

## CELL BIOLOGY

# Reprogramming competence of OCT factors is determined by transactivation domains

Kee-Pyo Kim<sup>1</sup>, You Wu<sup>2</sup>, Juyong Yoon<sup>1\*</sup>, Kenjiro Adachi<sup>1†</sup>, Guangming Wu<sup>1,3‡</sup>, Sergiy Velychko<sup>1</sup>, Caitlin M. MacCarthy<sup>1</sup>, Borami Shin<sup>1</sup>, Albrecht Röpke<sup>4</sup>, Marcos J. Arauzo-Bravo<sup>5,6</sup>, Martin Stehling<sup>7</sup>, Dong Wook Han<sup>8</sup>, Yawei Gao<sup>2</sup>, Johnny Kim<sup>9</sup>, Shaorong Gao<sup>2</sup>, Hans R. Schöler<sup>1,10§</sup>

OCT4 (also known as POU5F1) plays an essential role in reprogramming. It is the only member of the POU (Pit-Oct-Unc) family of transcription factors that can induce pluripotency despite sharing high structural similarities to all other members. Here, we discover that OCT6 (also known as POU3F1) can elicit reprogramming specifically in human cells. OCT6-based reprogramming does not alter the mesenchymal-epithelial transition but is attenuated through the delayed activation of the pluripotency network in comparison with OCT4-based reprogramming. Creating a series of reciprocal domain-swapped chimeras and mutants across all OCT factors, we clearly delineate essential elements of OCT4/OCT6-dependent reprogramming and, conversely, identify the features that prevent induction of pluripotency by other OCT factors. With this strategy, we further discover various chimeric proteins that are superior to OCT4 in reprogramming. Our findings clarify how reprogramming competences of OCT factors are conferred through their structural components.

## INTRODUCTION

Somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs) by ectopic expression of OCT4, SOX2, KLF4, and MYC (hereafter referred to as O4SKM) (1). While SKM can be replaced by a subset of their respective family members (2), OCT4 is the only factor that cannot be replaced by any of its family members (2–5), despite their profound sequence conservation. In mice, several other genes, which are evolutionarily unrelated to Oct4, can, however, substitute for Oct4 in reprogramming (6–11). Specifically, Nr5a1, Nr5a2, Tet1, Sall4/Nanog, or Nkx3-1 together with SKM can induce pluripotency by directly regulating endogenous Oct4 expression (6–9). In addition, Gata6, Gata3, Sox7, Pax1, Gata4, Cebpa, Hnf4a, or Grb2 can also elicit reprogramming together with SKM (10, 11). Their expression indirectly regulates endogenous Oct4 expression (11). These previous studies imply that any factor, which is capable of activating endogenous Oct4, directly or indirectly, can potentially replace Oct4 in reprogramming. Thus far, the possibility of replacing Oct4 with these factors has been assessed mostly in

mice, and profound interest lies in testing whether they can function similarly in other species and especially humans. Until now, engineered GATA3 fused with the VP16 transactivation domain (TAD) and NKX3-1 are the only known factors that can functionally replace OCT4 in inducing pluripotency in humans (9, 12).

Despite the exceptional significance of OCT4 in reprogramming, how its reprogramming competence and pioneering function are actually mediated still remains unknown. Moreover, the question arises why all other OCT factors do not have reprogramming function even though they harbor profound similarities with OCT4 at the level of both primary and secondary structure. Like all OCT factors, OCT4 harbors a DNA binding domain (DBD) and two intrinsic TADs (13–15). Few available studies have clearly attributed the importance of OCT4 DBD toward the reprogramming process, which is required to exert its pioneering function by binding to target gene loci in closed chromatin and determine its reprogramming competence (3, 5, 16–19). However, how transactivation of OCT4's target genes is conferred is largely unknown. As such, little attention has been given to the functions of the TADs within OCT4. Whether or how they influence its pioneering activity and reprogramming competence has not been investigated.

Insights into reprogramming biology have been predominantly obtained in murine systems. Consequently, molecular mechanisms underlying human-specific iPSC generation remain much less understood, although few studies have indicated significant species-specific differences in the reprogramming process (18–22). Reprogramming biology provides an ever more promising avenue for both disease modeling and regenerative medicine. Therefore, profound interest lies in fully understanding the molecular mechanisms of the reprogramming process per se but also in defining these specifically in human cells. Here, we found that OCT6 could induce pluripotency specifically in humans. Because OCT6 and OCT4 displayed different reprogramming competences, we created a series of domain-swapped chimeras that enabled us to decipher specific TADs as crucial elements in reprogramming. Isolating these TADs, we further engineered an additional series of chimeras that rendered almost all OCT factors competent in inducing pluripotency and outperformed OCT4 in reprogramming.

<sup>1</sup>Department of Cell and Developmental Biology, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, Münster 48149, Germany. <sup>2</sup>Clinical and Translational Research Center of Shanghai First Maternity and Infant Hospital, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China. <sup>3</sup>Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, 190 Kai Yuan Avenue, Science Park, Guangzhou 510530, China. <sup>4</sup>Institute of Human Genetics, University of Münster, Vesaliusweg 12-14, Münster 48149, Germany. <sup>5</sup>Group of Computational Biology and Systems Biomedicine, Biodonostia Health Research Institute, San Sebastian 20014, Spain. <sup>6</sup>IKERBASQUE, Basque Foundation for Science, Bilbao 48011, Spain. <sup>7</sup>Flow Cytometry Unit, Max Planck Institute for Molecular Biomedicine, Röntgenstrasse 20, Münster 48149, Germany. <sup>8</sup>School of Biotechnology and Healthcare, Wuyi University, Jiangmen 529020, China. <sup>9</sup>Department of Cardiac Development and Remodeling, Max-Planck-Institute for Heart and Lung Research, Bad Nauheim 61231, Germany. <sup>10</sup>University of Münster, Medical Faculty, Domagkstrasse 3, Münster 48149, Germany.

\*Present address: Department of Early Discovery, Ksilink, 16 rue d'Ankara, 67000 Strasbourg, France.

†Present address: Astellas Pharma Inc., 5-2-3 Tokodai, Tsukuba-shi, Ibaraki 300-2698, Japan.

‡Present address: Guangzhou Regenerative Medicine and Health Guangdong Laboratory LEDY Building 3, Kaiyuan Avenue 188, Hangpu District, Guangzhou 510320, P.R. China.

§Corresponding author. Email: office@mpi-muenster.mpg.de

**RESULTS****OCT6 induces pluripotency specifically in humans**

Discovering factors that can functionally replace OCT4 might enhance our understanding of the reprogramming process and facilitate our definition of its role. So far, attempts to achieve this goal have been performed mostly in mice (7, 8, 10, 11). As significant species-specific differences in reprogramming exist (18–22), profound interest lies in discovering factors that work differently between species. To this end, we performed a screen of 100 candidate genes (table S1) to test their potential of inducing pluripotency in conjunction with SKM in human fibroblasts. We selected 46 transcription factors (e.g., HNF4A and GATA3) that are involved in lineage specification, as counteracting two lineages has been shown to induce pluripotency (10–12). We also selected four epigenetic modifiers (e.g., TET1 and MBD3) that are crucial for the reprogramming process (7, 23–26). We further selected 18 transcription factors that are associated with induction and maintenance of pluripotency (e.g., *KLF2* and *NANOG*) (27–29). In addition, we selected 30 genes related to germ cells (e.g. *SALL1* and *TFAP2A*), as germ cells and pluripotent stem cells share gene expression profiles (30–32). Last, we included OCT1 (also known as POU2F1) and OCT6, which belong to the same protein family as OCT4 (13–15). We then transduced human fibroblasts with each candidate virus along with SKM viruses, cultured them for 21 days, and stained the resulting cells with TRA-1-60 antibody to score for the emergence of putative iPSC colonies (Fig. 1A).

NR5A2, TET1, and GATA3, each of which can induce pluripotency in mice (7, 8, 10, 11), failed to yield iPSC colonies in humans (Fig. 1B), indicating their mouse-specific reprogramming activity. As the transgenes were appropriately expressed in the transduced cells (Fig. 1C), their failure was not due to inadequate transgene expression. GATA3 fused with VP16 TAD can induce pluripotency in humans (12), but its wild-type version could not (Fig. 1B), suggesting that its own transactivation activity is insufficient for human reprogramming. OCT6, which fails to induce pluripotency in mice (2–5), produced distinct iPSC colonies in humans (Fig. 1B). OCT6 was not expressed in human embryonic stem cells (ESCs) and iPSCs (Fig. 1D), but it is expressed in specific tissues including brain, testis, and skin (33–38), suggesting that this nonpluripotency gene acts as a pluripotency inducer. All the tested genes besides OCT4 and OCT6 failed to yield iPSC colonies (fig. S1A). Together, these data demonstrate decisive differences between murine and human reprogramming and identify OCT6 as a human-specific pluripotency inducer.

Protein clustering analysis indicated that mouse and human OCT6 shared 98.89% identity, differing only in five amino acids (fig. S1B). Reprogramming of human fibroblasts with human/mouse OCT6 resulted in iPSC colony formation with similar efficiencies (Fig. 1E). In contrast, both human and mouse OCT6 failed to reprogram mouse embryonic fibroblasts (MEFs). These data indicate that the difference of five amino acids between human and mouse OCT6 is not a determinant for their functionality. Although human and mouse OCT4 (82.5%) shared less identity than human/mouse OCT6 (fig. S1C), human/mouse OCT4 could induce pluripotency in both mouse and human cells (Fig. 1F). Together, these data show that interspecies variations of orthologous OCT proteins do not fully account for their species-dependent reprogramming competence.

From the screening plates, wherein nine iPSC colonies had emerged from OCT6/SKM-transduced cells (fig. S1D), two iPSC lines were established. These iPSC lines fulfilled all hallmarks of

pluripotency, as determined by morphological assessment, gene expression profiling, transgene integration, bisulfite sequencing, karyotyping, and teratoma formation (fig. S2, A to G).

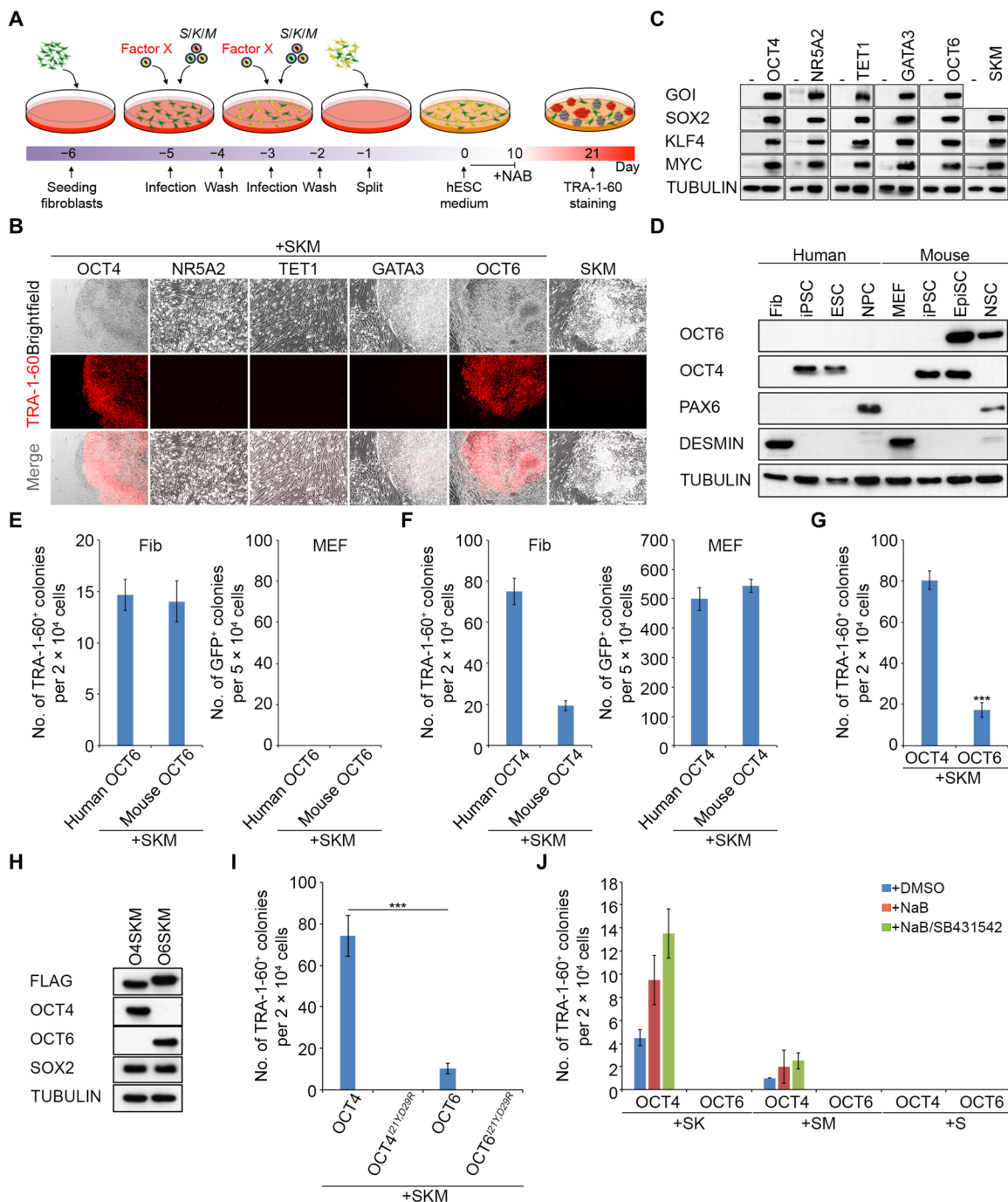
**Common and distinct characteristics of OCT6 and OCT4 in reprogramming**

Because both OCT6 and OCT4 elicited iPSC formation, we next compared characteristics of these OCT proteins in reprogramming. OCT6-based reprogramming was less efficient than that of OCT4 (Fig. 1G), although the transgenes displayed stoichiometrically equivalent expression levels (Fig. 1H). This lower efficiency was maintained upon extended culture (fig. S3A), and switching the donor cell type to neonatal human foreskin keratinocytes (NHFK) and human umbilical vein endothelial cells (HUVEC), which display higher cellular plasticity than fibroblasts (39, 40), did not change this lower efficiency (fig. S3B). Furthermore, increased viral levels of OCT6 did not improve this lower efficiency but rather negatively influenced the reprogramming process (fig. S3C). Overall, OCT4 and OCT6 have different reprogramming competences, although they are functionally interchangeable in inducing pluripotency.

