KLF4 N-Terminal Variance Modulates Induced Reprogramming to Pluripotency

Shin-Il Kim,1,2 Fabian Oceguera-Yanez,1 Ryoko Hirohata,1 Sara Linker,1,3,6 Keisuke Okita,1 Yasuhiro Yamada,1,2 Takuya Yamamoto,1,2 Shinya Yamanaka,1,2,4 and Knut Woltjen1,2,5,*  
1Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto 606-8507, Japan  
2Institute for Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Kyoto 606-8507, Japan  
3John P. Hussman Institute for Human Genomics, University of Miami, Miami, FL 33136, USA  
4Gladstone Institute of Cardiovascular Disease, San Francisco, CA 94158, USA  
5Hakubi Center for Advanced Research, Kyoto University, Kyoto 606-8501, Japan  
6Present address: Salk Institute for Biological Studies, La Jolla, CA 92037, USA  
*Correspondence: woltjen@cira.kyoto-u.ac.jp  
http://dx.doi.org/10.1016/j.stemcr.2015.02.004  
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

As the quintessential reprogramming model, OCT3/4, SOX2, KLF4, and c-MYC re-wire somatic cells to achieve induced pluripotency. Yet, subtle differences in methodology confound comparative studies of reprogramming mechanisms. Employing transposons, we systematically assessed cellular and molecular hallmarks of mouse somatic cell reprogramming by various polycistronic cassettes. Reprogramming responses varied in the extent of initiation and stabilization of transgene-independent pluripotency. Notably, the cassettes employed one of two KLF4 variants, differing only by nine N-terminal amino acids, which generated dissimilar protein stoichiometry. Extending the shorter variant by nine N-terminal amino acids or augmenting stoichiometry by KLF4 supplementation rescued both protein levels and phenotypic disparities, implicating a threshold in determining reprogramming outcomes. Strikingly, global gene expression patterns elicited by published polycistronic cassettes diverged according to each KLF4 variant. Our data expose a Klf4 reference cDNA variation that alters polycistronic factor stoichiometry, predicts reprogramming hallmarks, and guides comparison of compatible public data sets.

INTRODUCTION

Cellular identity can be guided by ectopic expression of master regulators (Graf, 2011). Deriving induced pluripotent stem cells (iPSCs) through the activities of OCT3/4, SOX2, KLF4, and c-MYC (Takahashi and Yamanaka, 2006) provides a potent model in which to study the role of transcription factor coordination in driving somatic cells toward pluripotency. Early mechanistic studies using mouse embryonic fibroblasts (MEFs) were conducted through de novo introduction of viral vectors, each expressing an individual (monocistronic) reprogramming factor (Brambrink et al., 2008; Stadtfeld et al., 2008a), where modulation of factor levels by viral titration led to altered reprogramming characteristics (Yamaguchi et al., 2011). Monocistronic reprogramming allows for variation in copy number and integration site, and as a result, stoichiometry is inconsistent on a cell-to-cell level. Therefore, this method was succeeded by the development of polycistronic factor cassettes that can produce multiple proteins from one single transcript (Kaji et al., 2009; Sommer et al., 2009). Although such fixed polycistronic stoichiometry revealed the importance of relative factor ratios in determining the quality of reprogramming (Carey et al., 2011), the principles that establish optimal stoichiometry remain undefined. Studies of the mechanisms that underlie somatic cell reprogramming have revealed multi-step processes involving proliferation and cell-cell adhesion, along with molecular changes such as downregulation of lineage-specific genes and eventual upregulation of pluripotency markers (Plath and Lowry, 2011). Cell-surface markers were associated with reprogramming stages such as emergence of the embryonic stem cell (ESC) marker SSEA-1 (stage-specific embryonic antigen 1) (Polo et al., 2012; Stadtfeld et al., 2008a). Secondary reprogramming systems (Woltjen et al., 2009) helped define initiation, maturation, and stabilization as key stages in reprogramming toward pluripotency (David and Polo, 2014). Proliferation, colony formation, and a mesenchymal-to-epithelial transition (MET) define the initiation phase (Samavarchi-Tehrani et al., 2010), while stabilization is characterized by transgene independence and activation of pluripotency reporters such as Nanog and Dppa4 (Golipour et al., 2012). Thus, changes in global gene expression and epigenetics were associated with the progression of reprogramming through these stages (Theunissen and Jaenisch, 2014). However, discrepancies in reprogramming platforms influence reprogramming hallmarks, the severity of MET responses, lineage-specific gene repression and ectopic activation, the timing of cell-surface marker presentation, and the frequency of partial and complete reprogramming (Golipour et al., 2012; Mikkelsen et al., 2008; O’Malley et al., 2004).
et al., 2013; Polo et al., 2012; Samavarchi-Tehrani et al., 2010; Wernig et al., 2008).

In order to clarify such method-dependent reprogramming hallmarks, we applied standard assays to compare polycistronic cassettes (constructed in-house or obtained from public resources) in a drug-inducible piggyBac (PB) transposon reprogramming system. The induced expression of various polycistronic cassettes in mouse fibroblasts evoked phenotypic and gene expression changes that we divided into two basic behavioral classes. An examination of individual factor sequences across the panel of polycistronic constructs revealed a re-occurring discrepancy in the Klf4 cDNA defining two coding region length variants. Increased length at the KLF4 N terminus was associated with higher protein levels and thus altered relative factor stoichiometry, ultimately impacting both the initiation and stabilization phases of iPSC derivation. Here, we report the consequences of KLF4 N-terminal variation in mono- and polycistronic reprogramming experiments, and apply these findings to recognize and reconcile differences in reprogramming characteristics implied hitherto.

RESULTS

A Transposon System for Collating Polycistronic Reprogramming Cassettes

Reprogramming studies in mouse have made use of unique polycistronic factor arrangements and delivery vectors. For uniform evaluation of factor-order effects, we employed a fundamental reprogramming scheme based on factor transposition in MEFs (Woltjen et al., 2009). The PB transposon vector (PB-TAC) employs doxycycline (dox)-responsive reprogramming cassette expression co-incidently with a mCherry reporter (Figure 1A). ROSA-rtTA; Nanog-GFP MEFs combine the m2-rtTA transactivator (Ohnishi et al., 2014) with a Nanog-GFP reporter (Okita et al., 2007). Thus, dox-responsive, PB-TAC-transgenic cells can be monitored throughout reprogramming initiation and maturation (day 2 [d2]–14) by mCherry fluorescence, while stabilization of pluripotency (d14–18) is indicated by activation of Nanog-GFP. Gain of factor independence through autonomous transgene silencing despite continued dox treatment is an established hallmark of the stabilization phase (Golipour et al., 2012), signaled here by a decrease in mCherry expression. For all polycistronic reprogramming cassettes tested, we routinely passaged populations on d8 and d18 using equal cell numbers without fractionation for extended culture and fluorescence-activated cell sorting (FACS) analysis. Dox-independent maintenance of iPSCs was verified by culture until d24, after which pluripotency was assayed by gene expression array and chimera contribution.

Initially, we used an alkaline phosphatase-positive (AP+) colony formation assay on d10 to compare various polycistronic cassettes with two previously published versions: OSKM (Carey et al., 2009) and MKOS (Kaji et al., 2009). Across the range of vectors tested, OKMS, constructed through the combination of OSKM and pCX-OKS-2A (Okita et al., 2008), displayed a reprogramming phenotype

Figure 1. KLF4 Isoforms Underlie Phenotypic Differences in Reprogramming

(A) Diagram of the PB reprogramming system and analysis steps. The PB-TAC transposon delivers inducible, reporter-linked (mCherry), polycistronic reprogramming cassettes (Poly) into ROSA-rtTA; Nanog-GFP MEFs (d-1). Cultures are induced with dox (d0, dox; filled arrows) and harvested on d8 for FACS and passage for late-stage analysis (d18). For AP staining, cultures are maintained without passage until d10. Full reprogramming and transgene independence is assessed by dox withdrawal from d18 to d24 (open arrows). Predicted expression periods for the mCherry and GFP reporters during iPSC derivation are shown with red and green gradated bars. Blue polygons represent PB 3' (left) and 5’ (right) inverted terminal repeats. tet0, dox-responsive promoter; IRES, internal ribosome entry signal; pA, polyadenylation signal.

(B) AP staining on d10 of OSKM and OKMS reprogramming cultures. Scale bar, 4,000 μm.

(C) Day 18 fluorescence microscopy of entire wells (composite 10 × 10 fields) or selected insets (3 × 3 subfields) for mCherry* and Nanog-GFP* (left). Scale bars, 4,000 μm (full well) and 1,000 μm (inset). FACS analysis of mCherry and Nanog-GFP expression in d18 Total and SSEA-1* populations (right).

(D) Polycistronic cassette structure and sequence of the 2A-Klf4 N-terminal cleavage junctions. Cloned cDNA is compared with murine Klf4 GenBank mRNA sequences (U20344.1 and U70662.1), indicating the position of the predicted initiation codons and amino acid translations. The N-terminal 9aa of U20344.1 were introduced into OKMS (Klf4j) to produce OK*MS (Klf4j).

(E and F) The OK*MS construct was evaluated according to the assays outlined in (A)–(C). The results in (B), (C), (E), and (F) are representative of the results from at least three independent experiments (summarized in Figures S1A and S1B).

