Stem Cell Reports
Article

ISSCR

OPEN ACCESS

Dynamics of Telomere Rejuvenation during Chemical Induction to Pluripotent Stem Cells

Haifeng Fu,[1,2,4] Cheng-lei Tian,[1,2,4] Xiaoying Ye,[1,2,4] Xiaoyan Sheng,[1,2] Hua Wang,[1,2] Yifei Liu,[3] and Lin Liu[1,2,*]

[1]State Key Laboratory of Medicinal Chemical Biology, Nankai University, Tianjin 300071, China
[2]Department of Cell Biology and Genetics, College of Life Sciences, Key Laboratory of Bioactive Materials of Ministry of Education, Nankai University, Tianjin 300071, China
[3]Department of Obstetrics, Gynecology & Reproductive Sciences, Yale School of Medicine, New Haven, CT 06511, USA
[4]Co-first author
*Correspondence: liulin@nankai.edu.cn
https://doi.org/10.1016/j.stemcr.2018.05.003

SUMMARY

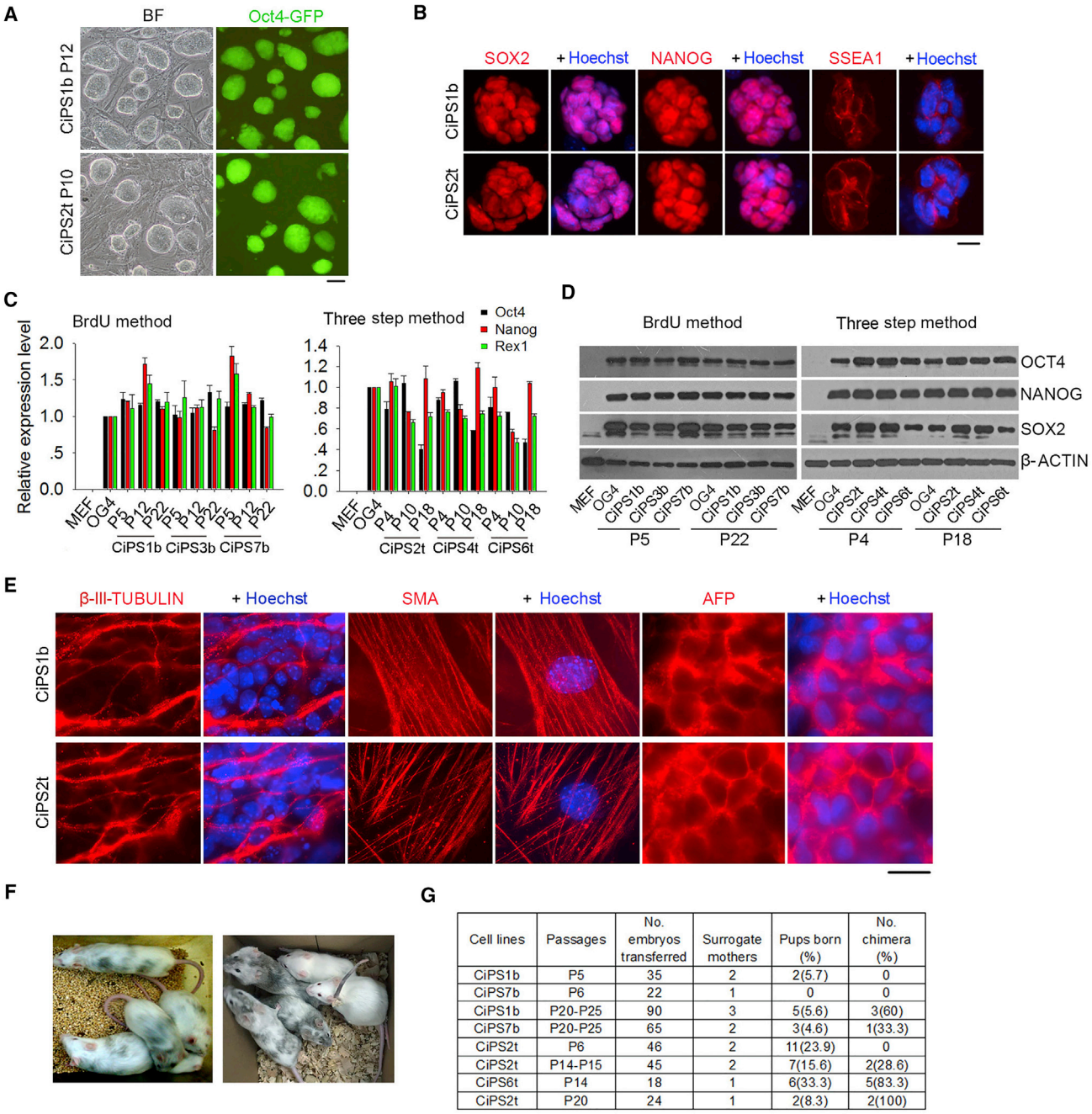
Chemically induced pluripotent stem cells (CiPSCs) may provide an alternative and attractive source for stem cell-based therapy. Sufficient telomere lengths are critical for unlimited self-renewal and genomic stability of pluripotent stem cells. Dynamics and mechanisms of telomere reprogramming of CiPSCs remain elusive. We show that CiPSCs acquire telomere lengthening with increasing passages after clonal formation. Both telomerase activity and recombination-based mechanisms are involved in the telomere elongation. Telomere lengths strongly indicate the degree of reprogramming, pluripotency, and differentiation capacity of CiPSCs. Nevertheless, telomere damage and shortening occur at a late stage of lengthy induction, limiting CiPSC formation. We find that histone crotonylation induced by crotonic acid can activate two-cell genes, including *Zscan4*; maintain telomeres; and promote CiPSC generation. Crotonylation decreases the abundance of heterochromatic H3K9me3 and HP1α at subtelomeres and *Zscan4* loci. Taken together, telomere rejuvenation links to reprogramming and pluripotency of CiPSCs. Crotonylation facilitates telomere maintenance and enhances chemically induced reprogramming to pluripotency.

INTRODUCTION

Somatic cells can be successfully induced by forced expression of four key transcription factors into induced pluripotent stem cells (iPSCs) that resemble embryonic stem cells (ESCs) (Takahashi and Yamanaka, 2006; Wernig et al., 2007; Zhao et al., 2009). The induction efficiency and quality of iPSCs can be significantly improved by various approaches, notably by small molecules (Huangfu et al., 2008a, 2008b; Li et al., 2013; Shi et al., 2008). In combination with small molecules or bone morphogenetic proteins, only Oct4 is needed for generation of iPSCs (Chen et al., 2011; Li et al., 2011; Zhu et al., 2010). High-content chemical screening has been attempted to achieve a goal of each transgenic factor being replaced with a small molecule that induces reprogramming (Ichida et al., 2009). Finally, somatic cells are successfully reprogrammed into pluripotent stem cells (PSCs) by pure small chemicals (Hou et al., 2013). The chemically induced PSCs (CiPSCs) presumably provide an alternative source with greater potential in stem cell-based therapy. Subsequently, two reports demonstrated successful generation of CiPSCs at higher efficiency by pure chemical compounds with slight modification of the recipe (Long et al., 2015; Zhao et al., 2015). Nevertheless, the reprogramming by complete chemicals is still at low efficiency and takes much longer than does OSKM-induced reprogramming (Maherali et al., 2007; Takahashi and Yamanaka, 2006).

Telomeres consist of repeated guanine-rich sequences and associated protein complexes known as shelterin that cap the end of chromosomes to maintain genomic stability (Blackburn, 2001; Palm and de Lange, 2008). Short telomeres can lead to cell senescence or tumorigenesis (Collado et al., 2007). Sufficiently long telomeres are important for unlimited self-renewal and pluripotency of PSCs (Huang et al., 2011; Marion et al., 2009). Short telomeres impair the differentiation capacity of ESCs (Pucci et al., 2013). Telomere length is maintained primarily by the ribonucleoprotein telomerase, a complex of a reverse transcriptase encoded by three core components: *Tert* (telomerase reverse transcriptase), template RNA *Terc* (essential RNA component), and *dyskerin* (Palm and de Lange, 2008). Telomerase is strongly expressed and required for telomere maintenance of mouse and human PSCs (Huang et al., 2011, 2014; Marion et al., 2009; Teichroeb et al., 2016; Wang et al., 2012). However, it remains elusive whether telomeres are appropriately reprogrammed and sufficiently elongated in CiPSCs.

