

Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse

Matthias Stadtfeld,^{1,2} Nimet Maherali,^{1,2,3} David T. Breault,^{4,5} and Konrad Hochedlinger^{1,2,*}

¹Massachusetts General Hospital Cancer Center and Center for Regenerative Medicine

²Harvard Stem Cell Institute

185 Cambridge Street, Boston, MA 02114, USA

³Department of Molecular and Cellular Biology, Harvard University, 7 Divinity Avenue, Cambridge, MA 02138, USA

⁴Division of Endocrinology, Children's Hospital

⁵Department of Pediatrics, Harvard Medical School

Boston, MA 02115, USA

*Correspondence: khochedlinger@helix.mgh.harvard.edu

DOI 10.1016/j.stem.2008.02.001

SUMMARY

Ectopic expression of the transcription factors Oct4, Sox2, c-Myc, and Klf4 in fibroblasts generates induced pluripotent stem (iPS) cells. Little is known about the nature and sequence of molecular events accompanying nuclear reprogramming. Using doxycycline-inducible vectors, we have shown that exogenous factors are required for about 10 days, after which cells enter a self-sustaining pluripotent state. We have identified markers that define cell populations prior to and during this transition period. While downregulation of Thy1 and subsequent upregulation of SSEA-1 occur at early time points, reactivation of endogenous Oct4, Sox2, telomerase, and the silent X chromosome mark late events in the reprogramming process. Cell sorting with these markers allows for a significant enrichment of cells with the potential to become iPS cells. Our results suggest that factor-induced reprogramming is a gradual process with defined intermediate cell populations that contain the majority of cells poised to become iPS cells.

INTRODUCTION

Nuclear reprogramming allows the study of the nature and reversibility of genomic modifications that are imposed on cells during differentiation and development (Hochedlinger and Jaenisch, 2006). In addition, the reprogramming of adult cells into embryonic stem (ES) cells enables the generation of patient-specific stem cells and thus has enormous potential for the treatment and analysis of degenerative diseases (Yamanaka, 2007). Reprogramming can be achieved by nuclear transfer into oocytes (Wakayama et al., 1998; Wilmut et al., 1997), by cell fusion between ES cells and somatic cells (Cowan et al., 2005; Tada et al., 2003), and by the ectopic expression of transcription factors in somatic cells (Takahashi et al., 2007; Takahashi and Yamanaka, 2006; Yu et al., 2007). In the latter approach, viral

expression of the transcription factors Oct4 and Sox2, combined with Klf4 and c-Myc (Maherali et al., 2007; Okita et al., 2007; Park et al., 2008; Takahashi et al., 2007; Wernig et al., 2007) or Lin28 and Nanog (Yu et al., 2007), generates iPS cells from mouse and human fibroblast cultures.

iPS cells were initially isolated using drug selection for the reactivation of ES cell-specific genes, including *Fbx15* (Takahashi and Yamanaka, 2006), *Oct4*, or *Nanog* (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). Curiously, iPS cells produced with *Fbx15* selection were less potent than ES cells, while iPS cells produced with either *Oct4* or *Nanog* selection appeared functionally and molecularly indistinguishable from ES cells, suggesting that *Fbx15* is a less stringent selection marker than *Oct4* and *Nanog*. The similarity between iPS cells and ES cells and the ease with which iPS cells can be generated compared with nuclear transfer or cell fusion make this approach a powerful tool for further studying the process of nuclear reprogramming and for potential clinical applications. Indeed, iPS cells have recently been shown in a proof-of-principle experiment to treat sickle cell anemia in mice (Hanna et al., 2007).

Little is known about the molecular and cellular events accompanying nuclear reprogramming. The generation of iPS cells from fibroblasts is a gradual process that takes between 15 and 20 days upon infection of somatic cells with retroviruses expressing Oct4, Sox2, Klf4, and c-Myc, giving rise to iPS cells at a frequency of less than 0.1% (Maherali et al., 2007; Takahashi and Yamanaka, 2006; Wernig et al., 2007). Omission of c-Myc from the reprogramming cocktail further reduces the efficiency and delays the process (Nakagawa et al., 2008; Wernig et al., 2008). Established iPS cells show silencing of retroviral genes and the re-expression of endogenous pluripotency genes such as *Oct4* and *Nanog* (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). Moreover, iPS cells reactivate the silenced X chromosome in female cells, restore telomerase activity, and re-establish a genome-wide histone methylation pattern characteristic of ES cells (Maherali et al., 2007; Takahashi and Yamanaka, 2006). It is not known, however, whether these events take place in a sequential order and which events coincide with the time point when somatic cells become independent of exogenous factor expression. These questions could not be fully

addressed in previous experiments, mainly because constitutively active viruses expressing the reprogramming factors had been used. We have therefore generated a doxycycline-inducible viral system, which allows temporal control of factor expression, and have used it to reprogram fibroblasts harboring reporters for pluripotency genes and retroviral gene activity. With these reagents, we have determined the temporal requirement for the four factors and have defined molecular cornerstones during the reprogramming of fibroblasts into iPS cells. Our results provide the basis for attempts to replace viral gene delivery systems with small compounds or protein transduction systems and a platform for further dissecting the mechanisms of nuclear reprogramming.

RESULTS

Temporal Requirement for Exogenous Factor Expression during In Vitro Reprogramming of Fibroblasts

To generate an inducible system for factor-mediated reprogramming of fibroblasts into iPS cells, we created doxycycline-controlled lentiviral vectors by replacing the ubiquitin promoter of FUGW (Lois et al., 2002) with a tetOP sequence, followed by the coding regions for Oct4, Sox2, c-Myc, and Klf4, respectively (Figure 1A). The majority of fibroblasts containing an optimized reverse tetracycline transactivator driven off the constitutively expressed ROSA26 promoter (R26-M2rtTA) (Hochedlinger et al., 2005) that were infected with a GFP-carrying variant of this virus became GFP⁺ within 12–24 hr after doxycycline administration, thus validating the system (see Figure S1 available online). R26-M2rtTA heterozygous fibroblasts with a knockin of GFP into the endogenous Oct4 locus (Oct4-GFP) (Lengner et al., 2007) were infected with viruses containing the four reprogramming factors. Two days after the addition of doxycycline to the culture medium, morphological changes were seen in infected fibroblast cultures, and microcolonies appeared 1 day later, which by day 6 had grown into large colonies of loosely attached cells (Figures 1B–1D). By day 9, tightly packed colonies that were weakly GFP⁺ were seen (Figures 1E and 1F). When picked and cultured in the absence of doxycycline, these cells gave rise to strongly GFP⁺ colonies that grew into stable iPS lines and had the potential to contribute to all germ layers in chimeric animals (Figure 1G and data not shown). These observations demonstrate the functionality of our inducible lentiviral system to produce bona fide iPS cells.

To determine the temporal requirement for the four transcription factors during reprogramming, infected fibroblasts were treated with doxycycline for up to 13 days to induce factor expression, followed by doxycycline withdrawal and culture in standard ES cell medium until day 20, when the total number of iPS colonies was scored based on morphology and GFP expression (Figure 1H). Colonies that formed in cultures treated with doxycycline for 7 days or less consistently regressed within 48 hr of doxycycline withdrawal, indicating that no stably reprogrammed cells were present until that time point (Figure 1I and data not shown). However, fibroblasts treated with doxycycline for 8 or more days gave rise to colonies that continued to grow in the absence of doxycycline and formed clearly visible iPS clones that homogeneously expressed GFP at day 20 (Figures

1J and 1K). At days 8 and 9, the frequency of cells that gave rise to stably reprogrammed iPS colonies at day 20 was about 0.016%. Addition of doxycycline until day 10 increased this frequency ~4-fold to 0.06%, while only a modest increase in the number of iPS colonies was seen thereafter (Figure 1I). Efficiencies were determined based on the theoretical number of cells infected with all four viruses (Takahashi and Yamanaka, 2006), which we calculated to be ~40%. The vast majority of colonies scored at day 20 were GFP⁺, had an ES-like morphology, and could be expanded into stable lines, whereas all GFP⁻ colonies had a differentiated morphology and could not be expanded (Figures 1L and 1M). These results show that reprogramming is a gradual process that requires the expression of the four factors for at least 8 days before cells begin to enter a stable self-sustaining pluripotent state.