Sodium butyrate [NaB, a histone deacetylase (HDAC) inhibitor] supply enhances OCT4-based reprogramming by increasing its transactivation activity (41, 42). This raises the question whether OCT6-based reprogramming is similarly amenable to acceleration. NaB supply markedly enhanced the reprogramming efficiency of both OCT4 and OCT6 (fig. S3E). The enhanced reprogramming efficiency went along with elevated expression of pluripotency genes including, but not limited to, *NANOG* and *LIN28A* (fig. S3F), which was mediated by the increased transactivation activity of both OCT proteins (fig. S3G). The binding ability of OCT4 and OCT6 to pluripotency gene enhancers remained unchanged upon the NaB supply (fig. S3H), suggesting that DNA binding of OCT proteins alone does not fully account for their reprogramming capability, but target gene activation through TADs is rather crucial for reprogramming.

The partnership between OCT4 and SOX2 is essential to establish the fundamental feature of the pluripotency network (43), raising the question whether OCT6 is subject to a similar criterion. We found that both OCT4<sup>I21Y/D29R</sup> and OCT6<sup>I21Y/D29R</sup>, which are defective in OCT-SOX binding (44), failed to yield any iPSC colonies (Fig. 1I), suggesting that SOX2 is an essential patterner for OCT6 to elicit reprogramming. In addition, we found that OCT4/SOX2 additionally required at least either KLF4 or MYC to elicit iPSC formation (Fig. 1J). In contrast, OCT6/SOX2 required both KLF4 and MYC, suggesting that in comparison to OCT4, the reprogramming competence of OCT6 depends more strongly on additional factors that facilitate the reprogramming process.

In mice, OCT6 is not expressed in naïve pluripotent iPSCs, but it is expressed in epiblast stem cells (EpiSCs) (Fig. 1D) (45, 46), which are commonly referred to as primed pluripotent stem cells. Because OCT6 can produce human iPSCs, which are thought to represent a primed pluripotent state (28, 29), we speculated that OCT6 might not induce a naïve but a primed state of pluripotency in mice. In contrast to OCT4, however, OCT6 failed to yield mouse EpiSC-like colonies (fig. S3D), demonstrating that neither naïve nor primed pluripotency can be achieved by OCT6 in mice. Together, these data reveal commonalities but decisive differences between OCT4- and OCT6-based reprogramming: They are both amenable to marked enhancement by reduction in an epigenetic blockade but display



**Fig. 1. Factor screening identifies OCT6 as a pluripotency inducer in humans.** (A) Schematic representation of the screening procedure. We transduced fibroblasts with each candidate virus along with SKM viruses. After two serial infections over the course of 3 days, the transduced cells were plated onto the plates pre-coated with feeder cells and cultured in hESC medium for 21 days. NaB (an HDAC inhibitor) was supplied to the medium for the first 10 days. The resulting cells were then stained with TRA-1-60 antibody. One hundred candidate genes (see table S1) were used as factor X. (B) Phase-contrast and fluorescence images of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transduced with the indicated factors. Scale bar, 250 μm. (C) Western blot for GOI (OCT4, NR5A2, TET1, GATA3 and OCT6), SOX2, KLF4, MYC, and TUBULIN in fibroblasts that were transduced with the indicated factors. GOI, gene of interest. (D) Western blot for OCT6, OCT4, PAX6, DESMIN, and TUBULIN in the indicated samples. Fib, fibroblasts; NPC, neural precursor cells that were differentiated from HUES6; MEF, mouse embryonic fibroblasts; EpiSC, epiblast stem cells; NSC, neural stem cells. (E to G) Quantification of TRA-1-60<sup>+</sup> colonies that had emerged from cells transduced with the indicated factors (n = 3). (H) Western blot for FLAG, OCT4, OCT6, SOX2, and TUBULIN in O4SKM- and O6SKM-transduced cells. FLAG tag was fused to N-TADs of OCT4 and OCT6. (I) Quantification of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transduced with the indicated factors (n = 3). (J) Quantification of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transduced with the indicated factors (n = 3). The transduced cells were cultured with dimethyl sulfoxide (DMSO) or the indicated chemicals. \*\*\*P < 0.001.



significant differences in reprogramming competence and dependence on other reprogramming factors.

### OCT6-based reprogramming is attenuated through delayed activation of pluripotency network

The difference of reprogramming efficiency between OCT4 and OCT6 is of particular interest because it provides a unique access point to determine differential reprogramming features between these OCT proteins. More specifically, these differences may highlight the features that render OCT4 as a strong pluripotency inducer. To this aim, we first sought to elucidate the molecular mechanisms underlying the OCT6-based reprogramming process in comparison with that of OCT4. We performed RNA-sequencing (RNA-seq) on FACS (fluorescence-activated cell sorting)-sorted CD13<sup>-</sup>/TRA-1-60<sup>+</sup> cells that had emerged from O4SKM- and O6SKM-transduced cells over time (fig. S4A). Up to six replicates for eight to nine individual time points over the reprogramming process were processed (table S2). Replicates displayed high concordance (fig. S4B), and each sample yielded an average of ~19.4 million mapped reads (table S2). The abundance of CD13<sup>-</sup>/TRA-1-60<sup>+</sup> O6SKM-transduced cells on day 2 was extremely low (fig. S4A). Consequently, RNA-seq of these samples led to significantly lower coverage compared with all other samples, and these were, therefore, excluded from comparative downstream analyses. As expected, both O4SKM- and O6SKM-transduced cells triggered marked transcriptomic changes over time in a progressive fashion and converged together toward an end point state of pluripotency (Fig. 2, A and B, and fig. S4B). Both O4SKM- and O6SKM-transduced cells displayed two major trajectories of transcriptome transition: an initial transition phase (days 2 to 6) indicative of reprogramming induction and a late transition phase (days 8 to 16) indicative of reprogramming maturation (fig. S4B). Differentially expressed genes between time points of each condition (>2-fold;  $P < 0.01$ ) were subdivided into activated, repressed, or transiently changed genes (fig. S4C). This subdivision revealed a notably high concordance of transcriptome changes over time between O4SKM- and O6SKM-transduced cells (fig. S4, D to F), suggesting that OCT4 and OCT6 engage similar molecular events to achieve reprogramming.

O6SKM-transduced cells underwent a markedly slower transition toward pluripotency in comparison with O4SKM-transduced cells (Fig. 2, A and B, and fig. S4B). For example, O6SKM-transduced cells on days 4 to 6 still clustered with fibroblasts (fig. S4B), and O6SKM-transduced cells on days 8 to 10 clustered with O4SKM-transduced cells on days 2 to 4 (Fig. 2, A and B). A mesenchymal-to-epithelial transition is required for the formation of iPSCs from fibroblasts, which is associated with the progressive down-regulation and up-regulation of fibroblast- and epithelial-related genes, respectively (47, 48). Both O4SKM and O6SKM equally induced down-regulation of fibroblast genes, including, but not limited to, *VIM*, *CLOL1A2*, and *FNI*, and up-regulation of epithelial genes, such as *EPCAM*, *CLND7*, and *CRB3* (Fig. 2C). In contrast, while pluripotency genes (*GDF3*, *NODAL*, and *DPPA2*) were activated as early as day 2 in O4SKM-transduced cells, O6SKM-induced activation of those genes was first apparent after day 8 (Fig. 2C and fig. S4C). This delayed activation of pluripotency genes was confirmed by immunofluorescence. OCT4<sup>(endo+exo)</sup><sup>+</sup> and NANOG<sup>+</sup> cells appeared on day 0 (day 5 after infection) to day 2 of O4SKM-transduced cells, but were only detected from day 10 onward in O6SKM-transduced cells (fig. S4G).

Transgene silencing is an important feature for successful reprogramming (49). Immunofluorescence analysis revealed silencing

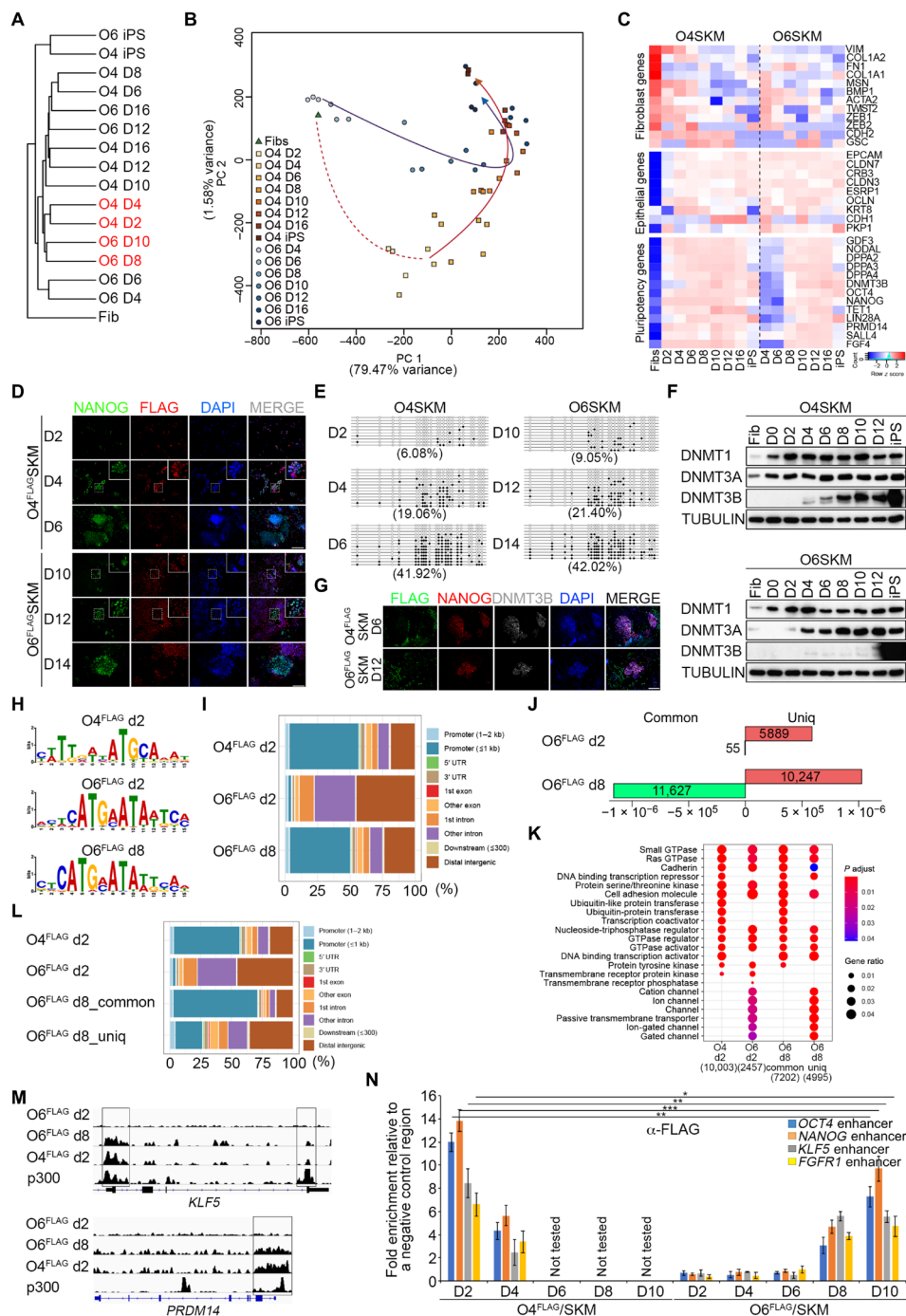
of OCT4<sup>FLAG</sup> and OCT6<sup>FLAG</sup> transgenes in emerging NANOG<sup>+</sup> cells (Fig. 2D). However, transgene silencing was delayed in O6SKM-transduced cells, which occurred on day 14, while O4SKM-transduced cells underwent transgene silencing as early as day 6 (Fig. 2D). Transgene silencing was the result of a gain of methylation at CpG sites of the 5' long terminal repeat (LTR) promoter of the viral vector (Fig. 2E) and was likely mediated by the timely induction of a de novo methyltransferase DNMT3B (Fig. 2, F and G). Consistent with the delayed acquisition of pluripotency, the induction of DNMT3B expression was markedly delayed in OCT6-based reprogramming that resulted in the delayed transgene silencing (Fig. 2, D to G).