(G and H) Time-course analysis of mCherry cell expansion (G) and SSEA-1* fractions (H) from OSKM, OKMS, and OK*MS transfections cultured for the indicated number of days before (2, 4, 6, 8) and after (10, 14, and 18) passage. Means ± SE for three independent experiments.

See also Figure S1.
contradictory to OSKM and MKOS (Figure 1B and data not shown). Despite similar transfection efficiencies (~3%-5% mCherry+ cells at d2 for both OSKM and OKMS), OKMS-induced AP+ colonies were clearly more abundant by d10, indicating that OKMS initiated somatic reprogramming more robustly than OSKM (Figures 1B and S1A).

To evaluate iPSC stabilization, we assayed activation of the Nanog-GFP pluripotency reporter and silencing of factor-linked mCherry after d8 passage of OSKM and OKMS reprogramming cultures. Whole-well microscopic evaluation revealed that most OKMS-induced colonies were unable to activate Nanog-GFP by d18 (Figure 1C, left), whereas mCherry expression remained high. In contrast, OSKM cultures showed a majority of GFP+ colonies with silenced mCherry. FACS on d18 confirmed the proportions of reporter-positive cells (Figure 1C, right). Both the Total GFP+ and SSEA-1+ GFP+ fractions on d18 were consistently higher for OSKM reprogramming (mean 15.6% and 39.5%, respectively, versus 5.4% and 9.9% for OKMS; Figure S1B). Conversely, overrepresentation of Total mCherry+ and SSEA-1+ mCherry+ populations was characteristic of OKMS (mean 53.1% and 58.8%, respectively, versus 21.0% and 27.0% for OSKM; Figure S1B). GFP+ cells emerged only in the SSEA-1+ fraction (data not shown), and only in mCherrylow or mCherry- cells, such that double-positive GFP+ mCherry+ cells were not observed under either condition (Figure 1C, right). Reporter ratios remained mostly unchanged after d18 passage and an additional 6 days of culture in dox (d24; Figure S1C, left). Failure to activate Nanog-GFP was not associated exclusively with transgene interference, as dox withdrawal led to a complete loss of mCherry, but no increase in GFP for either the OSKM or OKMS populations (Figures S1C and S1D). For both vectors, dox-independent, Nanog-GFP+ iPSC clones contributed to chimera formation and clustered with ESCs in microarray transcriptome analyses, validating pluripotency (Figures S1E–S1G).

Extended FACS analysis revealed the kinetics of mCherry+ cell expansion throughout the early phase (d2–8), equating OKMS with OSKM (Figure 1G). From our d8 FACS data, we found that a large proportion of OKMS mCherry+ cells were positive for SSEA-1. Therefore, we traced SSEA-1 dynamics from d2 to d8 (mCherry+ fraction; Figure 1H, left) and d10 to 18 post-passage (Total population; Figure 1H, right). Notably, only OKMS cultures were distinguished by rapid and robust SSEA-1 activation as early as d2 (29.4% versus 4.8% for OSKM, and 8.5% for OKMS), which continued through early reprogramming. For all constructs, SSEA-1- cells remained mCherry+ through d8, at which point no Nanog-GFP activation could be detected. An examination of the increasing proportion of SSEA-1- cells over d10–18 (Figure 1H, right) showed that activation of Nanog-GFP and coincident suppression of mCherry by OKMS was more consistent with OKMS, whereas for OKMS, both mCherry silencing and Nanog-GFP activation were less frequent (Figures 1I and 1J). Overall, these data indicate that elongation of Klf4S to Klf4L in Ok9MS subdues early reprogramming hallmarks but positively regulates stabilization of pluripotency.

**KLF4 Isoforms Underlie Phenotypic Differences in Polycistronic Cassette Reprogramming**

We next sought to expose key differences underlying the distinct phenotypes observed for OSKM and OKMS, and initially examined the reprogramming factors at the DNA sequence level. Of particular interest, the 2A peptide cleavage site preceding Klf4 in OSKM demarcated an ATG initiation codon that included an additional nine N-terminal amino acids (9aa) compared with that for OKMS (Figure 1D). When we examined public databases, we found that these two distinct Klf4 isoforms correspond to open reading frames (ORFs) predicted from full-length cDNA cloning (GenBank accession numbers U70662.1 and U20344.1) (Garrett-Sinha et al., 1996; Shields et al., 1996), anticipating proteins of 474aa and 483aa (GenBank accession numbers AAC52939.1 and AAC04892.1; henceforth referred to as KLF4S and KLF4L). To examine the effects of KLF4S and KLF4L on reprogramming while ruling out differences in factor order and 2A cleavage peptides, we introduced the N-terminal 9aa noted in OKMS (MRQPGGESD) directly into OKMS, producing OK+9MS (where K+9 is equivalent to KLF4S).

MEFs transposed with OK+9MS displayed a complete phenotypic switch. AP+ colony formation resembled that of OSKM (Figures 1E and S1A). Whole-field imaging on d18 strikingly revealed that OK+9MS colonies were primarily GFP+ and there were few mCherry+ colonies (Figure 1F, left). The Total and SSEA-1+ fractions on d18 were predominantly GFP+ with extensive mCherry silencing, mirroring OKSM reprogramming behavior (Figures 1F, right, and S1B). Dox-independent, GFP+ OK+9MS clones were chimera competent and shared a gene expression profile with OKSM and OKMS iPSCs, and ESCs (Figures S1E–S1G).

To determine whether the observed cassette-specific phenotypes were due to functional differences between KLF4S and KLF4L, we performed a series of monocistronic Klf4 expression experiments. We found that the transcriptome profiles of d6 mCherry+ cells arising from PB-TAC-Klf4S or -Klf4L MEF transfections were highly correlated (R = 0.999; Figure S2A). Gene enrichment compared with...
mock-transfected MEFs revealed a common set of 476 genes enriched for Gene Ontology (GO) terms associated with extracellular matrix, skin development, and cell adhesion (Figure S2B). Activation of a KLF4-responsive *Lefty1*-luciferase reporter (Nakatake et al., 2006) was similar irrespective of the KLF4 variant (Figure S2C), indicating that the capacity for regulation of this ESC promoter was unchanged. Moreover, both KLF4s and KLF4l could similarly rescue leukemia inhibitory factor (LIF)-independent culture of 1A2 Nanog-GFP reporter ESCs (data not shown). Thus, independent expression of either Klf4 variant supported a preservation of functional capacity.

Considering the remarkable reprogramming phenotypes observed with polycistronic cassettes, we next addressed whether monocistronic expression of Klf4s or Klf4l in combination with polycistronic Oct3/4, Sox2, and c-Myc (OMS) had any detectable effect on iPSC derivation. OMS alone did not foster expansion of mCherry+ cells, which remained at a steady ~6% through d18, nor did it activate SSEA-1 or Nanog-GFP expression (data not shown). Interestingly, expression of either Klf4 variant in combination with OMS resulted in nearly indistinguishable reprogramming phenotypes using the assays outlined in Figure 1 (see also Figures 2A, S2D, and S2E). These data demonstrate that the phenotypes associated with Klf4s and Klf4l arise mainly in the context of polycistronic factor linkage.

KLF4 Levels Are Diminished in OKMS and Rescued by N-Terminal Elongation

In order to rationalize the phenotypic differences observed between polycistronic cassettes, we evaluated transgene expression at the transcriptional and translational levels. Quantification of polycistronic mRNA revealed similar expression levels for OSKM and OKMS (Figure S2F). Strikingly, western blotting on d2 revealed that dox-induced KLF4 was expressed at distinctly higher levels in OSKM compared with OKMS, whereas the levels of OCT3/4, SOX2, and c-Myc (OMS) were generally comparable (Figure 2B). High KLF4 stoichiometry was also observed in MKOS, which, like OSKM, encodes KLF4l (Figure S2G). Consistent with the PB transposon experiments, MEFs from gene-targeted *Col1a1*-OSKM (Ohnishi et al., 2014) and -OKMS ESC chimeras faithfully reproduced cassette-specific protein levels (data not shown). Moreover, OKMS resulted in a marked increase in KLF4 levels compared with OKMS (Figure S2H). Surprisingly, N-terminal elongation of KLF4s in OKMS with a hemagglutinin (HA) epitope also rescued protein levels (OKN-HAMS; Figure S2H) and even reprogramming phenotypes (Figure S2J), whereas N-terminal HA extension of KLF4l in OK+9MS did not result in any additional phenotypic change. Concordant with monocistronic phenotypes, transfection of PB-TAC-

KLF4 Stoichiometry Regulates Reprogramming

To verify the reprogramming outcomes from adjusted factor stoichiometry, we supplemented OKMS with monocistronic factors in a PB-TAB transposon (Figure 2C). Supplementation of OKMS with *Klf4s* or *Klf4l* (Figure S2J) markedly reduced d10 AP+ colony numbers (Figures 2D and 2G), reminiscent of the reduced colony formation efficiencies observed when OSKM or OK+9MS was employed. Sox2, but not Oct3/4 or c-Myc, also had a significant effect on reducing colony formation frequencies compared with OKMS alone. Thus, both KLF4 isoforms modulated OKMS factor stoichiometry to subdue reprogramming initiation.

In the late phase of reprogramming, we gauged stabilization via mCherry silencing and Nanog-GFP activation (Figures 2E, 2F, and 2H). Supplementation of OKMS with either Klf4s or Klf4l was capable of activating Nanog-GFP with co-incident repression of mCherry on d18. Interestingly, although Sox2 supplementation also showed a repressive effect on initiation, d18 mCherry* and GFP* proportions remained mostly unchanged, indicating that additional Sox2 fails to rescue the OKMS stabilization phase. Likewise, Oct3/4 and c-Myc failed to promote Nanog-GFP reporter activation or mCherry silencing. Thus, only Klf4 variants were capable of augmenting the OKMS phenotype, implying a sensitivity of reprogramming pathways to KLF4 expression thresholds.

