npg

ORIGINAL ARTICLE

Pluripotent stem cells induced from mouse neural stem cells and small intestinal epithelial cells by small molecule compounds

Junqing Ye[1, 2, *], Jian Ge[1, 2, *], Xu Zhang[1, 2], Lin Cheng[1, 2], Zhengyuan Zhang[1, 2], Shan He [1, 2], Yuping Wang[4], Hua Lin[4], Weifeng Yang[5], Junfang Liu[1, 2], Yang Zhao[1, 2, 3], Hongkui Deng[1, 2]

[1]Shenzhen Stem Cell Engineering Laboratory, Key Laboratory of Chemical Genomics, Peking University Shenzhen Graduate School, Shenzhen, Guangdong 518055, China; [2]The MOE Key Laboratory of Cell Proliferation and Differentiation, College of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China; [3]Department of Cell Biology, School of Basic Medical Sciences, Peking University Stem Cell Research Center, State Key Laboratory of Natural and Biomimetic Drugs, Peking University Health Science Center, Beijing 100191, China; [4]Department of Gynecology and Obstetrics, China-Japan Friendship Hospital, Beijing 100029, China; [5]BeijingVitalstar Biotechnology Co., Ltd., Beijing 100012, China

Recently, we reported a chemical approach to generate pluripotent stem cells from mouse fibroblasts. However, whether chemically induced pluripotent stem cells (CiPSCs) can be derived from other cell types remains to be demonstrated. Here, using lineage tracing, we first verify the generation of CiPSCs from fibroblasts. Next, we demonstrate that neural stem cells (NSCs) from the ectoderm and small intestinal epithelial cells (IECs) from the endoderm can be chemically reprogrammed into pluripotent stem cells. CiPSCs derived from NSCs and IECs resemble mouse embryonic stem cells in proliferation rate, global gene expression profile, epigenetic status, self-renewal and differentiation capacity, and germline transmission competency. Interestingly, the pluripotency gene *Sall4* is expressed at the initial stage in the chemical reprogramming process from different cell types, and the same core small molecules are required for the reprogramming, suggesting conservation in the molecular mechanism underlying chemical reprogramming from these diverse cell types. Our analysis also shows that the use of these small molecules should be fine-tuned to meet the requirement of reprogramming from different cell types. Together, these findings demonstrate that full chemical reprogramming approach can be applied in cells of different tissue origins and suggest that chemical reprogramming is a promising strategy with the potential to be extended to more initial types.
Keywords: reprogramming; CiPSC; cell type; fine-tuning

Introduction

Somatic cells can be reprogrammed into pluripotent stem cells by ectopic expression of defined transcription factors [1-3]. Induced pluripotent stem cells (iPSCs) are commonly generated by *Oct4*, *Sox2*, *Klf4* and *c-Myc* (OSKM) overexpression. Although several studies have

proposed that iPSCs are predominantly derived from a rare cell population [4-6], numerous advances have been made in extending the application of iPS technology to other defined cell types, including keratinocytes [7], neural stem cells [8-10], hepatocytes [11] and hematocytes [12-15].

Recently, we have developed a novel approach to reprogram mouse somatic cells into pluripotent stem cells using only small molecules [16]. This chemical approach is different from the traditional genetic approach, as it offers more flexible control of the complex signaling networks and epigenetic status in the cells during reprogramming [17]. Small molecules are cell permeable; their effect is functionally reversible, and the treatment-

can be easily standardized. Thus, precise cell fate control can be rendered by manipulating the concentrations, treatment duration, and combinations of chemical compounds. Because of these advantages, the chemically induced pluripotent stem cell (CiPSC) approach provides a unique way to generate iPSCs and is a promising strategy for widespread use in the future.

In the previous report, we generated CiPSCs from fibroblasts, including mouse embryonic fibroblasts (MEFs), neonatal skin fibroblasts and adult lung fibroblasts [16]. As fibroblasts are a heterogeneous population of cells [18], and the reprogramming efficiency of our approach was relatively low [16], the possibility remains that there exists within fibroblast culture a specific cell subpopulation with propensity for reprogramming. Moreover, MEFs, neonatal skin fibroblasts and adult lung fibroblasts are cells from the mesoderm. It is unknown whether other cell types, including cells derived from ectoderm and endoderm, can also be reprogrammed by entirely chemical conditions. It has been reported that different cell types require modulations of different signaling pathways to achieve efficient reprogramming, possibly due to their intrinsic properties [19]. It is thus interesting to explore if fundamentally different small molecules are required to reprogram different cell types, which is a key question for the extension of the chemical reprogramming approach.

In this study, we tested the chemical reprogramming strategy on different initial cell types. We show that neural stem cells (NSCs) from the ectoderm and small intestinal epithelial cells (IECs) from the endoderm can also be reprogrammed by small molecules. Moreover, the same core small molecules can be applied for chemical reprogramming of different cell types, and fine-tuning of these small molecules is critical to achieve pluripotency.

Results

Lineage-tracing verifies generation of CiPSCs from fibroblasts

To determine whether CiPSCs are derived from a specific cell type within fibroblast culture, we performed *fibroblast specific protein 1 (Fsp1)* tracing during the chemical reprogramming process. We crossed mice carrying an *Fsp1*-Cre (*Fsp1* promoter-driven CRE expression) transgene with mice carrying a conditional Rosa26[LoxP] reporter locus containing a "floxed stop cassette" located in front of the *Tomato* gene (Rosa26R[tdTomato]). The offspring were crossed with *Oct4*-GFP (*Oct4* promoter-driven GFP expression, OG) transgenic mice to create *Fsp1*-Cre: Rosa26R[tdTomato]: *Oct4*-GFP progeny mice. As *Fsp1* is specifically expressed in fibroblasts, the initial

MEFs were separated into two subpopulations of tdTomato-positive and non-fluorescent cells. After treatment with the chemical cocktail as we previously reported [16], both tdTomato-positive and non-fluorescent CiPSCs were generated, suggesting that both fibroblasts and non-fibroblast cells can yield CiPSCs (Figure 1A-1D). Chimeric mice were obtained from fibroblast-derived CiPSCs (Figure 1E).

Generation of CiPSCs from NSCs

To test whether chemical reprogramming is applicable to cell types from other germ layers, we applied our chemical cocktail to NSCs from the ectodermal lineage, which exhibit distinct gene expression profiles from MEFs. NSCs were isolated from embryonic mouse brains of *Oct4*-GFP transgenic mice using the standard protocol and then cultured in standard NSC medium with bFGF and EGF in low attachment plates [20-22]. Established NSC lines could self-renew and form neurospheres for up to 8 passages *in vitro* (4-5 days/ passage) (Figure 2A). Cultured NSCs expressed neural stem cell markers SOX2 and NESTIN at high levels (Figure 2B). Gene ontology analysis showed that the most highly expressed genes in cultured NSCs were enriched in neural lineages (Supplementary information, Figure S1A, Table S1) [23].

