Induction of Pluripotency in Mouse Somatic Cells with Lineage Specifiers

Jian Shu,1,3,9 Chen Wu,3,6,9 Yetao Wu,3,6,9 Zhiyuan Li,2,7,9 Sida Shao,1 Wenhui Zhao,1,3 Xing Tang,4 Huan Yang,3,6 Lijun Shen,1 Xiaohan Zuo,3,6 Weifeng Yang,1 Yan Shi,6 Xiaochun Chi,5 Hongquan Zhang,5 Ge Gao,4 Youmin Shu,8 Kehu Yuan,8 Weiwu He,8 Chao Tang,2,3,* Yang Zhao,1,3 and Hongkui Deng1,3,6,*

1The MOE Key Laboratory of Cell Proliferation and Differentiation, College of Life Sciences
2Center for Quantitative Biology
3Peking-Tsinghua Center for Life Sciences
4College of Life Sciences, State Key Laboratory of Protein and Plant Gene Research, Center for Bioinformatics
Peking University, Beijing 100871, China
5Laboratory of Stem Cells, Development and Reproductive Medicine, School of Basic Medical Sciences, Peking University, Beijing 100191, China
6Laboratory of Chemical Genomics, School of Chemical Biology and Biotechnology, Peking University Shenzhen Graduate School, Shenzhen 518055, China
7Biophysics Graduate Program, University of California, San Francisco, San Francisco, CA 94158, USA
8OriGene Technologies, 9620 Medical Center Drive, Rockville, MD 20850, USA
9These authors contributed equally to this work
*Correspondence: tangc@pku.edu.cn (C.T.), hongkui_deng@pku.edu.cn (H.D.)
http://dx.doi.org/10.1016/j.cell.2013.05.001

SUMMARY

The reprogramming factors that induce pluripotency have been identified primarily from embryonic stem cell (ESC)-enriched factors, pluripotency-associated factors, or maternal factors such as Oct4, Sox2, Klf4, c-Myc, Nanog, PRDM14, Sall4, Esrrb, Utf1, Tet2, and Glis1. Our study suggests a "seesaw model" in which a balance that is established using pluripotency factors and/or counteracting lineage specifiers can facilitate reprogramming.

INTRODUCTION

Understanding the establishment of cellular identity in programming and reprogramming is a major goal of modern biology. For years, pluripotency-associated factors and their rivals, lineage specifiers, have generally been considered to determine the identities of pluripotent and differentiated cells, respectively (Jaenisch and Young, 2008; Young, 2011). Accordingly, the reprogramming factors that induce pluripotency have been identified primarily from embryonic stem cell (ESC)-enriched factors, pluripotency-associated factors, or maternal factors such as Oct4, Sox2, Klf4, c-Myc, Nanog, PRDM14, Sall4, Esrrb, Utf1, Tet2, and Glis1 (Aoi et al., 2008; Buganim et al., 2012; Chia et al., 2010; Doege et al., 2012; Maekawa et al., 2011; Maherali et al., 2008; Takahashi and Yamanaka, 2006; Yu et al., 2007). Likewise, the direct reprogramming of differentiated cells into other differentiated cell types has been successfully demonstrated by several lineage specifiers, such as Gata4 and Hnf4a (Vierbuchen and Wernig, 2011). Thus, the perspective that the direct conversion of cell state A to cell state B should be realized by a set of master regulatory factors of cell type B has been a prevailing strategy (Graf and Enver, 2009; Jopling et al., 2011); however, whether this is the only strategy for cell fate conversion is unclear.

Recent data indicate that the most critical reprogramming factor, Oct4, which inhibits the expression of differentiation-related genes in ESCs (Kim et al., 2008; Pardo et al., 2010), is sufficient to direct the reprogramming of somatic cells into induced pluripotent stem cells (iPSCs) (Li et al., 2011; Zhu et al., 2010). Therefore, it is of great significance to find substitutes for Oct4 to elucidate its physiological role and gain a better understanding of the reprogramming mechanisms, which remain largely unknown. The ESC-enriched factor Nr5a2 has been identified as an Oct4 substitute (Heng et al., 2010). However, the physiological role of Oct4 remains unclear because Nr5a2 directly regulates Oct4 and binds to the upstream promoter region of Oct4 (Gu et al., 2005; Guo and Smith, 2010). Therefore, extensively screening for novel Oct4 substitutes among factors including but not limited to ESC-related factors may cast light on the molecular mechanisms that underlie reprogramming and pluripotency, thereby facilitating the development of safer and more efficient reprogramming strategies.
Figure 1. GATA3 Can Substitute for OCT4 to Induce Pluripotency in Mouse Somatic Cells

(A) Schematic representation of the gene-screening process. For the screen, a plasmid library of 10,080 human genes was analyzed. An OCT4 plasmid and empty vectors (EV) were used as positive and negative controls in the first and second column, respectively.