KLF4 Stoichiometry Is Reflected in Gene Expression Analyses

Global gene expression has been used to define reprogramming paths. In order to link gene expression to initial reprogramming behaviors, we subjected d6 mCherry+ cells from PB-TAC-OSKM, -OKMS, -OK+9MS, and -MKOS to microarray analysis (Figures 3A and 3B). Based on OSKM and OKMS GO term enrichment (data not shown) and results from previous reports, we established sample indexing criteria (Figure 3A, left). *Krt14* and *Krt17* (encoding structural proteins), as well as *Sfi1* (a regulator of epithelial cell growth), are claimed to activate in reprogramming populations with preferred iPSC fate (O’Malley et al., 2013). Expression of *Ocln* and *Cdh1* indicates MET responses (Samavarchi-Tehrani et al., 2010), while additional cell-surface changes are reflected by upregulation of *Cadm1* and *Fut9* (the fucosyltransferase responsible for producing SSEA-1). The developmental regulators *Tbx21* and *Bcl11a* (transiently upregulated non-ESC markers; Polo et al., 2012) and the pre-iPSC marker *Nr0b1* (Mikkelsen et al.,
show the extent of reprogramming. As early as d6, these ten marker genes were >2-fold differentially expressed between OK+MS (green series) and OKMS (red series; Figure 3A, left), and <2-fold differentially expressed in a pairwise correlation of OSKM to OK+MS (R = 0.993; Figure 3A, top right), certifying our indexing for further pairwise comparisons. Correlation and indexing of OSKM to OKMS verified the influence of KLF4 on early gene expression responses (Figure 3A, bottom right).

GO term analysis of commonly upregulated gene expression versus mock-transfected MEFs (PB-TAC-lacZ) was performed for cassettes containing Klf4L and Klf4S (Figure 3B and data not shown). For Klf4S vectors (216 genes), only a few GO terms related to the biological processes “regulation of cell adhesion (lamins)” and “tissue morphogenesis,” as well as the molecular function “phospholipase C activity,” were enriched with p values ≤ 1.0E-3 (data not shown). Nearly double the number of genes were activated by Klf4L constructs (512 genes; Figure 3B; Table S1). The common gene set for polycistronic Klf4L constructs significantly enriched GO terms for cellular components of the “plasma membrane” and “cell-cell junction” (consistent with an expected MET response), as well as distinct enrichment for “keratinocyte differentiation” and “epidermal cell differentiation” biological processes (p = 2.3E-11 and 5.4E-11; Figure 3B, right) with Benjamini scores = 3.6E-8 and 4.1E-8, respectively. These data are consistent with monocistronic Klf4 expression (Figure S2B).

**KLF4 Isoforms Divide Reprogramming Cassettes into Two Major Classes**

In order to more broadly understand the impact of KLF4-mediated reprogramming outcomes, we extended our analysis of publically available polycistronic constructs. DNA sequence surveys revealed that Klf4S and Klf4L reference cDNAs have been used interchangeably in the reprogramming field (Table 1). To confirm that public polycistronic vectors also display incongruous reprogramming behaviors, we selected MKOS (Kaji et al., 2009), STEMCCA (Sommer et al., 2009), EB-C5 (Chou et al., 2011), and WTSI (Yusa et al., 2009), and applied our battery of phenotyping tests (Figure S3). mCherry+ fractions and AP+ colony numbers varied among cassettes but were highest for OKMS and EB-C5 (Figures S3A and S3C). Notably, all four Klf4S cassettes (OKMS, STEMCCA, EB-C5, and WTSI) showed large proportions of mCherry+ SSEA-1+ cells on d8 (>40%, compared with <21% for Klf4L cassettes OSKM, OK+MS, and MKOS; Figure S3B), a characteristic that was consistently associated with Klf4S (Figure 1H). In accordance with Klf4L construction and western bloting data (Figure S2G), MKOS AP+ colony numbers were low on d10, and became mostly GFP+ mCherry+ colonies on d18 (Figures S3D and S3E). On the other hand, the Klf4L cassettes STEMCCA, EB-C5, and WTSI all perpetuated mCherry expression after d8 passage (74.8%, 80.9%, and 79.6%, respectively), and activated Nanog-GFP in <3.0% of SSEA-1+ cells (Figure S3E).

Thus, the phenotypes elicited by these additional public cassettes predictably correlated reprogramming hallmarks with Klf4 coding length.

Finally, gene indexing confirmed that MKOS (Klf4L) is distinct from STEMCCA (Klf4S; Figure 3C), substantiating the phenotypes reported in Figure S3. Moreover, unsupervised clustering of d6 mCherry+ cell gene expression profiles from all eight constructs (PB-TAC-OSKM, -OKMS, -OK+MS, -OKN-HAMS, -MKOS, -STEMCCA, -EB-C5, and -WTSI) generated a dendogram (Figure 3D) that distinctly

---

**Figure 2. KLF4 Stoichiometry Affects Reprogramming Phenotypes**

(A) Reprogramming with polycistronic OMS + monocistronic Klf4S or Klf4L leads to nearly identical phenotypes, as analyzed by FACS for mCherry or SSEA-1 on d8 (left) and mCherry versus Nanog-GFP on d18 (right).

(B) Western blot analysis of OCT3/4, SOX2, c-MYC, and KLF4 in OSMK or OKMS transfected MEFs cultured for 2 days with or without dox treatment. Actin was used as a loading control. Arrows show the two KLF4 isoforms. Note that SOX2 and c-MYC also differ in size based on their positions relative to 2A peptides in the polycistronic cassette (Figure 1D). Uncropped data are provided in Figure S2G.

(C–F) Supplementation of PB-TAC-OKMS was performed by co-transfection of additional factors in PB-TAB transposons. β-geo, lacZ-neo fusion gene.

(D) AP staining on d10 after transfection with OKMS alone or in combination with Oct3/4, Sox2, c-Myc, Klf4S, or Klf4L. Scale bar, 4,000 μm.

(E) Day 18 fluorescence microscopy of entire wells (composite 10 × 10 fields) or selected insets (3 × 3 subfields) for mCherry and Nanog-GFP. Scale bars, 4,000 μm (full well) and 1,000 μm (inset).

(F) FACS analysis of mCherry and Nanog-GFP fractions in the d18 SSEA-1+ population of the cultures shown in (E). The results in (D)–(F) are representative of the results of at least three independent experiments (summarized in G and H).

(G) Quantification of the effects of factor supplementation on AP+ colony formation (D). Means ± SE for three independent experiments. n.s., not significant. *p < 0.05, Student’s t test.

(H) Percentages of mCherry+ and Nanog-GFP+ populations from FACS analysis of factor supplementation (F). Means ± SE for three independent experiments. n.s., not significant. *p < 0.05, Student’s t test.

See also Figure S2.
bifurcated based on the Klf4 variant employed in each vector (Table 1). Thus, early gene expression patterns disclosed the nature of cloned Klf4. Intriguingly, despite unrelated N-terminal sequence identity (yet predicted by phenotype), OKN-HAMS activated 463/512 (≈90%) of the core Klf4L genes (data not shown) and clustered with Klf4L vectors (Figure 3D). A standardized comparison of polycistronic reprogramming cassettes thus implicates KLF4 isoforms in differential modulation of the reprogramming transcriptome.

DISCUSSION

Principally, we have uncovered a fundamental characteristic of cloned Klf4 that governs relative factor stoichiometry and reprogramming responses during mouse iPSC derivation with polycistronic constructs. At the root of our discovery lies the disparate coding prediction of full-length Klf4 cDNA (Garrett-Sinha et al., 1996; Shields et al., 1996), resulting in two proteins with distinct N termini defined herein as KLF4S and KLF4L. These two
Table 1. Survey of Publically Available Reprogramming Cassettes