To induce pluripotency from NSCs, we initially treated NSCs by using the reported chemical protocol for generation of MEF-derived CiPSCs. As the original chemical reprogramming cocktail "VC6TF" (V, VPA; C, CHIR99021 or CHIR; 6, 616452; T, tranylcypromine; F, forskolin) [16] was unable to induce pluripotency from NSCs, we attempted to boost the reprograming by adding EPZ 004777 (EPZ, E), an inhibitor of H3K79 histone methyltransferase DOT1L [24], and Ch 55 (5), a retinoid acid receptor (RAR) agonist (Supplementary information, Figure S1B). Although neither epithelioid colonies nor CiPSC colonies were generated with the new cocktail "VC6TFE5" with the same drug concentrations as we reported for MEF, we found, surprisingly, that after reducing the concentration of 616452 from 10 μM to 2 μM, epithelial clusters emerged on day 8 (Figure 2C), and proliferated with time. After the addition of DZNep (Z), an S-adenosylhomocysteine (SAH) hydrolase inhibitor, from day 20, compact and epithelioid colonies were observed on day 32 (Figure 2C). The *Oct4*-GFP reporter was gradually turned on after the colonies were switched to 2i-medium with dual inhibition (2i) of mitogen-activated protein kinase (MAPK) signaling and glycogen synthase kinase-3 (GSK3) [25] from day 40, and mouse embryonic stem cell (mESC)-like OG-positive colonies with clear-cut edges were observed (Figure 2D and 2I).
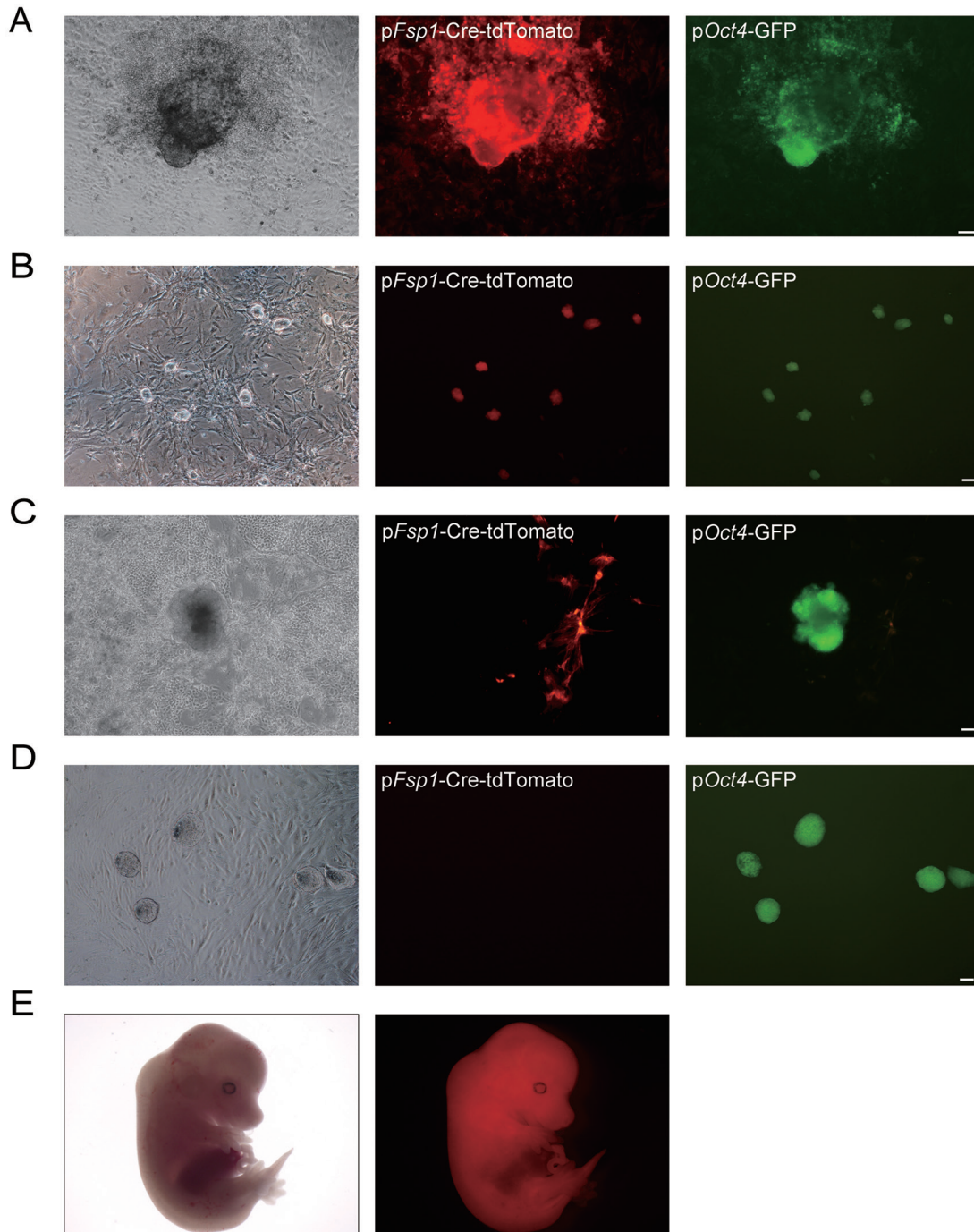
Figure 1 Generation of CiPSCs from both fibroblasts and non-fibroblast cells. **(A**, **B)** A primary CiPSC colony generated from *Fsp1*-Cre-tdTomato fluorescent MEFs **(A)**, and passaged CiPSCs **(B)**. **(C**, **D)** A primary CiPSC colony generated from *Fsp1*-Cre-tdTomato non-fluorescent MEFs **(C)**, and passaged CiPSCs **(D)**. **(E)** A chimeric mouse embryo (E13.5) derived from *Fsp1*-Cre-tdTomato fluorescent CiPSCs. For **A**-**D**, scale bar, 100 μm.