We attempted to investigate telomere dynamics of CiPSCs generated based on the methods described recently (Long et al., 2015; Zhao et al., 2015). We found that CiPSCs acquire telomere lengthening with increasing passages. Surprisingly, telomeres suffer from erosion at late stages during extended periods of chemical induction, limiting reprogramming efficiency. We searched for compounds that can reduce telomere damage and shortening and



Cell lines	Passages	No. embryos transferred	Surrogate mothers	Pups born (%)	No. chimera (%)
CiPS1b	P5	35	2	2(5.7)	0
CiPS7b	P6	22	1	0	0
CiPS1b	P20-P25	90	3	5(5.6)	3(60)
CiPS7b	P20-P25	65	2	3(4.6)	1(33.3)
CiPS2t	P6	46	2	11(23.9)	0
CiPS2t	P14-P15	45	2	7(15.6)	2(28.6)
CiPS6t	P14	18	1	6(33.3)	5(83.3)
CiPS2t	P20	24	1	2(8.3)	2(100)

Figure 1. Pluripotency and Differentiation Capacity of CiPSCs

(A) Representative morphology of CiPSCs generated by two methods (BrdU method, CiPSC1b at P12, and three-step method, CiPSC2t at P10) under bright field (BF) with phase contrast optics and expression of Oct4-GFP fluorescence. Scale bar represents 100 μm.

(B) Immunofluorescence microscopy of pluripotent markers SOX2, NANOG, and SSEA1. Scale bar represents 10 μm.

(C) mRNA expression levels by qPCR of *Oct4*, *Nanog*, and *Rex1* in CiPSCs at various passages, compared with isogenic ESCs (OG4) and progenitor MEFs. Data represent mean ± SEM from three independent experiments.

(D) Protein levels of OCT4, NANOG, and SOX2 by western blot analysis of CiPSCs at earlier and advanced passages.

(E) Differentiation capacity *in vitro* of CiPSCs by immunofluorescence microscopy of three germ layer markers. Scale bar represents 10 μm.

(F) Left photo represents chimeras generated from the BrdU method and the right from the three-step method.

(legend continued on next page)

thus improve chemical reprogramming. Promisingly, histone crotonylation induced by crotonic acid can alleviate telomere damage and shortening, enhancing the chemical induction efficiency.

RESULTS

Generation of CiPSCs

We attempted to generate CiPSCs following two recently published methods. One used a combination of seven small-molecule compounds (Hou et al., 2013), and bromodeoxyuridine (BrdU) (Long et al., 2015), herein referred to as the BrdU method. The other requires three stages to complete induction of CiPSCs, which undergo an extra-embryonic endoderm (XEN)-like state as an intermediate, and differs from the pathway of transcription factor-induced reprogramming, so is referred to as the three-step method (Zhao et al., 2015). Mouse embryonic fibroblasts (MEFs) were isolated from Oct4-GFP (OG2) transgenic mice harboring a GFP reporter driven by the distal Oct4 promoter and enhancer, activation of which indicates a naive state of pluripotency (Bao et al., 2009; De Los Angeles et al., 2015; Tang et al., 2010; Yeom et al., 1996). We successfully generated CiPSCs from OG2-MEFs following either the BrdU method (the randomly selected cell lines for further studies were CiPS1b, 3b, and 7b) or three-step method (cell lines named as CiPS2t, 4t, and 6t) (Figure S1A).

Continuous passages of ESC-like primary colonies established stable CiPSC lines that resembled typical ESC colonies in morphology, exhibiting large nuclei and nucleoli and clear compact clonal boundaries and expression of Oct4-GFP (Figures 1A and S1A), distinct from feeder fibroblasts. Colonies were stochastically picked and six established CiPSC lines chosen for further characterization of their pluripotency. By direct comparison with OG4 ESC lines established simultaneously from syngeneic background (Supplemental Experimental Procedures), CiPSCs exhibited pluripotency, as shown by expression at similarly high levels of key pluripotency factors OCT4, NANOG, SOX2, and *Rex1*, and surface marker SSEA1 by immunofluorescence microscopy, RT-qPCR, and western blot (Figures 1B, 1D, and S1C). Expression levels of pluripotency genes did not differ in CiPSCs at various passages (passage 4 [P4] to P22), like those of ESCs (OG4) at P4–5. CiPSCs also expressed high alkaline phosphatase activity (Figure S1D).

These CiPSCs were able to differentiate into three embryonic germ layers by embryoid body formation *in vitro*, namely ectoderm, mesoderm, and endoderm (Figure 1E). We also tested the developmental pluripotency *in vivo* by injecting the CiPSCs into four- to eight-cell recipient albino embryos followed by embryo transfer. CiPSC1b and 7b cell lines and CiPSC2t and 6t cell lines at advanced passages efficiently generated chimeras by coat color (Figures 1F and 1G), but chimeras from these CiPSCs lines (n = 4 for BrdU method, and n = 10 for three-step method) failed to produce germline transmission following breeding with albino ICR mice for more than two rounds. Nevertheless, CiPSCs at earlier passages (P4 or P5) failed to form chimeras (Figure 1G).

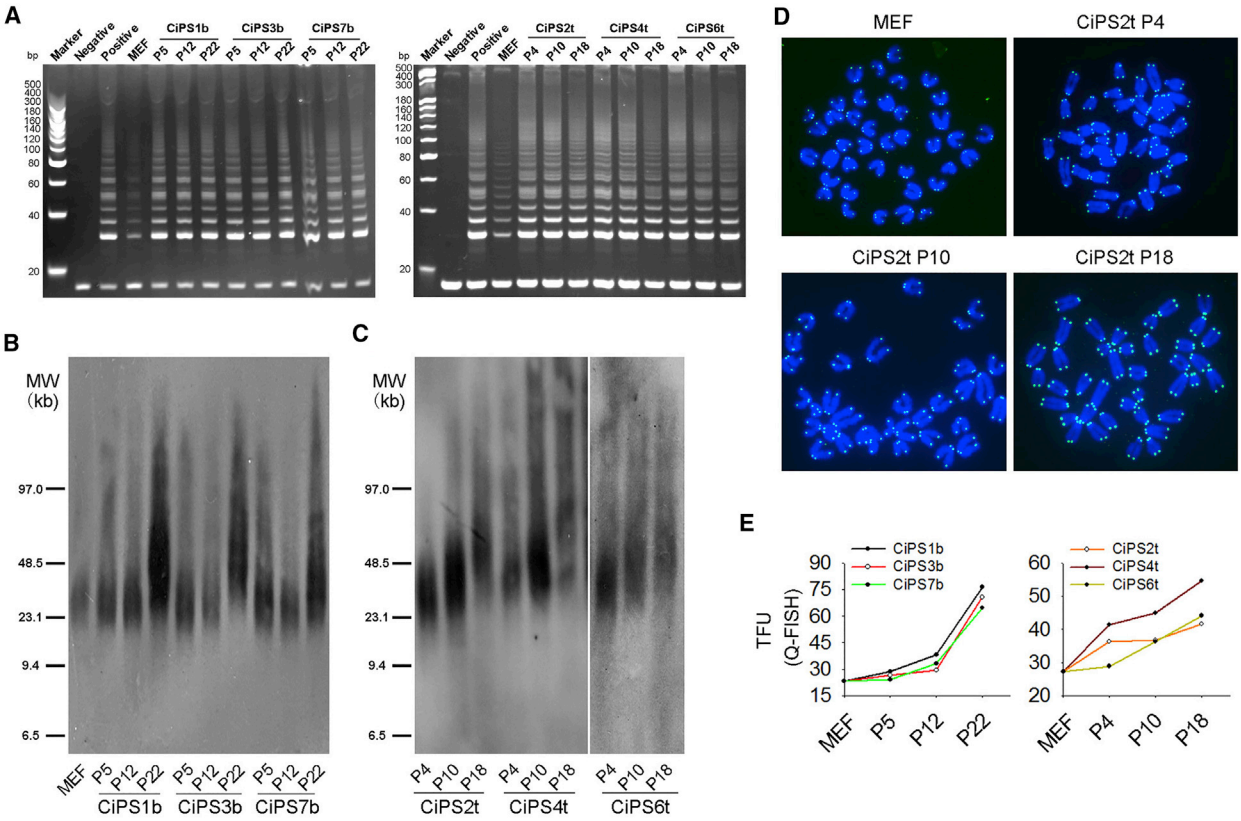
These results validate that the CiPSCs do exhibit pluripotency *in vivo* and differentiation capacity *in vitro*. However, acquisition of developmental pluripotency of CiPSCs depends on the extended passages.

