eminli, S., Utikal, J., Arnold, K., Jaenisch, R. & Hochedlinger, K. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 26, 2467–2474 (2008).
30. Hadjantonakis, A.K., Gertsenstein, M., Ikawa, M., Okabe, M. & Nagy, A. Non-invasive sexing of preimplantation stage mammalian embryos. *Nat. Genet.* 19, 220–222 (1998).
31. Wu, S., Wu, Y. & Capecchi, M.R. Motoneurons and oligodendrocytes are sequentially generated from neural stem cells but do not appear to share common lineage-restricted progenitors in vivo. *Development* 133, 581–590 (2006).