These cells are thus referred to as NSC-CiPSCs. Using the same small molecule combination, we also derived CiPSCs from NSCs of postnatal mice (Supplementary information, Figure S1E and S1F). Primary NSC-CiPSC colonies passaged in 2i-medium with mouse leukemia inhibitory factor (LIF) for more than 20 passages maintained stable ESC-like morphology and *Oct4*-GFP expression (Figure 2E). To confirm that the CiPSCs were
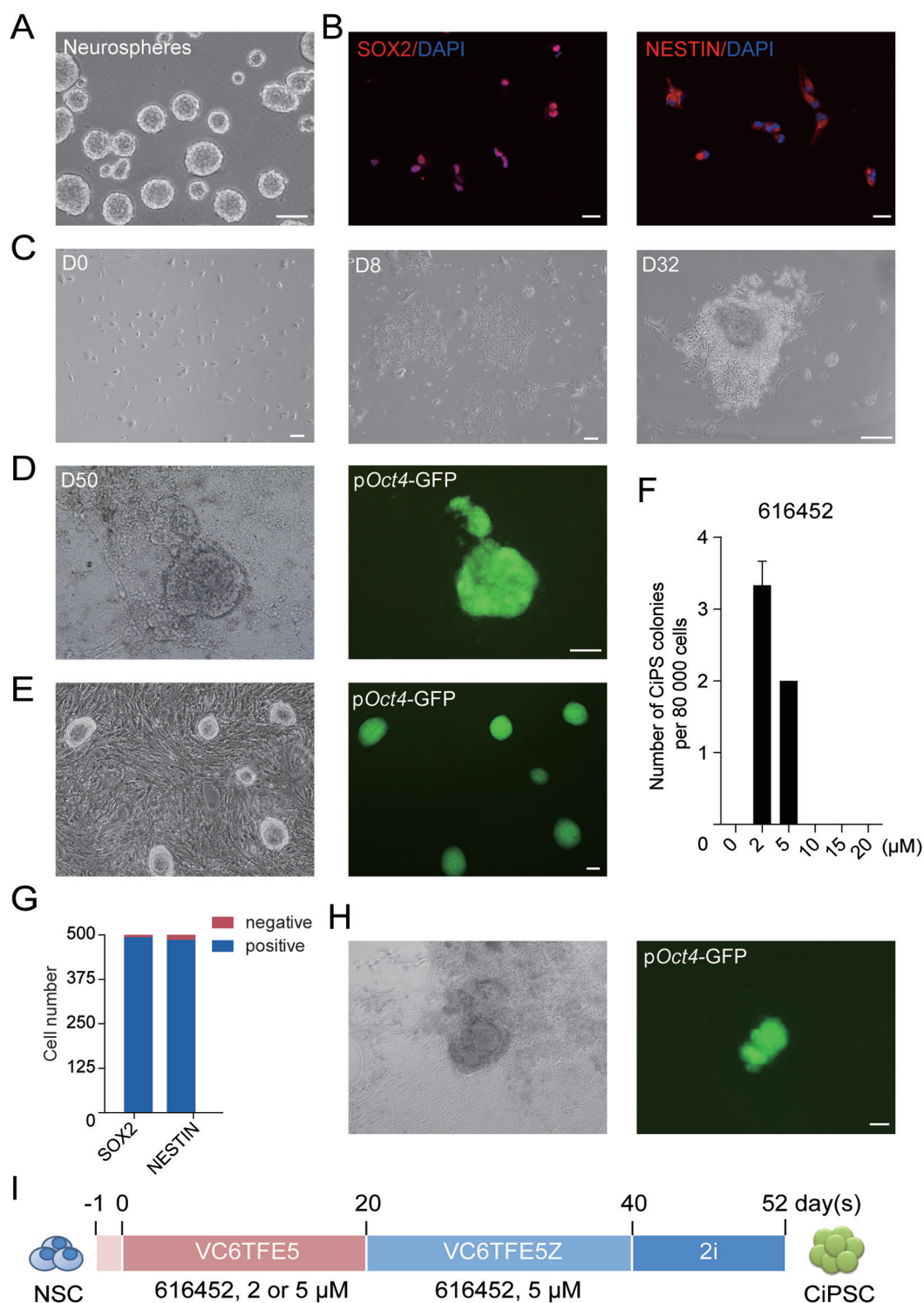
Figure 2 Generation of CiPSCs from mouse NSCs. **(A)** Neurosphere formation of isolated NSCs. **(B)** Immunofluorescence staining of SOX2 and NESTIN in embryonic NSCs (passage 2). **(C-E)** Epithelial clusters at day 8 (D8) after chemical treatment **(C)**, a compact, epithelioid colony at day 32 (D32) after treatment **(C)**, a primary CiPS colony on day 50 (D50) **(D)**, and passaged CiPSC colonies (P1) **(E)**. **(F)** Optimization of 616452 concentration in the first 20 days during chemical reprogramming (error bars, mean ± SD, *n* = 3). **(G)** Numbers of SOX2- and NESTIN-positive cells in 500 FACS-sorted SSEA1-positive NSCs. **(H)** A primary CiPSC colony generated from SSEA1-positive NSCs. **(I)** Schematic diagram for the chemical reprogramming of NSCs. V, VPA; C, CHIR 99021; 6, 616452; T, tranylcypromine; F, forskolin; E, EPZ004777; 5, Ch 55; Z, DZNep. For **A** and **C-E**, scale bar, 100 μm. For **B**, scale bar, 20 μm. See also Supplementary information, Figure S1.

derived from NSCs, we performed fluorescence-activated cell sorting (FACS) to enrich cells positive for a well-known NSC surface marker, SSEA1 [26]. Immunofluorescent staining showed that the sorted SSEA1-positive NSCs were SOX2- and NESTIN-positive (Figure 2G and Supplementary information, Figure S1G). CiPSCs could also be derived from these SSEA1-positive NSCs with a comparable efficiency (Figure 2H).

We optimized the concentrations of small molecules in our cocktail and found that the concentration of 616452 is critical for the early stage of chemical reprogramming from NSCs (Figure 2F). At the late stage, the concentration of 616452 should be increased to 5 μM from day 20 (data not shown).

Generation of CiPSCs from IECs

To determine whether chemical reprogramming is also applicable to cell types from endoderm, we tried to reprogram IECs isolated from small intestine of *Oct4*-GFP transgenic mice at embryonic day 13.5 [27]. Isolated IECs exhibited epithelial cell morphology, and expressed KERATIN 20 (KRT20), an IEC-specific marker [28] at a high level (Figure 3A). RNA-seq also revealed that IECs, in contrast to MEFs and NSCs, expressed a panel of IEC markers including *Epcam*, *Krt18*, *Cdh1*, *Glp1*, and *Smoc2* [29-33] (Supplementary information, Figure S2A).

Isolated IECs were cultured with the previously reported chemical reprogramming cocktail VC6TF and another chemical reprogramming booster, a RAR agonist AM 580 (A) [16]. DZnep was then added to the cocktail from day 16. During this process, epithelioid clusters were observed from day 4 to day 8, and formed colonies from day 15 (Figure 3B). After switching to 2i-medium on day 40, compact, epithelioid, ESC-like OG-positive colonies with clear-cut edges developed (Figure 3B and 3G). Primary CiPSC colonies were counted and picked up between day 52 and day 56, and cultured in 2i-medium with mouse LIF for more than 20 passages, maintaining ESC-like morphology (Figure 3C). We thus named these cells IEC-CiPSCs.

We next performed FACS for EpCAM-positive cells in the primary IEC culture. The EpCAM-positive cells formed classical epithelial colonies, and homogenously expressed IEC markers CDX2 and KRT20 (Supplementary information, Figure S2B and S2C). After small molecule treatment, CiPSCs were obtained from the sorted EpCAM-positive cells (Supplementary information, Figure S2D and S2E). Furthermore, we labeled the cells expressing *Villin*, an IEC marker gene [34], with tdTomato using a widely used lineage tracing system (Figure 3D). After exposure to the chemical cocktail, tdTomato-pos-

itive CiPSC colonies were generated from IECs (Figure 3E and Supplementary information, Figure S3C-S3E). Taken together, these results indicate that CiPSCs can be generated from IECs.

Interestingly, we found that increasing the concentration of 616452 up to 20 μM during the first 12 days best supported IEC reprograming (Figure 3F), whereas a decreased concentration of 616452 is beneficial for NSC reprogramming.

Characterization of NSC- and IEC-derived CiPSCs

The established NSC-CiPSC lines and IEC-CiPSC lines were further characterized. The doubling time of the established CiPSC lines was 18-24 h, similar to mouse ESCs (Supplementary information, Figure S3C). As detected by quantitative real-time PCR and immunofluorescence staining, pluripotency genes in NSC- and IEC-derived CiPSC lines were expressed at comparable levels to mESCs (Figure 4A-4D). Global gene expression analysis by RNA-seq revealed that NSC-CiPSCs and IEC-CiPSCs clustered with ESCs, in contrast to initial cells (Supplementary information, Figure S3A). To detect the epigenetic status of CiPSCs, DNA methylation analysis was performed at the promoters of two core pluripotency genes, *Oct4* and *Nanog*, by bisulfite sequencing. The result showed that *Oct4* and *Nanog* promoters were hypomethylated in NSC-CiPSCs, IEC-CiPSCs and mESCs (Supplementary information, Figure S3B). In addition, normal karyotypes were maintained for up to seven passages (Figure 4F and Supplementary information, Figure S3D).