| Source | Addgene Collection | Depositing Scientist | Repository ID | Plasmid Name | Gene/Insert Name | Species | Type | Klf4s | Klf4l | AddGene | Depositor | Direct | PMID | Reference |
|--------|--------------------|----------------------|---------------|--------------|------------------|---------|------|-------|-------|---------|-----------|--------|-------|------------|
| Addgene stem cell; Klf4/KLF4 | Shinya Yamanaka | 13370 | pMXs-Klf4 | Kruppel-like factor 4 (gut) | M. musculus | mono | ● | – | ○ | ○ | – | 16904174 | Takahashi and Yamanaka, 2006 |
| Addgene stem cell; Klf4/KLF4 | Shinya Yamanaka | 15920 | pMXs-Klf4-IP | Kruppel-like factor 4 (gut) | M. musculus | mono | ● | – | ○ | ○ | 16904174 | Takahashi and Yamanaka, 2006 |
| Addgene | Klf4/KLF4 | Miguel Ramalho-Santos | 15950 | pLOVE-Klf4 | Kruppel-like factor 4 | M. musculus | mono | n/d | n/d | n/c | – | – | 18371358 | Blelloch et al., 2007 |
| Addgene stem cell; Klf4/KLF4 | Shinya Yamanaka | 17219 | pMXs-hKLF4 | KLF4 | H. sapiens | mono | ● | – | ○ | ○ | – | 18035408 | Takahashi et al., 2007 |
| Addgene stem cell; Klf4/KLF4 | George Daley | 17227 | pMIG-hKLF4 | Kruppel-like factor 4 (gut) | H. sapiens | mono | ● | ○ | – | – | – | 18157115 | Park et al., 2008 |
| Addgene stem cell; Klf4/KLF4 | Kathrin Plath | 17967 | pMXs-hKLF4 (Plath) | Kruppel-like factor 4 (gut) | H. sapiens | mono | ● | ○ | ○ | – | – | 18287077 | Lowry et al., 2008 |
| Addgene stem cell; Klf4/KLF4 | Konrad Hochedlinger | 19764 | pLV-tetO-Klf4 | Klf4 | M. musculus | mono | ● | n/c | ○ | – | – | 18371448 | Stadtfeld et al., 2008a |
| Addgene stem cell; Klf4/KLF4 | Konrad Hochedlinger | 19770 | pAd-Klf4 | Klf4 | M. musculus | mono | ● | n/c | ○ | – | – | 18818365 | Stadtfeld et al., 2008b |
| Addgene stem cell; Klf4/KLF4 | Konrad Hochedlinger | 19777 | FU-tet-o-hKLF4 | KLF4 | H. sapiens | mono | ● | n/c | ○ | – | – | 18786420 | Maherali et al., 2008 |
| Addgene stem cell; Klf4/KLF4 | Juan Belmonte | 20074 | pMSCV-Flag-hKlf4 | hKlf4 | H. sapiens | mono | ● | ○ | – | – | – | 18931654 | Aasen et al., 2008 |
| Addgene stem cell; Klf4/KLF4 | Rudolf Jaenisch | 20322 | Tet0-FUW-Klf4 | Klf4 | M. musculus | mono | ● | ○ | ○ | – | – | 18371436 | Brambrink et al., 2008 |
| Addgene stem cell; Klf4/KLF4 | Rudolf Jaenisch | 20725 | FUW-tetO-hKLF4 | KLF4 | H. sapiens | mono | ● | ○ | ○ | – | – | 18786421 | Hockemeyer et al., 2008 |
| Addgene Klf4/KLF4 | Rudolf Jaenisch | 20727 | FUW-tetO-lox-hKLF4 | KLF4 | H. sapiens | mono | ● | ○ | ○ | – | – | 19269371 | Soldner et al., 2009 |
| Addgene Klf4/KLF4 | Michel Sadelain | 23243 | pLM-mCherry-Klf4 | mCherry_2A_KLF4 | H. sapiens | mono | ● | n/c | ○ | – | 19549847 | Papapetrou et al., 2009 |

(Continued on next page)
| Source       | Addgene Collection | Depositing Scientist | Repository ID | Plasmid Name | Gene/Insert Name | Species | Type | Klf4<sub>G</sub> | Klf4<sub>L</sub> | AddGene<sup>i</sup> | Depositor | Direct<sup>j</sup> | PMID      | Reference                   |
|--------------|--------------------|----------------------|---------------|--------------|------------------|---------|------|----------------|----------------|----------------|-----------|----------------|-----------|--------------------------------|
| Addgene      | Klf4/KLF4          | Rudolf Jaenisch      | 26274         | pBRPyCAG-fmKlf4-DsRed-1p | Klf4             | M. musculus | mono | n/c            |               |                 |           |                 | 20442331 | Hanna et al., 2010            |
| Addgene      | stem cell; Klf4/KLF4| Derrick Rossi        | 26815         | pcDNA3.3_KLF4  | 5′UTR-KLF4-3′UTR | H. sapiens | mono |               |               |                 |           |                 | 20888316 | Warren et al., 2010          |
| Addgene      | stem cell; Klf4/KLF4| Steven Dowdy         | 36118         | pCX4-KLF4     | KLF4             | H. sapiens | mono | n/d            | n/d            | n/c            |           |                 | 22278060 | Israel et al., 2012           |
| Addgene      | stem cell          | Rudolf Jaenisch      | 20331         | FUW-KO        | Klf4-F2A-Oct4    | M. musculus | bi   |               |               | o              |           |                 | 19109433 | Carey et al., 2009           |
| Addgene      | stem cell          | Rudolf Jaenisch      | 20332         | FUW-KM        | Klf4-F2A-cMyc    | M. musculus | bi   |               |               | o              |           |                 | 19109433 | Carey et al., 2009           |
| Addgene      | stem cell          | James Thomson        | 21164         | pSIN4-CMV-K2M | Klf4, cMyc       | H. sapiens | bi   |               |               | n/c            |           |                 | 19325077 | Yu et al., 2009              |
| Addgene      | stem cell          | Shinya Yamanaka      | 27078         | pCXLE-hSK     | SDX2, KLF4       | H. sapiens | bi   |               |               | o              |           |                 | 21460823 | Okita et al., 2011           |
| Addgene      | stem cell; Klf4/KLF4| Shinya Yamanaka      | 41814         | pCE-hSK       | SDX2, KLF4       | H. sapiens | bi   |               |               | o              |           |                 | 23193063 | Okita et al., 2013           |
| Addgene      | Klf4/KLF4          | Shinya Yamanaka      | 19771         | pCX-OKS-2A    | Oct3/4-2A-Klf4-2A-Sox2 | M. musculus | poly |               |               | n/c            |           |                 | 18845712 | Okita et al., 2008           |
| Addgene      | stem cell; Klf4/KLF4| Rudolf Jaenisch      | 20321         | TetO-FUW-OSKM | Oct3/4-Sox2-Klf4-cMyc | M. musculus | poly |               |               | n/c            |           |                 | 19109433 | Carey et al., 2009           |
| Addgene      | stem cell; Klf4/KLF4| Rudolf Jaenisch      | 20325         | FUW-SOKM      | Sox2-P2A-Oct4-T2A-Klf4-E2A-cMyc | M. musculus | poly |               |               | n/c            |           |                 | 19109433 | Carey et al., 2009           |
| Addgene      | stem cell; Klf4/KLF4| Rudolf Jaenisch      | 20327         | FUW-SOK       | Sox2-F2A-Oct4-T2A-Klf4 | M. musculus | poly |               |               | n/c            |           |                 | 19109433 | Carey et al., 2009           |
| Addgene      | stem cell          | Rudolf Jaenisch      | 20328         | FUW-OSKM      | Oct4-P2A-Sox2-T2A-Klf4-E2A-cMyc | M. musculus | poly |               |               | n/c            |           |                 | 19109433 | Carey et al., 2009           |
| Addgene      | stem cell          | Keisuke Kaji         | 20865         | pCAGMKOSIE    | c-Myc-F2A-Klf4-T2A-Oct4-E2A-Sox2 (Alt name) | M. musculus | poly |               |               | o              |           |                 | 19252477 | Kaji et al., 2009            |

(Continued on next page)
| Source              | Addgene Collection | Depositing Scientist | Repository ID | Plasmid Name | Gene/Insert Name | Species | Type | Klf4S | Klf4L | AddGene | Depositor | Direct | PMID       | Reference          |
|---------------------|--------------------|----------------------|---------------|--------------|-----------------|---------|------|------|------|---------|-----------|--------|------------|-------------------|
| Addgene stem cell   | Keisuke Kaji       | 20866                | pCAG2LMKOSim0D| c-Myc-F2A-Klf4-T2A-Oct4-E2A-Sox2-ires-mOrange | *M. musculus* | poly   |      |      |      |         |           |        | 19252477   | Kaji et al., 2009  |
| Addgene stem cell   | James Thomson      | 20923                | pEP4 E02S EM2K| Oct4 and Sox2; Myc and Klf4 | *H. sapiens* | poly   | n/c  |      |      |         |           |        | 19325077   | Yu et al., 2009    |
| Addgene stem cell   | James Thomson      | 20924                | pEP4 E02S CK2MEN2L| Oct4 and Sox2; Nanog and Lin28; Klf4 and Myc | *H. sapiens* | poly   | n/c  |      |      |         |           |        | 19325077   | Yu et al., 2009    |
| Addgene stem cell   | James Thomson      | 20925                | pEP4 E02S EN2K| Oct4 and Sox2; Nanog and Klf4 | *H. sapiens* | poly   | n/c  |      |      |         |           |        | 19325077   | Yu et al., 2009    |
| Addgene stem cell   | James Thomson      | 20927                | pEP4 E02S ET2K| Oct4 and Sox2; SV40LT and Klf4 | *H. sapiens* | poly   | n/c  |      |      |         |           |        | 19325077   | Yu et al., 2009    |
| Addgene stem cell   | Tim Townes         | 21627                | pKP332 (Lenti-OSK1) | hOCT4 - 2A - hSOX2 - 2A - hKLF4 | *H. sapiens* | poly   | n/c  |      |      |         |           |        | 19415770   | Chang et al., 2009 |
| Addgene stem cell   | Jose Cibelli       | 24603                | OKSIM         | Oct4 KLF4 Sox2 c-Myc | *H. sapiens* | poly   | n/c  |      |      |         |           |        | 20030562   | Ross et al., 2010  |
| Addgene stem cell   | Michel Sadelain    | 27512                | pLM-fSV2A     | OCT4_T2A_KLF4_ P2A_cMYC_ E2A_SOX2 | *H. sapiens* | poly   |      |      |      |         |           |        | 21151124   | Papapetrou et al., 2011 |
| Addgene stem cell   | Linzhao Cheng      | 28213                | pEB-C5        | Oct4, Sox2, Klf4, c-Myc, Lin28 | *M. musculus* | poly   |      |      |      |         |           |        | 21243013   | Chou et al., 2011  |
| Kotton Lab          | Darrell Kotton, Gustavo Mostoslavsky | — | — | pHAGE-Tet-STEMCCA | Oct3/4-Klf4-ires-Sox2-cMyc | *M. musculus* | poly   |      |      |      |         |           |        | 19096035   | Sommer et al., 2009 |
| WTSI                | Alan Bradley       | —                    | pPB-CAG.OSKMPuAtk | 0-T2A-S-T2A-K-F2A-M | *M. musculus* | poly   |      |      |      |         |           |        | 19337237   | Yusa et al., 2009  |
| RIKEN BRC            | Knut Weltjen       | RDB13133             | PB-TAC-OKMS   | Oct3/4-Klf4-cMyC-Sox2 | *M. musculus* | poly   |      |      |      |         |           |        | this study  |                   |
| RIKEN BRC            | Knut Weltjen       | RDB13132             | PB-TAC-OSKM   | Oct3/4-Sox2-Klf4-cMyc (from 20321) | *M. musculus* | poly   |      |      |      |         |           |        | this study  |                   |