(B) Reprogramming assay that determines the ability of the lineage specifier GATA3 to enhance reprogramming in the absence of OCT4, SOX2, KLF4, or c-MYC. The Oct4-GFP-positive colonies were counted at 9 days postinfection (dpi). SKM, OKM, OSM, OSK, and SK plus EV were used as negative controls. OSKM was used as a positive control. Error bars indicate the SD (n = 3).

(C) The kinetics of reprogramming using different combinations of genes. For the 50,000 MEFs seeded per well, the GFP fluorescence was monitored every 24 hr. The orange bars indicate that no Oct4-GFP-positive cells were observed, whereas the green bars indicate the emergence of Oct4-GFP-positive cells. Two representative independent experimental sets are shown.

(legend continued on next page)
Here, we identified eight lineage specifiers as OCT4 substitutes, including GATA3, GATA6, and SOX7, which are involved in mesendodermal (ME) lineage specification. OCT4 and its substitutes attenuated the elevated expression of a group of ectodermal (ECT) genes, such as the ECT lineage specifier Dlx3, that were collectively triggered by SOX2, KLF4, and c-MYC (SKM). Knockdown of Dlx3 enhanced reprogramming in the absence of OCT4, rather than in the absence of SOX2. In addition, SOX2 can be replaced by lineage specifiers involved in ECT lineage specification, such as GMNN. Surprisingly, the two counteracting lineage specifiers can synergistically induce pluripotency in the absence of both OCT4 and SOX2. We suggest a “seesaw model” in which the balance that is established by pluripotency factors and/or lineage specifiers facilitates the induction of pluripotency in somatic cells. This model could shed light on fundamental questions regarding the establishment of cellular identity during programming and reprogramming.

RESULTS

Induction of Pluripotency in Mouse Somatic Cells with GATA3

In our study, we screened an initial plasmid library of 10,080 human genes for their ability to replace OCT4 when introduced together with virally expressed SKM to direct the reprogramming of mouse embryonic fibroblasts (MEFs) containing a green fluorescent protein (GFP) reporter driven by an Oct4 promoter and enhancer. Reprogramming efficiency was evaluated by determining the number of Oct4-driven GFP-positive colonies. We found that GATA3 had the most significant effect in the primary hits (Figure 1A and Table S1 available online). Interestingly, GATA3 is not enriched in ESCs and is an important regulator of development and differentiation (Figure S1D and Table S4) (Ting et al., 1996).

We further validated the ability of GATA3 to replace OCT4 during the reprogramming of MEFs, mouse adult dermal fibroblasts (MDFs), mouse gastric epithelial cells (GECs), and mouse keratinocytes using viral vectors (Figures 1B, S1A, and S1B). The expression of exogenous genes was verified (Figure S1E). We found that GATA3 achieved a reprogramming efficiency that was comparable to or even higher than that of OCT4. We subsequently evaluated the ability of GATA3 to enhance reprogramming in the absence of SOX2, KLF4, or c-MYC. We observed that GATA3 was also able to enhance reprogramming in the absence of c-MYC, but it was unable to substitute for SOX2 or KLF4 (Figure 1B). Next, we monitored the kinetics of GATA3- and OCT4-induced reprogramming. We found that Oct4-GFP-positive cells emerged at 6–7 days postinduction (dpi) during both GATA3- and OCT4-induced reprogramming (Figure 1C). We found that GATA3 may mainly function at 4–7 dpi (Figure 1D), which corresponds to the period during which the pluripotency circuitry is reconstructed (Polo et al., 2012).

iPSCs Generated with GATA3 Are Pluripotent

The iPSCs generated using GATA3, SOX2, KLF4, and c-MYC (G3SKM) had morphology similar to mouse ESCs (Figures 1E and S2A). The G3SKM-induced iPSCs were stable during long-term passaging and stained positive for alkaline phosphatase (AP), SSEA-1, UTF1, and NANOG (Figures 1E, S2A, and S2B). The methylation levels of the Nanog and Oct4 promoters were similar to the methylation levels in mouse ESCs (Figure S2C). Genomic integrations of the viruses into the genomic DNA were confirmed in iPSCs, teratomas, and tissues from chimeric mice, and showed no OCT4 transgene integration (Figure S2D). The expression of endogenous pluripotency-associated genes was activated, and the expression of exogenous GATA3, SOX2, KLF4, and c-MYC was silenced in these cells (Figure S2E), which indicates that they were fully reprogrammed. G3SKM-induced iPSCs produced germline-competent chimeras (Figures 1F and 1G), and these iPSCs were further validated by the characterization of teratoma formation, gene expression profiles, and other assays (Figures S2F and S2G and Tables S2).