To characterize binding events of OCT4 and OCT6, we next performed chromatin immunoprecipitation followed by high-throughput sequencing [chromatin immunoprecipitation sequencing (ChIP-seq)] for FLAG in O4<sup>FLAG</sup>SKM-transduced cells (day 2), O6<sup>FLAG</sup>SKM-transduced cells (day 2), and O6<sup>FLAG</sup>SKM-transduced cells (day 8). Despite binding sites of OCT6<sup>FLAG</sup> (day 2) and OCT6<sup>FLAG</sup> (day 8) were enriched in the same OCT motif, distribution patterns of OCT6<sup>FLAG</sup> binding sites onto genomic regions at the respective time points were largely different (Fig. 2, H and I). Along this line, only few binding sites of OCT6<sup>FLAG</sup> (day 2) (O6\_day2 common: 55) were overlapped with those of OCT4<sup>FLAG</sup> (day 2) (Fig. 2J). In contrast, we found a large proportion of OCT6<sup>FLAG</sup> (day 8) binding sites (O6\_day 8 common: 11,627) that was overlapped with OCT4<sup>FLAG</sup> (day 2) binding sites. As expected, gene ontology (GO) terms of nearest genes of O6\_day 8 common sites and their genomic distribution profiles were highly similar to those of OCT4<sup>FLAG</sup> (day 2) binding sites (Fig. 2, K and L). However, GO terms of nearest genes, which were uniquely bound by OCT6<sup>FLAG</sup> (day 8), and their genomic distribution profiles were rather similar to those of OCT6<sup>FLAG</sup> (day 2) binding sites. Genome browser view and ChIP-quantitative polymerase chain reaction (qPCR) further confirmed that OCT6<sup>FLAG</sup> displayed differential binding abilities to enhancers of pluripotency genes during reprogramming (Fig. 2, M and N). No obvious binding of OCT6<sup>FLAG</sup> to these enhancers was observed on days 2 to 6 (Fig. 2N). However, significant binding of OCT6<sup>FLAG</sup> to the enhancer regions appeared first on day 8 (Fig. 2, M and N) in concomitance with their activation (Fig. 2C), and its binding further increased on day 10 (Fig. 2N). Together, these data clearly indicate that OCT6 has differential binding abilities to loci where OCT4 normally bind to during the reprogramming process, and the delayed O6SKM-based reprogramming is at least in part due to attenuated binding kinetics of OCT6 to regulatory regions of pluripotency genes.

### Transactivation domains influence the reprogramming process

The DNA binding targets of OCT6 were similar to those of OCT4, but its binding kinetics and activation of the target genes were delayed (Fig. 2, C and M). Furthermore, TADs of OCT4 and OCT6 proteins exhibited different transactivation activities (Fig. 3, A and B). Specifically, the overall transactivation activity of OCT4, when both N- and C-TADs were probed together, was >1.4-fold higher than that of OCT6. Furthermore, OCT6 N-TAD displayed a >1.4-fold higher activity than OCT4 N-TAD, while OCT4 C-TAD displayed a marked >21-fold higher activity than OCT6 C-TAD. To understand how functional features of two structural components (DBD and TAD) endow two OCT proteins with different reprogramming competences, we created a series of domain-swapped chimeras in which N-, C-, or N-/C-TAD of OCT4 or OCT6 was introduced into

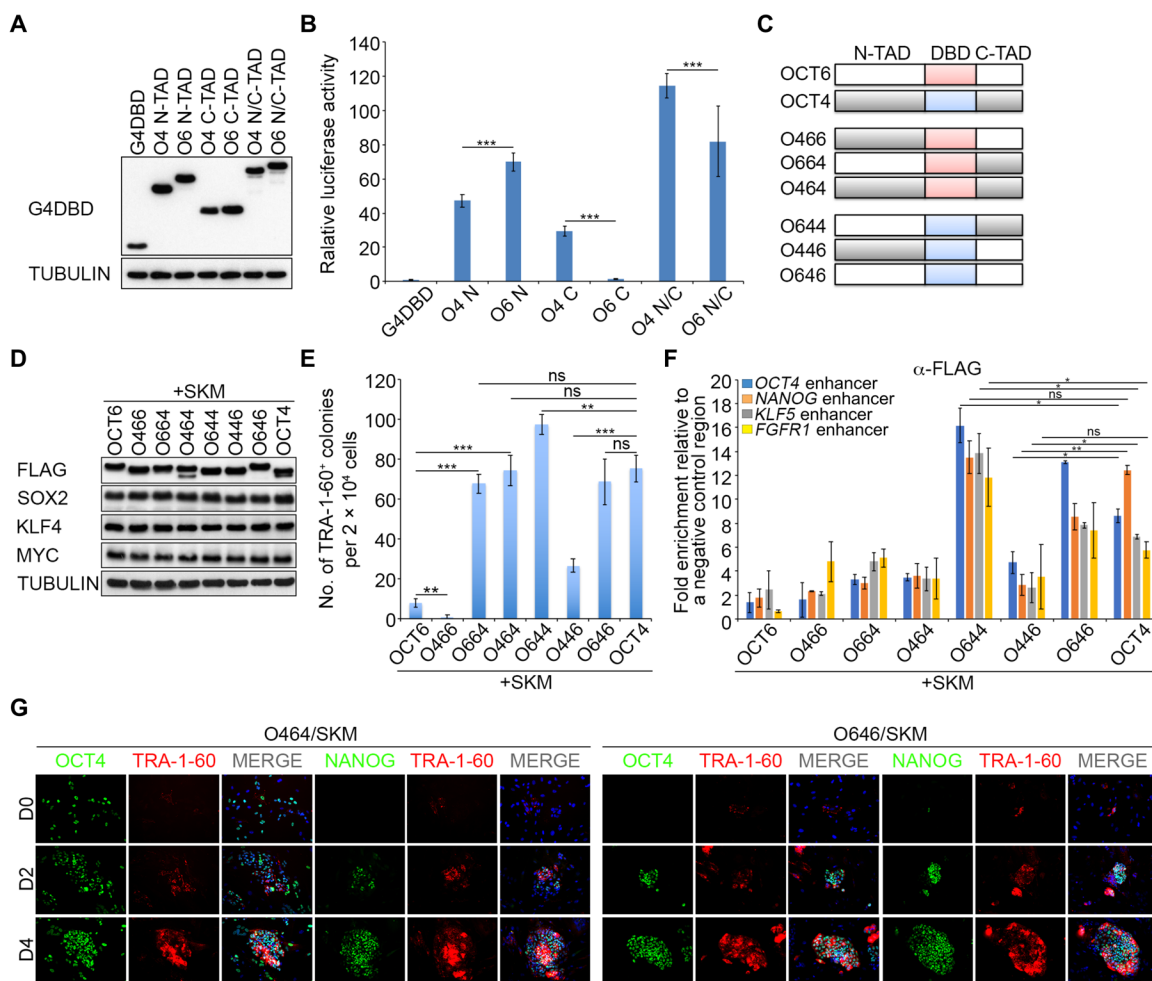




**Fig. 2. Characterization of OCT4- and OCT6-based reprogramming.** (A) Unsupervised hierarchical clustering analysis of the indicated samples. The list of samples we used for RNA-seq can be found in table S2. (B) Principal components analysis of the indicated samples. The solid and dashed red lines indicate O4SKM-reprogramming route (O4). The solid blue line indicates O6SKM-reprogramming route (O6). (C) Heatmap analysis of the indicated samples. (D) Transgene expression was monitored by NANOG and FLAG staining. FLAG tag was fused to N-TADs of OCT4 and OCT6. DAPI was used as a nuclear counterstain. Scale bars, 250  $\mu$ m. (E) Bisulfite sequencing analysis of the 5' LTR promoter of pMXs in the indicated samples. The percentages of methylation are given. (F) Western blot for DNMT1, DNMT3A, DNMT3B, and TUBULIN in the indicated samples. (G) Fluorescence images of iPSC colonies that were stained with FLAG, NANOG, and DNMT3B antibodies. DAPI was used as a nuclear counterstain. Scale bar, 250  $\mu$ m. (H) Motif analysis of OCT4<sup>FLAG</sup> and OCT6<sup>FLAG</sup> binding sites. While OCT4<sup>FLAG</sup> predominantly bound to the SOX-OCT motif, OCT6<sup>FLAG</sup> bound to the MORE motif. (I) Genomic distribution patterns of OCT4<sup>FLAG</sup> and OCT6<sup>FLAG</sup> binding sites. (J) Quantification of common and unique binding sites of OCT6<sup>FLAG</sup> and OCT4<sup>FLAG</sup>. (K) GO terms of common and unique binding sites of OCT6<sup>FLAG</sup> and OCT4<sup>FLAG</sup>. (L) Genomic distribution patterns of common and unique binding sites of OCT6<sup>FLAG</sup> and OCT4<sup>FLAG</sup>. UTR, untranslated region. (M) Representative regions cobound by OCT4<sup>FLAG</sup>/OCT6<sup>FLAG</sup> and p300. (N) ChIP assay of pluripotency gene enhancers in the indicated samples ( $n = 3$ ). The fold enrichment was calculated by the standard curve method and normalized to the value obtained at a negative control region. d or D, day. \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , and \* $P < 0.05$ .

reciprocal sites of OCT6 or OCT4, respectively, and tested their reprogramming capacity (Fig. 3, C and D). The chimeras were named on the basis of the origin of each functional unit in order: N-TAD, DBD, and then C-TAD. For example, O446 consisted of OCT4 N-TAD, OCT4 DBD, and OCT6 C-TAD (Fig. 3C). O664 and O464 showed a significant increase in reprogramming efficiency compared with OCT6, and both produced iPSC colonies as efficiently as OCT4 (Fig. 3E). Moreover, O466 yielded fewer iPSC colonies than OCT6. These data show that OCT4 N-TAD and OCT6 C-TAD negatively influence reprogramming or, conversely, that OCT4 C-TAD and OCT6 N-TAD are responsible for their positive impact on reprogramming. ChIP assays revealed that the binding abilities of O466, O664, and O464 to *OCT4*, *NANOG*, *KLF5*, and *FGFR1* enhancers were similar (Fig. 3F). Thus, the increased reprogramming efficiency of O664 and O464 over OCT6 was due to the domain-swapped OCT4 C-TAD, which displayed the higher transactivation

activity (Fig. 3B). Consistently, O644 and O646 produced iPSC colonies as efficiently as OCT4, but O446 significantly lost its reprogramming competence (Fig. 3E), further confirming that OCT4 N-TAD and OCT6 C-TAD negatively influence reprogramming. The loss of reprogramming competence of O446 appeared to be a consequence of its diminished binding ability to the pluripotency gene enhancers (Fig. 3F), further indicating that the OCT4 C-TAD is involved in DNA binding. O646, which did not contain the OCT4 C-TAD, produced iPSCs as efficiently as OCT4 (Fig. 3E), and its binding ability to the *OCT4*, *NANOG*, *KLF5*, and *FGFR1* enhancers was comparable to that of OCT4 (Fig. 3F), suggesting that similar to the OCT4 C-TAD, the OCT6 N-TAD is involved in DNA binding. Consequently, O644, which contained both the OCT4 C-TAD and the OCT6 N-TAD, showed the highest binding ability to the pluripotency gene enhancers and produced the most iPSC colonies (Fig. 3, E and F). The increased reprogramming efficiency of O464



**Fig. 3. Identification of transactivation domains negatively or positively influences the reprogramming process.** (A) Western blot for G4DBD and TUBULIN in HeLa cells transfected with the indicated luciferase constructs. G4DBD, DNA binding domain of the yeast transcription factor Gal4. (B) Relative luciferase activity of the indicated constructs ( $n = 3$ ). The luciferase activity of the indicated constructs was normalized to that of G4DBD. (C) Schematic representation of OCT4-OCT6 chimeras. (D) Western blot for FLAG, SOX2, KLF4, MYC, and TUBULIN in fibroblasts that were transduced with the indicated factors. FLAG tag was fused to N-TADs of each factor. (E) Quantification of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transduced with the indicated factors ( $n = 3$ ). (F) ChIP assay of pluripotency gene enhancers in cells that were transduced with the indicated factors ( $n = 3$ ). ChIP assay was carried out on day 2 of reprogramming. The fold enrichment was calculated by the standard curve method and normalized to the value obtained at a negative control region. (G) Expression of OCT4, NANOG, and TRA-1-60 in cells that were transduced with the indicated factors. The OCT4 antibody recognizes OCT4 N-TAD so that it can detect both endogenous OCT4 and exogenous O464 proteins in O464/SKM-transduced cells. DAPI was used as a nuclear counterstain. Scale bar, 100  $\mu$ m. \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , and ns  $P > 0.05$ .

and O646 over OCT6 was largely associated with shorten reprogramming kinetics, as the first OCT4<sup>+</sup> and NANOG<sup>+</sup> cells/colonies appeared on day 0 or 2 (Fig. 3G). Overall, these findings unequivocally delineate TADs within OCT factors as critical elements that render different reprogramming outcomes.