Surface Markers Thy1 and SSEA-1 Distinguish Early Intermediates of the Reprogramming Process

In order to further characterize subpopulations of cells before and after they become doxycycline independent, we used flow cytometry to follow the change in expression of the surface antigens Thy1, which is expressed at high levels on fibroblasts and other differentiated cell types (Rege and Hagoood, 2006), and SSEA-1, which is expressed on ES cells (Cui et al., 2004). FACS analysis showed that the majority (~70%–80%) of cells in untreated fibroblast cultures were Thy1⁺, while no SSEA-1 expression above background levels could be detected (Figure 2A). Established iPS cells, on the other hand, were uniformly Thy1⁻ and predominantly SSEA-1⁺ (Figure 2A). The observed small fraction of SSEA-1⁻ iPS cells (~10%) reflects heterogeneous expression of this marker, which is consistent with previous observations in undifferentiated ES cells (Cui et al., 2004).

To start with a cell population homogeneous for Thy1 expression, we FACS sorted Thy1⁺ fibroblasts to purity (Figure 2B) and infected them with inducible lentiviruses. In the absence of doxycycline, these cells stably maintained a Thy1⁺SSEA-1⁻ phenotype in culture (Figure 2C). In contrast, 3 days of doxycycline induction resulted in the downregulation of Thy1 expression in about a third of the cells (Figure 2D). No SSEA-1⁺ cells could be detected at this time point. At day 5, roughly half of the cultured cells shut down Thy1 expression and a small fraction of Thy1⁻ cells activated SSEA-1 expression (1%–5% of total cells). By day 7, about 70% of cells had become Thy1⁻ and about 5%–10% of total cells showed an SSEA-1⁺Thy1⁻ phenotype. This pattern did not significantly change until day 12, when doxycycline was withdrawn from the cultures. One week after doxycycline withdrawal, however, an SSEA-1^{high}Thy1⁻ population was detected, which likely represented successfully reprogrammed and hence doxycycline-independent cells that had expanded. The observation that the number of Thy1⁻SSEA-1⁻ cells was greatly reduced at this time compared to day 12 suggested that most of the Thy1⁻SSEA-1⁻ cells represent an unstable reprogramming intermediate. Collectively, these data indicate that Thy1 downregulation is an early event during the reprogramming process that occurs in the majority of cells. A subset of Thy1⁻ cells then initiates SSEA-1 expression, first at low and eventually at high levels. Figure 2F summarizes the reciprocal pattern of Thy1 downregulation and SSEA-1 upregulation during the reprogramming of purified Thy1⁺ fibroblasts.

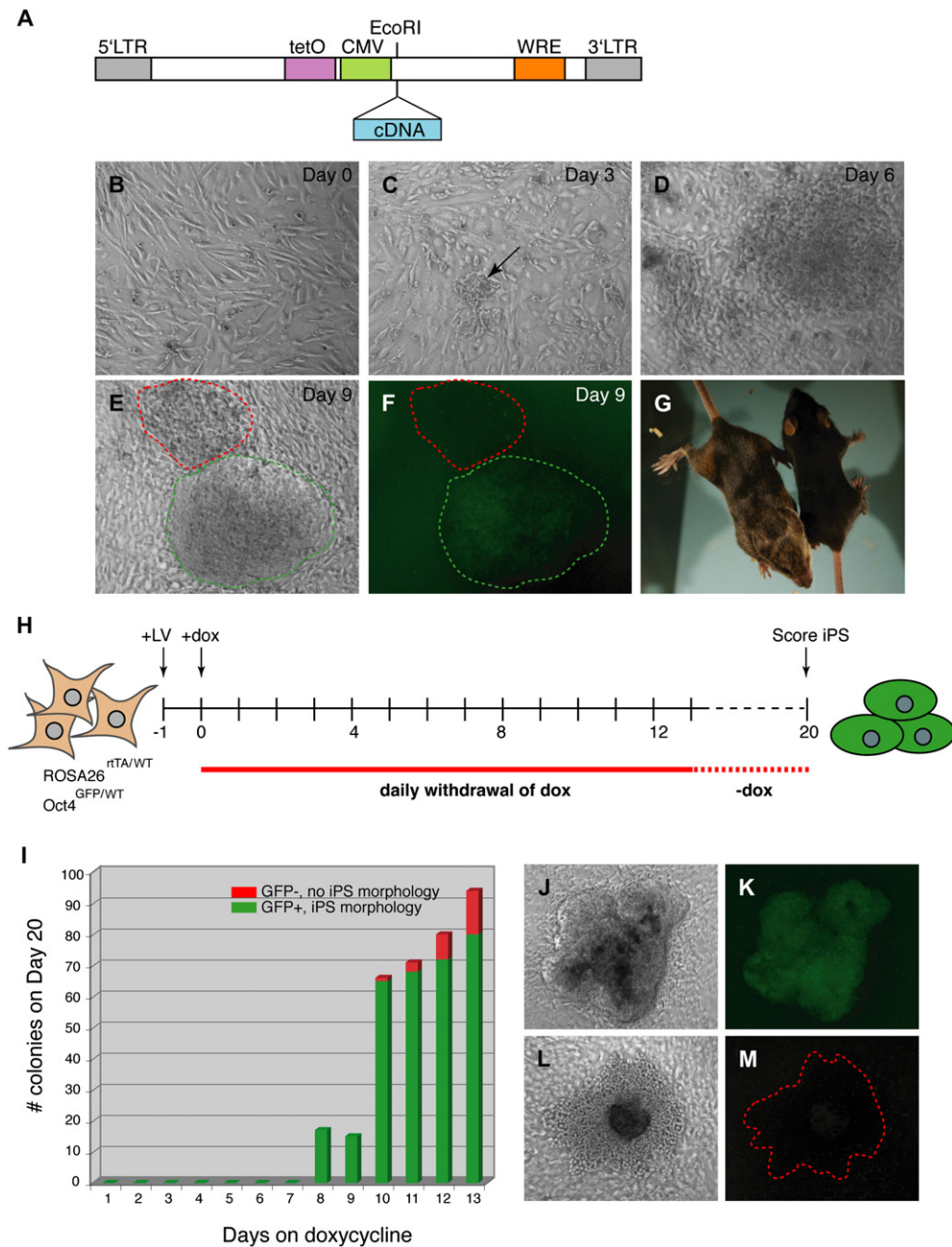


Figure 1. Development of a Doxycycline-Inducible Lentiviral System and Assessment of the Temporal Requirement for Reprogramming Factors

(A) Schematic drawing of the inducible lentivirus containing a doxycycline-controllable promoter (tetOP-CMV) and a unique EcoRI restriction site to insert cDNAs.

(B–E) Morphology of infected Oct4-GFP tail-tip fibroblasts cultured with doxycycline for 0, 3, 6, and 9 days. Arrow in (C) indicates a microcolony.

(F) GFP fluorescence image of colonies shown in (E). Note weakly GFP⁺ colony (outlined in green) with smooth borders next to a GFP⁻ colony with differentiated appearance (outlined in red).

(G) Three-week-old coat color chimera generated from blastocyst injection of iPS cells made with inducible lentiviruses shown left of a nonchimeric littermate.

(H) Experimental outline to determine temporal requirement of reprogramming factors. Fibroblasts heterozygous for ROSA26-rTA and Oct4-GFP alleles were infected with the four lentiviruses (LV), doxycycline (dox) was added for a period of 0–13 days, and GFP⁺ iPS colonies were scored on day 20.

(I) Effect of duration of doxycycline expression on the number of GFP⁺ (green bars) and GFP⁻ (red bars) colonies present at day 20.

(J–M) Bright-field and fluorescence images of a representative GFP⁺ iPS colony (J and K) as well as a GFP⁻ differentiated colony (L and M) at day 20.