GATA3 Has Little Effect on the Events Noted in Previous Studies

To identify the potential mechanisms by which GATA3 could replace OCT4, we examined several pathways that have been reported to facilitate or inhibit reprogramming. To test whether GATA3 could activate endogenous Oct4 to a high level shortly after induction so that it was the activated endogenous Oct4 plus SKM that induced pluripotency, we monitored the endogenous Oct4 expression levels using doxycycline (dox)-inducible G3SKM-secondary MEFs that induced pluripotency, we monitored the endogenous Oct4 response levels using doxycycline (dox)-inducible G3SKM-secondary MEFs. We found that GATA3 did not significantly activate Oct4 expression up until Oct4-GFP-positive cells had begun to emerge (Figure 1H). There was also no significant activation of the pluripotency-associated factor Nanog until Oct4-GFP-positive cells began to emerge (Figure S1C). We found that GATA3 had little effect on mesenchymal-epithelial transition (MET) (Figure 2A) (Li et al., 2010;...
In addition, p53 and p21 expression was not affected by GATA3 (Figure 2A) (Zhao et al., 2008). Moreover, we did not observe any significant changes in cell proliferation (Figure 2B), which suggests that GATA3 does not contribute to reprogramming by accelerating cell proliferation. These results suggest that GATA3 functions in a manner different from the mechanisms noted in previous studies to promote reprogramming.

Figure 2. OCT4 and GATA3 Can Inhibit a Group of ECT Genes that Are Elevated by SKM during Reprogramming
(A) RT-qPCR analysis of the expression of endogenous Cdh1, Cdh2, Snai1, p53, and p21 relative to their expression in MEFs. The samples were tested at 4 dpi. Error bars indicate the SD (n = 3).
(B) Cell proliferation curves. SKM was used as a control, and 50,000 MEFs were seeded per well. Three wells were harvested every 24 hr to count the number of cells. The cell populations were analyzed using a paired two-tailed Student’s t test. No significant difference was observed between G3SKM- and SKM-infected MEFs (p = 0.7803). Error bars indicate the SD (n = 3).
(C) A Venn diagram illustrating the overlap between the expression changes identified in cells induced with G3SKM and OSKM. The blue circle represents the DE lists (expression change > 2-fold, p < 0.05, FDR < 0.01) in G3SKM-induced cells compared with SKM-induced cells. The purple circle represents the DE lists of OSKM-induced cells compared with SKM-induced cells.
(D) Gene ontology (GO) analysis of the overlapping genes in (C). The GO analysis was based on the DE list. The p values represent the Bonferroni-corrected EASE score.
(E) GO analysis of genes regulated by SKM, G3SKM, and OSKM at 7 dpi. The GO analysis was based on the DE list. The p values represent the Bonferroni-corrected EASE score.
(F) Heatmaps depicting the relative fold change of gene expression at 7 dpi based on the ECT-related genes by RNA-seq. Red and green indicate increased and decreased expression, respectively.
OCT4 and GATA3 Inhibit the ECT-Related Genes that Are Elevated by SKM

To further determine the roles of OCT4 and GATA3 in inducing pluripotency, we performed RNA-sequencing (RNA-seq) assays at 7 dpi based on our kinetics test (Figures 1C and 1D). The shared targets of OCT4 and GATA3 were involved in several important processes, such as cell adhesion and the regulation of transcription, based on gene ontology (GO) function enrichment analysis (Figures 2C and 2D). Interestingly, we observed that, compared with MEFs, the upregulated genes that were most significantly enriched were related to ectoderm and epidermis development following the introduction of SKM. When OCT4 and GATA3 were introduced along with SKM, the most significantly enriched downregulated genes were also related to ectoderm and epidermis development (Figures 2E and 2F). These results suggest that SKM may have the potential to direct cells toward an ECT state and that the inhibition of ECT-related genes by OCT4 and GATA3 may facilitate the induction of pluripotency.