### Intrinsic properties within POU III factors that are essential or detrimental to reprogramming

Eight proteins within the POU (Pit-Oct-Unc) family have been classified as OCT proteins (Fig. 4A) (13, 14), of which only OCT4 is highly expressed in both naïve and primed pluripotent stem cells (Fig. 4B). The class III POU factors OCT6, OCT7 (also known as POU3F2 and BRN2), OCT8 (also known as POU3F3 and BRN1), and OCT9 (also known as POU3F3 and BRN4) are all expressed in brain cells (13–15, 36, 38) and have been shown to display functional redundancy in vivo (34, 35, 37). High sequence similarities between POU III factors explain their similar DNA binding characteristics as a predominant reason for their functional similarities (3, 50, 51). Despite these similarities, OCT6 was the only POU III factor that could induce pluripotency by activating *OCT4* and *NANOG* on day 8 (Fig. 4, C to E), raising the question as to what makes OCT6 unique among the POU III factors. The protein sequence alignment revealed that OCT6 differed from other POU III factors by seven amino acids within the DBD domain (Fig. 4F). Mutations of these amino acids (OCT6<sup>T74A</sup>, OCT6<sup>G113S</sup>, and OCT6<sup>G135N</sup>) significantly reduced its reprogramming competence (Fig. 4G and fig. S5A). Of those three, OCT6<sup>G113S</sup> entirely lost its reprogramming competence, demonstrating that the residue Gly<sup>113</sup> is necessary for reprogramming competence of OCT6 or, conversely, that the residue Ser<sup>113</sup> is detrimental to OCT7-, OCT8-, and OCT9-based reprogramming. We next created a series of reverse mutants (OCT7<sup>S113G</sup>, OCT8<sup>S113G</sup>, OCT9<sup>S113G</sup>, OCT7<sup>T74A,S113G</sup>, OCT8<sup>T74A,S113G</sup>, OCT9<sup>T74A,S113G</sup>, OCT7<sup>T74A,S113G,G135N</sup>, OCT8<sup>T74A,S113G,G135N</sup>, and OCT9<sup>T74A,S113G,G135N</sup>) and tested their reprogramming capacity. None of them yielded iPSC colonies (Fig. 4H), suggesting that additional elements beside the residue Gly<sup>113</sup> are required to induce pluripotency by these POU III proteins.

Luciferase assays revealed that TADs of the POU III factors exhibited variable transactivation activities (Fig. 5A and fig. S5B). This finding raises the possibility that the TADs might be additional elements that endow the POU III factors with different reprogramming competences. To test this, we created 18 domain-swapped chimeras and tested their reprogramming ability (Fig. 4I and fig. S5C). All chimeras containing the DBDs of OCT7, OCT8, or OCT9 failed to induce pluripotency (Fig. 4I), confirming that the residue Ser<sup>113</sup> within the DBDs is determinantal to reprogramming. When both OCT6 N-TAD and OCT6 DBD were introduced into corresponding sites of OCT7, OCT8, and OCT9 (O667, O668, and O669), they could produce iPSC colonies (Fig. 4I). Together, these data demonstrate that the uniqueness of OCT6 among the POU III factors arises from two intrinsic properties (Gly<sup>113</sup> in its DBD and its N-TAD) and that the persistent deficiency of OCT7, OCT8, and OCT9 in reprogramming is largely due to both Ser<sup>113</sup> in their DBDs and their N-TADs.

### Strong reprogramming competence of OCT4 arises from its C-TAD

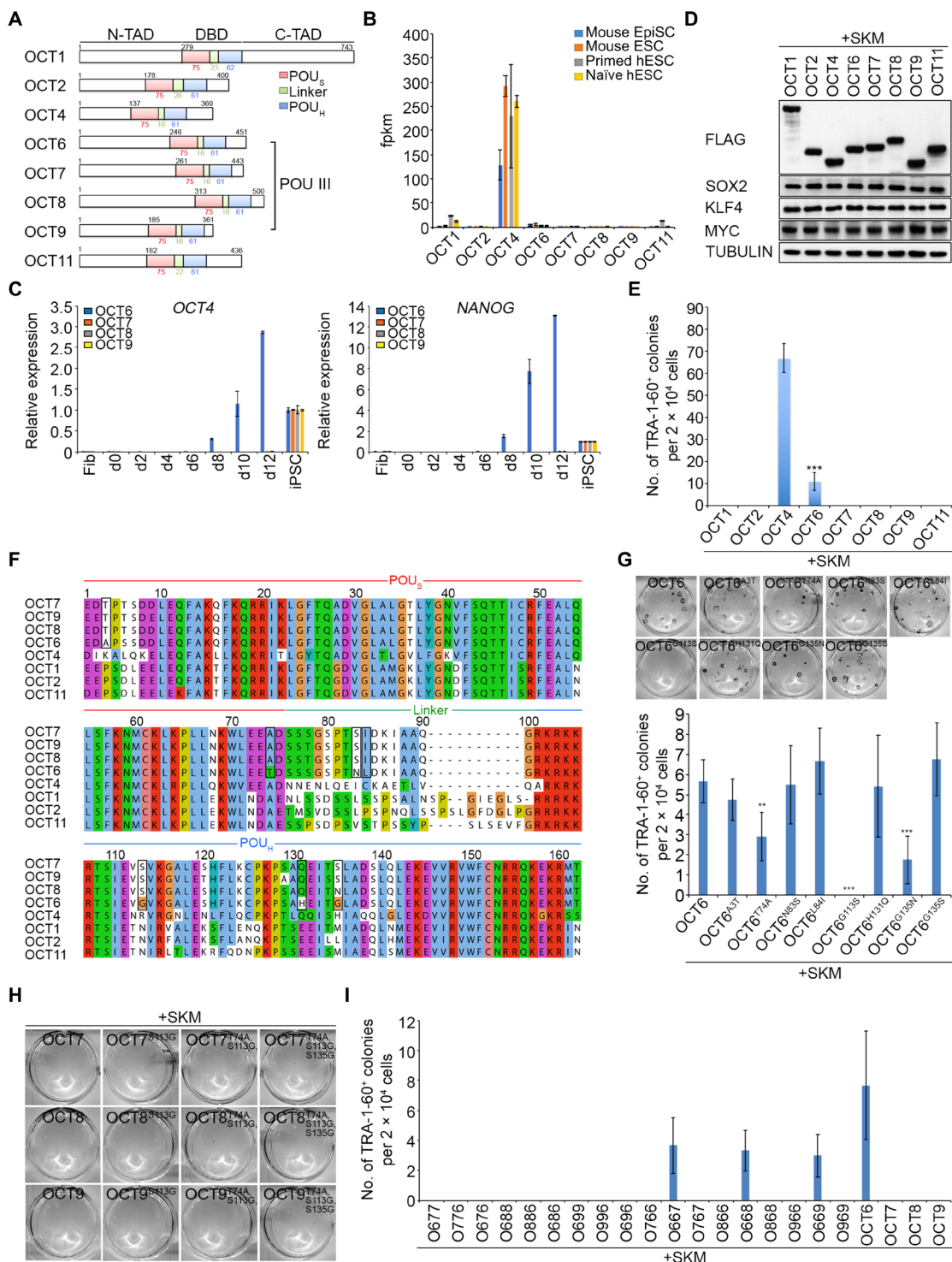
Thus far, our findings reveal that the OCT4 C-TAD is not only necessary for the strong reprogramming function of OCT4 itself but also sufficient to bestow superior reprogramming capacity when transferred to OCT6. In extension of these findings, we created an

additional series of domain-swapped chimeras (fig. S5, D and E) to test whether the C-TAD and/or N-TAD of OCT4 could bestow reprogramming competence to all other OCT family members which otherwise cannot induce pluripotency (Fig. 4E). As displayed in Fig. 5B, we found that (i) when N-TADs of all POU III factors were introduced into corresponding sites of OCT4 (O644, O744, O844, and O944), they produced iPSC colonies as efficient as OCT4, suggesting that the OCT4 N-TAD can be functionally replaced by N-TADs of the POU III factors; (ii) when C-TADs of all POU III factors were introduced into corresponding sites of OCT4 (O446, O447, O448, and O449), they significantly lost reprogramming competence, confirming that the OCT4 C-TAD is critical for reprogramming; and (iii) when both N- and C-TADs of all POU III factors were introduced into corresponding sites of OCT4 (O646, O747, O848, and O949), they produced iPSC colonies with variable efficiencies, which well correlated with the transactivation activities of their corresponding N-/C-TADs (Fig. 5A). As shown in Fig. 5C, we further found that (i) when the N-TAD of OCT4 was introduced into corresponding sites of the POU III factors (O466, O477, O488, and O499), they did not yield any iPSC colonies, like OCT7, OCT8, and OCT9, confirming that the OCT4 N-TAD does not play a critical role in reprogramming; (ii) when the C-TAD of OCT4 was introduced into corresponding sites of the POU III factors (O664, O774, O884, and O994), they efficiently produced iPSC colonies, confirming that the OCT4 C-TAD has a profound effect on reprogramming; and (iii) when both N- and C-TADs of OCT4 were introduced into corresponding sites of the POU III factors (O464, O474, O484, and O494), they produced iPSC colonies with variable efficiencies, confirming that OCT6 DBD and other POU III factor DBD have different abilities for reprogramming, which is largely due to the residue Gly/Ser<sup>113</sup> (Fig. 4, F and G). The stability of the various chimeric proteins did not appear to play a role in their reprogramming competence, as judged by their similar expression levels revealed by Western blot (fig. S5, D and E). Together, these results unequivocally demonstrate that the C-TAD of OCT4 is the unique functional entity across all OCT factors that make OCT4 a strong reprogramming inducer.

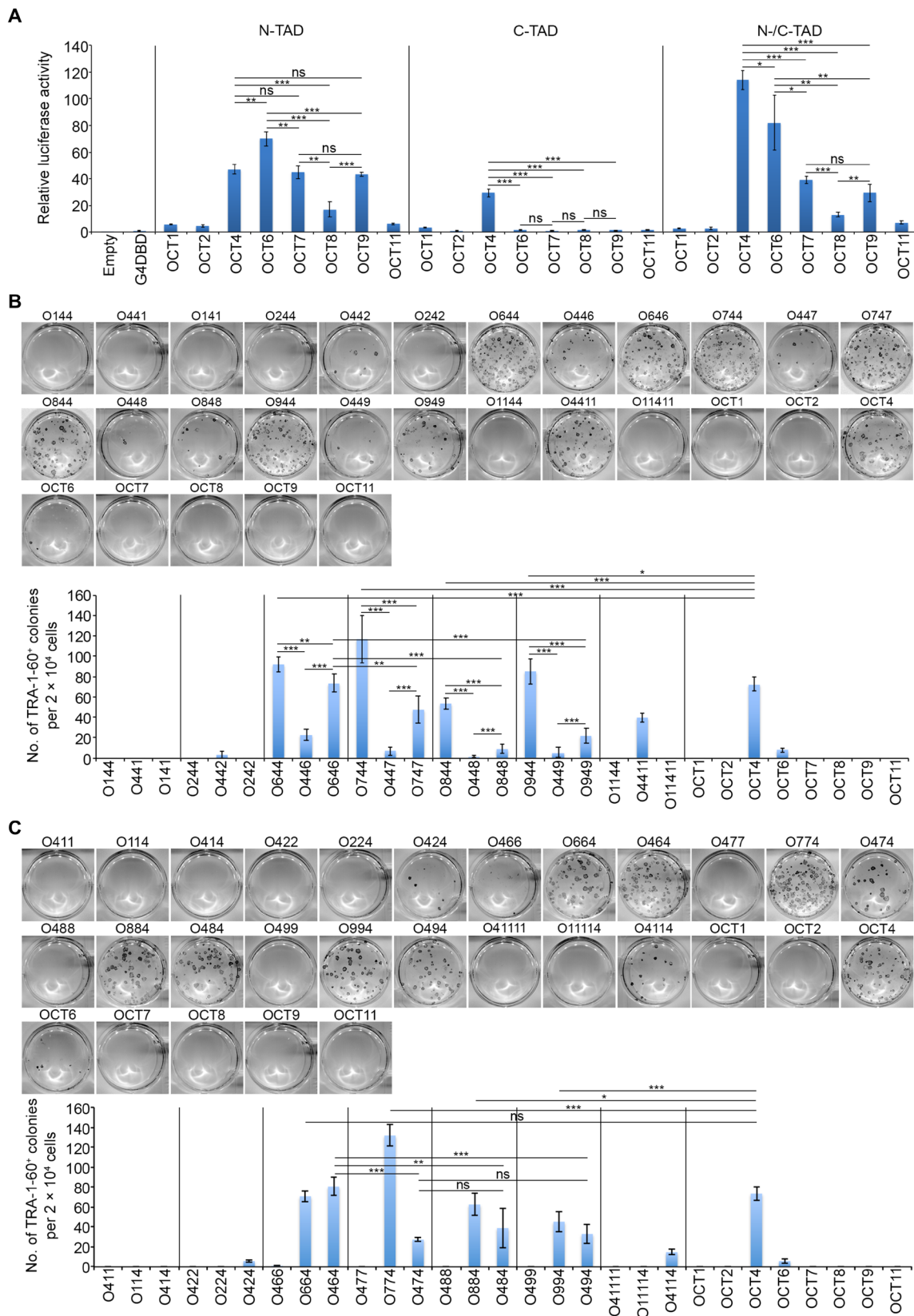
### Characterization of chimeras that are superior to OCT4

Our domain-swapping strategy resulted in the identification of four chimeras (O644, O744, O944, and O774) that outperformed OCT4 in reprogramming (Fig. 5, B and C). To elucidate how they achieve such a high reprogramming efficiency, we transduced fibroblasts with viruses containing O644, O744, O944, O774, or OCT4 (Fig. 6A) and monitored cell fate transition toward pluripotency by flow cytometry. At the early phase of reprogramming (days –4 to 6), O644, O744, and O944 produced more TRA-1-60<sup>+</sup> cells than OCT4 (Fig. 6B and fig. S6A). Furthermore, they elicited reprogramming faster than OCT4, as evidenced by significantly elevated expression of pluripotency genes (Fig. 6C). This accelerated reprogramming process was mediated by the increased binding ability of O644, O744, and O944 to the pluripotency gene enhancers (Fig. 6D), suggesting that the superiority of O644, O744, and O944 in reprogramming is by virtue of their positive regulation onto key pluripotency genes. In contrast to these, O774, however, displayed lower binding ability to these enhancers, mediated slower reprogramming process, and yielded fewer TRA-1-60<sup>+</sup> cells than OCT4 (Fig. 6, B to D, and fig. S6A), indicating that O774 engages in a different route to achieve such a high reprogramming efficiency.