To evaluate the pluripotency of CiPSCs derived from NSCs and IECs, their *in vivo* developmental potential was examined. Both CiPSC lines could form well differentiated teratomas containing tissues from all three germ layers after injection into immunodeficient (NOD/SCID) mice (Figure 4E). When injected into blastocysts, CiPSCs were able to generate chimeric mice with germline transmission competency (Figure 4G and Supplementary information, Table S3). Thus, CiPSC lines derived from NSCs and IECs are pluripotent and fully reprogrammed.

Similar initial gene activation profile in chemical reprogramming from different cell types

At the initial stage of chemically induced reprogramming to pluripotency, NSCs and IECs were transformed into highly refractive phase-bright and epithelial-like cells, which shared similar morphology of partially reprogrammed colonies during MEF chemical reprogramming. Compact epithelioid colonies were observed during the period from day 20 to day 32 (Figures 2B and
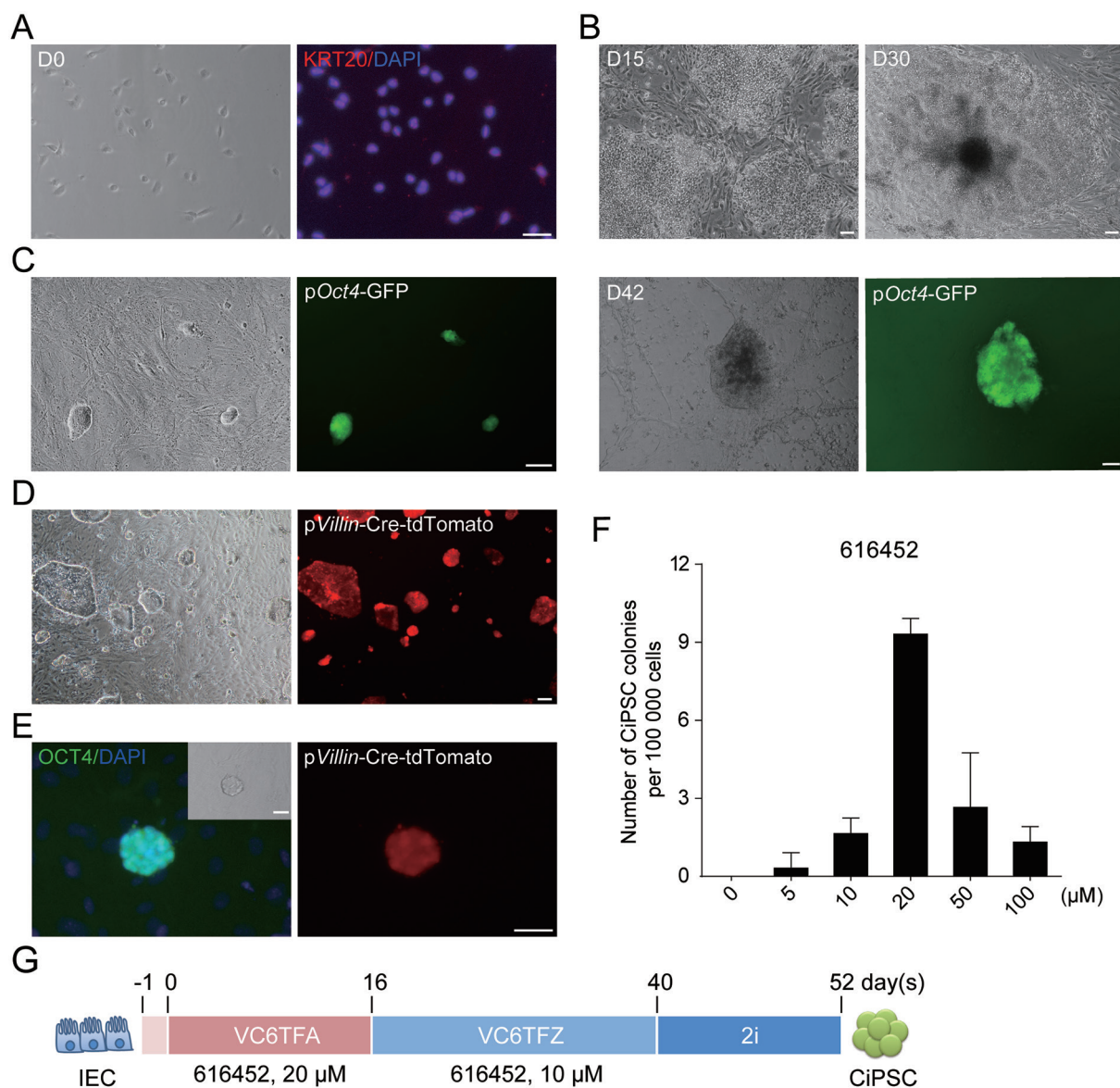
Figure 3 Generation of CiPSCs from mouse IECs. **(A)** Immunofluorescence staining of intestinal epithelial cell marker KERA-TIN20 (KRT20) expression in primary embryonic IECs. **(B**, **C)** Morphology of colonies during chemical reprogramming process, and a primary CiPS colony on day 42 (D42) **(B)**, and passaged CiPSC colonies (P1) **(C)**. **(D**, **E)** Lineage tracing of the isolated IECs (red) from p*Villin*-Cre: Rosa26R^{tdTomato} mice **(D)**, and a passaged CiPSC colony derived from p*Villin*-Cre-tdTomato-positive IECs **(E)**. **(F)** Optimization of 616452 concentration in the first 12 days during chemical reprogramming (error bars, mean ± SD, *n* = 3). **(G)** Schematic diagram for the chemical reprogramming of IECs. V, VPA; C, CHIR 99021; 6, 616452; T, tranylcypromine; F, forskolin; A, AM 580; Z, DZNep. For **A**-**E**, scale bar, 100 μm. See also Supplementary information, Figure S2.