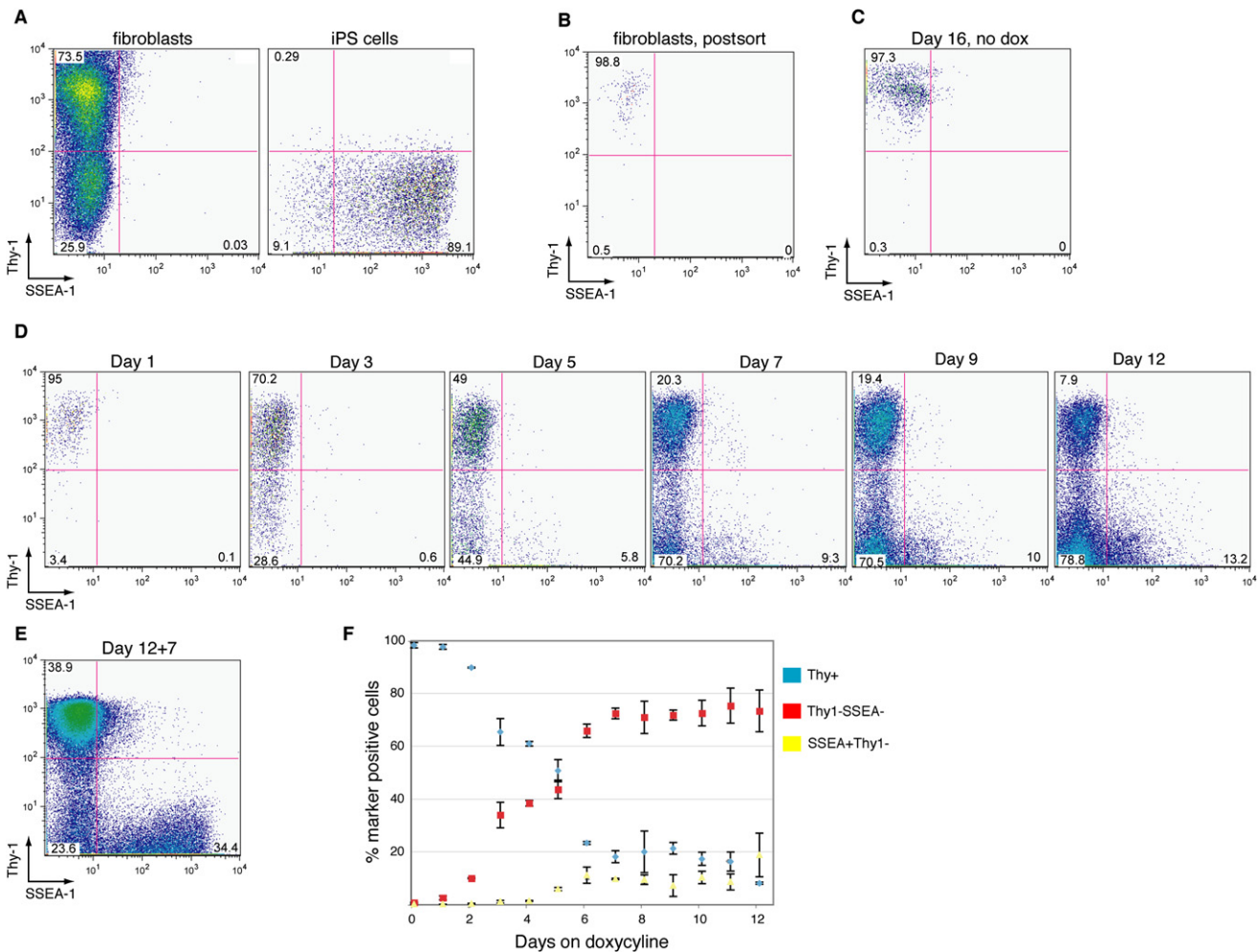


Figure 2. Kinetics of Thy1 and SSEA-1 Surface Antigen Expression during Reprogramming

(A) FACS plots showing Thy1 and SSEA-1 expression in an established iPS line, respectively.

(B and C) (B) FACS plot of Thy1⁺ fibroblasts immediately after sorting and (C) after lentiviral infection and culture in the absence of doxycycline for 16 days, demonstrating sort purity and stability of Thy1 expression in the absence of viral gene expression.

(D) Representative FACS plots of Thy1 and SSEA-1 expression in infected fibroblasts on days 1, 3, 5, 7, 9, and 12 after doxycycline induction, showing the gradual downregulation of Thy1 and upregulation of SSEA-1.

(E) FACS plot of infected fibroblasts cultured for 12 days in the presence of doxycycline and analyzed at 7 days postwithdrawal (12+7). Note the emergence of an SSEA^{high}Thy1⁻ population resembling the phenotype of stable iPS cells shown in (A).

(F) Summary of kinetic changes of Thy1 and SSEA-1 expression, indicating the fractions of Thy1⁺SSEA-1⁻, Thy1⁻SSEA-1⁻, and SSEA-1⁺Thy1⁻ cells. Voltage settings of the FACS analyzer were not identical between (A) and (C–E), leading to slight differences in the appearances of the depicted populations. Error bars indicate range of marker-positive cells in two independent experiments.

Reactivation of Pluripotency Genes, Telomerase, and the Silent X Chromosome Are Late Events in the Reprogramming Process

Successfully reprogrammed cells show reactivation of endogenous pluripotency genes and the silent X chromosome (Maherli et al., 2007; Okita et al., 2007; Wernig et al., 2007) and express telomerase (Takahashi and Yamanaka, 2006). We tested whether activation of these markers occurs early or late during the reprogramming process and determined how they intersect with the kinetics of SSEA-1 upregulation. To this end, we FACS sorted Thy1⁺ fibroblasts isolated from mice carrying GFP reporters for Sox2 (Ellis et al., 2004), mTert (Armstrong et al., 2000; D.T.B., I.M. Min, D.L. Carlone, L.G. Far-

illa, D.E. Henderson, S. Algra, R.K. Montgomery, A.J. Wagers, and N. Hole, unpublished data), or the silent X chromosome (X-GFP) (Hadjantonakis et al., 1998) and infected them with inducible lentiviruses. Female X-GFP fibroblasts were presorted by FACS for GFP⁻ cells in which the GFP-carrying X chromosome had been inactivated. Established iPS cells showed strong expression of Sox2 and mTert and reactivation of the silent X chromosome (Figures 3A–3F). Both the mTert-GFP and X-GFP transgenes, however, had heterogeneous expression patterns similar to ES cells (data not shown), possibly due to transgene silencing, loss of the transgene-carrying chromosome (Robertson et al., 1983), or heterogeneity of ES cells (Chambers et al., 2007).