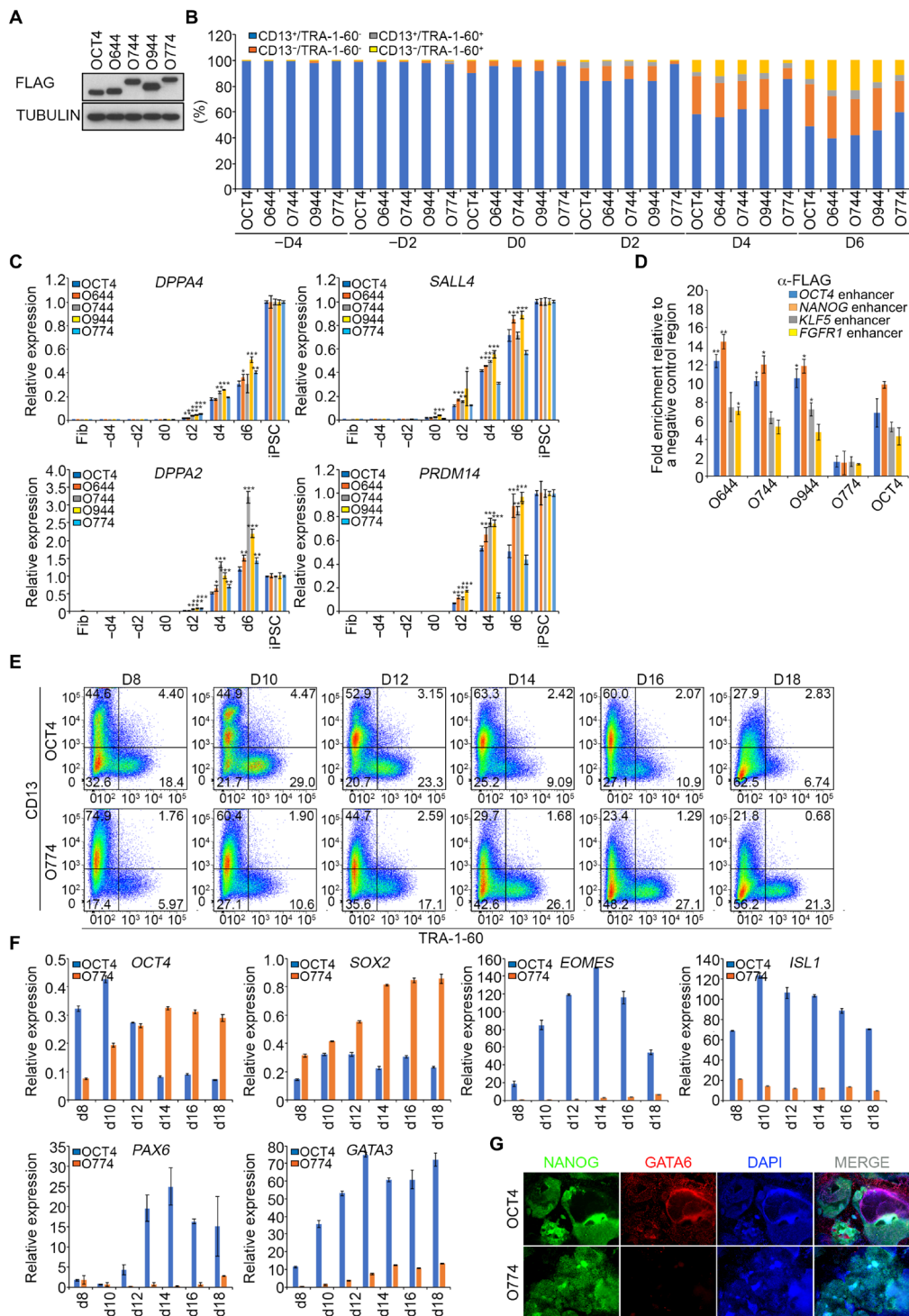




**Fig. 4. Intrinsic properties that are essential or detrimental to POU III factor-based reprogramming.** (A) Schematic representation of eight OCT proteins. The numbers indicate length of amino acids. N-TAD, N-terminal transactivation domain; DBD, DNA binding domain; C-TAD, C-terminal transactivation domain. The DBD has a bipartite structure with two subdomains, POU-specific domain (red) and POU-homeodomain (blue), which are connected by a linker (green). (B) Expression levels of all OCT factors in primed and naïve pluripotent stem cells ( $n = 2$  to 3). The data were obtained from ArrayExpress (E-MTAB-2857, E-MTAB-2856, and E-MTAB-2856). fpm, fragments per kilobase million. (C) Relative expression of *OCT4* and *NANOG* in the indicated samples ( $n = 3$ ). Expression values of these genes were normalized by expression value of *RPL37A*. (D) Western blot for FLAG, SOX2, KLF4, MYC, and TUBULIN in fibroblasts that were transfected with the indicated factors. FLAG tag was fused to N-TADs of all OCT factors. (E) Quantification of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transfected with the indicated factors ( $n = 3$ ). (F) Protein sequence alignment of OCT factor's DBD. (G to I) Representative images and quantification of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transfected with the indicated mutants and chimeras ( $n = 6$ ). \*\*\* $P < 0.001$  and \*\* $P < 0.01$ .



**Fig. 5. Strong reprogramming competence of OCT4 arises from its C-TAD.** (A) Relative luciferase activity of the indicated constructs ( $n = 3$ ). The luciferase activity of the indicated constructs was normalized to that of G4DBD. (B and C) Representative images and quantification of TRA-1-60<sup>+</sup> colonies that had emerged from fibroblasts transduced with the indicated chimeras ( $n = 6$ ). \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , and ns  $P > 0.05$ .



**Fig. 6. Characterization of chimeras that are superior to OCT4.** (A) Western blot for FLAG and TUBULIN in cells that were transduced with the indicated factors. FLAG tag was fused to N-TADs of each chimera. (B) Quantification of cells that had immunoreactivity to CD13 and TRA-1-60 by flow cytometry (see fig. S6A). (C) Relative expression of *DPPA4*, *SALL4*, *DPPA2*, and *PRDM14* in the indicated samples ( $n = 3$ ). Expression values of these genes were normalized by expression value of *RPL37A*. (D) ChIP assay of *OCT4*, *NANOG*, *KLF5*, and *FGFR1* enhancers in cells that were transduced with the indicated factors ( $n = 3$ ). ChIP assay was carried out on day 2 of reprogramming. The fold enrichment was calculated by the standard curve method and normalized to the value obtained at a negative control region. (E) FACS plots for CD13 and TRA-1-60 in the indicated samples. O4SKM- and O774/SKM-transduced cells were cultured and stained with CD13 and TRA-1-60 antibodies at the indicated days. Their expression was analyzed by flow cytometry. (F) Relative expression of *OCT4*, *SOX2*, *EOMES*, *ISL1*, *PAX6*, and *GATA3* in the indicated samples ( $n = 2$  to  $3$ ). Expression values of these genes were normalized by expression value of *RPL37A*. (G) Fluorescence images NANOG<sup>+</sup> and GATA6<sup>+</sup> cells in O4SKM- and O774/SKM-transduced cells. The cells were stained on day 18 of reprogramming. DAPI was used as a nuclear counterstain. Scale bar, 500  $\mu$ m. \*\*\* $P < 0.001$ , \*\* $P < 0.01$ , \* $P < 0.05$ , and ns  $P > 0.05$ .



Further investigations on the late phase of reprogramming (days 8 to 18) revealed that the number of TRA-1-60<sup>+</sup> cells continuously increased in O774-transduced cells but greatly decreased in OCT4-transduced cells on day 12 and beyond (Fig. 6E). Moreover, pluripotency genes were continuously up-regulated and/or maintained their expression in O774-transduced cells over time but were markedly down-regulated in OCT4-transduced cells on day 12 and beyond (Fig. 6F). The down-regulation of pluripotency genes was accompanied by the up-regulation of *EOMES* (an early mesendoderm maker), *ISL1* (a cardiac mesoderm maker), *PAX6* (a neuroectoderm maker), and *GATA3* (a mesoderm marker), indicative of spontaneous differentiation (Fig. 6F). We further confirmed that a large proportion of GATA6<sup>+</sup> (an endoderm maker) cells within NANOG<sup>+</sup> colonies had emerged from OCT4-transduced cells (Fig. 6G). However, GATA6<sup>+</sup> cells were barely detected in NANOG<sup>+</sup> colonies that had emerged from O774-transduced cells, indicating that iPSC colonies generated by OCT4 have a higher propensity toward differentiation compared with ones generated by O774. Both OCT4 and O774 transgenes were silenced at the early phase of reprogramming (fig. S6B), suggesting that transgene expression does not fully account for the stability of iPSC colonies generated from OCT4- and O774-transduced cells. Together, these data demonstrate that while O644, O744, and O944 mediate their superiority by directly regulating key pluripotency genes at early stages of reprogramming, O774 mediates its superiority by stabilizing the pluripotency state at late stages of reprogramming.

## DISCUSSION

Given the tremendous potential of iPSCs for regenerative medicine and disease modeling, a comprehensive understanding of reprogramming mechanisms in a human-specific context has become increasingly more important. In this study, we discover that OCT6 can induce pluripotency specifically in humans. This discovery provides a means to compare between OCT4- and OCT6-mediated reprogramming in unprecedented detail and prompts us to dissect and determine the functionality of individual domains across all OCT proteins. Thereby, we answer the long-standing open question as to what makes OCT4 such a strong reprogramming factor, e.g., what features does OCT4 have that other OCT factors do not have? In addition, how do those features convey reprogramming competence?

OCT4 is a pioneer transcription factor (18, 19) and has strong reprogramming competence compared with its family members (Fig. 4E), but exactly how its reprogramming competence and pioneering function are achieved has essentially remained unknown. Few available studies have attributed the importance of OCT4 DBD toward reprogramming, which is required to exert its pioneering function and determine its reprogramming competence (3, 5, 16–19). However, our current study unequivocally shows that the strong reprogramming competence of OCT4 is achieved through its C-TAD. The absence of the OCT4 C-TAD abolishes its binding ability to regulatory regions of pluripotency genes that results in a significant loss of reprogramming competence. Conversely, introducing the OCT4 C-TAD into corresponding sites of other OCT factors endows almost all OCT factors competent in inducing pluripotency. In contrast to OCT4, the reprogramming competence of OCT6 arises from its N-TAD. The OCT4 C-TAD can be functionally replaced by the OCT6 N-TAD. For instance, the chimera O646, which does not

contain the OCT4 C-TAD, can still produce iPSC colonies as efficiently as OCT4, and its binding ability to the pluripotency gene enhancers is still comparable to that of OCT4. This is largely due to the presence of the OCT6 N-TAD in this chimera. Thus, our findings clearly pinpoint that TADs are critical elements that are essential to the reprogramming competence of OCT factors. Similarly, it has recently shown that TADs of other transcription factors like GATA3, EBF, and FOXA1 influence their pioneering function and reprogramming competence, which supports our conclusion (52–54).

The functional role of OCT6 appears to be similar in neural and epidermal lineages in both mice and humans (33–37). Therefore, the finding that OCT6 can induce pluripotency in humans but not in mice struck us with surprise (2–4). It has been shown that OCT6 and OCT4 exhibit different DNA-dependent binding propensities (3, 50, 51). While OCT4 preferentially forms heterodimers with SOX2 on the canonical SoxOct motif, OCT6 preferentially forms homodimers on the MORE motif (3, 50, 51). A point mutation (151S) within OCT6 DBD diminishes its preference for homodimerization through the MORE motif (3), but it fails to increase its binding propensity to the SoxOct element (3, 5). As such, this mutated OCT6 fails to elicit reprogramming in mice (3). Instead, altering additional sites, SOX2 interaction surface (7K,22T) or 7K,22T/OCT4 linker, within OCT6<sup>151S</sup> DBD enables iPSC generation in mice (3). However, the reprogramming efficiency with these OCT6 mutants was extremely low (>6 colonies out of  $2.5 \times 10^4$  starting cells). We tested the reprogramming capability of our OCT6-OCT4 chimeras in murine reprogramming. Most chimeras beside O466 could elicit reprogramming with variable efficiencies (fig. S7, A and B). Thus, these findings together with previous studies clearly underscore that the DNA-dependent binding propensity through DBD alone is not the sole barrier for OCT6-based reprogramming in mouse cells, but functional features of its TAD are also crucial for mouse reprogramming.

Chromatin configurations and epigenomes differ between human and mouse cells such that the accessibility of homologous transcription factors to their binding sites is also essentially different (5, 18, 19, 21, 55, 56). This difference likely accounts for the functional discrepancy of OCT6 between reprogramming mouse and human cells. Mouse cells might have higher epigenetic thresholds on its binding sites that are essential for reprogramming, and thus, wild-type OCT6 cannot access these sites where OCT4 can. Thereby, OCT6 additionally requires a strong TAD to dismantle these epigenetic barriers to elicit reprogramming in mice (fig. S7, A and B). Conceivably, a forced elimination of these epigenetic blockades through either genetic depletion or chemical inhibition might enhance intrinsic TAD features of OCT6 that enables iPSC generation in mice. In this context, it is tempting to speculate that acquiring an epigenetic state that is unfavorable for OCT6-dependent reprogramming may provide a specific means to inhibit transformation of neural and epidermal lineages where and when OCT6 is normally expressed. If true, it will be interesting to test whether reducing specific epigenetic barriers could elicit Oct6-dependent reprogramming in mouse cells.