3B), and all CiPSC colonies were derived from these epithelioid colonies in at least ten independent experiments. As we reported before that the expression of *Sall4*, *Gata4*, *Gata6* and *Sox17* were most significantly induced at the early stage of MEF chemical reprogramming [16], we asked whether the same genes were activated during chemical reprogramming of NSCs and IECs. Indeed, time-course quantitative real-time PCR revealed that during chemical reprogramming of NSCs and IECs, the pluripotency gene *Sall4*, and differentiation-associated genes *Gata4*, *Gata6* and *Sox17*, were activated as early as day 4 and their expression levels increased with time in the early stage, while other pluripotency genes such as *Lin28a*, *Dppa2*, *Esrrb* and *Oct4* were activated much later (Figure 5A, 5C and 5E), which is similar to what we have observed during the chemical reprogramming
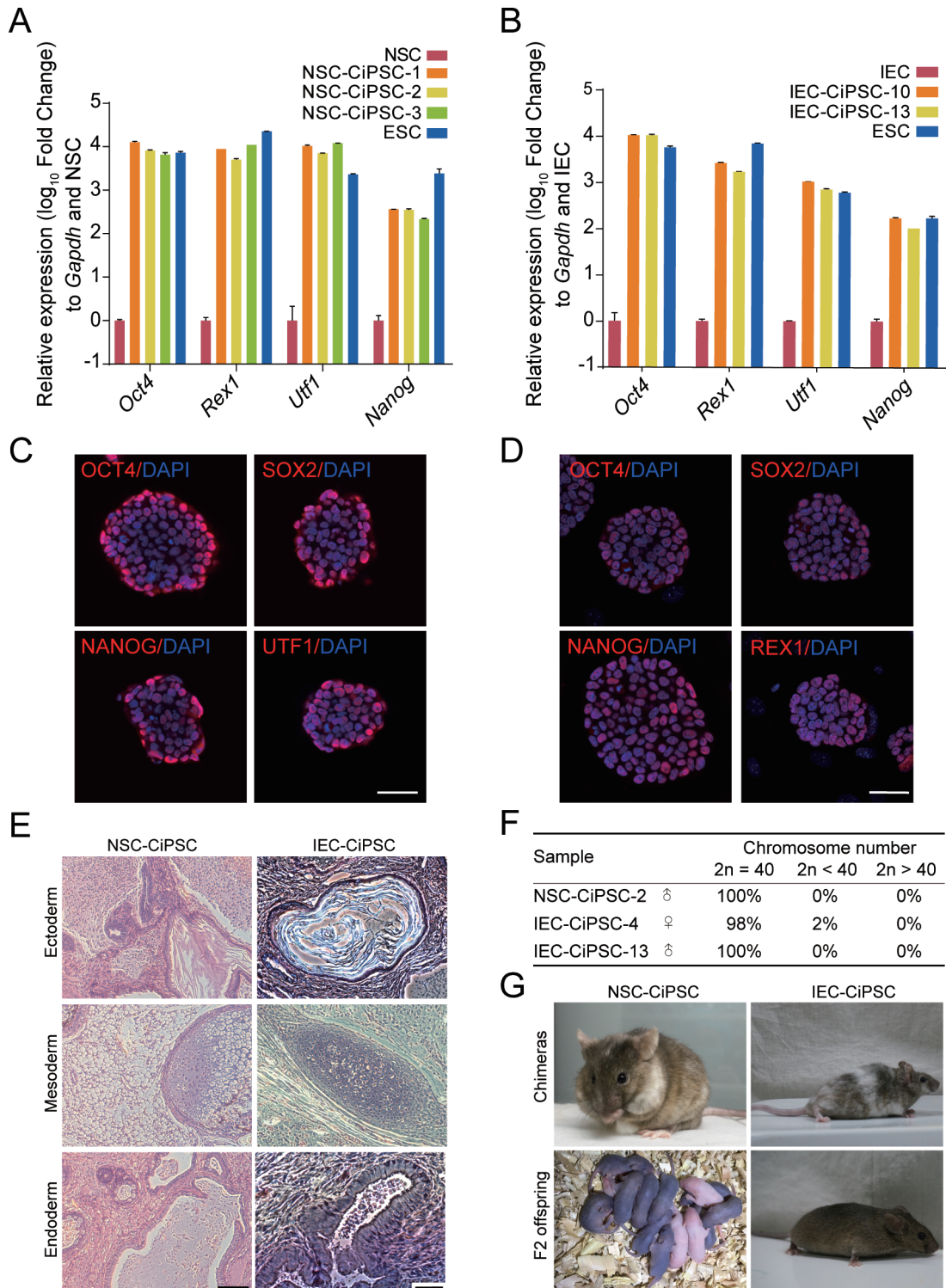
Figure 4 Characterization of CiPSCs derived from NSCs and IECs. (**A**, **B**) Pluripotency gene expression profile of NSC-CiP-SCs (**A**) and IEC-CiPSCs (**B**) as revealed by quantitative real-time PCR (error bars, mean ± SD, n = 3). (**C**, **D**) Immunofluorescence staining for the indicated pluripotency markers in CiPSCs derived from NSCs (**C**) and IECs (**D**). (**E**) Hematoxylin and eosin staining of teratoma derived from NSC-CiPSCs (left) and IEC-CiPSCs (right). (**F**) Number of chromosomes in NSC-CiPSCs and IEC-CiPSCs by karyotype analysis (clone IEC-CiPSC-4 is derived from p*Villin*-Cre-tdTomato-positive IECs). (**G**) Chimeric mice and F2 offspring derived by NSC-CiPSCs (left) and IEC-CiPSCs (right). For **C** and **D**, scale bar, 50 μm. For **E**, scale bar, 100 μm. See also Supplementary information, Figure S3.
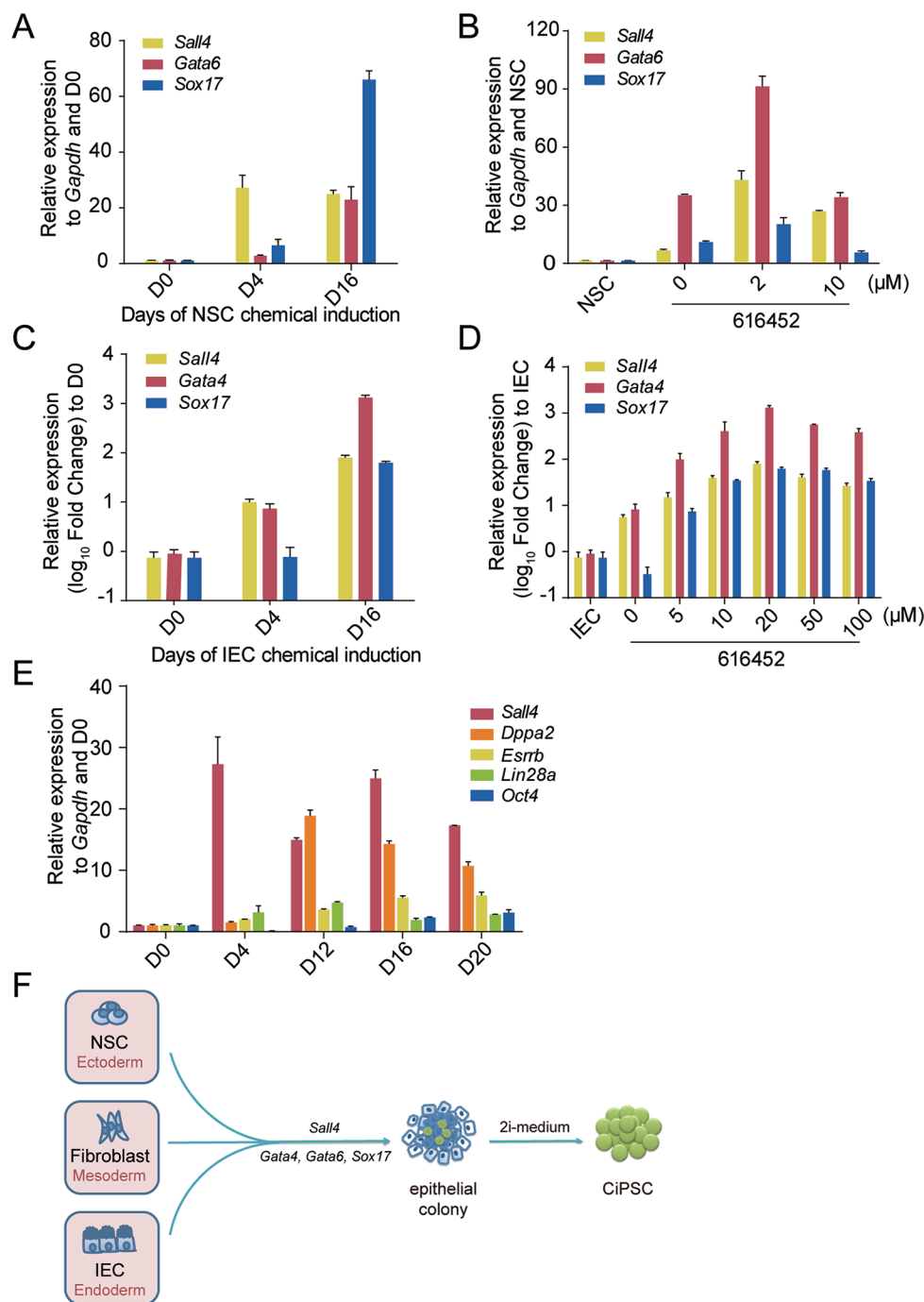
Figure 5 Initial gene activation was conserved in chemical reprogramming from different cell types. **(A)** Expression of *Sall4*, *Gata6* and *Sox17* at the early stage of chemical reprogramming from NSCs (day 0 (D0), day 4 (D4) and day 16 (D16), respectively) measured by quantitative real-time PCR (error bars, mean ± SD, *n* = 3). **(B)** Expression of *Sall4*, *Gata6* and *Sox17* genes at day 16 of the reprogramming by the chemical cocktail with different concentration of 616452 (error bars, mean ± SD, *n* = 3). **(C)** Expression of *Sall4*, *Gata4* and *Sox17* genes at the early stage of chemical reprogramming from IECs (day 0 (D0), day 4 (D4) and day 16 (D16), respectively) measured by quantitative real-time PCR (error bars, mean ± SD, *n* = 3). **(D)** Expression of *Sall4*, *Gata4* and *Sox17* genes at day 20 of the reprogramming by the chemical cocktail with different concentration of 616452 (error bars, mean ± SD, *n* = 3). **(E)** Expression of pluripotency genes *Sall4*, *Lin28*, *Esrrb*, *Dppa2* and *Oct4* during the chemical reprogramming from NSCs (day 0 (D0), day 4 (D4), day 12 (D12), day 16 (D16) and day 20 (D20)) measured by quantitative real-time PCR (error bars, mean ± SD, *n* = 3). **(F)** Schematic diagram illustrating the step-wise chemical reprogramming of different cell types.