Telomeres of CiPSCs Rejuvenate with Increasing Passages

To investigate whether telomeres are elongated and associated with pluripotency of CiPSCs with increasing passages, three CiPSC lines generated by either method were analyzed for their telomere lengths at various passages, P4∼P6 (arbitrarily considered as early passages), P10∼P12 (mid passages), and P18∼P22 (late or advanced passages). Key telomerase genes *Terc* and *Tert* were expressed at higher levels in all CiPSC lines than in MEFs, and comparable with those of ESCs regardless of passages (Figure S2A). Higher expression levels of telomerase genes corroborated with higher telomerase activity in CiPSCs compared with MEFs (Figure 2A). TRF1, *Pot1a/b*, *Rap1*, and *Tin2* components of the shelterin complex to safeguard robust telomeres (de Lange, 2005) also were upregulated in the CiPSCs (Figures S2B and S2C).

Moreover, telomeres lengthened gradually in the formed CiPSCs and elongated further with increasing passages, estimated by telomere quantitative fluorescent in situ hybridization (Q-FISH) and also validated by telomere restriction fragment (TRF) analysis (Figures 2B and S2D). Interestingly, CiPSCs generated by the BrdU method initially exhibited only minor telomere elongation from P5 to P12, comparable with that of MEFs, but achieved robust telomere elongation by continuous passage and culture to P22 (Figures 2B, 2E, and S2D). CiPSCs obtained by the three-step method showed longer telomeres at early

(G) Summary table showing percentage of chimeras generated from CiPSCs at different passages compared with OG4 ESCs. Chimeras (black and albino coat) were initially identified by coat color and some confirmed by microsatellite genotyping. CiPSC1b and 7b were generated using the BrdU method and CiPSC2t and the 6t by three-step method.
See also Figure S1.



(legend on next page)

passage (e.g., P4), relative to those of MEFs, and gradually elongated telomeres from P4 to P18 (Figure 2C). Telomeres were modestly extended in most CiPSCs at earlier passage CiPSCs (CiPSC1b, 3b, 7b at P5 and CiPSC2t and 6t at P4) in comparison with those of the precursor MEFs, except for one cell line, CiPS4t, that possessed moderately longer telomere at P4 (Figures 2C, 2E, and S2D). Telomere lengths and the trend in their elongation estimated by the Q-FISH method corresponded with those measured by TRF.

These data suggest that telomeres of somatic cells are rejuvenated in the formed CiPSCs following passage and culture *in vitro*. Achievement of developmental pluripotency of CiPSCs at advanced passages is closely associated with sufficiently elongated telomeres and this also can be suggestive of full reprogramming.

CiPSCs Express Two-Cell Genes, Including *Zscan4*

By RNA-sequencing (RNA-seq) analysis, genome-wide gene expression profile differed remarkably between CiPSCs and MEFs but was similar between CiPSCs and ESCs (Figure 2F). CiPSCs highly expressed naive state marker genes, such as *Zfp42/Rex1*, *Esrrb*, *Dppa3*, and *Prdm14*, like ESCs (Nichols and Smith, 2009; Valamehr et al., 2014) (Figure 2F). Furthermore, CiPSCs expressed *Zscan4* at higher levels but *Twist2*, *Thy1*, and *Fbn1* genes for lineage differentiation at lower levels, in contrast to MEFs (Figure 2F). *Zscan4* and other 2C-genes, including *MuERVL*, *Tcstv1*, and *Tcstv3*, were expressed at similarly high levels in CiPSCs and ESCs in contrast to MEFs (Figures 2F, 2G, and S2E). Proportion of Zscan4^{+} cells was comparable in CiPSCs and ESCs (P15) by both immunofluorescence microscopy and flow cytometry (Figures S2F–S2H). *Zscan4* and *MuERVL* are expressed specifically in two-cell (2C) embryos and transiently in sporadic ESCs (1%–5%) at any given time, marking a transient 2C state of mouse ESCs (Macfarlan et al., 2012; Zalzman et al., 2010). Telomeres lengthen rapidly in 2C embryos and ZSCAN4 elongates telomeres promptly by recombination-based mechanisms and maintains genomic stability (Liu et al., 2007; Zalzman et al., 2010).

The single-stranded telomeric (CCCTAA)n DNA circles (C-circles) assay detects recombination-based alternative lengthening of telomeres (ALT) (Henson et al., 2009, 2017). We showed that C-circle levels were high in CiPSCs and ESCs, comparable with U2OS serving as ALT control, but low or minimal in Hela as negative controls, and in MEFs (Figure 2H). In addition, telomere sister chromatid exchange (T-SCE) has been implicated as telomere recombination-based ALT (Bailey et al., 2004; Bechter et al., 2004; Liu et al., 2007; Londono-Vallejo et al., 2004; Zalzman et al., 2010). Like iPSCs generated by the conventional induction method (Wang et al., 2012), CiPSC lines also displayed increased frequency of T-SCE at a level similar to that of ESCs, in contrast to MEFs (Figures 2I and 2J). These data suggest that the telomere recombination-based ALT mechanism is involved in telomere elongation of CiPSCs.

Although telomerase is activated earlier during induction, it is unclear whether levels of telomerase activity indeed influence induction of CiPSCs. To understand this, we attempted to generate CiPSCs using a three-step method from telomerase *Terc*-deficient (knockout, *Terc*$^{-/-}$) MEFs, compared with heterozygous (HT) *Terc*$^{+/-}$ and wild-type (WT) MEFs. CiPSC primary colonies were obtained, and the efficiency of induction did not differ among *Terc*$^{-/-}$, *Terc*$^{+/-}$, and WT MEFs, based on alkaline phosphatase activity assay (Figures S3A and S3B). These data suggest that telomerase is dispensable for induction of CiPSCs. Despite lack or haploinsufficiency of telomerase activity, these CiPSCs still expressed pluripotency genes (e.g., *Oct4*, *Sox2*, *Nanog*, *Zscan4*, and SSEA1) at high levels, as did WT CiPSCs (Figures S3C–S3G). However, the telomerase-deficient CiPSCs showed telomere shortening, telomere loss, and chromosomal fusions, compared with

Figure 2. Telomere Rejuvenates in CiPSCs with Passages

(A) Telomerase activity by telomeric repeat amplification protocol (TRAP) assay of CiPSCs during passages compared with progenitor cell MEFs. Lysis buffer served as negative control and OG4 ESCs as positive control.

(B and C) Telomere length distribution shown as TRF by Southern blot analysis of MEFs and CiPSCs following passages.

(D) Representative images displaying telomere FISH of MEF and CiPSCs at various passages. Blue, chromosomes stained with DAPI; green dots, telomeres.

(E) Line progression plot showing dynamics of relative telomere length by Q-FISH shown as telomere fluorescence unit (TFU) of MEFs and CiPSCs for each cell line from early to late passage. Left, BrdU method; right, three-step method.