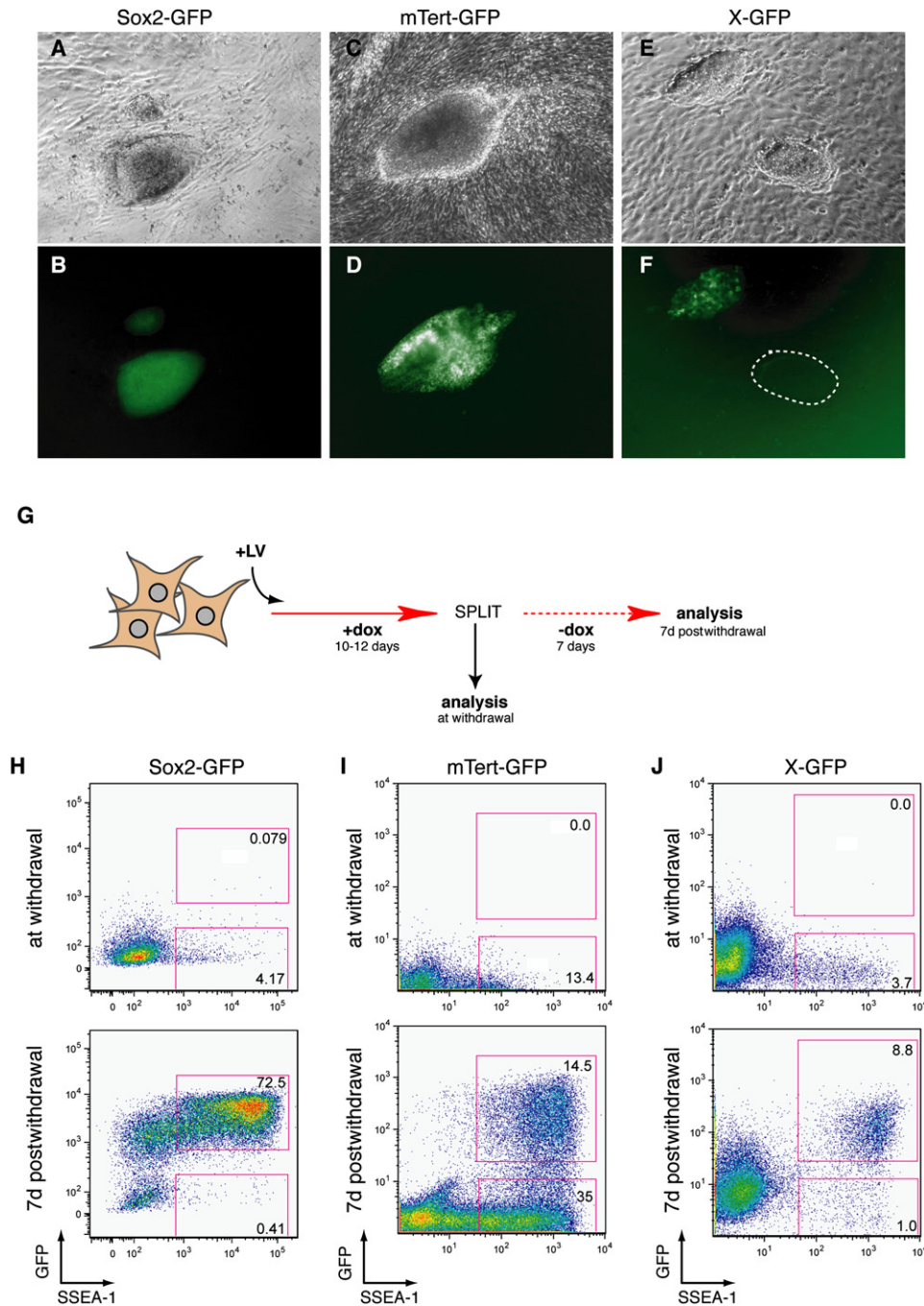


Figure 3. Timing of Sox2-GFP, mTert-GFP, and X-GFP Reactivation during Reprogramming

(A–F) Bright-field and GFP images of iPS cells derived from Sox2-GFP (A and B), mTert-GFP (C and D), and X-GFP (E and F) fibroblasts. The dotted circle in (F) indicates a colony entirely GFP⁻, possibly due the loss of one X chromosome.

(G) Experimental outline to determine temporal relationship of SSEA-1 expression, GFP marker activation, and the appearance of doxycycline-independent iPS cells. Fibroblasts from the different GFP reporter lines were infected with the inducible lentiviruses, cultured in the presence of doxycycline for 10–12 days, and then split equally into two fractions: one fraction was immediately analyzed by flow cytometry, and another fraction was analyzed after 7 days of culture without doxycycline.

(H–J) FACS plots showing GFP and SSEA-1 expression in the cultures at the time of doxycycline withdrawal (upper panels) and at 7 days postwithdrawal (lower panels). Sox2-GFP cells were analyzed with a FACSAria, while mTert-GFP and X-GFP cells were analyzed with a FACSCalibur.

To assess the kinetics of reactivation of these markers, infected fibroblasts were cultured for 10–12 days in the presence of doxycycline, at which time no GFP⁺ colonies could be detected by microscopic examination (data not shown). The cultures were then harvested and split into two fractions (Figure 3G). The first fraction was analyzed immediately by flow cytometry for GFP and SSEA-1 expression, which revealed a significant number of SSEA-1⁺ cells, and a lack of GFP⁺ cells in the mTert-GFP and X-GFP cultures, with rare SSEA-1⁺GFP⁺ cells being detectable in the Sox2-GFP cultures (Figures 3H–3J, upper panels). The other fraction was analyzed in a similar manner at 7 days after withdrawal of doxycycline. At that time, SSEA-1⁺GFP⁺ cells were readily detectable in all reporter lines (Figures 3H–3J, lower panels). As expected from the microscopic inspection of doxycycline-independent iPS colonies (Figures 3A–3F), some SSEA-1⁺ cells were negative for mTert-GFP and X-GFP, respectively. These observations indicate that reactivation of the pluripotency transcription factor Sox2, the silenced X chromosome, and telomerase are late events during the reprogramming process and correlate with the time period when cells start to become independent of doxycycline and thus exogenous factor expression.

Retroviral Silencing Is a Gradual Process that Is Completed in Established iPS Cells

One of the characteristics of pluripotent cells that distinguishes them from somatic cells is their ability to silence retroviruses (Barklis et al., 1986; Teich et al., 1977) by de novo DNA methylation (Okano et al., 1999) and other mechanisms (Wolf and Goff, 2007). Consistent with this observation, faithfully reprogrammed cells show silencing of the retroviral *Oct4*, *Sox2*, *c-myc*, and *Klf4* genes (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007), whereas partially reprogrammed cells show incomplete silencing and persistent expression of the viral factors (Takahashi and Yamanaka, 2006).

To determine the kinetics of retroviral silencing, we infected fibroblasts carrying the Sox2-GFP reporter with a pMX retrovirus expressing the red fluorescent protein tdTomato (short “Tomato”). Thy1⁺Tomato⁺ cells were FACS sorted (Figure S2A), infected with lentiviruses, induced with doxycycline, and analyzed at different time points for the percentages of Tomato⁺ cells (Figure 4A). We focused on four different populations, representing different stages of the reprogramming process: Thy1⁺ fibroblasts, Thy1[−]SSEA-1[−] early intermediate cells, SSEA-1⁺Thy1[−]GFP[−] late intermediate cells, and SSEA-1⁺Thy1[−]GFP⁺ cells, in which Sox2-GFP expression had just become detectable (Figure S2B). Flow cytometric analyses of the four different populations at day 14 (Figure 4B) showed >80% Tomato labeling in the Thy1⁺ fibroblast fraction, consistent with strong viral activity and a low degree of silencing, which was comparable to that of infected fibroblasts grown in the absence of doxycycline (Figure S2D). Early intermediate cells contained roughly 50% Tomato-positive cells, while only 20% of the late intermediate cells and less than 5% of GFP⁺SSEA-1⁺ cells retained retroviral gene expression at day 14. FACS analysis at different time points after induction confirmed these observations, showing stable viral gene expression in the Thy1⁺SSEA-1[−] population, 50% silencing in the Thy1[−]SSEA-1[−] fraction, and progressive silencing in the SSEA-1⁺Thy1[−]GFP[−] and SSEA-1⁺Thy1[−]GFP⁺ populations

(Figure 4C). In agreement with these observations, microscopic inspection of individual colonies derived from SSEA-1⁺Tomato⁺GFP[−] cells isolated by FACS from day 8 cultures and imaged after 5 additional days in culture in the presence of doxycycline showed speckled patterns of Tomato⁺GFP[−] cells, where silencing had not yet occurred, and Tomato[−]GFP⁺ cells, where successful silencing and concomitant reactivation of the Sox2 locus had occurred. Tomato/GFP double-positive cells were never observed (Figures 4D and 4E).

To test whether the early silencing of retroviruses may be due to the upregulation of de novo methyltransferases or other known silencing factors, we examined the expression of *Dnmt3a*, *Dnmt3b* (Okano et al., 1999), and *TRIM28* (Wolf and Goff, 2007) by quantitative RT-PCR. All three genes showed dynamic expression patterns, with *Dnmt3a* levels peaking at day 3 in the Thy1[−] fraction and *Dnmt3b* and *TRIM28* levels gradually increasing with highest expression in Oct4-GFP⁺ cells at day 13 (Figure S2). Together, these results argue that retroviral silencing is a gradual process that is initiated early but completed late in the reprogramming process, coinciding with the activation of epigenetic regulators and pluripotency genes.

Identification of Cell Populations with Increased Reprogramming Potential by Marker Gene Expression

The observation that surface and genetic markers identify phenotypic intermediates of the reprogramming process implies that the marked cell populations should have different potentials to give rise to iPS cells. To test this notion, we FACS sorted different populations from lentivirally infected fibroblasts at days 3, 8, and 13 after induction based on the expression of Thy1, SSEA-1, and either Sox2-GFP (day 8) or Oct4-GFP (days 3 and 13) (see Figures S3A–S3C for gating logic) and plated equal numbers of cells in the presence of doxycycline. As outlined in Figure 5A, doxycycline was withdrawn from all cultures at day 13 and iPS colonies were scored at day 20. The kinetics of Sox2-GFP and Oct4-GFP activation during reprogramming were similar and will therefore only be referred to as GFP. At day 3, all cells with iPS-forming potential were found in the Thy1[−] fraction, which was calculated to be enriched 3.3-fold over an equal number of unfractionated cells (Figure 5B and Figure S3D). By day 8, most cells with iPS-forming potential had moved into the SSEA-1⁺Thy1[−] fraction, which had a 20-fold enrichment of iPS-forming cells over the total population. At day 13, cells with the highest iPS-forming potential had shifted from the SSEA-1⁺Thy1[−] fraction to the SSEA-1⁺GFP⁺ fraction and showed a 5-fold enrichment of iPS-forming cells over total cells. The low enrichment of iPS-forming cells in the latter fraction compared with the SSEA-1⁺ fraction at day 8 is due to a significant expansion of GFP⁺ cells in the total culture, thus lowering the overall enrichment.