## MATERIALS AND METHODS

### Cell culture

CRL-2097 fibroblasts were purchased from the American Type Culture Collection. Fibroblasts, MEFs, Platinum-E (PLAT-E) cells,

and human embryonic kidney (HEK) 293 cells were cultured in KnockOut Dulbecco's modified Eagle's medium (DMEM; Invitrogen) supplemented with 10% fetal bovine serum (FBS; Biochrom Ltd), 1× non-essential amino acids (NEAA) (Sigma-Aldrich), 1× GlutaMax (Invitrogen), and 1× penicillin/streptomycin (P/S; 100 U/ml each; Sigma-Aldrich). Human iPSCs and ESCs were cultured either on feeder layers with hESC medium or on Matrigel-coated plates with MEF-conditioned medium (MEF-CM). The hESC medium consisted of DMEM/F12 (Invitrogen) supplemented with 20% KSR; Invitrogen), 1× NEAA, 1× GlutaMax, 1× P/S (100 U/ml each), 100 μM 2-mercaptoethanol (Invitrogen), and fibroblast growth factor (5 ng/ml; Peprotech). H1 and HUES6 hESC lines were obtained from WiCell Research Institute Inc. Mouse iPSC cells and ESCs were cultured in mESC medium consisting of KnockOut DMEM supplemented with 5% FBS, 10% KSR, 1× NEAA, 1× GlutaMax, 1 mM sodium pyruvate (Sigma-Aldrich), 1× P/S (100 U/ml each), 100 μM 2-mercaptoethanol, and leukemia inhibitory factor (LIF; prepared in-house). Mouse EpiSCs were cultured in MEF-CM. NaB (Sigma-Aldrich) was used at a final concentration of 250 μM. SB431542 (Cayman Chemical) was used at a final concentration of 2 μM.

### Plasmid

The coding region of genes described in table S1 was cloned into pMXs or pMXs-gw. pMXs was a gift from T. Kitamura. pMXs-gw (#18656), pMXs-hOCT4 (#17217), pMXs-hSOX2 (#17218), pMXs-hKLF4 (#17219), pMXs-hc-MYC (#17220), and pLenti6/Ubc/mSlc7a1 (#17224) were obtained from Addgene. Site mutagenesis was performed as previously described (3). For the luciferase assay, each TAD of OCT factors was amplified by PCR and cloned into pPyCAG-G4DBD-IP. For instance, OCT4 N-TAD was amplified by PCR using the oligonucleotide primers OCT4 N-TA Bgl II F and OCT4 N-TA Spe I R and cloned into the Bgl II and Spe I sites of pPyCAG-G4DBD-IP. The OCT4 C-TAD was amplified by PCR using the oligonucleotide primers OCT4 C-TA Mlu I F and OCT4 C-TA Not I R and cloned into the Mlu I and Not I sites of pPyCAG-G4DBD-IP. By using above methods, all other luciferase constructs were cloned. To create domain-swapped chimeras, multiple PCRs were performed, and the resulting PCR products were cloned into pMXs. For instance, to create pMXs-O466, OCT4 N-TAD was first amplified by PCR using the oligonucleotide primers OCT4 Eco N-3×Flag F and OCT4N-OCT6M R. Then, O466 was amplified from the template, pMXs-hOCT6, by PCR using the above PCR product as a forward primer and OCT6 Spe I/Sal I R. Then, the final PCR product was digested with Eco RI and Sal I and ligated into the Eco RI and Xho I sites of pMXs. Using the above methods, all other domain-swapped chimeras were created. All constructs were verified extensively by restriction enzyme digestion and sequencing. Primers used for cloning are listed in table S3.

### Virus production

To produce the retrovirus, PLAT-E cells were transfected with 9 μg of the retroviral vector using 27 μl of FuGENE 6 transfection reagent (Promega) in 600 μl of Opti-MEM (Invitrogen) per 10-cm dish. Virus-containing supernatants were collected at 48 and 72 hours after transfection and filtered through a 0.4-μm polyvinylidene difluoride (PVDF) filter (Millipore). To produce the lentivirus, HEK293 cells were transfected with 3 μg of psPAX2 (Addgene), 1.5 μg of pMD2.G (Addgene), and 4.5 μg of lentiviral vector using 27 μl of FuGENE 6 transfection reagent in 600 μl of Opti-MEM. Virus-

containing supernatants were collected at 48 hours, filtered through a 0.4-μm PVDF filter, concentrated, and resuspended in KnockOut DMEM and stored at −80°C until use.

### Reprogramming

For the human reprogramming, fibroblasts were plated at a density of  $8 \times 10^5$  cells per 10-cm dish. On the next day, the cells were infected with the Slc7a1 lentivirus in the presence of protamine sulfate (8 μg/ml; Sigma-Aldrich) overnight. The cells were then washed three times with phosphate-buffered saline (PBS) and cultured with the medium containing blasticidin S (10 μg/ml; InvivoGen) to select the cells expressing Slc7a1. The cells expressing Slc7a1 were then plated at a density of  $1.3 \times 10^5$  cells per well of six-well plates. On the next day, the cells were infected with retroviruses containing each tested combination in the presence of protamine sulfate (8 μg/ml; Sigma-Aldrich) overnight. Then, the cells were washed three times with PBS and incubated with fresh medium overnight. The cells were then infected again overnight, washed three times with PBS, and incubated with fresh medium overnight. The cells were then dissociated with trypsin (Invitrogen) and plated at a density of  $2 \times 10^4$  to  $4 \times 10^4$  cells per well of six-well plates precoated with CF1 MEF feeder cells or Matrigel. On the next day, the cells were washed once with PBS and incubated in hESC medium or MEF-CM supplemented with 250 μM NaB. NaB was supplied to the medium for the first 10 days. For the mouse reprogramming, MEFs carrying green fluorescent protein under *Oct4* promoter (OG2 MEFs) were then plated at a density of  $5 \times 10^4$  cells per well of six-well plates. On the next day, the cells were infected with retroviruses containing each tested combination in the presence of protamine sulfate (8 μg/ml) overnight. Then, the cells were washed three times with PBS and incubated with fresh medium overnight. The cells were then infected again overnight, washed three times with PBS, and incubated with fresh medium overnight. The cells were then cultured in mESC medium. For mouse iEpiSC generation, the transduced MEFs were cultured in MEF-CM supplemented with LIF antibody (MAB449; R&D Systems).

### Immunofluorescence

The cells were fixed with 4% paraformaldehyde (PFA; Sigma-Aldrich) for 15 min, incubated with 0.1% Triton X-100/PBS for 15 min, and blocked in 5% BSA/PBS for 1 hour. The cells were then incubated with the following primary antibodies overnight at 4°C: mouse monoclonal anti-TRA-1-60 (1:100, MAB4360; Millipore), mouse monoclonal anti-TRA-1-81 (1:100, MAB4381; Millipore), mouse monoclonal anti-SSEA-4 (1:100, 330402; BioLegend), rabbit monoclonal anti-NANOG (1:1000, 5232; Cell Signaling Technology), rabbit monoclonal anti-OCT4 (1:1000, 5677; Cell Signaling Technology), rabbit monoclonal anti-SOX2 (1:1000, 5024; Cell Signaling Technology), rabbit monoclonal anti-DNMT3B (1:1000, 67259; Cell Signaling Technology), mouse monoclonal anti-FLAG (1:1000, F1804; Sigma-Aldrich), and goat polyclonal anti-GATA6 (1:500, AF1700; R&D Systems). The cells were then washed three times with PBS and incubated with appropriate fluorescently labeled Alexa Fluor secondary antibodies (1:1000; Invitrogen) for 1 hour. The cells were then washed three times with PBS, incubated with 4',6-diamidino-2-phenylindole (DAPI; 0.5 μg/ml; Molecular Probes) for 10 min, and washed once with PBS. Images were acquired using the Leica DMI6000B inverted fluorescence microscope equipped with the Hamamatsu ORCA-R2 charge-coupled device camera and analyzed with the Leica application suite advanced fluorescence software.

**DAB staining**

The cells were fixed with 4% PFA for 15 min, incubated with 0.1% Triton X-100/PBS for 15 min, and blocked with 5% BSA/PBS for 1 hour. The cells were incubated with mouse monoclonal anti-TRA-1-60-HRP (horseradish peroxidase) conjugated (1:1000, MA1-023-HRP; Thermo Fisher Scientific) overnight at 4°C. The cells were then washed three times with PBS, and TRA-1-60-positive cells were visualized using the 3,3'-diaminobenzidine (DAB) peroxidase staining kit (Vector Laboratories). The images were scanned using the Epson Perfection V370 photo scanner.

**Luciferase assay**

The luciferase constructs were transiently transfected into HeLa cells, along with pGL4.75[hRluc/CMV] and 5xUAS-luc2 constructs. At 24 hours after transfection, the luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega). The luciferase activity (luc/hRluc) of each construct was normalized to that of G4DBD.

**Flow cytometry**

Antibodies were diluted in 50  $\mu$ l of 3% FBS/PBS solution and incubated with the cells on ice for 15 min. The cells were then washed three times in 500  $\mu$ l of 3% FBS/PBS and used for flow cytometry analysis. The cells were separated from debris and aggregates by forward scatter/side scatter (FSC/SSC) gating. Single cells were identified by plotting FSC area versus FSC width. Dead cells were excluded by staining with DAPI and gating on DAPI-negative cells. Unstained cells and isotype controls were used as staining controls. hESC cells (TRA-1-60<sup>+</sup>, CD13<sup>-</sup>) and fibroblasts (TRA-1-60<sup>-</sup>, CD13<sup>+</sup>) served as biological controls for gating. Mouse monoclonal anti-human CD13 APC (1:50, 301706; BioLegend), mouse monoclonal anti-human TRA-1-60 PE (1:25, 330610; BioLegend), mouse monoclonal anti-mouse immunoglobulin G1 (IgG1) allophycocyanin (APC; 1:50, 400120; BioLegend), and mouse monoclonal anti-mouse IgM PE (1:25, 401609; BioLegend) were used for the analysis. Fluorescence was measured using a FACSAria IIu cell sorter (BD Biosciences). Flow cytometry data were processed using FlowJo software (Tree Star Inc.).

**Polymerase chain reaction**

Total RNA was isolated using the RNeasy kit (Qiagen). First-strand complementary DNA (cDNA) was synthesized using Oligo(dT)<sub>12-18</sub> and M-MLV Reverse Transcriptase (USB). qPCR was performed using iTaq Universal SYBR Green Supermix (Bio-Rad). Relative gene expression levels were calculated by the  $2^{-\Delta\Delta Ct}$  method, normalized to an endogenous control gene, and presented as fold change over control samples. For the ChIP-qPCR, the fold enrichment was calculated by the standard curve method and normalized to the value obtained at a negative control region. Primers are listed in table S3.

**Bisulfite sequencing**

Bisulfite conversion was performed using the EZ DNA methylation kit (Zymo Research). PCR was performed using HotStarTaq DNA Polymerase (Qiagen). The resulting PCR products were cloned into the pCRII TOPO vector (Invitrogen). Individual clones were sequenced with the M13 reverse primer. The resulting sequencing data were analyzed using the Quantification Tool for Methylation Analysis. Primers are listed in table S3.

**Microarray**

Total RNA (300 ng) was used as input for labeling. T7-linked double-stranded cDNA was synthesized, and in vitro transcription incorporating biotin-labeled nucleotides was performed using the Premier RNA Amplification Kit (Thermo Fisher Scientific). Purified and labeled cRNA was then hybridized onto MouseRef-8 v2 Expression BeadChips (Illumina). After washing, the chips were stained with streptavidin-Cy3 (GE Healthcare) and scanned using the iScan reader (Illumina) and its software. Bead intensities were mapped using BeadStudio 3.2 (Illumina). Background correction was performed using the Affymetrix Robust Multiarray Analysis background correction model. Variance stabilization was performed using the log<sub>2</sub> scaling, and gene expression normalization was calculated with the method implemented with the lumi package of R-Bioconductor. Data postprocessing and graphics were performed with in-house-developed functions. Hierarchical clustering was performed with one minus correlation metric and the unweighted average distance (also known as group average) linkage method. The data were deposited in the National Center for Biotechnology Information's (NCBI's) Gene Expression Omnibus with accession number GSE95608.

**Teratoma**

iPSCs were resuspended in 50  $\mu$ l of medium, mixed with 50  $\mu$ l of Matrigel (BD Biosciences), and injected into the hindlimb femoral muscle of immunodeficient SCID (severe combined immunodeficient) mice according to the approved institutional animal protocol. After 10 weeks, teratomas were harvested and fixed in the Bouin's solution overnight. The fixed teratomas were embedded in paraffin wax and serially sectioned using the microtome (Thermo Fisher Scientific). The sections were stained with hematoxylin and eosin using a standard protocol. Images were acquired using the Axio Imager M1 microscope (Zeiss) equipped with the Hamamatsu ORCA-ER digital camera and analyzed by the Volocity software (Improvision).

**Karyotyping**

The cells were cultured with 0.3  $\mu$ g/ml of KaryoMAX Colcemid solution (Invitrogen) for 3 hours. The cells were then washed once with PBS, dissociated with 0.25% trypsin, resuspended in 75 mM KCl solution, and incubated at room temperature for 7 min. The cells were then centrifuged and resuspended in ice-cold methanol:glacial acetic acid (3:1) fixative solution (3:1 methanol/acetic acid) with shaking of the cell suspension. The cells were then centrifuged, resuspended in the fresh fixative solution, incubated for 20 min at 4°C, and dropped onto glass slides (Menzel Gläser, Thermo Fisher Scientific). The chromosomes were GTG-banded using a standard procedure. Metaphase spreads were analyzed on the Zeiss AxioScop microscope. Ten metaphase spreads were analyzed for each sample using the CytoVision software (Applied Imaging Corporation).