(F) Scatterplots showing comparison of genome-wide transcription profile between CiPS1b or CiPS2t and ESC (OG4) or MEF. Parallel diagonal lines indicate 2-fold threshold in expression difference. Red, upregulated genes; green, downregulated genes.

(G) ZSCAN4 protein levels by western blot analysis. MEF and ESCs (OG4) served as negative and positive controls, respectively. β-ACTIN served as loading control.

(H) C-circle assay. Hela and U2OS served as negative and positive controls, respectively.

(I) Representative micrographs showing telomere sister chromatid exchange (T-SCE, white arrows) by chromosome orientation FISH analysis.

(J) Frequency of T-SCE. n, number of spread counted. ***$p < 0.001$, compared with MEFs. Two independent experiments.

See also Figures S2 and S3.

WT CiPSCs generated simultaneously (Figures S3H and S3I), indicating that telomerase is essential for telomere maintenance of CiPSCs. This is also consistent with human iPSC induction, where telomerase gene mutation causes telomere shortening during induction, but activation of telomerase *TERC* by OCT4 acquires telomere rejuvenation and maintenance during passages after iPSC formation (Agarwal et al., 2010). Collectively, telomeres rejuvenate in CiPSCs following culture *in vitro* by passaging. Both telomerase and telomere recombination-based mechanism are required for telomere elongation in CiPSCs, like ESCs.

CiPSCs with Longer Telomeres Exhibit Higher Differentiation Potential

Sufficiently elongated telomeres may indicate the achievement of full reprogramming and differentiation capacity of CiPSCs. To test whether telomere lengths indicate differentiation capacity of CiPSCs *in vivo*, we performed a teratoma formation test by injection of CiPSCs at early and advanced passages under the skin of immunodeficient mice (Maherali and Hochedlinger, 2008). CiPSCs derived by either the BrdU or three-step method had comparatively short telomeres at early passages, slightly longer than those of progenitor cells, and achieved longer telomeres at relatively late passages (Figures 2B–2E). Remarkably, teratomas derived from CiPSCs at late passages with longer telomeres were larger and heavier than those of CiPSCs at early passages with shorter telomeres (Figures 3A and 3B). Linear regression analysis showed high correlation between telomere length and teratoma weight of CiPSCs ($R^2 = 0.662$ for BrdU method and $R^2 = 0.526$ for three-step method) (Figure 3C).

Moreover, although the representative three embryonic germ layers, including epidermis and neural ectoderm, muscle (mesoderm), and gland epithelium (endoderm), were found in the teratomas formed from CiPSCs, the structures of the characteristic germ layers generally were smaller from CiPSCs at early passages than those of CiPSCs at late passages (Figure 3D). Immunofluorescence microscopy further validated that these teratomas expressed molecular markers representative of three germ layers, including NESTIN (ectoderm), smooth muscle actin (SMA) (mesoderm), and alfa-fetoprotein (AFP) (endoderm). However, these germ layers developed more extensively in teratomas formed from CiPSCs at late passages than did those at early passages (Figure 3E). Hence, by both histology and immunofluorescence microscopy, CiPSCs that acquire telomere lengthening and longer telomeres show higher differentiation potential *in vivo*.

Telomere Shortening during Late Stage of Lengthy Induction by Small Molecules

To understand telomere changes during long period of chemical induction, we collected cell samples on days 0 (MEFs), 15, 30, and 40 and compared them with CiPSCs at early passage, which served as positive control, and analyzed expression of genes important for pluripotency and telomerase, telomerase activity, and telomere lengths. Formation of the primary CiPSC colonies can be obtained 40 days after induction by complete small molecules, consistent with previous reports (Long et al., 2015; Zhao et al., 2015). The MEFs emerged as aggregates and achieved a XEN-like state by expression of marker genes *Gata4*, *Gata6*, and *Sall4* 15 days following induction (Figures 4A and S4A), in agreement with the findings that the chemical reprogramming process requires the early formation of XEN-like cells (Zhao et al., 2015). By the BrdU method, mRNA levels of *Oct4*, *Sox2*, and *Zscan4* slightly increased by days 30 and 40, and *Nanog* was not readily detectable. By the three-step method, *Oct4*, *Sox2*, *Nanog*, and *Zscan4* were expressed at variable levels but much lower than those of the formed CiPSCs (Figure 4B). Western blot confirmed that protein levels of OCT4, NANOG, SOX2, and ZSCAN4 were minimal or undetectable during induction, in contrast to the formed CiPSCs after primary clonal selections (Figure S4B). Notably, expression levels of telomerase genes *Tert* and *Terc* were already elevated by day 15 of induction and continued to increase to day 30, like those of the formed CiPSCs (Figure 4C). Thus telomerase genes are activated earlier than the pluripotent genes *Oct4*, *Sox2*, and *Nanog* (Figures 4B and 4C). The dynamics are consistent with those of conventional iPSC induction despite much less time required by OSKM-induced reprogramming (Wang et al., 2012).

However, expression levels of both *Tert* and *Terc* declined by day 40 of induction (Figure 4C). Correspondingly, telomerase activity was increased at day 15 and high at day 30 but decreased by day 40 (Figure 4D). In addition, TRF1 protein expression was gradually increased but also slightly reduced by day 40. TRF1 expression is increased during iPSC induction by the conventional method (Schneider et al., 2013). Moreover, levels of histone H3K9ac, H3K9me3, and H3K27me3 were increased at early stage of chemical induction, and H3K9ac and H3K27me3 were reduced at day 40 (Figure S4B).

Telomeres maintained and elongated during the early to mid stage of chemical induction (days 0–30) (Figure 4E) coincided with increased telomerase activity (Figure 4D). Unexpectedly, telomeres shortened during the late stage of CiPSC induction (days 30–40), when the telomerase activity also was decreased (Figure 4D). These outcomes prompted us to investigate whether reprogramming cells undergo telomere damage after prolonged induction by examining telomere dysfunction-induced damage foci (TIFs) indicated by co-localized foci of phosphorylated H2AX with TRF1 at telomeres (Dan et al., 2014; Takai et al., 2003). Indeed, co-localized foci of γH2AX and TRF1