These results show that the combination of surface markers and reporters allows for a prospective enrichment for cells with the potential to become iPS cells. The observation that the fraction of cells with the highest iPS-forming potential changes over time suggests that the phenotype of reprogramming cells moves in a linear fashion from an early Thy1[−] phenotype via an intermediate SSEA-1⁺ phenotype to a late Oct4⁺/Sox2⁺ phenotype, which is closest to the final reprogrammed state. To determine the cellular phenotype, which is indicative of cells that have

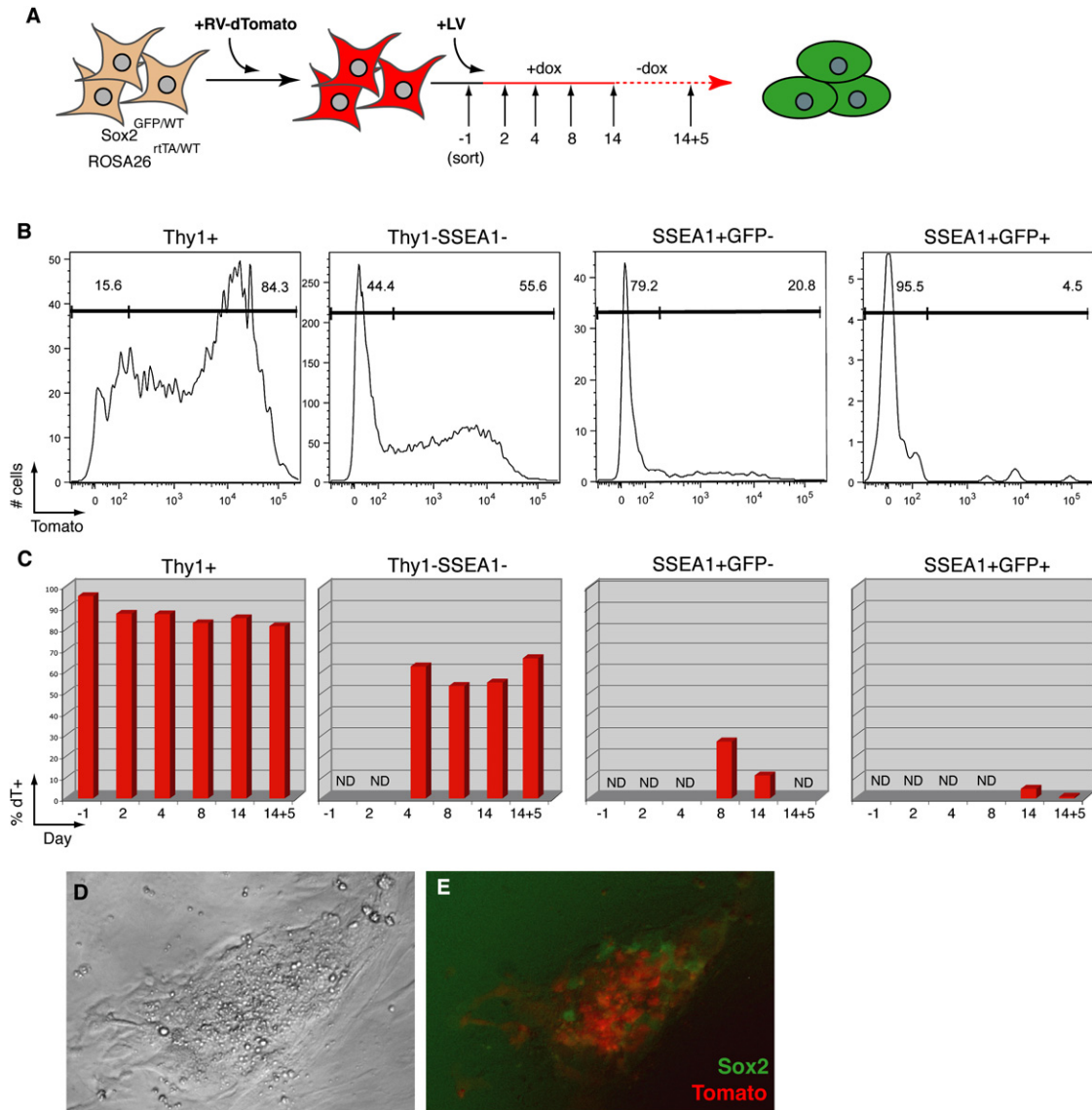


Figure 4. Kinetics of Retroviral Silencing during Reprogramming

(A) Experimental outline. Tail fibroblasts heterozygous for ROSA26-M2rtTA and Sox2-GFP alleles were infected with pMX retrovirus expressing the fluorescent protein tdTomato (RV-tdTomato). Thy1⁺Tomato⁺ cells were purified by FACS, infected with the four lentiviral factors, cultured in the presence of doxycycline, and reanalyzed by FACS at the indicated time.

(B) FACS histograms showing the percentages of Tomato⁺ cells in fibroblasts (Thy1⁺) and intermediate populations (Thy1-SSEA1⁻, SSEA1⁺GFP⁻ and SSEA1⁺GFP⁺) 14 days after viral induction.

(C) Percentages of Tomato⁺ cells in fibroblasts and infected subpopulations in the same subpopulations shown in (B) but analyzed at different time points after induction. ND (not determined) indicates lack of that cell population at the respective time point.

(D and E) Bright-field image and GFP/Tomato overlay of an early iPS colony that formed after sorting SSEA1⁺Tomato⁺Thy1⁻GFP⁻ cells at day 8 and seeding them on feeder cells at clonal density. Note the speckled pattern and lack of overlap between Sox2-GFP and retroviral Tomato expression.

reached a self-sustaining pluripotent state, we cultured the different populations sorted at day 13 for another 3 days in the presence of doxycycline before doxycycline withdrawal. As shown in Figure 6A, colonies were scored at day 20 and results were compared with those obtained after immediate doxycycline withdrawal at day 13 (see Figure 5B). We consistently detected more colonies in the fractions that were treated with doxycycline for 16 days compared with fractions treated for only 13 days (Figures 6B and 6C). However, there was a significant

difference in the ratio of doxycycline-dependent versus doxycycline-independent colonies, depending on the cellular phenotype. For example, extension of the doxycycline treatment from day 13 to day 16 increased the numbers of colonies obtained from the Thy1⁻ population 15-fold, from the SSEA1⁺GFP⁻ fraction 3- to 4-fold and from the SSEA1⁺GFP⁺ fraction ~2-fold (Figure 6C), indicating that there are qualitative differences in the reprogramming status of the individual populations, with Thy1⁻ cells being the least reprogrammed population and