process of MEFs [16]. In addition, it is important to note that *Sall4*, *Gata4*, *Gata6* and *Sox17* were not expressed in NSCs (Supplementary information, Figure S1C and S1D). Interestingly, fine-tuning the concentration of 616452 to a level that best supported the reprogramming of different cell types resulted in higher gene expression levels of *Sall4, Gata4* and *Sox17* (Figure 5B and 5D). These results suggest that the chemical reprogramming in different cell types uses similar initial gene activation program, regardless of germ layer and inherent cellular features. Consistently, in another study of our group, a conserved intermediate extraembryonic endoderm (XEN)-like state is uncovered in the process of chemical reprogramming from somatic cells to pluripotent stem cells [35].

Discussion

Our data demonstrate that our previously reported chemical reprogramming approach is reproducible in fibroblasts, NSCs and IECs, which are cell types from mesoderm, ectoderm and endoderm, respectively. Furthermore, a similar combination of core small molecule compounds is required for the reprogramming of different cell types from all three germ layers. Our results demonstrate that the previously reported chemical reprogramming approach is cell type-independent and imply a general underlying molecular mechanism of chemical reprogramming.

Indeed, the chemical reprogramming of NSCs and IECs shares certain commonalities with the chemical reprogramming of fibroblasts. First, they all require the same core small molecules, VPA, CHIR, 616452, forskolin, and DZNep. Second, the expression of *Sall4*, *Gata4*, *Gata6* and *Sox17* is activated during the early stage of chemical reprogramming of the different cell types. As reported in our previous studies, *Sall4*, *Gata4*, *Gata6* and *Sox17* are critical for chemical reprogramming of MEFs [16]. Interestingly, in other studies by our group, two of these genes, *Gata4* and *Gata6*, can substitute for *Oct4* in transgenic reprogramming in the presence of exogenous *Sox2*, *Klf4* and *c-Myc* [36], partially through activating the expression of the pluripotency-associated gene, *Sall4* [37]. These results, taken together, suggest a conserved and unique molecular roadmap in chemical reprogramming. It would be particularly interesting to further investigate this molecular roadmap underlying chemical reprogramming.

Importantly, although a similar chemical cocktail is required for different cell types, fine-tuning the regimen of small molecule treatment is critical for reprogramming initiation in different cell types. Especially for the chemical reprogramming of NSCs, the concentration of the TGFβ inhibitor, 616452, must be reduced. Notably, the reduced 616452 concentration resulted in enhanced expression of *Sall4* and *Gata6* genes (Figure 5B), which further supports that high expression of these genes is critical for the initiation of chemical reprogramming. Thus, successful chemical reprogramming of a cell type should depend on a unique regimen of same small molecules that has to be empirically established.

Our study shows that 616452 exerts different effects in chemical reprogramming and traditional transcription factor-induced reprogramming. In traditional transcription factor-induced reprogramming, ectopic expression of *Oct4* and *Klf4* or *c-Myc* is sufficient to generate iPS cells from NSCs, in the absence of exogenous *Sox2* [10]. Besides, 616452, also known as RepSox, has been reported to substitute for *Sox2* overexpression in traditional transcription factor-induced reprogramming [38]. However, to our surprise, in chemical reprogramming from NSCs, 616452 is indispensable, despite the high endogenous expression of *Sox2*. Intriguingly, in chemical reprogramming, 616452 facilitates the early expression of *Sall4*, *Gata4*, *Gata6* and *Sox17* and the subsequent epithelioid colony formation (Figure 5B and 5D). These findings indicate that chemical reprogramming is a different process from transcription factor-induced reprogramming.

Interestingly, the reprogramming kinetics and frequency of NSCs and IECs are distinct from each other. Chemical reprogramming of NSCs has a longer early stage, with less epithelioid colonies generated, but almost 100 percent of the colonies can be converted into CiPSC colonies. In contrast, IECs could readily form more epithelioid colonies, but only 20-30 percent of these colonies are converted into CiPSC colonies. A possible explanation for this difference is that the inherent expression of *Gata4* and *Gata6* in IECs may facilitate colony formation, while in NSCs, the high endogenous expression of *Sox2* may antagonize early *Gata4* and *Gata6* gene activation [36]. Moreover, the higher expression of the pluripotency gene *Sox2* in NSCs may facilitate the establishment of a pluripotent network at the late stage of chemical reprogramming [16, 39]. Our data imply that an optimized chemical approach can be achieved by temporally and precisely manipulating the small molecules during the two stages of chemical reprogramming.

Overall, this study demonstrates that our previous discovery of chemical reprogramming is a general and reproducible approach for cells from mesodermal, endodermal and ectodermal lineages. Moreover, this study demonstrates the chemical reprogramming of cells from different germ layers initiates with a conserved gene expression pattern (Figure 5F), and fine-tuning the con-

centrations of chemical compounds is sufficient to reprogram different cell types using the same core chemical cocktail. In summary, our study provides a blueprint for reprogramming different cell types to pluripotency by using small molecules, and suggests chemical reprogramming is a general and novel way to establish and understand pluripotency.

Materials and Methods

Mice

Mouse strains C57BL/6J-Tg(GOFGFP)11Imeg/Rbrc (*Oct4*-GFP, OG) and ICR [17] were crossed to provide primary mouse embryonic fibroblasts (MEFs), mouse small intestinal epithelial cells (IECs), mouse neural stem cells (NSCs) and mouse postnatal NSCs. The postnatal mice used were 3-7 days old.

For lineage tracing, mouse strain B6/J-Tg(p*Villin*-Cre)Nju (p*Villin*-Cre) was purchased from Nanjing Biomedical Research Institute of Nanjing University. Mouse strains BALB/c-Tg(S100a4-cre)1Egn/YunkJ (*Fsp*1-Cre) and B6.Cg-Gt(RO-SA)26-Sor^tm14(CAG-tdTomato)Hze/J (Rosa26R^tdTomato) were purchased from the Jackson Laboratory. p*Villin*-tdTomato mice were generated by crossing p*Villin*-Cre mice with Rosa26R^tdTomato mice [40]. Animal experiments were performed according to the Animal Protection Guidelines of Peking University, China.