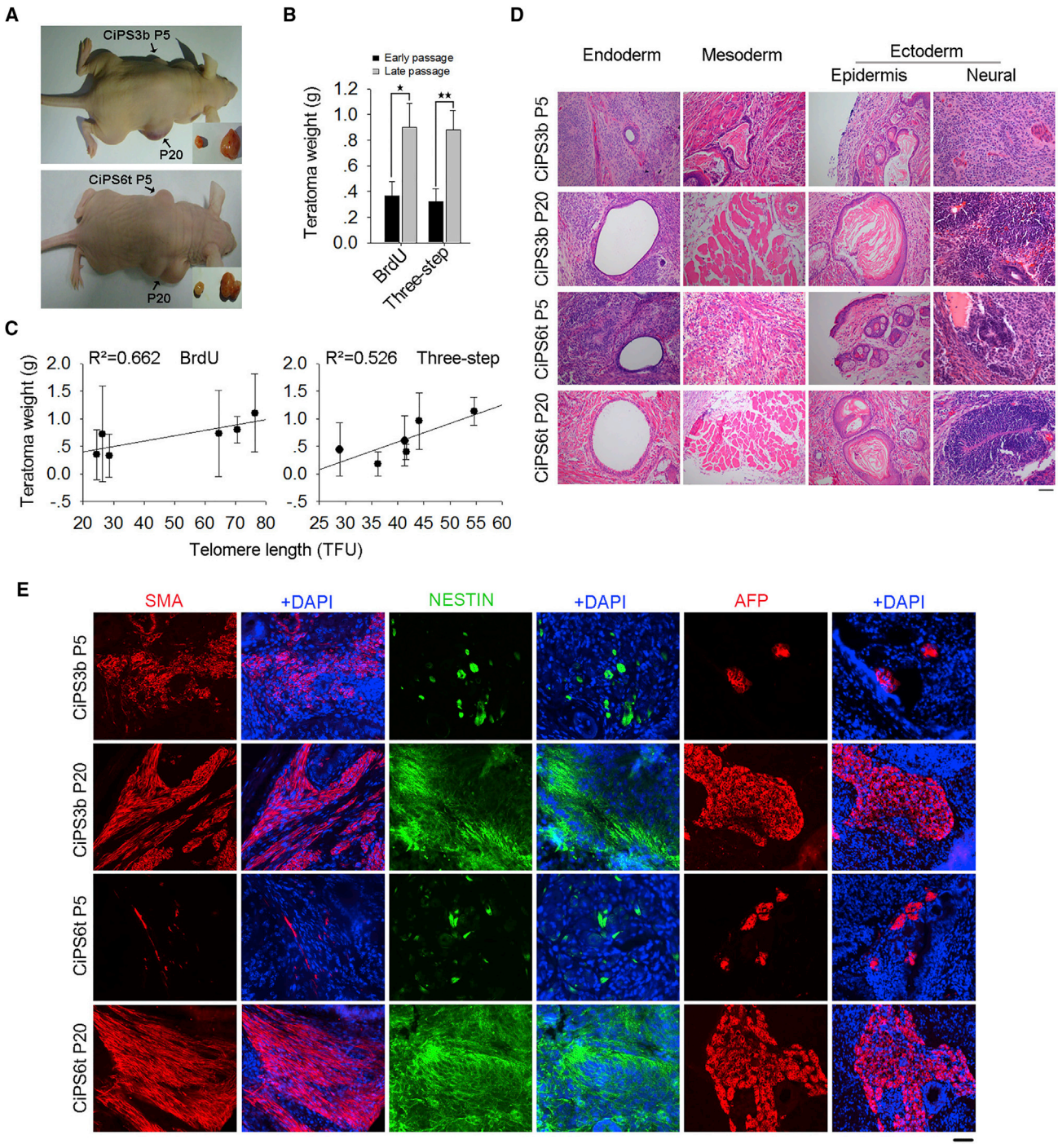


Figure 3. Telomere Length Indicates Differentiation Capacity by Teratoma Formation

(A) Differentiation *in vivo* of CiPSCs at early and late passages by teratoma formation test.

(B) Weight of teratomas formed from CiPSCs, Mean ± SEM (n = 6 mice). *p < 0.05, **p < 0.01.

(C) Linear regression analysis showing high correlation between telomere length (TFU by Q-FISH) and teratomas weight of CiPSCs.

(D) Histology by H&E staining of teratoma tissues derived from early and late passage CiPSCs. Scale bar represents 50 μm.

(E) Immunofluorescence of the teratomas showing markers representative of three germ layers: NESTIN (ectoderm), SMA (mesoderm), and AFP (endoderm). Scale bar represents 100 μm.

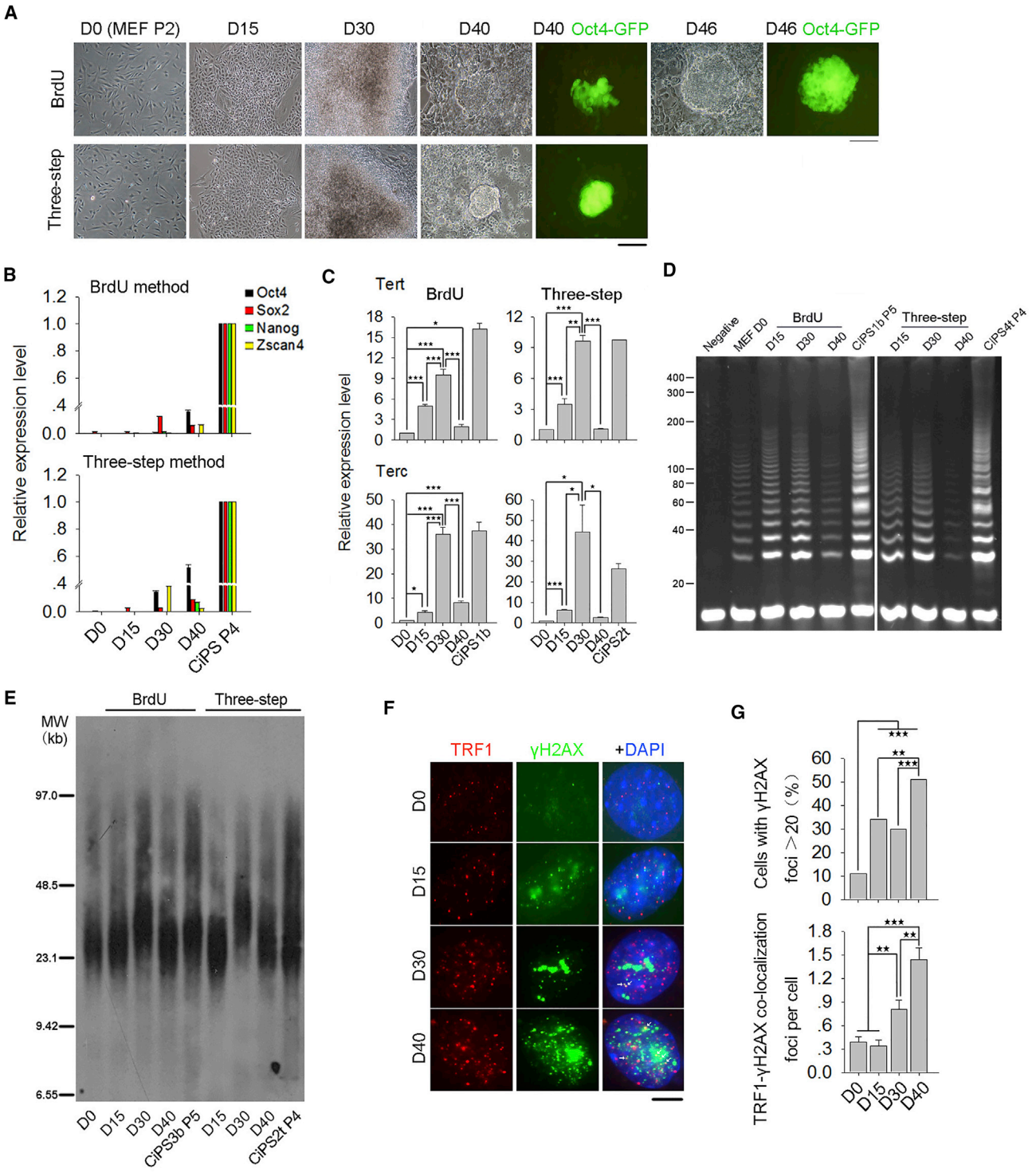


Figure 4. Telomere Dynamics during CiPSC Induction

(A) Morphological changes of MEFs during chemical induction under bright field with phase contrast optics or Oct4-GFP fluorescence. Compact colonies were formed at day 40 of induction by three-step method, earlier than with the BrdU method. Scale bar represents 100 μm.

(B) mRNA expression levels of *Oct4*, *Sox2*, *Nanog*, and *Zscan4* during chemical induction.

(C) Relative expression levels of *Tert* and *Terc* during chemical induction.

(legend continued on next page)

were increased by day 40, and the proportion of the cells exhibiting more than 20 γH2AX foci also was elevated (Figures 4F and 4G). Thus, DNA and telomere damage may occur during the late stage of extended induction.

Few apoptotic cells were found on day 0 and day 15, but number of apoptotic cells with fragmented nuclei increased from day 30 to day 40 (Figures S4C and S4D). These data suggest that reduced telomerase, increased telomere damage, and apoptosis together may be responsible for the telomere erosion during prolonged periods of chemical induction, and this can limit CiPSC formation.