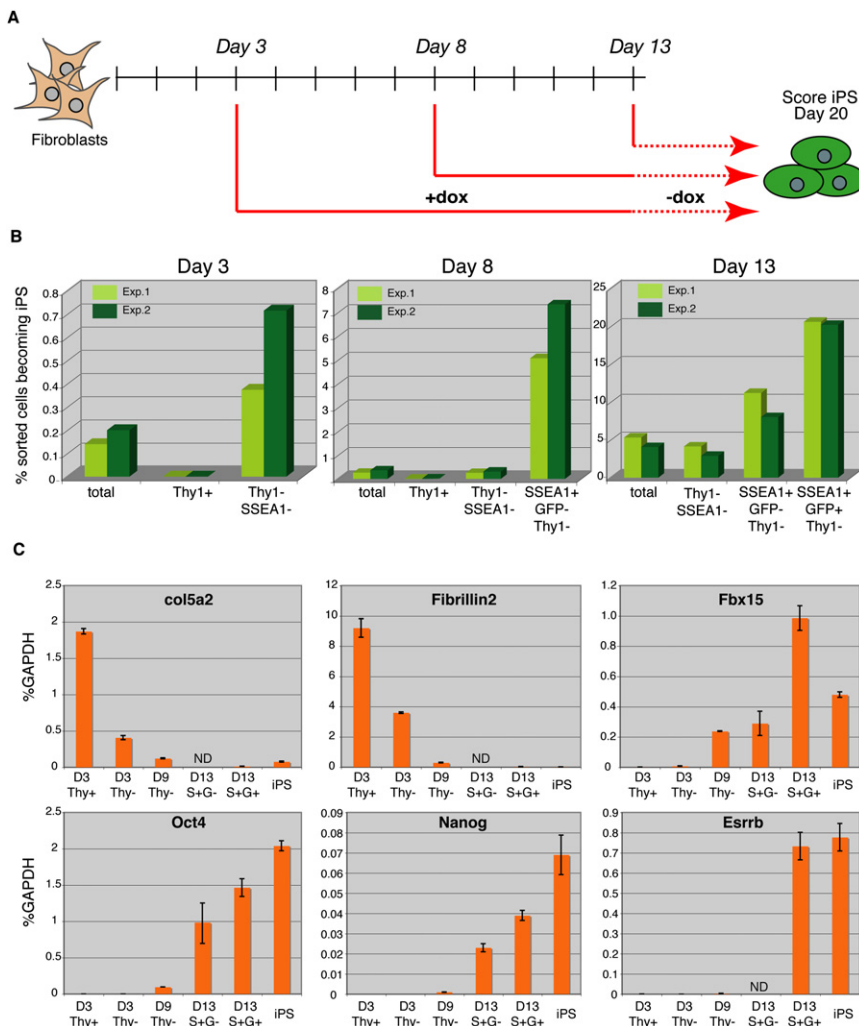


Figure 5. Phenotypically Defined Subpopulations Have Different iPS-Forming Potentials

(A) Experimental outline. Different subpopulations of cells, defined by the expression of Thy1, SSEA-1, and Oct4-GFP or Sox2-GFP, were FACS sorted from fibroblast cultures at the indicated times after factor induction and seeded at equal numbers onto feeder cells. Doxycycline was withdrawn at day 13, and doxycycline-independent GFP⁺ colonies were scored at day 20.

(B) Seeding efficiencies of the different subpopulations and of unfractionated cultures sorted at the indicated time points are plotted as percentages of sorted cells that formed iPS colonies at day 20. Data from two independent experiments are shown.

(C) Graphs showing expression levels of candidate fibroblast (*col5a2* and *Fibrillin2*) and ES cell (*Fbx15*, *Oct4*, *nanog*, and *Esrrb*) markers by qPCR relative to *GAPDH*. D, day; ND, not determined; S⁺G⁻, SSEA-1⁺GFP⁻; S⁺G⁺, SSEA-1⁺GFP⁺; Thy⁻, Thy1⁻SSEA-1⁻. Error bars indicate one standard deviation.

explain the lower plating efficiency of the former cells and suggests that the pluripotent transcriptome is further finetuned after the initial reactivation of pluripotency genes.

DISCUSSION

We have generated a doxycycline-inducible lentiviral system to transiently express the four reprogramming factors c-Myc, Klf4, Oct4, and Sox2 in fibroblasts. iPS cells produced with this system are pluripotent and give rise to chimeras when injected into blastocysts, similar to retrovirally produced iPS cells. We and others have recently shown that drug selection is not required to obtain iPS cells (Blelloch et al., 2007; Maherali et al., 2007; Meissner et al., 2007). However, omission of drug selection increases the number of false positive colonies, such as transformed cells or cells that failed to receive all four viruses, and necessitates a trained eye to identify ES cell-like colonies. A major advantage of the inducible system over constitutive expression systems is that it allows for the “self-selection” of reprogrammed cells in the absence of drug selection and obviates the need for ES cell expertise. After stably reprogrammed cells have been generated and doxycycline has been withdrawn, cells that survive go on to reactivate the endogenous pluripotency program, while unstable reprogramming intermediates and transformed colonies disappear, likely through differentiation or apoptosis.

We have used the doxycycline-dependent system to assess the temporal requirement for exogenous factor expression to produce iPS cells. Our results show that the four factors are required for only about 10 days, at which point the somatic genome becomes poised for conversion into a pluripotent state. This has been consistently seen in independent experiments

GFP⁺ cells being the most reprogrammed population. When compared to the plating efficiency of established iPS cells, however, day 13 GFP⁺ cells were still about 60% less efficient in making iPS colonies in the absence of doxycycline, while addition of doxycycline almost closed this gap (20% for GFP⁺ cells versus 30% for iPS cells).

To further characterize the different intermediate populations at the molecular level, we analyzed the expression of candidate fibroblast- and ES cell-specific markers (Figure 5C). Consistent with the early downregulation of Thy1 expression, we observed rapid reduction in the levels of the fibroblast genes *Col5a2* and *Fibrillin2* in the Thy1⁻ fraction at day 3, which were further downregulated in later fractions. Conversely, expression of the ES cell-specific marker *Fbx15* became weakly detectable by day 3 and was further increased in the day 9 Thy1⁻ and day 13 SSEA-1⁺GFP⁻ fractions before being massively upregulated in the SSEA-1⁺GFP⁺ fraction by day 13. This is in contrast to *Nanog* and *Oct4* expression, which became detectable only at around day 9. *Esrrb* was strongly expressed only in the SSEA-1⁺GFP⁺ fraction, at similar levels as seen in established iPS cells. The observation that the day 13 SSEA-1⁺GFP⁺ fraction expresses lower levels of *Oct4* and *Nanog* than established iPS cells might

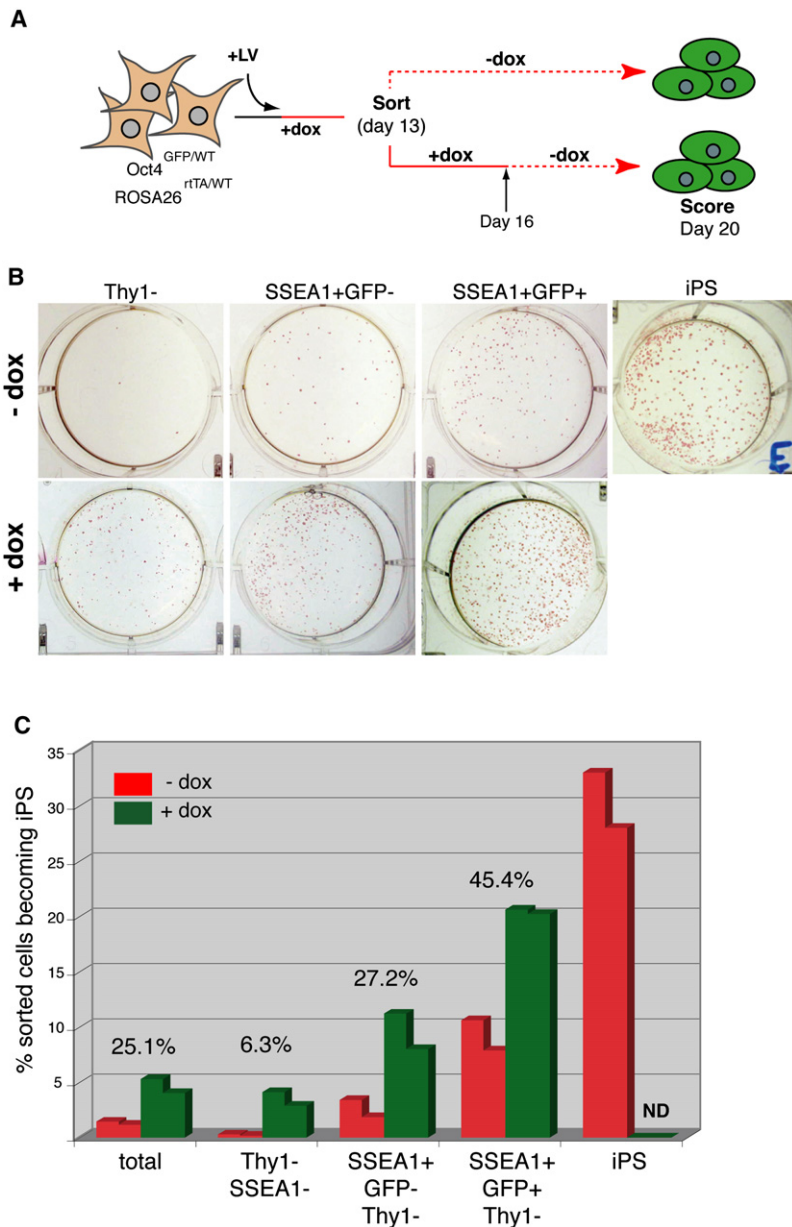


Figure 6. Correlation between Cellular Phenotype and Degree of Reprogramming

(A) Experimental outline. Thy1⁺, Thy1⁻SSEA-1⁻, SSEA-1⁺GFP⁻Thy1⁻ as well as SSEA-1⁺GFP⁺Thy1⁻ cells were sorted from Oct4-GFP fibroblast cultures at day 13 after factor induction. Twenty-five hundred cells from each fraction were cultured either in the absence or presence of doxycycline for another 3 days. Doxycycline-independent GFP⁺ iPS colonies that had formed under either condition were scored at day 20.