Screening of Other Lineage Specifiers as Substitutes for OCT4

The current model proposes that ESCs are maintained by a shield of pluripotency factors that collaboratively prohibit differentiation into any lineage to preserve an undifferentiated state (Boyer et al., 2005; Silva et al., 2009). A more recent and provocative perspective is that pluripotency factors function as classical lineage-specifying factors to direct the differentiation of ESCs into a specific lineage while inhibiting their commitment to mutually exclusive lineages (Loh and Lim, 2011). In ESCs, Oct4 promotes ME and primitive endoderm differentiation and suppresses ECT differentiation (Niwa et al., 2000; Thomson et al., 2011; Wang et al., 2012). Interestingly, the overexpression of Gata3 in ESCs directs primitive endodermal differentiation, and similar to Oct4, Gata3 also functions in some ME lineages (Nishiyama et al., 2009). The similarity of OCT4 and GATA3 in lineage specification suggests that other lineage specifiers may also replace OCT4. To test this possibility, we screened several additional lineage specifiers using viral overexpression (Table S3). The results indicated that the lineage specifiers GATA6, SOX7, PAX1, GATA4, CEBPa, HNF4a, and GRB2 were able to substitute for OCT4 to induce pluripotency (Figures 3A, 3B, and S1E). These lineage specifiers have been shown to function mainly during multiple stages of ME differentiation and in early embryonic patterning (Table S4). Most of these substitutes for OCT4 are not enriched in ESCs (Figure S1D). However, the lineage specifiers that are primarily involved in ectodermal lineage specification (e.g., SOX1 and ASCL1) were unable to replace OCT4 (Table S3).

Other OCT4 Substitutes Also Inhibit ECT-Related Genes Elevated by SKM

Because both OCT4 and GATA3 attenuated the elevated expression of the group of ECT genes that was collectively triggered by SKM, we asked whether the other substitutes for OCT4 could also attenuate their expression. We performed microarray assays to analyze the roles of other OCT4 substitutes. In addition, we analyzed another lineage specifier, MIXL1, as a negative control, which could not replace OCT4 despite its involvement in ME lineage specification. Consistent with our expectations, the other OCT4 substitutes, but not MIXL1, attenuated the upregulation of ECT genes by SKM (Figure 3C). GO analysis also showed that the most significantly enriched downregulated genes after the introduction of OCT4 substitutes other than MIXL1 were associated with ectoderm and epidermis development and other terms related to ECT specification (Figure S3).

Next, to validate the tendency apparent from the RNA-seq and microarray analyses (Figures 2F and 3C), we performed real-time quantitative PCR (RT-qPCR) assays. We observed that typical ECT marker genes, such as Dlx3 and Lhx5, were upregulated by SKM and significantly downregulated after the introduction of OCT4 and its substitutes (Figures 3D). To further validate these findings, we performed single-colony assays. We acquired AP-positive, Oct4-GFP-negative colonies in different conditions, including controlled SKM conditions. Then, we mechanically split each colony into two subcolonies. We analyzed one subcolony for gene expression while tracing the other half subcolony in culture to monitor whether it became Oct4-GFP positive and dox independent (Figure S4). In this case, we could analyze those reprogramming-competent colonies that were headed toward pluripotency. AP-positive colonies grown in SKM conditions failed to become Oct4-GFP positive, suggesting that they are reprogramming refractory. We found a concomitant downregulation of Dlx3 and Lhx5 after the introduction of OCT4 and its substitutes (Figure 3E). This finding is consistent with the results from mixed cell populations, supporting the lineage specification tendency observed in the bulk population. These results suggest that OCT4 and its substitutes can repress the upregulation of ECT-related genes that are triggered by SKM.

Suppression of Dlx3, a Master ECT-Related Gene, Facilitates Reprogramming in the Presence of SKM

To test whether the inhibition of ECT-related genes by OCT4 and its substitutes correlated with the promotion of reprogramming, we further evaluated whether the knockdown of the master ECT lineage-related genes that are upregulated by SKM could also promote reprogramming in the absence of OCT4. We found that the knockdown of the key ECT marker Dlx3, which subsequently resulted in the downregulation of several other ECT genes, promoted reprogramming in the absence of OCT4, rather than in the absence of SOX2 (Figures 4 and S5), functionally supporting the data observed by gene profiling (Figures 2E, 2F, and 3C–3E). Furthermore, iPSCs generated with Dlx3 shRNAs were proven to be pluripotent by stringent characterization assays (Figures 4E, 4F, and Table S2). We found no significant binding of Oct4 and Gata3 to the Dlx3 promoter region in published ChIP data from “hmChIP,” which suggests that Oct4 and Gata3 may indirectly regulate Dlx3 (Figure S5D). Together, our data indicate that OCT4 and its substitutes attenuate the upregulation of ECT-related genes that is triggered by the SKM combination; this inhibition correlates with facilitating the induction of pluripotency.