**RNA sequencing**

CD13<sup>-</sup>/TRA-1-60<sup>+</sup> cells sorted by FACS at different time points were washed three times in 0.5% BSA/PBS. Then, 50 cells were transferred into the PCR tube containing 4.45  $\mu$ l of lysis buffer using a mouth pipette. Reverse transcription was performed directly on the cytoplasmic lysate. Terminal deoxynucleotidyl transferase (Thermo Fisher Scientific) was then used to add a poly(A) tail onto the 3' end of the first-strand cDNAs. The total cDNA library was then amplified by PCR (18 to 20 cycles). The amplified cDNA library was



fragmented using the Covaris S220 sonicator (Covaris). The sequence libraries were prepared using KAPA HyperPlus Library Preparation kit (KAPA Biosystems). Single-end 50–base pair (bp) sequencing was performed on HiSeq 2500 (Illumina) at BGI and Berry Genomics Corporation. RNA-seq sequencing reads were first mapped to hg19 reference genome using Tophat (v 2.1.1) with default parameters. Gene expression levels of each sample were quantified to fragments per kilobase million (FPKM) using Cufflinks (v 2.2.1) to eliminate the effects of sequencing depth and transcript length. Principal component analysis was implemented using R function `prcomp`. Differential expression analyses were conducted by R package `limma` (v 3.30.7). For each comparison, genes with a *P* value <0.01 and a mean fold change of >2 were considered to be differentially expressed. All analyses were done using customized R scripts. The data were deposited in NCBI's Gene Expression Omnibus with accession number GSE93706.

### Western blot

Total protein lysates were extracted in radioimmunoprecipitation assay (RIPA) buffer containing 1× Complete Protease inhibitor (Roche), quantified by the Bradford assay (Bio-Rad), separated by 5 to 15% SDS–polyacrylamide gel electrophoresis, and transferred onto nitrocellulose membranes (GE Healthcare). The membranes were blocked in 5% skim milk in tris-buffered saline with 0.1% Tween 20 (TBST) for 1 hour and incubated with the following primary antibodies overnight at 4°C: rabbit polyclonal anti-OCT4 (1:5000, ab19857; Abcam), mouse monoclonal anti-NR5A2 (1:1000, PP-H2325-00; R&D Systems), rabbit polyclonal anti-TET1 (1:1000, GTX125888; GeneTex), mouse monoclonal anti-GATA3 (1:5000, 653802; BioLegend), mouse monoclonal anti-OCT6 (1:5000, MABN738; Millipore), rabbit monoclonal anti-SOX2 (1:5000, 5024S; Cell Signaling Technology), goat polyclonal anti-KLF4 (1:1000, AF3640; R&D Systems), rabbit polyclonal anti-MYC (1:1000, sc-764; Santa Cruz), rabbit polyclonal anti-PAX6 (1:5000, 901301; BioLegend), rabbit polyclonal anti-DESMIN (1:1000, ab15200; Abcam), mouse monoclonal anti-FLAG (1:5000, F1804; Sigma-Aldrich), mouse monoclonal anti-G4DBD (1:1000, sc-510; Santa Cruz), rabbit monoclonal anti-DNMT1 (1:5000, 5032; Cell Signaling Technology), rabbit polyclonal anti-DNMT3A (1:5000, sc-20703; Santa Cruz), rabbit monoclonal anti-DNMT3B (1:5000, 67259; Cell Signaling Technology), and mouse monoclonal anti-TUBULIN (1:5000, T6199; Sigma-Aldrich). The membranes were then washed three times with TBST and incubated with the following horseradish peroxidase–conjugated secondary antibodies for 1 hour: chicken polyclonal anti-goat IgG HRP (1:20,000, HAF019; R&D Systems), donkey polyclonal anti-rabbit IgG HRP (1:20,000, NA934; GE Healthcare), and goat polyclonal anti-mouse IgG + IgM (H + L) HRP (1:20,000, 115-035-044; Dianova). The membranes were washed three times with TBST, incubated with enhanced chemiluminescence (ECL) solution (GE Healthcare), and exposed to x-ray films (GE Healthcare).

### Chromatin immunoprecipitation sequencing

The cells were cross-linked with 1% formaldehyde for 10 min at room temperature and quenched with 0.125 M glycine for 5 min at room temperature. The cross-linked cells were incubated with lysis buffer 1 [50 mM Hepes-KOH (pH 7.5), 140 mM NaCl, 1 mM 0.5 M EDTA, 10% glycerol, 0.5% IGEPAL CA630, 0.25% Triton X-100, and 1× Complete Protease inhibitor] for 30 min at 4°C and washed in lysis buffer 2 [10 mM tris-HCl (pH 8.0), 200 mM NaCl, 1 mM

EDTA, 0.5 mM EGTA, and 1× Complete Protease inhibitor] for 10 min at 4°C. The cells were then resuspended in sonication buffer [50 mM tris-HCl (pH 8.0), 10 mM EDTA, 0.5% SDS, and 1× Complete Protease inhibitor] and sonicated using the Diagenode Bioruptor (high power, 30 cycles of 30 s on and 30 s off). Sonicated chromatin (100 μg) was then incubated with Dynabeads Protein G (Invitrogen) coupled to 15 μg of mouse monoclonal anti-FLAG (F1804; Sigma-Aldrich) in 4× volume of ChIP dilution buffer [10 mM tris-HCl (pH 8.0), 125 mM NaCl, 0.125% sodium deoxycholate, 1.25% Triton X-100, and 1× Complete Protease inhibitor] at 4°C overnight. Beads were washed once with low-salt buffer [20 mM, tris-HCl (pH 8.0), 150 mM NaCl, 2 mM EDTA, 0.1% SDS, and 1% Triton X-100], twice with high-salt buffer [20 mM, tris-HCl (pH 8.0), 500 mM NaCl, 2 mM EDTA, 0.1% SDS, and 1% Triton X-100], twice with RIPA buffer [50 mM, Hepes-KOH (pH 7.6), 250 mM LiCl, 1 mM EDTA, 1% IGEPAL CA630, and 0.7% sodium deoxycholate], and once with TE buffer [1 mM EDTA, 10 mM tris-HCl (pH 8.0)] containing 50 mM NaCl. Elution was performed with elution buffer [10 mM tris-HCl (pH 8.0), 300 mM NaCl, 5 mM EDTA, and 0.5% SDS] for 15 min at 65°C. DNA was extracted by reverse cross-linking at 65°C overnight with proteinase K (20 μg/μl) and RNase A (20 μg/μl) and purified using phenol/chloroform extraction. DNA concentration was measured by Qubit (Thermo Fisher Scientific), and sequencing libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit (NEB). Amplified libraries were sequenced on NextSeq 500 (Illumina) as 75–bp pair-end reads. Sequencing reads were aligned to hg19 reference genome using bowtie2 (v 2.2.9) with default parameters. Input DNA from each sample was used as controls. The peak calling was performed using MACS14 (v 1.4.2) with parameter: `macs14 -t Sample.bam -c input.bam -g hs -p 1e-7`. The peak annotation, overlap of the nearest genes, peak distribution, and Kyoto Encyclopedia of Genes and Genomes pathway enrichment were analyzed by ChIPseeker (57). The promoter region was defined as –2 to +2 kb of transcription start site (TSS). The enhancer region was defined using p300 hESC ChIP-seq (GSE24447) without TSS –3 to +3 kb. The peak-covered enhancer was defined as enhancer overlapped with at least 1 bp of a peak calculated by bedtools. Motif finding is calculated by MEME Suite (v 4.11.2) with the range of 25 bp from the peak summit site. The data were deposited in NCBI's Gene Expression Omnibus with accession number GSE93706.

### Statistical analysis

The statistical differences between two groups were analyzed by two-tailed Student's *t* tests (ns *P* > 0.05, \**P* < 0.05, \*\**P* < 0.01, and \*\*\**P* < 0.001). All data are presented as means ± SD of at least triplicates.

### SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at <http://advances.sciencemag.org/cgi/content/full/6/36/eaaz7364/DC1>

[View/request a protocol for this paper from Bio-protocol.](#)

### REFERENCES AND NOTES

1. K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, S. Yamanaka, Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
2. M. Nakagawa, M. Koyanagi, K. Tanabe, K. Takahashi, T. Ichisaka, T. Aoi, K. Okita, Y. Mochizuki, N. Takizawa, S. Yamanaka, Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat. Biotechnol.* **26**, 101–106 (2008).