(B) Images of representative day 20 cultures stained with alkaline phosphatase to demonstrate the colony-forming ability of the different subfractions in the absence (upper panel) or presence (lower panel) of doxycycline. A stable iPS line that has been passaged several times is shown for comparison.

(C) Graphic representation of the iPS-forming potential of the different day 13 subpopulation in the absence (red) or presence (green) of doxycycline, showing results from two independent experiments for each condition. Percentages above the bars indicate the fraction of colonies that formed in the absence of doxycycline compared with the fraction of colonies that formed in the presence of doxycycline for the different subpopulations.

using fibroblasts from different reporter strains, suggesting that reprogramming follows a defined rather than a random sequence of molecular events that transform a differentiated cell into a pluripotent cell. In agreement with this notion, we have identified and characterized cell populations that appear prior to and during this transition point by surface marker expression, reporter gene activity, and transcriptional analyses (Figure 7). Based on these results, we conclude that downregulation of fibroblast markers such as Thy1 is the earliest detectable change in cells undergoing reprogramming, followed by modest activation of the ES cell marker SSEA-1, which is further upregulated at the time when cells become doxycycline independent. At this transition point, other major modifications take place, including the reactivation of Oct4 and Sox2, telomerase, and the silenced X chromosome. Importantly, reactivation of these markers is not

yet indicative of a fully reprogrammed phenotype because continuous expression of exogenous factors can still increase the number of iPS colonies. Interestingly, the ES cell-specific *Fbx15* gene is reactivated at day 3, before other pluripotency genes, including *Nanog*, *Oct4*, and *Esrrb*, become detectable. This may explain why the initially derived iPS cells made with an *Fbx15* selectable allele (Takahashi and Yamanaka, 2006) appeared incompletely reprogrammed compared with the subsequently reported *Nanog* selection approach (Maherali et al., 2007; Okita et al., 2007; Wernig et al., 2007). Support for the functional importance of the markers described here comes from the observation that FACS sorting of subpopulations defined by these markers allows for a significant enrichment for cells with the potential to become iPS cells.

Retroviral silencing during reprogramming appears to be a gradual process that occurs as early as 4 days after the induction of the transcription factors but is completed only in cells that have reached the transition point of doxycycline independence. In accordance with the observed early retroviral silencing is the transient upregulation of the de novo methyltransferase *Dnmt3a*, while silencing at later stages coincides with the upregulation of *Dnmt3b* (Okano et al., 1999) and the retroviral silencing factor TRIM28 (Wolf and Goff, 2007). This finding implies that the frequency of obtaining iPS cells with retroviruses should be reduced due to premature silencing in some cells. Indeed, we have observed that infection of fibroblasts with lentiviruses, which are not subject to silencing, gives rise to ES-like colonies roughly 5–10 times more efficiently compared with retroviruses (M.S., N.M., and K.H., unpublished data).

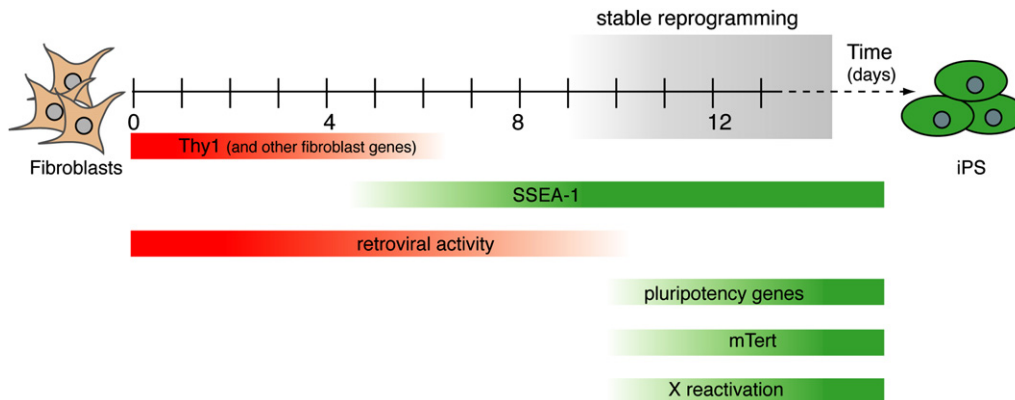


Figure 7. Kinetics of Marker Expression during Reprogramming

The molecular mechanism underlying nuclear reprogramming remains enigmatic. Due to its low efficiency, it has thus far been impossible to capture intermediate stages of the reprogramming process. Our identification of intermediate cell populations that are enriched for iPS-producing cells will help to solve some of the open questions surrounding nuclear reprogramming. For example, a genome-wide analysis of individual subpopulations at different stages of the reprogramming process will be instrumental for understanding how the three or four transcription factors induce pluripotency and epigenetic reprogramming. Lastly, the system described here provides the framework for the identification of genes and chemicals that replace or enhance the production of iPS cells.

EXPERIMENTAL PROCEDURES

Viral Production

A tet-inducible lentivirus called LV-tetO was generated by removing the EcoRI/PacI fragment containing the ubiquitin promoter elements from the FUΔGW vector (Lois et al., 2002) (modified from FUGW by removal of GFP with EcoRI-XbaI and re-creation of the EcoRI site) and replacing it with tetO sequences derived from pTet-Splice by EcoRI/XhoI digestion. cDNAs for c-Myc (T58A mutant), Klf4, Oct4, and Sox2 were isolated from pMIG or pMX vectors and cloned into LV-tetO using a unique EcoRI restriction site. To produce infectious viral particles, 293T cells cultured on 10 cm dishes were transfected with the LV-tetO vectors together with the packaging plasmids VSV-G and Δ8.9 using Fugene (Roche). To generate RV-tdTomato, the BamHI/NotI fragment from pIRES-tdTomato (a kind gift of Dr. Niels Geijsen) was cloned into the EcoRI site of pMX (Morita et al., 2000) and transfected into 293T cells together with packaging plasmids PCG-ori-gagpol and PCG-ori-VSVG using Fugene. Viral supernatants were harvested on 3 consecutive days starting 24 hr after transfection, concentrated by centrifugation at 20,000 rpm for 1.5 hr at 4°C, resuspended in PBS, and stored at -80°C.

Flow Cytometry

Cultures were harvested by incubation in 0.25% trypsin/1 mM EDTA for 5 min at 37°C, and single-cell suspensions were obtained by repetitive pipetting and transfer through a 40 μm cell strainer. Cells were incubated with PE-conjugated rat anti-Thy1 (53-2.1, eBiosciences) and unconjugated mouse anti-SSEA-1 (MC-480, Developmental Hybridoma Bank) antibodies, followed by incubation with APC-conjugated mouse anti-mouse IgM antibody (II/41, eBiosciences) and either analyzed on a FACSCalibur (BD Biosciences) or sorted/analyzed on a FACSAria (BD Biosciences). For staining of pMX-tdTomato infected cells, Thy1 expression was visualized by staining with biotinylated rat anti-Thy1 antibody, followed by Pacific blue-conjugated streptavidin (Invitrogen). Dead cells were excluded by staining with either propidium iodide or DAPI. Data were analyzed with FlowJo software (Tree Star).

Generation of Sox2-GFP Reporter Mice

V6.5 ES cells were targeted with a Sox2-GFP knockin construct described previously (Ellis et al., 2004). Linearized DNA was electroporated into ES cells, colonies were selected with G418, and expanded clones were screened for correct integration by standard Southern blot analysis. Correct clones were injected into tetraploid blastocysts to produce entirely ES cell-derived mice (Eggen et al., 2001).

Production of Chimeric Mice

Female BDF1 mice were superovulated with PMS and hCG and mated to BDF1 stud males. Zygotes were isolated from females with a vaginal plug 24 hr after hCG injection. After 3 days of in vitro culture in KSOM media, blastocysts were identified, injected with iPS cells, and transferred into pseudo-pregnant recipient females. Pups were delivered by Cesarean section at day 19.5 and nurtured by foster mothers.