SOX2 and Its Substitutes Attenuate the Expression of ME-Related Genes Elevated by OKM

Similar to Oct4, Sox2 also regulates lineage specification in ESCs. Sox2 inhibits mesendodermal differentiation and
Figure 3. Identification of Other Substitutes for OCT4 that also Inhibit ECT-Related Genes Elevated by SKM during Reprogramming

(A) Identification of additional lineage specifiers that are related to the ME lineage that can replace OCT4 during reprogramming. The Oct4-GFP-positive colonies were counted at 9 dpi. SKM plus EV was used as a control. Error bars indicate the SD (n = 3).

(B) GFP images of iPS colonies generated with SKM+GATA6 (G6SKM), SKM+GATA3 (G3SKM), SKM+SOX7 (S7SKM), SKM+PAI1 (P1SKM), SKM+GATA4 (G4SKM), SKM+CEBPα (CaSKM), SKM+HNF4α (H4SKM), and SKM+GRB2 (GrSKM). Scale bars, 100 μm.

(C) Heatmaps depicting the relative fold change of gene expression at 7 dpi based on the ECT-related genes by microarray. Red and green indicate increased and decreased expression, respectively.

(D) RT-qPCR analysis of the expression of endogenous Dlx3 (top) and Lhx5 (bottom) relative to the expression in SKM. The samples were tested at 7 dpi after removal of Oct4-GFP-positive cells. Significance was assessed compared with the controls using a one-tailed Student’s t test. ***p < 0.001. Error bars indicate the SD (n = 3).

(E) RT-qPCR analysis of the expression of endogenous Dlx3 (top) and Lhx5 (bottom) relative to the expression in SKM from a single colony; three colonies from each condition were tested. Significance was assessed compared with the controls using a one-tailed Student’s t test. *p < 0.05. Error bars indicate the SD (n = 3).

See also Figures S3 and S4 and Tables S2, S3, S4.

promotes neural ectodermal differentiation (Thomson et al., 2011; Wang et al., 2012). Intuitively, we asked whether the lineage specifiers involved in ECT lineage specification could substitute for SOX2 during reprogramming. Indeed, the previously identified SOX2 substitutes, Sox1, Sox3, and RCOR2, are regulators of ECT development or are particularly expressed in neural tissues (Nakagawa et al., 2008; Tontsch et al., 2001; Yang et al., 2011; Zeng et al., 2010). In addition, we demonstrated that GMNN, which is involved in ECT lineage specification (Seo et al., 2005), could substitute for SOX2 to promote reprogramming (Figures 5A, 5B, S1E, and Table S4).

To assess whether SOX2 and its substitutes could attenuate the upregulation of the ME-related genes that are induced by OCT4, KLF4, and c-MYC (OKM), we performed microarray and RT-qPCR analyses. The expression of a group of classical ME markers was elevated after the induction of OKM; these elevated expression levels were attenuated by SOX2 and its substitutes (Figures 5C and 5D). Furthermore, we performed similar single-colony assays as previously mentioned to test this tendency in reprogramming-competent cells. We observed that SOX2 and its substitutes inhibited the expression of T and Eomes in these colonies (Figure 5E). These data indicate that SOX2 and its substitutes attenuate the upregulation of ME-related genes that is induced by OKM during the reprogramming process.
A “Seesaw Model” Suggests that the Balancing of Counteracting Forces Facilitates the Induction of Pluripotency

A two-node model of cell fate determination has been studied in various instances of pluripotent stem or progenitor cells that undergo a binary cell fate decision (e.g., GATA1 and PU.1, RUNX2, and PPARg) (Huang, 2009). This circuit has already hinted at the concept of a “balanced pluripotent state” (Figures S6A and S6B). To better understand the induction of pluripotency by lineage specifiers, we propose a “seesaw model” (Figures 6A and 7) based on previous studies (Hanna et al., 2009; Loh and Lim, 2011; Niakan et al., 2010; Niwa et al., 2000; Polo et al., 2012; Rosa and Brivanlou, 2011; Teo et al., 2011; Thomson et al., 2011; Wang et al., 2012) and our current data (Figures 2, 3, 4, 5, S3, S4, and S5). This model consists of two coupled modules: the pluripotency module and the differentiation module. The pluripotency module is represented by the mutual activation of Oct4 and Sox2, whereas the differentiation module is modeled by mutually inhibiting the ME and ECT genes (Figure 6A).