Cell culture

Primary MEFs were isolated as previously described [1], with careful attention to the removal of the genital ridges. MEFs were cultured in high glucose Dulbecco's modified Eagle medium (DMEM, Hyclone).

Fetal NSCs were isolated from mouse forebrain at embryonic day 13.5 as previously described [41], and postnatal NSCs were isolated from the subventricular zone of the postnatal mouse brain [41, 42]. NSCs were cultured in NSC culture medium (DMEM/F-12 (1:1), DF12) containing N2 and B27 supplements, 1% non-essential amino acids (NEAA), 2 mM GlutaMAX, 10 U PS, 55 μM β-mercaptoethanol (β-me) (all from Invitrogen), 25 ng/ml basic fibroblast growth factor (bFGF, Origene), and 20 ng/ml epidermal growth factor (EGF, R&D), and passaged by accutase (Millipore) every 4-5 days. Single-cell suspension of NSCs were allowed to form neurospheres for 2-3 days.

Fetal small intestinal epithelial cells were isolated from mouse embryonic small intestine at embryonic 13.5 day as previously described [17], and cultured in DMEM/F-12 (1:1) (Invitrogen) containing N2 and B27 supplements, 10% knockout serum replacement (KSR, Invitrogene), 1% NEAA, 2 mM GlutaMAX-I (GlutaMAX), 10 U penicillin-streptomycin (PS), 55 μM β-me, 25 ng/ml Wnt3a (R&D), 20 ng/ml EGF, and 20 ng/ml Noggin (AC-ROBiosystems).

Mouse ESCs (R1) and CiPSCs were maintained on feeder layers of mitomycin C-treated MEFs in 2i-medium (Knockout DMEM containing 10% KSR, 10% fetal bovine serum (FBS, Pan-Biotech), 2 mM GlutaMAX, 1% NEAA, 55 μM β-me, 10 U PS (all from Invitrogen), 3 μM CHIR99021 (CHIR), 1 μM PD0325901 (PD03) and 10 ng/ml mouse leukemia inhibitory factor (mLIF; Millipore)). The medium was changed daily. ESCs and CiPSCs were passaged by trypsin-EDTA (Invitrogen).

For CiPSC induction, cells were cultured in Knockout DMEM containing 10% KSR, 10% FBS, 2 mM GlutaMAX, 1% NEAA, 55 μM β-me, 10 U PS and 25 ng/ml bFGF.

Small molecule compounds

The source and concentration of each small molecule compound are described in Supplementary information, Table S2.

CiPSC induction

MEF-CiPSC induction The initial MEFs were seeded at a density of 300 000 cells per well of a 10 cm culture dish. On the next day (day 0), the medium was replaced with chemical reprogramming medium containing the small molecule cocktail (0.5 mM VPA, 10 μM CHIR, 10 μM 616452, 10 μM tranylcypromine, 10 μM forskolin, 0.05 μM AM580) and changed every 4 days. From day 20, 0.05 μM DZNep was added into chemical reprogramming medium. The medium was replaced with 2i-medium (with DF12 containing N2 and B27 supplements) from day 40. After another 12 days, 2i-competent, ESC-like and OG-positive colonies were counted as primary CiPSC colonies. The primary CiPSC colonies were picked for expansion and further characterization.

NSC-CiPSC induction The initial NSCs were seeded at a density of 80 000 cells per well of a six-well plate. The original culture medium was replaced by chemical reprogramming medium containing the small-molecule cocktail (0.5 mM VPA, 15 μM CHIR, 2 μM or 5 μM 616452, 10 μM tranylcypromine, 20 μM forskolin, 1 μM Ch 55, 5 μM EPZ) and changed every 4 days. From day 20, 0.05 μM DZNep was added into the chemical reprogramming medium. The medium was replaced with 2i-medium (with DF12 containing N2 and B27 supplements) from day 40-44. After another 12 days, 2i-competent, ESC-like and OG-positive colonies were counted as primary CiPSC colonies. The primary CiPSC colonies were picked for expansion and further characterization.

IEC-CiPSC induction The initial IECs were seeded at a density of 100 000 cells per well in a 12-well plate. On the next day (day 0), the medium was replaced with chemical reprogramming medium containing the small-molecule cocktail (0.5 mM VPA, 10 μM CHIR, 10 μM or 20 μM 616452, 10 μM tranylcypromine, 10 μM forskolin, 0.05 μM AM 580) and changed every 4 days. From day 16, 0.05 μM DZNep was added into the chemical reprogramming medium, and AM 580 was withdrawn. The medium was replaced with 2i-medium (with DF12 containing N2 and B27 supplements) from day 40. After another 12 days, 2i-competent, ESC-like and OG-positive colonies were counted as primary CiPSC colonies. The primary CiPSC colonies were picked for expansion and further characterization.

Immunofluorescence staining

Immunofluorescence staining was performed as previously described [17]. Primary antibodies included OCT4 (Abcam ab18976), SOX2 (Santa Cruz sc-17320), NANOG (R&D AF2729), REX1 (Santa Cruz sc-99000), UTF1 (R&D AF3958), SALL4 (Santa Cruz sc-166033), GATA4 (Santa Cruz sc-25310), KERATIN 20 (Santa Cruz sc-17112), CDX2 (BioGenex antibody CDX2-88) and NESTIN (Santa Cruz sc-58813). Secondary antibodies were TRITC-conjugated or FITC-conjugated, including donkey-anti-mouse lgG (H+L), donkey-anti-rabbit lgG (H+L), and

donkey-anti-goat lgG (H+L) from Jackson ImmunoResearch.

Quantitative real-time PCR and RNA-seq

Total RNA was isolated from fresh cells using Direct-zol RNA MiniPrep Kit (ZYMO), and converted into cDNA using TransScript First-Strand cDNA Synthesis SuperMix (TransGen Biotech). Quantitative real-time PCR was conducted using Power SYBRR Green PCR Master Mix (Applied Biosystems) and performed on an ABI Prism 7300 Real-Time PCR System. Primers used are shown in Supplementary information, Table S4. The delta-delta Ct method was used for data analysis.

RNA sequencing libraries were constructed using the NEBNext Ultra RNA Library Prep Kit for Illumina (NEB). Fragmented and randomly primed 2×125 bp paired-end libraries were sequenced using Illumina HiSeq 2500. We applied kallisto to obtain estimation of raw counts of unique mapped RNA-seq reads. Gene ontology analysis was performed by DAVID online tools.

Teratoma formation

For this purpose, 1×10^6 CiPSCs were resuspended in 100 μl PBS and subcutaneously injected into the inguinal region of NOD/SCID mice.

Karyotype analysis

Standard G-banding chromosome analysis was performed as reported [17]. VIDEO TEST 3.0 software was used to analyze the karyotype.

Chimera construction

Chimeric mice were obtained by the injection of CiPSCs into blastocysts or eight-cell embryos using a XY Clone laser system (Hamilton Thorne Bioscience). For blastocyst injection, 10-15 CiPSCs were injected into one recipient embryo cavity of F2 (intercross of B6D2F1) or CD-1 (albino) female mouse at 3.5 days post coitum (dpc). Host eight-cell embryos were collected from female mice at 2.5 dpc, and 7-10 CiPSCs were injected into each embryo. After injection, blastocysts and eight-cell embryos were transferred into 2.5 dpc and 0.5 dpc pseudopregnant CD-1 females (6-8 embryos in each oviduct or horn of the uterus, respectively). Chimeric mice were identified by coat color and then mated with ICR mice for germline transmission.