3. S. Jerabek, C. K. Ng, G. Wu, M. J. Arauzo-Bravo, K.-P. Kim, D. Esch, V. Malik, Y. Chen, S. Velychko, C. M. MacCarthy, X. Yang, V. Cojocar, H. R. Schöler, R. Jauch, Changing POU dimerization preferences converts Oct6 into a pluripotency inducer. *EMBO Rep.* **18**, 319–333 (2016).
4. M. Maekawa, K. Yamaguchi, T. Nakamura, R. Shibukawa, I. Kodanaka, T. Ichisaka, Y. Kawamura, H. Mochizuki, N. Goshima, S. Yamanaka, Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225–229 (2011).
5. V. Malik, L. V. Glaser, D. Zimmer, S. Velychko, M. Weng, M. Holzner, M. Arend, Y. Chen, Y. Srivastava, V. Veerapandian, Z. Shah, M. A. Esteban, H. Wang, J. Chen, H. R. Schöler, A. P. Hutchins, S. H. Meising, S. Pott, R. Jauch, Pluripotency reprogramming by competent and incompetent POU factors uncovers temporal dependency for Oct4 and Sox2. *Nat. Commun.* **10**, 3477 (2019).
6. Y. Buganim, D. A. Faddah, A. W. Cheng, E. Itskovich, S. Markoulaki, K. Ganz, S. L. Klemm, A. van Oudenaarden, R. Jaenisch, Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* **150**, 1209–1222 (2012).
7. Y. Gao, J. Chen, K. Li, T. Wu, B. Huang, W. Liu, X. Kou, Y. Zhang, H. Huang, Y. Jiang, C. Yao, X. Liu, Z. Lu, Z. Xu, L. Kang, J. Chen, H. Wang, T. Cai, S. Gao, Replacement of Oct4 by Tet1 during iPSC induction reveals an important role of DNA methylation and hydroxymethylation in reprogramming. *Cell Stem Cell* **12**, 453–469 (2013).
8. J.-C. Heng, B. Feng, J. Han, J. Jiang, P. Kraus, J.-H. Ng, Y. L. Orlov, M. Huss, L. Yang, T. Lufkin, B. Lim, H.-H. Ng, The nuclear receptor Nr5a2 can replace Oct4 in the reprogramming of murine somatic cells to pluripotent cells. *Cell Stem Cell* **6**, 167–174 (2010).
9. T. Mai, G. J. Markov, J. J. Brady, A. Palla, H. Zeng, V. Sebastiano, H. M. Blau, NKX3-1 is required for induced pluripotent stem cell reprogramming and can replace OCT4 in mouse and human iPSC induction. *Nat. Cell Biol.* **20**, 900–908 (2018).
10. J. Shu, C. Wu, Y. Wu, Z. Li, S. Shao, W. Zhao, X. Tang, H. Yang, L. Shen, X. Zuo, W. Yang, Y. Shi, X. Chi, H. Zhang, G. Gao, Y. Shu, K. Yuan, W. He, C. Tang, Y. Zhao, H. Deng, Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* **153**, 963–975 (2013).
11. J. Shu, K. Zhang, M. Zhang, A. Yao, S. Shao, F. Du, C. Yang, W. Chen, C. Wu, W. Yang, Y. Sun, H. Deng, GATA family members as inducers for cellular reprogramming to pluripotency. *Cell Res.* **25**, 169–180 (2015).
12. N. Montserrat, E. Nivet, I. Sancho-Martinez, T. Hishida, S. Kumar, L. Miquel, C. Cortina, Y. Hishida, Y. Xia, C. R. Esteban, J. C. I. Belmonte, Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* **13**, 341–350 (2013).
13. X. He, M. N. Treacy, D. M. Simmons, H. A. Ingraham, L. W. Swanson, M. G. Rosenfeld, Expression of a large family of POU-domain regulatory genes in mammalian brain development. *Nature* **340**, 35–41 (1989).
14. M. G. Rosenfeld, POU-domain transcription factors: pou-er-ful developmental regulators. *Genes Dev.* **5**, 897–907 (1991).
15. H. R. Schöler, A. K. Hatzopoulos, R. Balling, N. Suzuki, P. Gruss, A family of octamer-specific proteins present during mouse embryogenesis: Evidence for germline-specific expression of an Oct factor. *EMBO J.* **8**, 2543–2550 (1989).
16. D. Esch, J. Vahokoski, M. R. Groves, V. Pogenberg, V. Cojocar, H. V. Bruch, D. Han, H. C. A. Drexler, M. J. Arauzo-Bravo, C. K. L. Ng, R. Jauch, M. Wilmanns, H. R. Schöler, A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nat. Cell Biol.* **15**, 295–301 (2013).
17. W. Jin, L. Wang, F. Zhu, W. Tan, W. Lin, D. Chen, Q. Sun, Z. Xia, Critical POU domain residues confer Oct4 uniqueness in somatic cell reprogramming. *Sci. Rep.* **6**, 20818 (2016).
18. A. Soufi, G. Donahue, K. S. Zaret, Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994–1004 (2012).
19. A. Soufi, M. F. Garcia, A. Jaroszewicz, N. Osman, M. Pellegrini, K. S. Zaret, Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).
20. D. Cacchiarelli, C. Trapnell, M. J. Ziller, M. Soumillon, M. Cesana, R. Karnik, J. Donaghey, Z. D. Smith, S. Ratanasirintrao, X. Zhang, S. J. H. Sui, Z. Wu, V. Akopian, C. A. Gifford, J. Doench, J. L. Rinn, G. Q. Daley, A. Meissner, E. S. Lander, T. S. Mikkelsen, Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. *Cell* **162**, 412–424 (2015).
21. C. Chronis, P. Fiziev, B. Papp, S. Butz, G. Bonora, S. Sabri, J. Ernst, K. Plath, Cooperative binding of transcription factors orchestrates reprogramming. *Cell* **168**, 442–459.e20 (2017).
22. J. M. Polo, E. Andersson, R. M. Walsh, B. A. Schwarz, C. M. Nefzger, S. M. Lim, M. Borkent, E. Apostolou, S. Alaei, J. Cloutier, O. Bar-Nur, S. Cheloufi, M. Stadtfeld, M. E. Figueroa, D. Robinton, S. Natesan, A. Melnick, J. Zhu, S. Ramaswamy, K. Hochedlinger, A molecular roadmap of reprogramming somatic cells into iPSCs. *Cell* **151**, 1617–1632 (2012).
23. K. Chen, Q. Long, T. Wang, D. Zhao, Y. Zhou, J. Qi, Y. Wu, S. Li, C. Chen, X. Zeng, J. Yang, Z. Zhou, W. Qin, X. Liu, Y. Li, Y. Li, X. Huang, D. Qin, J. Chen, G. Pan, H. R. Schöler, G. Xu, X. Liu, D. Pei, Gadd45a is a heterochromatin relaxer that enhances iPSC cell generation. *EMBO Rep.* **17**, 1641–1656 (2016).
24. R. L. dos Santos, L. Tosti, A. Radziszheuskaya, I. M. Caballero, K. Kaji, B. Hendrich, J. C. R. Silva, MBD3/NuRD facilitates induction of pluripotency in a context-dependent manner. *Cell Stem Cell* **15**, 102–110 (2014).
25. X. Hu, S.-Q. Mao, Z. Li, J. Chen, R.-R. Zhang, H.-P. Wu, J. Gao, F. Guo, W. Liu, G.-F. Xu, H.-Q. Dai, Y. G. Shi, X. Li, B. Hu, F. Tang, D. Pei, G.-L. Xu, Tet and TDG mediate DNA demethylation essential for mesenchymal-to-epithelial transition in somatic cell reprogramming. *Cell Stem Cell* **14**, 512–522 (2014).
26. Y. Rais, A. Zviran, S. Geula, O. Gafni, E. Chomsky, S. Viukov, A. A. Mansour, I. Caspi, V. Krupalnik, M. Zerbib, I. Maza, N. Mor, D. Baran, L. Weinberger, D. A. Jaitin, D. Lara-Astiaso, R. Blecher-Gonen, Z. Shipony, Z. Mukamel, T. Hagai, S. Gilad, D. Amann-Zalcenstein, A. Tanay, I. Amit, N. Novershtern, J. H. Hanna, Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65–70 (2013).
27. S.-J. Dunn, G. Martello, B. Yordanov, S. Emmott, A. G. Smith, Defining an essential transcription factor program for naive pluripotency. *Science* **344**, 1156–1160 (2014).
28. Y. Takashima, G. Guo, R. Loos, J. Nichols, G. Ficz, F. Krueger, D. Oxley, F. Santos, J. Clarke, W. Mansfield, W. Reik, P. Bertone, A. Smith, Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254–1269 (2014).
29. T. W. Theunissen, B. E. Powell, H. Wang, M. Mitalipova, D. A. Faddah, J. Reddy, Z. P. Fan, D. Maetzel, K. Ganz, L. Shi, T. Lungjangwa, S. Imsoonthornruksa, Y. Stelzer, S. Rangarajan, A. D'Alessio, J. Zhang, Q. Gao, M. M. Dawlaty, R. A. Young, N. S. Gray, R. Jaenisch, Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell* **15**, 471–487 (2014).
30. N. Mise, T. Fuchikami, M. Sugimoto, S. Kobayakawa, F. Ike, T. Ogawa, T. Tada, S. Kanaya, T. Noce, K. Abe, Differences and similarities in the developmental status of embryo-derived stem cells and primordial germ cells revealed by global expression profiling. *Genes Cells* **13**, 863–877 (2008).
31. D. Sabour, M. J. Arauzo-Bravo, K. Hübner, K. Ko, B. Greber, L. Gentile, M. Stehling, H. R. Schöler, Identification of genes specific to mouse primordial germ cells through dynamic global gene expression. *Hum. Mol. Genet.* **20**, 115–125 (2011).
32. F. Sugawa, M. J. Arauzo-Bravo, J. Yoon, K.-P. Kim, S. Aramaki, G. Wu, M. Stehling, O. E. Psathaki, K. Hübner, H. R. Schöler, Human primordial germ cell commitment in vitro associates with a unique PRDM14 expression profile. *EMBO J.* **34**, 1009–1024 (2015).
33. I. Faus, H. J. Hsu, E. Fuchs, Oct-6: A regulator of keratinocyte gene expression in stratified squamous epithelia. *Mol. Cell Biol.* **14**, 3263–3275 (1994).
34. R. P. Friedrich, B. Schlierf, E. R. Tamm, M. R. Bösl, M. Wegner, The class III POU domain protein Brn-1 can fully replace the related Oct-6 during schwann cell development and myelination. *Mol. Cell Biol.* **25**, 1821–1829 (2005).
35. M. Jaegle, M. Ghazvini, W. Mandemakers, M. Piirsoo, S. Driegen, F. Levavasseur, S. Raghoenath, F. Grosveld, D. Meijer, The POU proteins Brn-2 and Oct-6 share important functions in Schwann cell development. *Genes Dev.* **17**, 1380–1391 (2003).
36. M. Jaegle, W. Mandemakers, L. Broos, R. Zwart, A. Karis, P. Visser, F. Grosveld, D. Meijer, The POU factor Oct-6 and Schwann cell differentiation. *Science* **273**, 507–510 (1996).
37. J. Schreiber, J. Enderich, E. Sock, C. Schmidt, C. Richter-Landsberg, M. Wegner, Redundancy of class III POU proteins in the oligodendrocyte lineage. *J. Biol. Chem.* **272**, 32286–32293 (1997).
38. N. Suzuki, H. Rohdewohld, T. Neuman, P. Gruss, H. R. Schöler, Oct-6: A POU transcription factor expressed in embryonic stem cells and in the developing brain. *EMBO J.* **9**, 3723–3732 (1990).
39. T. Aasen, A. Raya, M. J. Barrero, E. Garreta, A. Consiglio, F. Gonzalez, R. Vassena, J. Bilić, V. Pekarik, G. Tiscornia, M. Edel, S. Boué, J. C. I. Belmonte, Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* **26**, 1276–1284 (2008).
40. A. D. Panopoulos, S. Ruiz, F. Yi, A. Herrerías, E. M. Batchelder, J. C. I. Belmonte, Rapid and highly efficient generation of induced pluripotent stem cells from human umbilical vein endothelial cells. *PLOS ONE* **6**, e19743 (2011).
41. P. Mali, B.-K. Chou, J. Yen, Z. Ye, J. Zou, S. Dowey, R. A. Brodsky, J. E. Ohm, W. Yu, S. B. Baylin, K. Yusa, A. Bradley, D. J. Meyers, C. Mukherjee, P. A. Cole, L. Cheng, Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells* **28**, 713–720 (2010).
42. Z. Zhang, D. Xiang, W.-S. Wu, Sodium butyrate facilitates reprogramming by derepressing OCT4 transactivity at the promoter of embryonic stem cell-specific miR-302/367 cluster. *Cell. Reprogram.* **16**, 130–139 (2014).
43. L. M. Julian, A. C. McDonald, W. L. Stanford, Direct reprogramming with SOX factors: Masters of cell fate. *Curr. Opin. Genet. Dev.* **46**, 24–36 (2017).
44. A. Reményi, K. Lins, L. J. Nissen, R. Reinbold, H. R. Schöler, M. Wilmanns, Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.* **17**, 2048–2059 (2003).
45. D. Acampora, D. Omodei, G. Petrosino, A. Garofalo, M. Savarese, V. Nigro, L. G. Di Giovannantonio, V. Mercadante, A. Simeone, Loss of the Otx2-Binding site in the nanog promoter affects the integrity of embryonic stem cell subtypes and specification of inner cell mass-derived epiblast. *Cell Rep.* **15**, 2651–2664 (2016).
46. D. Kurek, A. Neagu, M. Tastemel, N. Tüysüz, J. Lehmann, H. J. G. van de Werken, S. Philipsen, R. van der Linden, A. Maas, W. F. J. van Ucken, M. Drukker, D. T. Berge, Endogenous WNT signals mediate BMP-induced and spontaneous differentiation

- of epiblast stem cells and human embryonic stem cells. *Stem Cell Rep.* **4**, 114–128 (2015).
47. R. Li, J. Liang, S. Ni, T. Zhou, X. Qing, H. Li, W. He, J. Chen, F. Li, Q. Zhuang, B. Qin, J. Xu, W. Li, J. Yang, Y. Gan, D. Qin, S. Feng, H. Song, D. Yang, B. Zhang, L. Zeng, L. Lai, M. A. Esteban, D. Pei, A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell* **7**, 51–63 (2010).
  48. P. Samavarchi-Tehrani, A. Golipour, L. David, H.-K. Sung, T. A. Beyer, A. Datti, K. Woltjen, A. Nagy, J. L. Wrana, Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell* **7**, 64–77 (2010).
  49. Y. Buganim, D. A. Faddah, R. Jaenisch, Mechanisms and models of somatic cell reprogramming. *Nat. Rev. Genet.* **14**, 427–439 (2013).
  50. Y. K. Chang, Y. Srivastava, C. Hu, A. Joyce, X. Yang, Z. Zuo, J. J. Havranek, G. D. Stormo, R. Jauch, Quantitative profiling of selective Sox/POU pairing on hundreds of sequences in parallel by Coop-seq. *Nucleic Acids Res.* **45**, 832–845 (2017).
  51. T. K. Mistri, A. G. Devasia, L. T. Chu, W. P. Ng, F. Halbritter, D. Colby, B. Martynoga, S. R. Tomlinson, I. Chambers, P. Robson, T. Wohland, Selective influence of Sox2 on POU transcription factor binding in embryonic and neural stem cells. *EMBO Rep.* **16**, 1177–1191 (2015).
  52. S. Boller, S. Ramamoorthy, D. Akbas, R. Nechanitzky, L. Burger, R. Murr, D. Schübeler, R. Grosschedl, Pioneering activity of the C-terminal domain of EBF1 shapes the chromatin landscape for B cell programming. *Immunity* **44**, 527–541 (2016).
  53. L. A. Cirillo, F. R. Lin, I. Cuesta, D. Friedman, M. Jarnik, K. S. Zaret, Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9**, 279–289 (2002).
  54. M. Takaku, S. A. Grimm, T. Shimbo, L. Perera, R. Menafrá, H. G. Stunnenberg, T. K. Archer, S. Machida, H. Kurumizaka, P. A. Wade, GATA3-dependent cellular reprogramming requires activation-domain dependent recruitment of a chromatin remodeler. *Genome Biol.* **17**, 36 (2016).
  55. S. Cao, S. Yu, D. Li, J. Ye, X. Yang, C. Li, X. Wang, Y. Mai, Y. Qin, J. Wu, J. He, C. Zhou, H. Liu, B. Zhao, X. Shu, C. Wu, R. Chen, W. Chan, G. Pan, J. Chen, J. Liu, D. Pei, Chromatin accessibility dynamics during chemical induction of pluripotency. *Cell Stem Cell* **22**, 529–542.e5 (2018).
  56. H. Guo, B. Hu, L. Yan, J. Yong, Y. Wu, Y. Gao, F. Guo, Y. Hou, X. Fan, J. Dong, X. Wang, X. Zhu, J. Yan, Y. Wei, H. Jin, W. Zhang, L. Wen, F. Tang, J. Qiao, DNA methylation and chromatin accessibility profiling of mouse and human fetal germ cells. *Cell Res.* **27**, 165–183 (2017).
  57. G. Yu, L.-G. Wang, Q.-Y. He, ChIPseeker: An R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).

**Acknowledgments:** We thank B. Greber, T. Cantz, H. Zaehres, and V. Cojocaru for discussions and sharing materials. **Funding:** This work was supported by grants from the Max Planck Society. **Author contributions:** K.-P.K. conceived the study, performed most of the experiments, interpreted the results, and wrote the manuscript. Y.W., Y.G., M.J.A.-B., and S.G. contributed to the RNA-seq, ChIP-seq, and microarray. J.Y., C.M.M., S.V., D.W.H., and B.S. contributed to the plasmid construction and Western blots. G.W. contributed to the teratoma assay. K.A. contributed to the luciferase assay. A.R. contributed to the karyotyping. M.S. contributed to the flow cytometry. J.K. interpreted the results and wrote the manuscript. H.R.S. supervised the study, interpreted the results, and wrote the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Microarray, RNA-seq, and ChIP-seq data have been deposited in GEO under accession codes GSE93706 and GSE95608 (secure token: orsfocwkjdydvj). All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 4 October 2019

Accepted 20 July 2020

Published 2 September 2020

10.1126/sciadv.aaz7364

**Citation:** K.-P. Kim, Y. Wu, J. Yoon, K. Adachi, G. Wu, S. Velychko, C. M. MacCarthy, B. Shin, A. Röpke, M. J. Arauzo-Bravo, M. Stehling, D. W. Han, Y. Gao, J. Kim, S. Gao, H. R. Schöler, Reprogramming competence of OCT factors is determined by transactivation domains. *Sci. Adv.* **6**, eaaz7364 (2020).