Simulations of the model produced three cell states: the pluripotency state, the ME state, and the ECT state (Figures 6B and S6C). In somatic states, the pluripotent state is unreachable (Figure 6C, box i). However, overexpression of KLF4 and c-MYC (KM) may act as a driving force that reduces reprogramming barriers (Polo et al., 2012), so the pluripotency state becomes a local attractor that is reachable from a somatic state if properly perturbed (Figure 6C, box ii). Note that the pluripotent state is most easily reached (high pluripotency potential) if the expression strengths of the ME and ECT genes are forced to be in the blue region, i.e., if they are balanced.

A key result of this model is that pluripotency and reprogramming can be achieved by a variety of perturbations to the system, as illustrated in Figure 6D. For example, the overexpression of SOX2 alone will switch on the ECT genes that inhibit the pluripotency module, in turnbestowing the ECT state onto the cell, regardless of the initial cell state (Figures 6D, box i, S6D, and S6E). The overexpression of OCT4 (or the ME gene GATA3) in combination with SOX2 counteracts the inhibition of the ECT genes to the pluripotency module, which permits the cell to move into the pluripotency attractor region (Figures 6D, boxes ii and iii, S6D, and S6E). More strategies for restoring pluripotency as suggested by the model are shown in Figures 6D and S6 and Tables S5 and S6.

Some of the results from our model are consistent with previous studies (Table S6). Deviation from the balanced equilibrium for pluripotency directs cells to flow into divergent differentiated states. For example, the enforced expression of OCT4 results in cellular transdifferentiation into mesodermal hematopoietic cells (Szabo et al., 2010). SOX2, SKM with neural lineage specifiers, and reduced Oct4 expression in the OSKM cocktail can result in neuroectodermal transdifferentiation from fibroblasts (Han et al., 2012; Ring et al., 2012; Thier et al., 2012). Other strategies identified by our
Figure 5. SOX2 and Its Substitutes Attenuate the Expression of ME Genes that Are Elevated by OKM during Reprogramming

(A) Identification of lineage specifiers related to the ECT lineage that can facilitate reprogramming in the presence of OKM. The Oct4-GFP-positive colonies were counted at 9 dpi. OKM plus EV was used as a control. Error bars indicate the SD (n = 3).

(B) GFP images of iPS colonies generated with OKM+SOX1 (OS1KM), OKM+SOX3 (OS3KM), and OKM+GMNN (OGmKM). Scale bars, 100 μm.

(legend continued on next page)
model call for further investigation to validate or falsify the model.

One of the striking predictions of the model is that, when expressed together, an ME specifier and an ECT specifier can synergistically enhance reprogramming in the absence of both OCT4 and SOX2. The mutually antagonistic nature of the differentiation module makes the system behave like an easily tilted “seesaw”; when the system tends to fall into either the ME or ECT state, in turn suppressing the pluripotency module, the cell is left with no chance to enter the pluripotency state. However, when both ME and ECT specifiers are exogenously expressed at appropriate levels, their external expression may be unable to compensate for the decrease in their endogenous expression due to the mutual inhibition. The differentiation module may end up in the balanced region, wherein neither ME specifiers nor ECT specifiers are present at levels that are sufficiently high to suppress the pluripotency module. Then, a self-activating pluripotency module may switch on by itself, eliciting the serial activation of a pluripotency network to realize successful reprogramming (Figures 6D, box v, S6G, and S6R–S6T).

To test the feasibility of this idea, we screened different combinations of lineage specifiers. We found that either (1) GATA3, GATA6, PAX1, and SOX1 or SOX3 or (2) GATA6 plus GMNN can promote successful reprogramming without the two most critical pluripotency factors, OCT4 and SOX2 (Figures 6E, 6F, and S1E). Furthermore, these iPSCs were proven to be pluripotent (Figure S7). These lineage specifiers are not enriched in ESCs (Figure S1D), which indicates that the induction and maintenance of pluripotency may be two different scenarios.

DISCUSSION

Our data present the first evidence that the lineage specifiers GATA3, GATA6, SOX7, PAX1, GATA4, CEBPA, HNF4a, GRB2, and GMNN, which are generally considered as pluripotency rivals, can unexpectedly facilitate reprogramming and replace reprogramming factors of a corresponding lineage-specifying potential. We suggest a “seesaw model” in which lineage specifiers facilitate reprogramming when they are balanced with other mutually exclusive lineage specifiers. There are multiple ways to reach the pluripotency state, all involving balancing the forces that would otherwise drive the cell to a differentiated fate. Deviation from the balance directs cells to divergent differentiated states, reducing the possibility of restoring pluripotency. This hypothesis is complementary to previous studies demonstrating that lineage specifiers such as Gata6, a reported repressor of Nanog, hinder the reprogramming process (Chazaud et al., 2006; Mikkelsen et al., 2008).