(Continued on next page)
Table 1. Continued

| Source | Addgene Collectionb | Depositing Scientist | Repository ID | Construct Data | Gene/Insert Name | Speciesc | Typec | Klf4s | Klf4lc | Deposition ID | Depositor | Direct | PMID | Reference |
|--------|---------------------|---------------------|---------------|----------------|----------------|-----------|-------|-------|--------|--------------|-----------|--------|------|-----------|
| RIKEN BRC | — | Knut Woltjen | RDB13135 | PB-TAC-OK+9MS | Oct3/4-Klf4+9-cMyc-Sox2 | M. musculus | poly | • | — | — | C0/C0/C0 | B | this study |
| RIKEN BRC | — | Knut Woltjen | RDB13136 | PB-TAC-STEMCCA | Oct3/4-Klf4-ires-Sox2-cMyc (from Sommer et al., 2009) | M. musculus | poly | • | — | — | C0/C0/C0 | B | this study |
| RIKEN BRC | — | Knut Woltjen | RDB13134 | PB-TAC-MKOS | c-Myc-F2A-Klf4-T2A-Oct4-E2A-Sox2 (from 20866) | M. musculus | poly | • | — | — | C0/C0/C0 | B | this study |
| RIKEN BRC | — | Knut Woltjen | RDB13137 | PB-TAC-EB-C5 | Oct4, Sox2, Klf4, c-Myc, Lin28 (from 28213) | M. musculus | poly | • | — | — | C0/C0/C0 | B | this study |
| RIKEN BRC | — | Knut Woltjen | RDB13138 | PB-TAC-WTSI | 0-T2A-S-T2A-K-F2A-M (from Yusa et al., 2009) | M. musculus | poly | • | — | — | C0/C0/C0 | B | this study |

WTSI, Wellcome Trust Sanger Institute; RIKEN BRC, RIKEN Bio Resource Center DNA Bank. Closed circles indicate the Klf4 variant, and open circles denote sequence validation steps for each construct. See also Figures 3 and S3.

*aOnly vectors containing Klf4 are listed. Data from Addgene were current as of September 3, 2014. As Addgene is an independently curated and dynamically changing database, the authors do not warrant the accuracy or completeness of this resource beyond the date of data collection.

*bAddgene Collection lists can be found at http://www.addgene.org/stemcell/ and http://www.addgene.org/KLF/ or by browsing for Klf4 (gene 16600) or KLF4 (gene 9314).

*cCompared to mouse Klf4L, the human ORF is 87% and 90% similar at the nucleotide and amino acid levels, respectively, with 100% conservation across the first 35 amino acids. The effect of human Klf4S/L on reprogramming was not tested in this study.

*d“Type” refers to the number of cistrons or ORFs in the construct.

*Comparison of the ATG start site with U70662.1 (Klf4S) and U20344.1 (Klf4L) reference cDNAs. Note that for polycistronic vectors, this annotation does not take into account elongation by 2A peptides.

*n/d, form not determined by the available data.

*fValidation of Klf4 identity was from Addgene sequencing data. n/c, no coverage of the Klf4 start site, although available sequence confirmed the presence of Klf4.

*gCassette sequences were verified empirically by the authors.
variants have seen unprejudiced use among vectors constructed for reprogramming (Table 1 and associated references). OKMS (Klf4S) presented low KLF4 levels, which were restored along with Klf4L-related phenotypes by elongation of the KLF4s N terminus by 9aa. Thus, OK+9MS excluded the influence of 2A peptide choice and factor order variation. Moreover, high KLF4-associated phenotypes were reproduced by supplementation of OKMS with either monocistronic Klf4S or Klf4L, limiting the effect to a polycistronic context. Relative stoichiometry was explored in a recent study, and it was found that chemically controlled degradation of KLF4 resulted in premature termination of reprogramming (Nishimura et al., 2014). Although precise factor titration and more direct biochemical assays may be required, our monocistronic reprogramming data argue against modification of an N-terminal functional domain, and rather support the idea that threshold levels of KLF4 are important for minimizing the occurrence of partial reprogramming.

Polycistronic cassettes encoding KLF4S tend to permit rapid expansion of transgenic cells with robust SSEA-1 activation. Yet, this majority population is ultimately impeded in the acquisition and stabilization of pluripotency, as indicated by transgene dependence and high-level expression of markers associated with pre-iPSCs (Mikkelsen et al., 2008; Theunissen et al., 2011). Polycistronic reprogramming with Klf4L or OKMS supplementation with Klf4L improved reporter activation and dox independence. Nanog-GFP reporter activation is consistent with a central role for KLF4 in the core ESC pluripotency network, as ectopic expression of Klf4 promotes epiblast stem cell conversion to ground-state pluripotency (Guo et al., 2009) and prevents loss of pluripotency in the absence of LIF (Niwa et al., 2009). Transgene silencing is considered a hallmark of complete reprogramming (Goli-pour et al., 2012). How the silencing phenomenon is related to the acquisition of pluripotency, and whether it is coordinated directly or indirectly through KLF4 are issues of considerable interest.

Individual expression of either Klf4S or Klf4L in MEFs led to a potent epidermal gene response (Figure S2B). Moreover, a similar response was noted as a common feature of d6 reprogramming populations induced with Klf4L cassettes (Figure 3B). A retrospective examination of early reprogramming experiments recalls reports of epidermis-related gene activation (Mikkelsen et al., 2008). Klf4 knockout mice die postnatally from failed epidermal stratification (Segre et al., 1999). Reciprocally, ectopic Klf4 expression during embryogenesis results in premature barrier formation (Jaubert et al., 2003). Therefore, it is pertinent that high KLF4 stoichiometry induces a cascade of epithelialization and expression of epidermis-associated genes such as Krt6a, Krt17, Sprt1a, Cnfn, and Tgm1 during early reprogramming (Figures 3A and 3B; Table S1). Interestingly, many of these genes are not expressed in ESCs, nor are they maintained in GFP+ iPSCs (data not shown), which brings into question their relevance in iPSC derivation. To address potential cellular heterogeneity, it will be important to link gene expression responses and reprogramming outcomes to specific subpopulations of cells. Thus, it remains to be determined whether the epidermal response elicited by excess KLF4 in the early phase directly contributes to high-fidelity reprogramming or to an alternate cell fate.

On the path toward induced pluripotency, investigators rely on gene expression patterns as milestones to gauge appropriate progression, and comparative analysis of publicly available gene expression profiles is a common approach. Clearly, our data show that inconsistencies in the relative levels of KLF4 to OCT3/4, SOX2, and c-MYC can fundamentally confound comparative analysis. Application of monocistronic Klf4S or Klf4L leads to nearly indistinguishable reprogramming phenotypes. Thus, typical clonal isolation of fully reprogrammed iPSCs with monocistronic vectors, which is subject to phenotypic selection for appropriate factor expression levels, should be mostly unaffected. Mechanistic studies, on the other hand, are prone to transcriptional noise arising from variable factor expression between cells and among transgenes. We expect the KLF4 threshold effects observed herein to be pronounced when the relative factor stoichiometry is fixed, for polycistronic transfection or even in secondary reprogramming systems with pre-integrated transgenes (Wernig et al., 2008; Woltjen et al., 2009). As such, we must be mindful that it is difficult to make direct comparisons across distinct reprogramming systems without first defining the inherent factor stoichiometry.

**EXPERIMENTAL PROCEDURES**

**Plasmid Construction**

The PB transposon-based PB-TAC expression vector (RDB13131) was generated by standard cloning (restriction digestion and ligation). PB-TAB vectors containing the four Yamanaka factors were generated by Gateway Cloning (Invitrogen) as described previously (Woltjen et al., 2009). Monocistronic and polycistronic reprogramming cassettes constructed in-house or obtained from public resources (Table 1) were cloned into pDONR221 using PCR and attB/P Clonase (Invitrogen) or pENTR2B using standard restriction endonuclease cloning or InFusion (Takara Clontech). Note that the cassette from pPB-CAG.OSKMBpuTk (Yusa et al., 2009) is in the order O-S-K-M, similarly to OSKM (Carey et al., 2009), and thus is referred to in the text as WTSI to avoid confusion. Modifications such as N-terminal extension and HA tagging were produced using InFusion or the Q5 Site-Directed Mutagenesis Kit (New England Biolabs). Sequences of the main Gateway
Entry constructs (from attL1 to attL2) were obtained using Sanger sequencing with Big Dye Terminator Ver3.1 chemistry (ABI) and assembled as contigs using Sequencher (GeneCodes). Complete details regarding the cloning procedures used, including InFusion amplification and sequencing primers, are available upon request.