Bisulfite genomic sequencing

Genomic DNA was modified by bisulfite treatment and purified using the MethylCode Bisulfite Conversion Kit (Invitrogen). The primers are listed in Supplementary information, Table S4. The amplified fragments were cloned into the pEASY-blunt Vector (Transgene). Seven randomly picked clones from each sample were sequenced.

Accession number

The raw and processed RNA-seq data have been deposited in the Gene Expression Omnibus (GEO) database under the accession number GSE75522.

Acknowledgments

Author Contributions

Competing Financial Interests

References

1   Takahashi, K., Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**:663-676.

2   Takahashi K, Tanabe K, Ohnuki M, *et al*. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 2007; **131**:861-872.

3   Yu J, Vodyanik MA, Smuga-Otto K, *et al*. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 2007; **318**:1917-1920.

4   Yamanaka S. Elite and stochastic models for induced pluripotent stem cell generation. *Nature* 2009; **460**:49-52.

5   Wakao S, Kitada M, Kuroda Y, *et al*. Multilineage-differentiating stress-enduring (Muse) cells are a primary source of induced pluripotent stem cells in human fibroblasts. *Proc Natl Acad Sci USA* 2011; **108**:9875-9880.

6   Guo S, Zi X, Schulz VP, *et al*. Nonstochastic reprogramming from a privileged somatic cell state. *Cell* 2014; **156**:649-662.

7   Aasen T, Raya A, Barrero MJ, *et al*. Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotech* 2008; **26**:1276-1284.

8   Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K. Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 2008; **26**:2467-2474.

9   Kim JB, Sebastiano V, Wu G, *et al*. Oct4-induced pluripotency in adult neural stem cells. *Cell* 2009; **136**:411-419.

10   Kim JB, Zaehres H, Wu G, *et al*. Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 2008; **454**:646-650.

11   Aoi T, Yae K, Nakagawa M, *et al*. Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 2008; **321**:699-702.

12   Hanna J, Markoulaki S, Schorderet P, *et al*. Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* 2008; **133**:250-264.

13   Loh YH, Agarwal S, Park IH, *et al*. Generation of induced

pluripotent stem cells from human blood. *Blood* 2009; **113**:5476-5479.

14 Loh YH, Hartung O, Li H, *et al*. Reprogramming of T Cells from Human Peripheral Blood. *Cell Stem Cell* 2010; **7**:15-19.

15 Staerk J, Dawlaty MM, Gao Q, *et al*. Reprogramming of human peripheral blood cells to induced pluripotent stem cells. *Cell Stem Cell* 2010; **7**:20-24.

16 Hou P, Li Y, Zhang X, *et al*. Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science* 2013; **341**:651-654.

17 Li Y, Zhang Q, Yin X, *et al*. Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. *Cell Res* 2011; **21**:196-204.

18 Driskell RR, Watt FM. Understanding fibroblast heterogeneity in the skin. *Trends Cell Biol* 2015; **25**:92-99.

19 Vidal SE, Amlani B, Chen T, Tsirigos A, Stadtfeld M. Combinatorial modulation of signaling pathways reveals cell-type-specific requirements for highly efficient and synchronous iPSC reprogramming. *Stem Cell Reports* 2014; **3**:574-584.

20 McKay, R. Stem cells in the central nervous system. *Science* 1997; **276**:66-71.

21 Conti L, Pollard SM, Gorba T, *et al*. Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biol* 2005; **3**:e283.

22 Pollard SM, Conti L, Sun Y, Goffredo D, Smith A. Adherent neural stem (NS) cells from fetal and adult forebrain. *Cereb Cortex* 2006; **16**:i112-i120.

23 Episkopou, V. SOX2 functions in adult neural stem cells. *Trends Neurosci* 2005; **28**:219-221.

24 Onder TT, Kara N, Cherry A, *et al*. Chromatin-modifying enzymes as modulators of reprogramming. *Nature* 2012; **483**:598-602.

25 Ying QL, Wray J, Nichols J, *et al*. The ground state of embryonic stem cell self-renewal. *Nature* 2008; **453**:519-523.

26 Capela A, Temple S. LeX is expressed by principle progenitor cells in the embryonic nervous system, is secreted into their environment and binds Wnt-1. *Dev Biol* 2006; **291**:300-313.

27 Wernig M, Lengner CJ, Hanna J, *et al*. A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nat Biotechnol* 2008; **26**:916-924.

28 Chan CW, Wong NA, Liu Y, *et al*. Gastrointestinal differentiation marker Cytokeratin 20 is regulated by homeobox gene CDX1. *Proc Natl Acad Sci USA* 2009; **106**:1936-1941.

29 Stingl J, Raouf A, Emerman JT, Eaves CJ. Epithelial progenitors in the normal human mammary gland. *J Mammary Gland Biol Neoplasia* 2005; **10**:49-59.

30 Chou CF, Riopel CL, Rott LS, Omary MB. A significant soluble keratin fraction in "simple" epithelial cells. Lack of an apparent phosphorylation and glycosylation role in keratin solubility. *J Cell Sci* 1993; **105**:433-444.

31 Tang A, Eller MS, Hara M, Yaar M, Hirohashi S, Gilchrest BA. E-cadherin is the major mediator of human melanocyte adhesion to keratinocytes in vitro. *J Cell Sci* 1994; **107**:983-992.

32 Eissele R, Göke R, Willemer S, *et al*. Glucagon-like peptide-1 cells in the gastrointestinal tract and pancreas of rat, pig and man. *Eur J Clin Invest* 1992; **22**:283-291.

33 Muñoz J, Stange DE, Schepers AG, *et al*. The Lgr5 intestinal stem cell signature: Robust expression of proposed quiescent "+4" cell markers. *EMBO J* 2012; **31**:3079-3091.

34 Madison B, Dunbar L, Qiao XT, Braunstein K, Braunstein E, Gumucio DL. Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. *J Biol Chem* 2002; **277**:33275-33283.

35 Zhao Y, Zhao T, Guan JY, *et al*. A XEN-like state bridges somatic cells to pluripotency during chemical reprogramming. *Cell* 2015 Dec 10; doi:10.1016/j.cell.2015.11.017

36 Shu J, Wu C, Wu Y, *et al*. Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell* 2013; **153**:963-975.

37 Shu J, Zhang K, Zhang M, *et al*. GATA family members as inducers for cellular reprogramming to pluripotency. *Cell Res* 2015; **25**:169-180.

38 Ichida JK, Blanchard J, Lam K, *et al*. A small-molecule inhibitor of tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. *Cell Stem Cell* 2009; **5**:491-503.

39 Buganim Y, Faddah DA, Cheng AW, *et al*. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell* 2012; **150**:1209-1222.

40 Madisen L, Zwingman TA, Sunkin SM, *et al*. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* 2010; **13**:133-140.

41 Fischer J, Beckervordersandforth R, Tripathi P, Steiner-Mezzadri A, Ninkovic J, Götz M. Prospective isolation of adult neural stem cells from the mouse subependymal zone. *Nat Protoc* 2011; **6**:1981-1989.

42 Guo W, Patzlaff NE, Jobe EM, Zhao X. Isolation of multipotent neural stem or progenitor cells from both the dentate gyrus and subventricular zone of a single adult mouse. *Nat Protoc* 2012; **7**:2005-2012.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)