We suggest that pluripotency factors may direct lineage specification and that lineage specifiers may in turn influence the induction of pluripotency; this model provides the conceptually new perspective that one can reach the pluripotency “destination” by getting on cars that do not normally run at the destination.

Our “seesaw model” is in agreement with the recent perspective that pluripotency factors act as classical lineage-specifying factors to direct the differentiation of ESCs into a specific lineage while cross-inhibiting commitment to mutually exclusive lineages (Loh and Lim, 2011). The collaborative result is the temporary inhibition of all lineages to preserve an undifferentiated state. The self-activating pluripotency module could be activated when all of the lineage-specifying forces are counteracted, i.e., at the dynamic balanced point of the “seesaw,” wherein no particular lineage-specifying activity is dominant to inhibit the pluripotency module. Once the pluripotency module is activated, the ME and ECT lineage fates are blocked by SOX2 and OCT4, respectively, and the pluripotency state is maintained.

Most of these substitutes are usually depicted as typical ME markers (GATA6, GATA4, etc.) or ECT markers (SOX1, SOX3, etc.). To make our model succinct, these lineage specifiers are classified as either ME or ECT. However, this classification is not absolute, as indicated by our “seesaw model.” The ME and ECT lineages are the two major counteracting lineages, but in the reprogramming symphony, an enormously intricate self-conflicted coalition of different lineage specifiers that are not limited to the ME and ECT lineages orchestrates the concerted induction of pluripotency.

The “seesaw model” focuses on one major aspect in reprogramming: understanding the implications of the reprogramming landscape that is shaped by the interactions among the different cell states. Needless to say, reprogramming is not merely a “seesaw.” Several other important events, such as epigenetic changes, are also involved in this process. Furthermore, there are several important barriers to overcome for successful reprogramming. The various reprogramming factors execute their corresponding functions to facilitate the induction of pluripotency. Concomitantly, these factors may induce side effects, such as the stimulation of lineage specification. In addition to their reported functions in the activation of the pluripotency circuit and in epigenetic and cell-cycle regulation, KLF4 and c-MYC may also promote stimuli for particular lineages (Cao et al., 2012; Smith et al., 2010; Sridharan et al., 2009) to concert an equilibrium symphony. The function of these lineage specifiers may be involved in other processes and warrants further in-depth study.

We hope that our model can serve as a starting point to decipher the mysterious reprogramming code in a novel

---

(C) Heatmaps depicting the relative fold change of gene expression at 7 dpi based on the ME-related genes by microarray. Red and green indicate increased and decreased expression, respectively.

(D) RT-qPCR analysis of the expression of endogenous T (left) and Eomes (right) relative to the expression in OKM. Significance was assessed compared with the controls using a one-tailed Student’s t test. **p < 0.001. The samples were tested at 7 dpi. Error bars indicate the SD (n = 3).

(E) The expression of endogenous T (top) and Eomes (bottom) relative to the expression in OKM from a single colony; three colonies from each condition were tested. Significance was assessed compared with the controls using a one-tailed Student’s t test. *p < 0.05. **p < 0.001. Error bars indicate the SD (n = 3). See also Figure S4 and Tables S2 and S4.
Figure 6. Balanced Equilibrium Facilitates the Induction of Pluripotency

(A) A model for the coupled pluripotency module (self-activation of Oct4 and Sox2) and for the differentiation module (mutual antagonism between the ME and ECT lineages).

(B) The cell-state landscape obtained from the density of the trajectories. The blue color indicates deep “attractors” near which the cell states tend to stay. The x axis represents the difference between the two groups of lineage inducers, and the y axis indicates the “pluripotency” of the cell.

(legend continued on next page)
scenario. Our results suggest that a chemical screen for small molecules that can substitute for OCT4 and SOX2 in directing the corresponding lineage specification would be a novel and feasible strategy to generate iPSCs. Our findings also demonstrate the need for further investigation of the conventional perspectives regarding the functions of pluripotency factors and lineage specifiers and on the manner in which the pluripotency regulatory circuitry is established during reprogramming.

**EXPERIMENTAL PROCEDURES**

**Screen**
The 10,080 cDNAs that were screened for their ability to replace OCT4 were obtained from Origene Co., Ltd. Oct4-GFP MEFs were infected with SKM (SOX2, KLF4, and c-MYC) lentiviral supernatant (MOI: 4) supplemented with 10 ng/μl polybrene (Sigma) for 12 hr. Details are provided in the **Extended Experimental Procedures**.