MEF Isolation
Nanog-GFP (Okita et al., 2007) transgenic reporter mice were maintained as homozygotes on a C57BL/6 background. For ROSA-rTA; Nanog-GFP double transgenic MEFs, C57BL/6 homozygous ROSA26-rTA females (Ohnishi et al., 2014) and inbred homozygous Nanog-GFP males were crossed. Mice bearing Col1a1-targeted OSKM and OKMS cassettes were generated as described previously (Ohnishi et al., 2014). MEFs were isolated from 12.5–13.5 days postcoitum transgenic embryos. Embryos were decapitated, eviscerated, and macerated, and then dissociated with 0.25% trypsin/EDTA at 37°C for 15 min and plated in DMEM containing 10% fetal bovine serum (FBS), penicillin-streptomycin, and L-glutamine. The presence of the desired transgenes and male gender were determined by PCR genotyping of residual embryonic tissues. Cells from male embryos were expanded 1:5 once and then frozen in liquid nitrogen. Cells for reprogramming were defrosted, counted, and plated for transfection without additional expansion. The genotyping primers are available upon request. Animal care and experiments using animal tissues were approved by the RIKEN Animal Experiment Committee in accordance with Kyoto University guidelines.

PB Reprogramming and iPSC and ESC Culture
MEFs were seeded in DMEM containing 10% FBS, penicillin-streptomycin, and L-glutamine on gelatinized (0.1%) 6-well dishes at a density of 1.0 × 105 cells per 10 cm2. After a 24 hr culture, Fugene HD (Promega) was used to transfect cells with 500 ng of transposons (PB-TAG-OSKM, -OKMS, -OKkMS, -MKOS, -STEMCCA, -EB-C5, -WTSI, -OKN/AMMS, -OKkN/AMMS, and -lacZ) plus 1,000 ng of pCyl43 PB transposase plasmid at a Fugene/DNA ratio of 8. After transfection, cells were incubated for 24 hr, and then cultured for an additional 3 days in DMEM containing 10% FBS, penicillin-streptomycin, GlutaMAX, β-mercaptoethanol, sodium-pyruvate, non-essential amino acids, LIF, and dox (1 μg/mL). After transfection, cells were fed daily with dox-containing ESC medium. On d8, cells were trypsinized and re-seeded at 3.0 × 105 cells per 10 cm2 on gelatin for late-stage analysis (until d18). For dox withdrawal, cells were again harvested on d18 and re-seeded at 3.0 × 105 cells per 10 cm2 on gelatin in the presence or absence of dox, and analyzed on d24. For isolation of factor-independent iPSC clones, d18 cultures were reseeded on gelatin at limiting dilutions in the absence of dox, and colonies were picked on d24. Southern blot analysis using HindIII digestion and an mCherry probe identified genetically unique iPSC lines. Both dox-independent iPSCs and ESCs were maintained on mitomycin-c-inactivated DR4 MEF feeders and gelatin in ESC medium, with passage every 2–3 days. For all cells, total viable cell counts were performed using a TC10 Automated Cell Counter (Bio-Rad) with trypan blue exclusion.

HEK293T Cell Culture and PB Transfection
HEK293T cells were maintained in DMEM containing 10% FBS, penicillin-streptomycin, and L-glutamine. Then 3.0 × 106 cells per 10 cm2 were transfected with 500 ng of transposons (PB-TAG-OKN/AMMS and -OKkN/AMMS) plus 500 ng of PB-CAG-rTA (Woltjen et al., 2009) at a Fugene/DNA ratio of 8 μL:2 μg. After 24 hr, the culture medium was replaced with medium containing dox (1 μg/mL). After an additional 24 hr, transfected cells were harvested for western blot analysis.

Whole-Well Fluorescence Microscopy Imaging
Murine fibroblasts were plated on standard tissue culture 6-well plastic plates (BD Falcon Labware). Images were acquired with a Nikon BioStation CT (Nikon) equipped with GFP and mCherry fluorescence filters and phase contrast using ×2 lenses. The single-plane images of each channel were stitched automatically using the automated image analysis software CL-Quant 3.0 (Nikon).

AP Staining
Staining for AP activity was performed using the BCIP/NBT Alkaline Phosphatase Substrate Kit IV (SK-5400, Vector Laboratories) according to the manufacturer’s instructions. Images were acquired with a color camera (Canon IXY Digital 900IS), and colony count analysis was performed using a custom CL-Quant 3.0 macro developed in cooperation with Nikon, with AP-stained, mock-transfected MEF plates used to establish background.

Flow Cytometry and Cell Sorting
For FACS analysis, 3.0 × 105 cells were stained on ice for 30 min with Alexa Fluor 647 conjugated SSEA-1 (480, sc-21702, Santa Cruz Biotechnology) antibody or the corresponding isotype-matched control, mouse IgM (1:50 dilution). Control staining with the appropriate isotype-matched control monoclonal antibodies (BD Biosciences) was included to establish thresholds for positive staining. Samples were analyzed using a BD LSRFortessa Cell Analyzer (BD Biosciences) with BD FACSDiva software (BD Biosciences). Flow cytometry data were analyzed and generated by FlowJo software (Tree Star). For cell sorting, the mCherry+ cell population was collected on a BD FACSAria II cell sorter (BD Biosciences).

Microarray Analysis
Total RNA was prepared from harvested cells using the RNeasy Mini Kit (QIAGEN) according to the manufacturer’s instructions. cDNA synthesis and transcriptional amplification were performed using 200 ng of total RNA with the Whole Transcript (WT) Expression Kit (Ambion/Affymetrix). Fragmented and biotin-labeled cDNA targets were hybridized to GeneChip Mouse Gene 1.0 ST arrays (Affymetrix) according to the manufacturer’s protocols. Hybridized arrays were scanned using an Affymetrix GeneChip Scanner. Probe signal intensities were normalized with the RMA algorithm in GeneSpring. Quality, correlation, and cluster analyses were performed using GeneSpring GX software v12.1 (Agilent).
were used as described previously (Kaji et al., 2009). Mouse Anti-sc-17320, 1:1,000) and GKLF (H-180, sc-20691, 1:500) antibodies were obtained from Santa Cruz Biotechnology, respectively. HA-probe (F-7, sc-7392, 1:200) and anti-actin (A2066, 1:5,000) antibodies were obtained from BD Biosciences and Cell Signaling, respectively. Oct-3/4 (40/Oct-3, 611203, 1:1,000) and c-Myc (9402, 1:1,000) antibodies were obtained from Cell Signaling Technologies. For western analysis, we prepared total cell lysates by boiling 1 × 10^6 cells for 10 min in SDS sample buffer (25 mM Tris (pH 6.8), 2% β-mercaptoethanol, 3% SDS, 0.05% bromophenol blue, and 5% glycerol). Lysates were resolved on SDS-polyacrylamide gels, and proteins were detected by western blotting using ECL Plus (Amersham Pharmacia). For western analysis, Sox-2 (Y-17, sc-17320, 1:1,000) and GKLF (H-180, sc-20691, 1:500) antibodies were used as described previously (Kaji et al., 2009). Mouse Anti-Oct-3/4 (40/Oct-3, 611203, 1:1,000) and c-Myc (9402, 1:1,000) antibodies were obtained from BD Biosciences and Cell Signaling, respectively. HA-probe (F-7, sc-7392, 1:200) and anti-actin (A2066, 1:5,000) antibodies were obtained from Santa Cruz Biotechnology and Sigma-Aldrich, respectively.

Material Distribution
The PB-TAC reprogramming transposons are available through the RIKEN BRC DNA Bank (http://dna.brc.riken.jp) under accession numbers RDB13131–RDB13138. Gateway versions of the polycistronic reprogramming cassettes will be made available through AddGene (http://www.addgene.org). The PB transposase vector pCYL43 is available through the Wellcome Trust Sanger Center (http://www.sanger.ac.uk/form/Sanger_CloneRequests).

ACCESSION NUMBERS
Datasets for the d6 gene expression microarray analysis have been deposited in the Gene Expression Omnibus under accession number GSE65468.

SUPPLEMENTAL INFORMATION
Supplemental Information includes Supplemental Experimental Procedures, three figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.stemcr.2015.02.004.

ACKNOWLEDGMENTS
We thank Kazutoshi Takahashi, Masamitsu Sone, and Ren Shimamoto for their critical readings of the manuscript, and Katsunori Semi for bioinformatics advice. We also thank Chiho Sakurai, Michiko Nakamura, and Akiko Oishi for technical support; Toshi-iko Sato for microarray preparations; Akito Tanaka for blastocyst injections; and Kanako Asano and Kotaro Ohnishi for mouse husbandry. We appreciate the advice of Hiroaki Kii (Nikon, Japan) regarding image acquisition and analysis. We also thank Darrell N. Kotton and Gustavo Mostoslavsky (Boston University School of Medicine) for providing pHAGE-Tet-STEMCCA, Alan Bradley and Kosuke Yusa (Wellcome Trust Sanger Institute) for providing pPB-CAG.OSKMBratuTk, Hitoshi Niwa (RIKEN Center for Developmental Biology) for providing the pLefty1-luc reporter, and all Addgene contributors, especially Linzhao Cheng and Rudolf Jaenisch. This work was supported by the Cabinet Office, Government of Japan and the Japan Society for the Promotion of Science (JSPS) through the Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program), the Core Center for iPS Cell Research, the Research Center Network for Realization of Regenerative Medicine, and the Strategic International Collaborative Research Program of the Japan Science and Technology Agency (JST). K.W. is a Hakubi Center Special Project Researcher. S.Y. is a non-salaried scientific advisor of iPS Academia Japan. S.L. received funding from EPASI (JSPS/National Science Foundation, 2013). S.-I.K. is a JSPS Fellowship recipient (2011-2013) and JST Researcher.