Histone Crotonylation Activates *Zscan4* and Maintains Telomeres during Chemical Induction

Zscan4 is not effectively activated during chemical induction until in the formed CiPSCs following passages (Figures 4B and S4B), like conventional iPSC induction (Wang et al., 2012). Forced expression of *Zscan4*, in combination with Yamanaka factors, greatly reduces DNA damage response, and increases telomere elongation and quality of iPSCs (Jiang et al., 2013). Heterochromatic histone assembly, such as HP1 and H3K9me3 at telomeres/sub-telomeres, can regulate *Zscan4* expression (reviewed by Liu, 2017). We searched for a variety of histone modification by small molecules, including trichostatin A, valproic acid, and sodium butyrate, that can activate *Zscan4* expression in mESCs (Dan et al., 2014, 2015), but none of them can effectively activate *Zscan4* during chemical induction (data not shown). Excitingly, we found that a chemical compound, crotonic acid (Figure S5A), which mediates lysine crotonylation (Kcr) (Tan et al., 2011; Wei et al., 2017a, 2017b), can induce protein epigenetic modification and force *Zscan4* expression. Initially, mESCs were treated with 5 mM or 10 mM crotonic acid (Figure S5B). 2C genes, including *Zscan4*, *Tcstv1*, and *Tcstv3*, were rapidly upregulated following two passages (Figures S5C and S5D), while genes for pluripotency (*Oct4*, *Sox2*, and *Nanog*) and telomerase *Tert* were downregulated (Figures S5C and S5D). Proportion of Zscan4$^+$ cells also was increased (Figure S5E). Moreover, level of histone lysine crotonylation (Kcr), rather than other histone modifications (e.g., H3K27me3, H3K9me3, or H3K9ac), was elevated (Figures S5D and S5F).

Further, we tested whether crotonic acid could also activate *Zscan4* during chemical induction (Figure 5A). Initially, MEFs were treated with crotonic acid or vehicle controls for the whole period of induction using the three-step method because of its higher induction efficiency (Figure S1B), and were examined for colonies positive for Oct4-GFP fluorescence by the end of induction. Number of Oct4-GFP-positive colonies was not affected by crotonic acid. Then, we treated MEFs with crotonic acid at various periods, illustrated in Figure 5A, and observed that only the treatment for the period of stage II (days 16–28) effectively promoted Oct4-GFP colony formation. The Oct4-GFP colonies increased 3-fold in crotonic acid-treated-cells compared with non-treated cells (Figures 5A, 5B, S6A, and S6B). Expression of *Oct4*, *Nanog*, *Sox2*, *Lin28*, and *Terc* and *Tert* was not affected by crotonic acid during induction (Figures 5F and S6C), consistent with telomerase activity (Figure 5C). Remarkably, addition of crotonic acid from day 16 resulted in ZSCAN4 activation by day 28 (Figures 5D–5F and S6D). Histone lysine crotonylation (Kcr) was increased by crotonic acid treatment, while H3K9ac, H3K27me3, and H3K9me3 levels were not changed (Figures 5F and S6D). 2C genes TCSTV1/3, which can elongate telomeres in mESCs (Zhang et al., 2016), also were upregulated after crotonic acid treatment (Figure 5F).

Consistently, addition of crotonic acid at stage II rapidly lengthened telomeres following *Zscan4* upregulation, compared with the control, and maintained telomeres at the late stage of induction (Figures 5G and 5H). In addition, number of γH2AX-positive cells did not differ at stage II between crotonic acid-treated and non-treated cells. However, co-localization foci (TIFs) of γH2AX and TRF1 were much reduced on day 28 by crotonic acid (Figure 5I).

These data suggest that addition of crotonic acid at stage II (days 16–28) promotes CiPSC generation. Crotonic acid activates *Zscan4*, reduces telomeric damage, and maintains telomere length during chemically induced reprogramming.

Crotonic Acid Activates 2C Genes and Increases T-SCE during Chemical Induction

To better understand the mechanisms underlying crotonic acid in facilitating chemically induced reprogramming, we

(D) Telomerase activity by TRAP assay of MEFs and reprogramming cells during induction, compared with CiPSCs at P4 or P5. Lysis buffer as negative control.

(E) Telomere length distribution shown as TRF by Southern blot analysis during chemical induction (day 0–40) compared with CiPSCs. MW, molecular weight.

(F) Immunofluorescence showing co-staining of γH2AX (green) at telomeres (TRF1, red). Yellow (white arrows) in merged images indicates co-localized foci of TRF1 and γH2AX. Scale bar represents 5 μm.

(G) Quantification of TIFs. n = 100 cells counted.

Data represent mean ± SEM from three independent experiments. *p < 0.05, **p < 0.01, ***p < 0.001.

See also Figure S4.

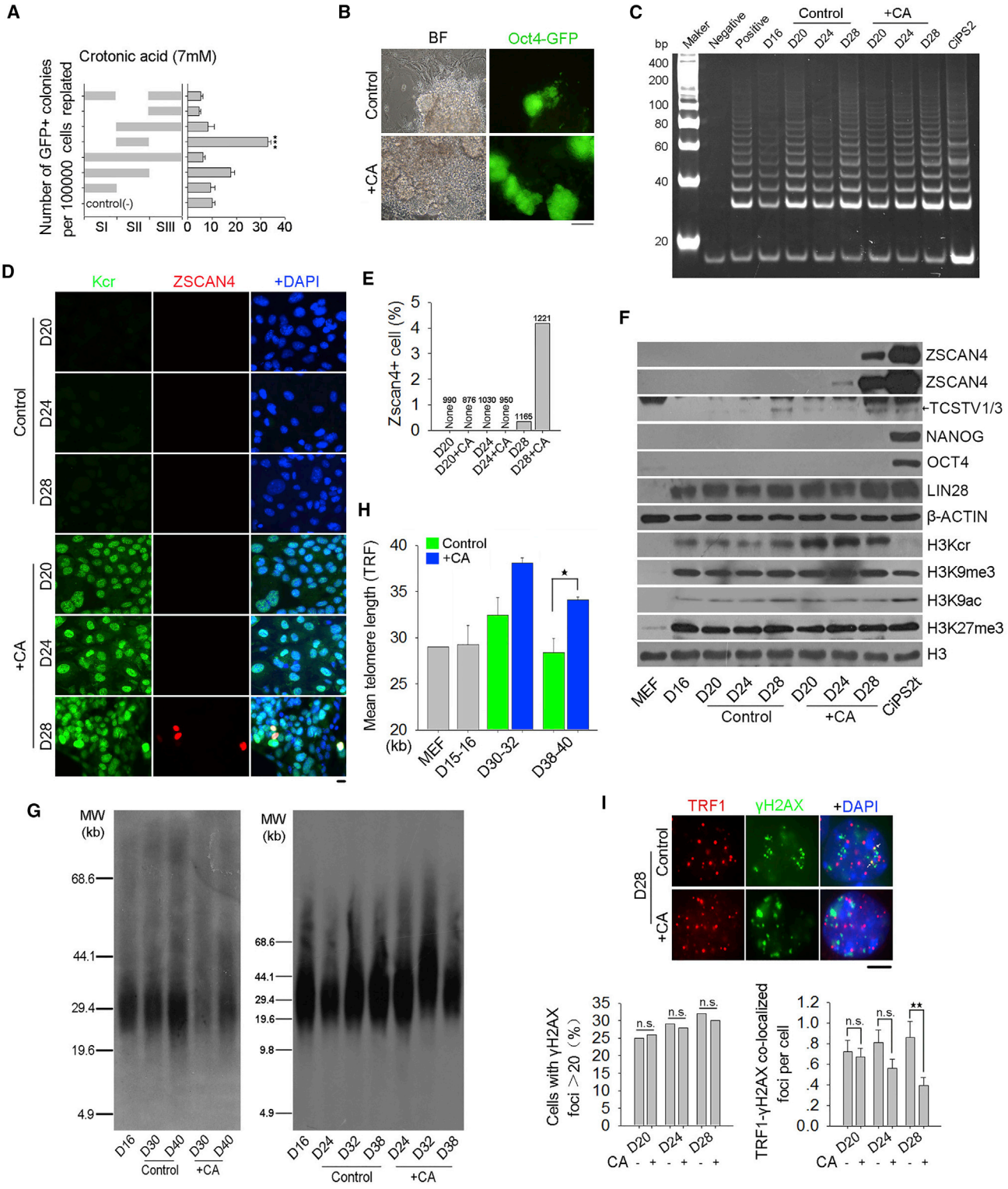


Figure 5. Crotonic Acid Addition at Stage II Activates *Zscan4*, Maintains Telomeres, and Promotes CiPSC Generation