**Cell Culture**
Primary MEFs, mouse GECs, keratinocytes, and MDFs were derived from ICR × Oct4-GFP transgenic mice. Mouse GECs were isolated and cultured as previously described (Aoi et al., 2008). Details are provided in the **Extended Experimental Procedures**.

**iPSCs Generation**
The dox-inducible lentiviral system used has been previously described (Maherali et al., 2008). Further details are provided in the **Extended Experimental Procedures**.

**Time Course of GATA3 Induction**
A dox-inducible lentiviral system was used, so the expression of tet-GATA3, but not pII3.7-ΔU6ΔGFP-SOX2, pII3.7-ΔU6ΔGFP-KLF4, or pII3.7-ΔU6ΔGFP-c-MYC (Zhao et al., 2008), could be induced by dox. Dox was added to the culture medium for different periods of time, and the number of Oct4-GFP-positive colonies was counted at 10 dpi.

**Characterization of iPSCs**
The teratoma formation and chimera experiments were performed as previously described (Li et al., 2011). The primers used for RT-PCR, genomic PCR, and RT-qPCR are listed in **Table S7**. Bisulfite treatment of DNA was performed using the CpGenome Fast DNA Modification Kit (Millipore). The primers used for promoter fragment amplification have been previously described (Takahashi and Yamanaka, 2006). The details for immunofluorescence and the detection of AP activity are provided in the **Extended Experimental Procedures**.

**Single-Colony Colony Analysis**
Five days after differentiation with reprogramming cocktails, Alkaline Phosphatase Live Stain (Invitrogen) was used for the detection of AP activity. AP-positive colonies were then picked, and each colony was divided into two parts. One part was kept culturing in induction medium for 2–4 days. The other part was frozen at −80°C. If Oct4-GFP turned on after an additional 2–4 days induction and became dox independent in 4 days without dox, total RNA was isolated from the frozen part using the RNeasy micro kit (QIAGEN) and then used for RT-qPCR analysis.

**Knockdowns**
Dlx3 knockdown was achieved using shRNA lentiviral vectors (Sigma) containing a puromycin resistance gene. Further details are provided in the **Extended Experimental Procedures**.

**Western Blotting**
Cells were washed with PBS, treated with 1 × SDS sample buffer, and then boiled. Additional details are provided in the **Extended Experimental Procedures**.

**DNA Microarray**
Total mRNA from iPSCs, MEFs, MEFs at 7 dpi, and R1 were labeled with Cy5 hybridized to a mouse oligo microarray (Phalanx Mouse Whole Genome OneArray; Phalanx Biotech). The data were analyzed according to the manufacturer’s protocol.

**RNA-Seq**
RNA-sequencing libraries were constructed using the Illumina mRNA-seq Prep Kit. The fragmented and randomly primed 200 bp paired-end libraries were prepared.

See also Figures S6 and S7 and Tables S5 and S6.
were sequenced using Illumina HiSeq 2000. Details are included in the Extended Experimental Procedures.

Statistical Analysis

Cell proliferation curves were analyzed using a paired two-tailed Student’s t test. The unpaired one-tailed t test was used for statistical analysis of other experiments. The standard deviation (SD) was used to quantify variability.

Model Construction

A coarse-grained model was built based on our results and the results of previous studies. The interactions involved in this model are illustrated in Figure 6A. Additional details are provided in the Extended Experimental Procedures.

ACCESSION NUMBERS

Microarray and RNA-seq data are available in the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE43995.

SUPPLEMENTAL INFORMATION

Supplemental Information includes the Extended Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.cell.2013.05.001.

ACKNOWLEDGMENTS

We thank Chengyong Wang and Xiaolei Yin for helpful discussions. We also thank Haisong Liu, Jiancheng Wang, Zhiyuan Hou, Ren Li, Fangfang Zhu, Dongyi Xu, and Xing Zhang for technical assistance. J.S., C.W., and Y.W. conceived of and supervised the study. This work was supported by grants from the Key New Drug Creation and Manufacturing Program (grant numbers 2012CB966401, 2010CB945204, and 2009CB918500), the National Natural Science Foundation of China (grant numbers 90919031 and 11021463), the NIH project, and NSF (CMII-0941355).

Received: October 26, 2012

Revised: April 15, 2013

Accepted: May 13, 2013

Published: May 23, 2013

REFERENCES