Received: August 8, 2014
Revised: February 6, 2015
Accepted: February 6, 2015
Published: March 12, 2015

REFERENCES
Aassen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, E., Vassena, R., Bilic, J., Pekarik, V., Tiscornia, G., et al. (2008). Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. Nat. Biotechnol. 26, 1276–1284.

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.

Brambrink, T., Foreman, R., Welstead, G.G., Lengner, C.J., Wernig, M., Suh, H., and Jaenisch, R. (2008). Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. Cell Stem Cell 2, 151–159.

Carey, B.W., Markoulaki, S., Hanna, J., Saha, K., Gao, Q., Mittal-pova, M., and Jaenisch, R. (2009). Reprogramming of murine and human somatic cells using a single polycistronic vector. Proc. Natl. Acad. Sci. USA 106, 157–162.

Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., Ganz, K., Steine, E.J., Cassady, J.P., Creyghton, M.P., et al. (2011). Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. Cell Stem Cell 9, 588–598.
Chang, C.W., Lai, Y.S., Pawlik, K.M., Liu, K., Sun, C.W., Li, C., Schoeb, T.R., and Townes, T.M. (2009). Polycistronic lentiviral vector for “hit and run” reprogramming of adult skin fibroblasts to induced pluripotent stem cells. Stem Cells Tor for “hit and run” reprogramming of adult skin fibroblasts to human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. Cell Res. (Amst.) 21, 518–529.

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

Garrett-Sinha, L.A., Eberspaecher, H., Seldin, M.F., and de Crombrugghe, B. (1996). A gene for a novel zinc-finger protein expressed in differentiated epithelial cells and transiently in certain mesenchymal cells. J. Biol. Chem. 271, 31384–31390.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., and Wrana, J.L. (2012). A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. Cell Stem Cell 11, 769–782.

Graf, T. (2011). Historical origins of transdifferentiation and reprogramming. Cell Stem Cell 9, 504–516.

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

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., Cassady, J.P., Muffat, J., Carey, B.W., and Jaenisch, R. (2010). Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. Proc. Natl. Acad. Sci. USA 107, 9222–9227.

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M., and Jaenisch, R. (2008). A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. Cell Stem Cell 3, 346–353.

Israel, M.A., Yuan, S.H., Bardy, C., Reyna, S.M., Mu, Y., Herrera, C., Heffner, M.P., Van Gorp, S., Nazor, K.L., Boscolo, E.S., et al. (2012). Probing sporadic and familial Alzheimer’s disease using induced pluripotent stem cells. Nature 482, 216–220.

Jaubert, J., Cheng, J., and Segre, J.A. (2003). Ectopic expression of kruppel like factor 4 (Klf4) accelerates formation of the epidermal permeability barrier. Development 130, 2767–2777.

Kaji, K., Norathy, K., Paca, A., Mileikovsky, M., Mohseni, P., and Woltjen, K. (2009). Virus-free induction of pluripotency and subsequent excision of reprogramming factors. Nature 458, 771–775.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc. Natl. Acad. Sci. USA 105, 2883–2888.

Maherali, N., Ahfeldt, T., Rigamonti, A., Uitkal, J., Cowan, C., and Hochedlinger, K. (2008). A high-efficiency system for the generation and study of human induced pluripotent stem cells. Cell Stem Cell 3, 340–345.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. (2008). Dissecting direct reprogramming through integrative genomic analysis. Nature 454, 49–55.

Nakatake, Y., Fukui, N., Iwamatsu, Y., Masui, S., Takahashi, K., Yagi, R., Yagi, K., Miyazaki, J., Matoba, R., Ko, M.S., and Niwa, H. (2006). KLF4 cooperates with Oct3/4 and Sox2 to activate the Left1 core promoter in embryonic stem cells. Mol. Cell. Biol. 26, 7772–7782.

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

Niwa, H., Ogawa, K., Shimosato, D., and Adachi, K. (2009). A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells. Nature 460, 118–122.

Ohnishi, K., Sem, K., Yamamoto, T., Shimizu, M., Tanaka, A., Mitsunaga, K., Okita, K., Osaune, K., Akioka, Y., Maeda, T., et al. (2014). Premature termination of reprogramming in vivo leads to cancer development through altered epigenetic regulation. Cell 156, 663–677.

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

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T., and Yamanaka, S. (2008). Generation of mouse induced pluripotent stem cells without viral vectors. Science 322, 949–953.

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., Hong, H., Nakagawa, M., Tanabe, K., Tezuka, K., et al. (2011). A more efficient method to generate integration-free human iPS cells. Nat. Methods 8, 409–412.

Okita, K., Yamakawa, T., Matsumura, Y., Sato, Y., Amano, N., Watanabe, A., Goshima, N., and Yamanaka, S. (2013). An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. Stem Cells 31, 458–466.

O’Malley, J., Skylaki, S., Iwabuchi, K.A., Chantzoura, E., Ruetz, T., Johnsson, A., Tomlinson, S.R., Linnarsson, S., and Kaji, K. (2013). High-resolution analysis with novel cell-surface markers identifies routes to iPSCs. Nature 499, 88–91.

Papapetrou, E.P., Tomishima, M.J., Chambers, S.M., Mica, Y., Reed, E., Menon, J., Tabar, V., Mo, Q., Studer, L., and Sadelain, M. (2009). Stoichiometric and temporal requirements of Oct4, Sox2, Klf4, and c-Myc expression for efficient human iPSC induction and differentiation. Proc. Natl. Acad. Sci. USA 106, 12759–12764.

Papapetrou, E.P., Lee, G., Malani, N., Setty, M., Riviere, I., Tirunagari, L.M., Kadota, K., Roth, S.L., Giardina, P., Viale, A., et al. (2011). Genomic safe harbors permit high beta-globin transgene expression in thalassemia induced pluripotent stem cells. Nat. Biotechnol. 29, 73–78.

Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Hsu, 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.
Plath, K., and Lowry, W.E. (2011). Progress in understanding reprogramming to the induced pluripotent state. Nat. Rev. Genet. 12, 253–265.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., Borkent, M., Apostolou, E., Alaei, S., Cloutier, J., et al. (2012). A molecular roadmap of reprogramming somatic cells into iPS cells. Cell 151, 1617–1632.

Ross, P.J., Suhr, S.T., Rodriguez, R.M., Chang, E.A., Wang, K., Siripataraprat, K., Ko, T., and Cibelli, J.B. (2010). Human-induced pluripotent stem cells produced under xeno-free conditions. Stem Cells Dev. 19, 1221–1229.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.K., Beyer, T.A., Datti, A., Woltjen, K., Nagy, A., and Wrana, J.L. (2010). Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. Cell Stem Cell 7, 64–77.

Segre, J.A., Bauer, C., and Fuchs, E. (1999). Klf4 is a transcription factor required for establishing the barrier function of the skin. Nat. Genet. 22, 356–360.

Shields, J.M., Christy, R.J., and Yang, V.W. (1996). Identification and characterization of a gene encoding a gut-enriched Kruppel-like factor expressed during growth arrest. J. Biol. Chem. 271, 20009–20017.

Soldner, F., Hockemeyer, D., Beard, C., Gao, Q., Bell, G.W., Cook, E.G., Hargus, G., Blak, A., Cooper, O., Mitalipova, M., et al. (2009). Parkinson’s disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. Cell 136, 964–977.

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

Stadtfeld, M., Maherali, N., Breault, D.T., and Hochedlinger, K. (2008a). Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. Cell Stem Cell 2, 230–240.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G., and Hochedlinger, K. (2008b). Induced pluripotent stem cells generated without viral integration. Science 322, 945–949.

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.

Theunissen, T.W., and Jaenisch, R. (2014). Molecular control of induced pluripotency. Cell Stem Cell 14, 720–734.

Theunissen, T.W., van Oosten, A.L., Castelo-Branco, G., Hall, J., Smith, A., and Silva, J.C. (2011). Nanog overcomes reprogramming barriers and induces pluripotency in minimal conditions. Curr. Biol. 21, 65–71.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.H., Li, H., Lau, F., Ebina, W., Mandal, P.K., Smith, Z.D., Meissner, A., et al. (2010). Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. Cell Stem Cell 7, 618–630.

Wernig, M., Lengner, C.J., Hanna, J., Lodato, M.A., Steine, E., Foreman, R., Staerk, J., Markoulaki, S., and Jaenisch, R. (2008). A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. Nat. Biotechnol. 26, 916–924.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hämäläinen, R., Cowling, R., Wang, W., Liu, P., Gertsenstein, M., et al. (2009). piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. Nature 458, 766–770.

Yamaguchi, S., Hirano, K., Nagata, S., and Tada, T. (2011). Sox2 expression effects on direct reprogramming efficiency as determined by alternative somatic cell fate. Stem Cell Res. (Amst.) 6, 177–186.