(A) Durations of crotonic acid (7 mM) addition at different time points during chemical induction and assessment of Oct4-GFP-positive colonies on day 40. Left panel shows different time points during reprogramming. Stage I (SI) represents days 0–15; stage II (SII)

(legend continued on next page)

performed RNA-seq analysis of cell samples by days 20, 24, and 28 without or with crotonic acid treatment. More genes were upregulated than downregulated in crotonic acid-treated cells compared with controls (Figure 6A). Numbers of differentially expressed genes gradually increased from day 20 to day 28 following induction (Figure 6A). These genes were enriched in regulation of cell differentiation, cell proliferation, and apoptosis (Figure 6B). RNA-seq data also verified that *Zscan4* was prominently upregulated at day 28 following addition of crotonic acid on day 16 (Figures 6A and 5D–5F).

We compared all 196 differentially expressed genes with single-cell RNA-seq data of mouse preimplantation embryos (Fan et al., 2015). Interestingly, by clustering of the differentially expressed genes to those expressed at cleavage stage of early embryo development, these genes were significantly enriched in the 2C (2C) embryo ($p = 1.90 \times 10^{-5}$), compared with other early embryonic development stages from oocyte to inner cell mass (ICM) of blastocysts (Figure 6C). Most of the 36 2C genes overlapped with the single-cell RNA-seq data were upregulated during induction by day 24 and noticeably by day 28 (Figure 6D). Additional comparison of our RNA-seq data with those of 2C genes (Macfarlan et al., 2012) substantiated that 2C genes were gradually upregulated following addition of crotonic acid during stage II of chemical induction (Figure S6E). The differentially expressed 2C genes and their gradual activation were further confirmed by qPCR (Figure S6F). However, crotonic acid reduced cell proliferation during the stage II reprogramming period (Figure S6G). To test whether crotonylation regulates local heterochromatic distribution at 2C gene regions, we performed a chromatin immunoprecipitation (ChIP)-qPCR experiment using H3K9me3 and HP1α antibodies and the primers specific for the promoter-proximal regions of *Zscan4* loci, subtelomere of chromosome 13, and the *ERVK10c* locus (Dan et al., 2014, 2015). Crotonylation decreased levels of H3K9me3 and HP1α at these loci following treatment with crotonic acid for 12 days (Figure 6E).

Moreover, crotonic acid increased the frequency of T-SCE at a level similar to that of CiPSCs, compared with untreated reprogramming cells at day 28 (Figures S6H and S6I). Addition of crotonic acid in CiPSCs further increased the incidence of T-SCE (Figures S6H and S6I). Increased ZSCAN4 expression levels are associated with higher frequency of T-SCE.

These data show that crotonic acid activates a series of 2C genes in the middle of chemical induction, and, together with crotonylation-mediated *Zscan4* activation, increases T-SCE and telomere maintenance, and promotes CiPSC generation (Figure 6F).

Crotonic Acid-Facilitated CiPSCs Possess High Pluripotency

The question was whether CiPSCs generated by addition of crotonic acid exhibit normal pluripotency. We characterized the CiPSCs derived by addition of crotonic acid (CaCiPSCs). Continuous passaging of the CaCiPSC primary colonies achieved stable CiPSC lines with Oct4-GFP fluorescence resembling typical ESC colonies in morphology, with large nuclei and nucleoli and clear, compact clonal boundaries (Figure 7A). Importantly, the majority of CaCiPSCs maintained normal karyotypes (Figure S7A). Moreover, CaCiPSCs expressed multiple pluripotent markers (*Oct4*, *Nanog*, *Sox2*, *Lin28*, *Zscan4*, and SSEA1) and also the naive state markers *Stella*, *Prdm14*, *Rex1*, and *Tbx3* (Nichols and Smith, 2009; Valamehr et al., 2014), comparable with those of ESCs (Figures S7B–S7D).

RNA-seq analysis revealed that the transcriptome profile of CaCiPS1t resembled that of ESCs (OG4), with a correlation coefficient (R^2) of 0.979 but was quite different from that of MEF (Figure 7B). Expression levels of pluripotency genes, naive marker genes, and *Zscan4* did not differ between CaCiPSCs and ESCs. Expression profile of CaCiPS1t and CaCiPS4t also was very similar to that of CiPS2t, and expression levels of pluripotency genes and naive marker genes did not differ between CaCiPSCs and CiPSCs (Figure 7B). All CiPSCs clustered with ESCs but not with the

represents days 16–28; stage III (SIII) represents days 29–40. Right panel shows number of Oct4-GFP-positive colonies on day 40; 100,000 cells were re-plated on day 12 following induction. ***$p < 0.001$, compared with control.

(B) Morphology by phase contrast optics and Oct4-GFP fluorescence of primary CiPSC colonies on day 40 induced with or without crotonic acid (CA) at stage II. Scale bar represents 100 μm.

(C) Telomerase activity by TRAP assay. Lysis buffer and ESCs served as negative and positive controls, respectively.

(D) Immunofluorescence microscopy of ZSCAN4 and Kcr (lysine crotonylation) expression. Scale bar represents 10 μm.

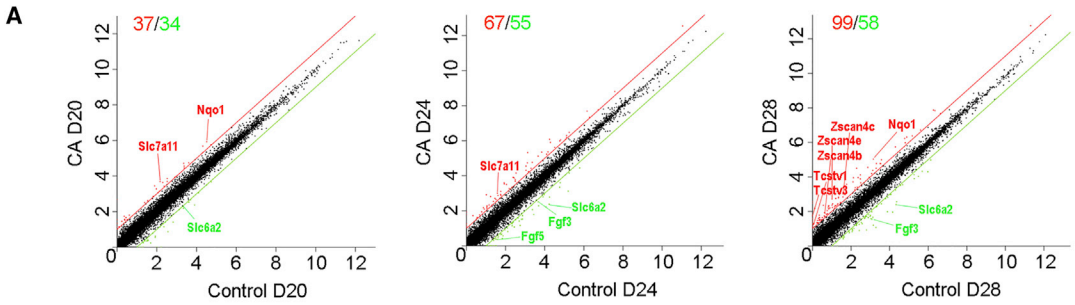
(E) Percentage of Zscan4-positive cells. Number of cells counted is shown above the bar.

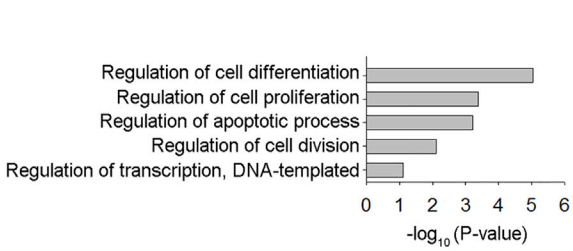
(F) Western blot analysis of protein levels. MEF and CiPSCs served as controls.

(G) Telomere length distribution shown as TRF by Southern blot analysis.

(H) Quantification of telomere length by TRF in three repeats using TeloTool (Gohring et al., 2014). *$p < 0.05$.