Fluorescence Microscopy

Cells were imaged using a Leica DMI4000B inverted fluorescence microscope equipped with a Leica DFC350FX camera. Images were processed and analyzed using Adobe Photoshop software.

RNA Isolation and qPCR

RNA was isolated from FACS-sorted cells using the RNeasy Plus Microkit, following the manufacturer's instructions to use spin columns to exclude genomic DNA, and cDNA was produced with the First Strand cDNA Synthesis Kit (Roche). Real-time quantitative PCR reactions were set up in duplicate or triplicate with the Brilliant II SYBR Green QPCR Master Mix (Stratagene) and run on an Mx3000P QPCR System (Stratagene). Primer sequences for qPCR were taken from PrimerBank (<http://pga.mgh.harvard.edu/primerbank/index.html>) (Wang and Seed, 2003) and are listed in Table S1.

Cell Culture

Fibroblasts were isolated from tail-tip biopsies of newborn (3–8 days of age) Sox2-GFP/R26-M2rtTA, Oct4-GFP/R26-M2rtTA, and mTert-GFP mice as well as from E14.5 female X-GFP/R26-M2rtTA embryos and expanded in fibroblast medium (DMEM with 10% FBS, L-glutamine, penicillin-streptomycin, nonessential amino acids, and β-mercaptoethanol). Viral infections were usually carried out at passage 2 or 3 in 35 mm plates at a density of ~200,000 fibroblasts/plate in fibroblast medium containing 5 μg/ml polybrene. mTert-GFP fibroblasts were coinfecting with a lentivirus constitutively expressing M2rtTA. The infection media was replaced after 12–24 hr with ES medium (DMEM with 15% FBS, L-glutamin, penicillin-streptomycin, nonessential amino acids, β-mercaptoethanol, and 1000 U/ml LIF) supplemented with 1 μg/ml doxycycline. Fresh ES medium with doxycycline was added every 2–4 days until either colonies were picked, cultures were harvested for FACS analysis, or doxycycline was withdrawn by completely aspirating the medium, washing twice with PBS, and then continuing the culture in standard ES medium. Alkaline phosphatase staining was performed using the Vector Red Substrate Kit (Vector Labs).

SUPPLEMENTAL DATA

Supplemental Data include three figures and one table and can be found with this article online at <http://www.cellstemcell.com/cgi/content/full/2/3/230/DC1/>.

ACKNOWLEDGMENTS

We thank Drs. Niels Geijsen for the IRES-tdTomato plasmid, Gustavo Mostoslavsky for retroviral packaging plasmids, Larysa Pevny for providing the Sox2-GFP targeting construct, Carlos Lois for providing the FUGW lentiviral construct, and Laura Prickett and Kat Folz-Donahue for expert help with flow cytometry. We are grateful to Raul Mostoslavsky and Adlen Foudi for critical reading of the manuscript. We thank members of the Hochedlinger lab for critical comments and suggestions. Support to K.H. came from the NIH Director's Innovator Award, the Harvard Stem Cell Institute, the Kimmel Foundation, and the V Foundation.

Received: January 20, 2008

Revised: February 4, 2008

Accepted: February 5, 2008

Published online: February 14, 2008

REFERENCES

- Armstrong, L., Lako, M., Lincoln, J., Cairns, P.M., and Hole, N. (2000). mTert expression correlates with telomerase activity during the differentiation of murine embryonic stem cells. *Mech. Dev.* 97, 109–116.
- Barklis, E., Mulligan, R.C., and Jaenisch, R. (1986). Chromosomal position or virus mutation permits retrovirus expression in embryonal carcinoma cells. *Cell* 47, 391–399.
- Blelloch, R., Venere, M., Yen, J., and Ramalho-Santos, M. (2007). Generation of induced pluripotent stem cells in the absence of drug selection. *Cell Stem Cell* 1, 245–247.
- Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* 450, 1230–1234.
- Cowan, C.A., Atienza, J., Melton, D.A., and Eggan, K. (2005). Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science* 309, 1369–1373.
- Cui, L., Johkura, K., Yue, F., Ogiwara, N., Okouchi, Y., Asanuma, K., and Sasaki, K. (2004). Spatial distribution and initial changes of SSEA-1 and other cell adhesion-related molecules on mouse embryonic stem cells before and during differentiation. *J. Histochem. Cytochem.* 52, 1447–1457.
- Eggan, K., Akutsu, H., Loring, J., Jackson-Grusby, L., Klemm, M., Rideout, W.M., III, Yanagimachi, R., and Jaenisch, R. (2001). Hybrid vigor, fetal overgrowth, and viability of mice derived by nuclear cloning and tetraploid embryo complementation. *Proc. Natl. Acad. Sci. USA* 98, 6209–6214.
- Ellis, P., Fagan, B.M., Magness, S.T., Hutton, S., Taranova, O., Hayashi, S., McMahon, A., Rao, M., and Pevny, L. (2004). SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev. Neurosci.* 26, 148–165.
- Hadjantonakis, A.K., Gertsenstein, M., Ikawa, M., Okabe, M., and Nagy, A. (1998). Non-invasive sexing of preimplantation stage mammalian embryos. *Nat. Genet.* 19, 220–222.
- Hanna, J., Wernig, M., Markoulaki, S., Sun, C.W., Meissner, A., Cassady, J.P., Beard, C., Brambrink, T., Wu, L.C., Townes, T.M., and Jaenisch, R. (2007). Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* 318, 1920–1923.
- Hochedlinger, K., and Jaenisch, R. (2006). Nuclear reprogramming and pluripotency. *Nature* 441, 1061–1067.
- Hochedlinger, K., Yamada, Y., Beard, C., and Jaenisch, R. (2005). Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* 121, 465–477.
- Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer, H.R., Tomilin, A., and Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic stem cell self-renewal. *Cell Stem Cell* 1, 403–415.
- Lois, C., Hong, E.J., Pease, S., Brown, E.J., and Baltimore, D. (2002). Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* 295, 868–872.
- Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1, 55–70.
- Meissner, A., Wernig, M., and Jaenisch, R. (2007). Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat. Biotechnol.* 25, 1177–1181.
- Morita, S., Kojima, T., and Kitamura, T. (2000). Plat-E: an efficient and stable system for transient packaging of retroviruses. *Gene Ther.* 7, 1063–1066.
- Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochizuki, 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.
- Okano, M., Bell, D.W., Haber, D.A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99, 247–257.
- Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313–317.
- Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* 451, 141–146.
- Rege, T.A., and Hagoood, J.S. (2006). Thy-1, a versatile modulator of signaling affecting cellular adhesion, proliferation, survival, and cytokine/growth factor responses. *Biochim. Biophys. Acta* 1763, 991–999.
- Robertson, E.J., Evans, M.J., and Kaufman, M.H. (1983). X-chromosome instability in pluripotential stem cell lines derived from parthenogenetic embryos. *J. Embryol. Exp. Morphol.* 74, 297–309.
- Tada, M., Morizane, A., Kimura, H., Kawasaki, H., Ainscough, J.F., Sasai, Y., Nakatsui, N., and Tada, T. (2003). Pluripotency of reprogrammed somatic genomes in embryonic stem hybrid cells. *Dev. Dyn.* 227, 504–510.
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872.
- Teich, N.M., Weiss, R.A., Martin, G.R., and Lowy, D.R. (1977). Virus infection of murine teratocarcinoma stem cell lines. *Cell* 12, 973–982.
- Wakayama, T., Perry, A.C., Zuccotti, M., Johnson, K.R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. *Nature* 394, 369–374.
- Wang, X., and Seed, B. (2003). A PCR primer bank for quantitative gene expression analysis. *Nucleic Acids Res.* 31, e154. 10.1093/nar/gng154.
- Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* 448, 318–324.
- Wernig, M., Meissner, A., Cassady, J.P., and Jaenisch, R. (2008). C-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2, 10–12.
- Wilmot, I., Schnieke, A.E., McWhir, J., Kind, A.J., and Campbell, K.H. (1997). Viable offspring derived from fetal and adult mammalian cells. *Nature* 385, 810–813.
- Wolf, D., and Goff, S.P. (2007). TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell* 131, 46–57.
- Yamanaka, S. (2007). Strategies and new developments in the generation of patient-specific pluripotent stem cells. *Cell Stem Cell* 1, 39–49.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920.