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., and Thomson, J.A. (2009). Human induced pluripotent stem cells free of vector and transgene sequences. Science 324, 797–801.

Yusa, K., Rad, R., Takeda, J., and Bradley, A. (2009). Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. Nat. Methods 6, 363–369.
KLF4 N-Terminal Variance Modulates
Induced Reprogramming to Pluripotency
Shin-Ill Kim, Fabian Oceguera-Yanez, Ryoko Hirohata, Sara Linker, Keisuke Okita,
Yasuhiro Yamada, Takuya Yamamoto, Shinya Yamanaka, and Knut Woltjen
Figure S1. The Efficiencies of Dox-Independent, Nanog-GFP iPS Cell Formation by Polycistronic Vectors Differs According to the N-terminus of KLF4, Related to Figure 1
(A) Quantification of AP positive colony formation using three different polycistronic reprogramming cassettes: OSKM, OKMS, and OK+9MS. Means ± SE for three independent experiments.

(B) Percentages of mCherry+ and Nanog-GFP+ populations from FACS analysis of reprogramming cultures. The reporter-positive fractions of the Total (left) and SSEA-1+ (right) populations are shown. Means ± SE for three independent experiments.

(C-D) FACS analysis (d24) showing the effect of continued dox treatment, or dox withdrawal on d18 reprogramming cultures. Relative mCherry+ and Nanog-GFP+ cell fractions remain essentially unchanged following extended dox treatment (compare d24 +dox to d18 levels, Figure 1C and 1F). On the other hand, mCherry+ cells are absent after 6 days of culture in the absence of dox. An ~30% decrease in GFP cell numbers is observed for all cassettes following dox withdrawal. The data in panel C is summarized for two experiments in panel D.

(E) Phase and fluorescence microscopy images of colony morphology and Nanog-GFP expression from representative dox-independent iPS cell clones, as compared to parental Nanog-GFP ES cells (1A2). Scale bar, 200 µm.

(F) Chimeric mice derived from iPS cell clones and parental ES cells (1A2) injected into ICR (albino) host embryos.

(G) Correlation plot of microarray transcriptome analysis for mock-transfected MEFs (lacZ), dox-independent iPS cell clones (4 clones each for OSKM, OKMS, and OK+9MS), grandparental RF8 and parental 1A2 ES cells (two independent passages), and d18 mCherry+ cells from OKMS and OK+9MS.
Figure S2. *Klf4* Variant-Associated Phenotypes are Linked to Protein Expression Levels, and are Not Exclusive to N-terminal Identity, Related to Figure 2

(A) Pairwise comparison of genes induced in *Klf4*$_S$ (x-axis) and *Klf4*$_L$ (y-axis) transfected MEFs on d6 of dox treatment (mCherry$^+$ fraction). Green lines indicate two-fold changes in log$_2$ signal intensity between the two data sets. Signal intensities were normalized with the RMA algorithm (GeneSpring).
(B) Venn diagram (left) illustrating genes expressed >2-fold higher in Klf4\textsubscript{S} and Klf4\textsubscript{L} transfected cells as compared to LacZ mock-transfected MEFs. Klf4\textsubscript{S} and Klf4\textsubscript{L} show a high degree of overlap. Gene Ontology (GO) analysis of 476 shared genes, arranged by p-value and indicating the proportion of genes represented for each enriched GO term (right).

(C) Activation of a Lefty1 promoter in HeLa cells transfected with PB-TAC expressing LacZ (Mock), Oct3/4, Sox2, or one of two Klf4 variants, as measured by luciferase activity in the presence or absence of dox. Means ± SE for three independent experiments. n.s., not significant. *, p<0.05. Student t-test.

(D) AP staining on d10 following transfection of OMS in combination with Klf4\textsubscript{S} or Klf4\textsubscript{L}. Scale bar, 4000 µm.

(E) Percentages of mCherry\textsuperscript{+} and Nanog-GFP\textsuperscript{+} populations in the SSEA-1\textsuperscript{+} fraction from FACS analysis of OMS + Klf4\textsubscript{S} or Klf4\textsubscript{L} reprogramming cultures on d18. Means ± SE for three independent experiments. *, p<0.05. Student t-test.

(F) qRT-PCR analysis of the common E2A region and total Oct3/4 mRNA in OSKM and OKMS transfected MEFs cultured for 2 days with (+) or without (-) dox treatment. cDNA synthesis without reverse transcriptase (-RT) was used as a negative control. All mRNA levels were normalized to Hprt. Means ± SE for three independent experiments. *, p<0.05. Student t-test.

(G) Western blot analysis of OCT3/4, SOX2, KLF4, and c-MYC in OSKM or OKMS transfected MEFs cultured on d2 without or d2, 4 with dox treatment. MKOS transfected MEFs were included as a reference (Kaji et al., 2009). ACTIN was used as a loading control. Note that proteins produced from positions 1, 2 and 3 in the quadracistronic vectors are C-terminally modified with residual 2A sequences, resulting in slightly higher molecular weights. Asterisk indicates presumptive mega-protein detected with the SOX2 antibody. Blue boxes indicate samples in which the mCherry positive cell numbers are most similar between OSKM and OKMS.

(H) Western blot analysis of protein levels in HEK293T cells following modification of Klf4 in OKMS. OKMS and OK\textsuperscript{N-HA}MS are shown in lanes 4 and 5. KLF4 from OKMS and OK\textsuperscript{N-HA}MS was N- (lanes 2 and 3) or C-terminally (lanes 6 and 7) tagged with HA and probed with a HA antibody (upper panels) to exclude the possibility of differences in native antibody recognition between KLF4\textsubscript{S} and KLF4\textsubscript{L} epitopes with the KLF4 antibody (lower panels). Non-transfected (NT) cells are included as negative controls (lanes 1 and 8). TUBULIN was used as a loading control.

(I) N-terminal HA tagging of KLF4 in OKMS (OK\textsubscript{N-HA}MS) and OK\textsuperscript{C-HA}MS (OK\textsuperscript{C-N-HA}MS) cassettes. FACS analysis of mCherry and Nanog-GFP expression in the d18 SSEA-1\textsuperscript{+} population (bottom).

(J) Western blot analysis of OKMS supplementation with Klf4\textsubscript{S} and Klf4\textsubscript{L} (lanes 1-3), compared to expression in MEFs (lanes 5, 6) on d2 after dox induction. Mock transfected MEFs (LacZ) are included as a control (lane 4). TUBULIN was used as a loading control.
**Figure S3. Comparative Evaluation of Polycistronic Vectors, Related to Table 1 and Figure 3**

(A) FACS analysis of mCherry expression in OSKM, OKMS, OK^MS, MKOS, STEMCCA, EB-C5, and WTSI transfected MEFs cultured for 8 days with dox treatment.

(B) Expression of SSEA-1 in mCherry^+ cell fractions from panel A.

(C) AP staining on d10. Scale bar, 4000 µm.

(D) Day 18 fluorescence microscopy of entire wells (composite 10 × 10 fields) for and mCherry^+ and Nanog-GFP^+. Scale bar, 4000 µm.

(E) FACS analysis of mCherry expression and Nanog-GFP in the SSEA-1^+ population of d18 cultures described in panel D. The results in panels A-E are representative of at least two to three independent experiments.

**Table S1. Genes Commonly Upregulated in the Klf4 Cluster, Related to Figure 3**

The table containing the gene list is provided as a separate Excel file.
Supplemental Experimental Procedures

Luciferase Assay
HeLa cells were seeded in DMEM containing 10% FBS, penicillin-streptomycin and L-glutamine on 96-well plate at a density of $5 \times 10^4$ cells per 0.32 cm$^2$ and transfected with 200 ng pLefty1-luc reporter plasmid (Nakatake et al., 2006), 200 ng PB-CAG-rtTA (Woltjen et al., 2009), 200 ng PB-TAC-LacZ, -Oct3/4, -Sox2, -Klf4$, or$ -Klf4$, and 20 ng of Renilla luciferase internal control plasmid (pGL4.74[hRluc/TK], Invitrogen) using Lipofectamine 2000 (Invitrogen). piggyBac transposase was not included for these transient expression assays. After 24 hr, culture media was replaced with medium containing doxycycline (1 μg/mL). After additional 24 hr, transfected cells were harvested and luciferase assays were performed using a Dual-Glo® Luciferase Assay System (Promega) and luminescence was measured by a 2104 EnVision Multilabel Reader (Perkin Elmer).

Chimera Generation
A total of twelve dox-independent iPS clones generated with OSKM (C4, G5, A6, B6), OKMS (A7, H7, A9, H9), or OK$^{+}\text{MS}$ (A10, D10, E10, H10) were grown under standard ES cell conditions. Cells were harvested as a feeder-depleted (pre-plated on non-treated dishes for 30-45 min) single cell suspension on the date of injection, and each ICR blastocyst received 8–20 iPS cells. After injection, blastocysts were cultured in KSOM and placed at 37°C until transferred to recipient females. About 20–27 injected blastocysts were transferred to each uterine horn of 2.5 day postcoitum pseudopregnant B6D2F1 female. Chimeras were maintained and documented by photography at 4-6 weeks of age. No incidence of mortality or tumor formation was noted in mice over 12 months of age. The care of the animals was in accordance with Kyoto University guidelines, and all animal experiments were approved by the CiRA Animal Experiment Committee.