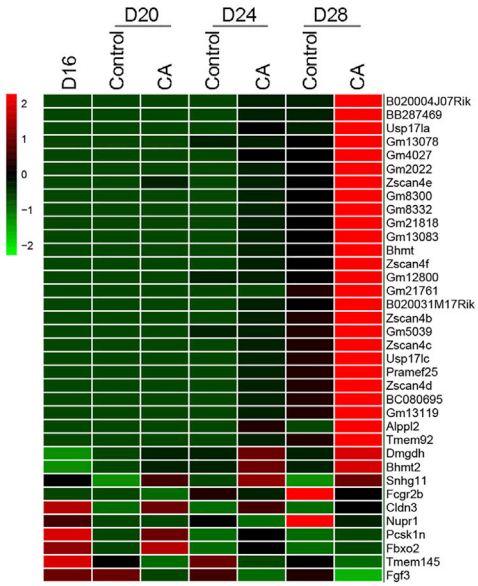
(I) Immunofluorescence of γH2AX (green) at telomeres (TRF1, red). Yellow foci in the merged images indicate TRF1-γH2AX co-localization. Scale bar represents 5 μm. Bottom panel, quantification of TIFs. n = 100 cells counted. **$p < 0.01$, n.s., not significant. Data represent mean ± SEM from three independent experiments. See also Figure S5.

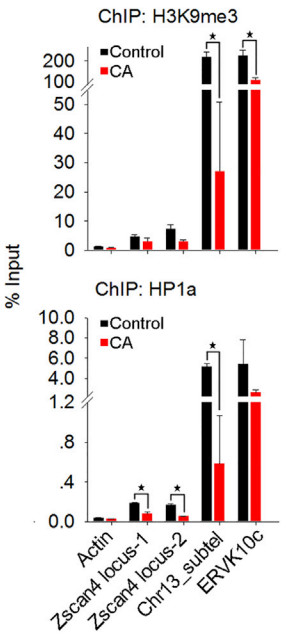


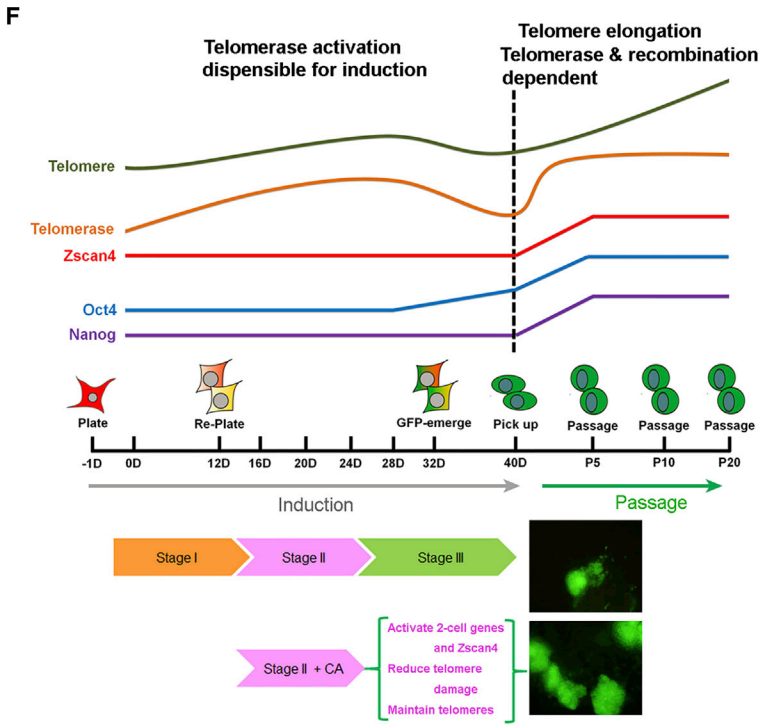


C

Stage	Oocyte VS Zygote	Zygote VS 2 cell embryo	2 cell embryo VS 4 cell embryo	4 cell embryo VS 8 cell embryo	8 cell embryo VS Morula	Morula VS ICM
P value	5.83e-01	1.90e-05	5.83e-01	1.45e-02	3.18e-01	3.96e-01







(legend on next page)

progenitor MEF cells (Figure 7C). Moreover, two CaCiPS cell lines, CaCiPS4t and CaCiPS1t, preferentially clustered with ESC OG4, relative to CiPS2t and CiPS1b (Figure 7C).

Furthermore, we tested the developmental potency by injection of CaCiPSCs into four- or eight-cell embryos from Balb/c albino mice. Chimeras were obtained with high efficiency by coat color (Figures 7D and 7E). However, cell lines varied in their contribution to chimeras. For instance, CaCiPSC2t generated chimeras at a very low rate. Chimerism derived from CaCiPSC4t was high based on coat color and microsatellite genotyping (Figures 7E and S7F). Although CaCiPSCs can contribute to gonad in the chimeras (Figure S7F), they also failed to produce germline transmission following breeding of chimeras (n = 10) with albino ICR mice for more than two rounds.

We also measured telomere lengths of CaCiPSCs at early and advanced passages. We picked up primary CaCiPSC and CiPSC colonies derived at the same time from the same batch of OG2-MEFs and passaged them *in vitro*. Telomere lengths varied among the cell lines but generally were not elongated much at P5 compared with MEFs (Figure S7E), and this may indicate heterogeneity of primary CiPSC clones. By Q-FISH, telomere lengths also were indistinguishable among CiPS1t, CiPS2t, CaCiPS1t, and CaCiPS2t at P5 (Figures 7F and 7G). Nevertheless, telomeres lengthened significantly in both CaCiPS cell lines with increasing passages from P5 to P15 (CaCiPS1t, from 46.00 ± 15.98 to 70.36 ± 23.13 telomere fluorescence units [TFU]; CaCiPS2t, from 42.18 ± 18.17 to 63.49 ± 19.08 TFU) (Figure 7G). Hence, CaCiPSCs manifested high pluripotency, in association with telomere rejuvenation with passages.

To further explore the effect on genome stability of BrdU or crotonic acid addition during chemical reprogramming, we performed the whole-exome sequencing analysis. We compared BrdU-treated MEF cells with control MEF cells following culture for 12 days simultaneously and did not observe somatic point mutations in protein-coding regions (Tables S4 and S5). Moreover, we characterized the protein-coding region of six CiPSC lines by three induction methods that generated chimeras at single-nucleotide resolution. The number of point mutations of these CiPSCs was similar to what had been previously reported for iPSCs (Figure S7G) (Gore et al., 2011). The CiPS1b and CiPS7b (BrdU method) cell lines had slightly more point mutations in coding regions (13 and 29) than did CiPS2t and CiPS6t (three-step method, 12 and 9) and CaCiPS2t and CaCiPS4t (crotonic acid method, 17 and 9) lines (Figure S7G, Tables S4 and S5). These mutated genes, however, did not cluster in a specific functional pathway.

DISCUSSION

We show that telomeres of CiPSCs elongate with passages after clonal formation. Telomere elongation with increasing passage depends on both telomerase and recombination-based ALT mechanisms (Figure 6F). Moreover, CiPSCs express 2C genes including *Zscan4*, *Tcstv1/3*, and *MuERVL*, like ESCs. Successful generation of CiPSCs further supports the notion that reprogramming and rejuvenation of telomeres are intimately linked to somatic reprogramming to induction of pluripotency. However, telomere damage occurs and telomeres shorten at late stage during prolonged chemical induction. Crotonylation induced by crotonic acid activates 2C genes, notably *Zscan4*, and increases T-SCE, which maintains telomeres and reduces telomere damage during chemical induction, improving induction efficiency (Figure 6F).

Currently, the low efficiency in achieving Oct4-GFP fluorescence colonies and particularly the lengthy induction still limit CiPSC generation. *Oct4* expression levels are very low during induction and can only be seen at late stage by qPCR analysis. OCT4 protein levels are too low to be detected by western blot during induction compared with the formed CiPSCs. Importantly, *Oct4* distal enhancer

Figure 6. RNA-Seq Reveals Crotonic Acid-Induced Activation of 2C Genes

(A) Scatterplots showing global differential gene expression profile (fold ≥ 2.0) of reprogramming cells on days 20–28 following treatment with CA in comparison with controls. Red, upregulated genes; green, downregulated genes in CA-treated cells. Average from two independent experiments.

(B) Gene ontology analysis of differentially expressed genes in cells treated with crotonic acid, compared with controls.

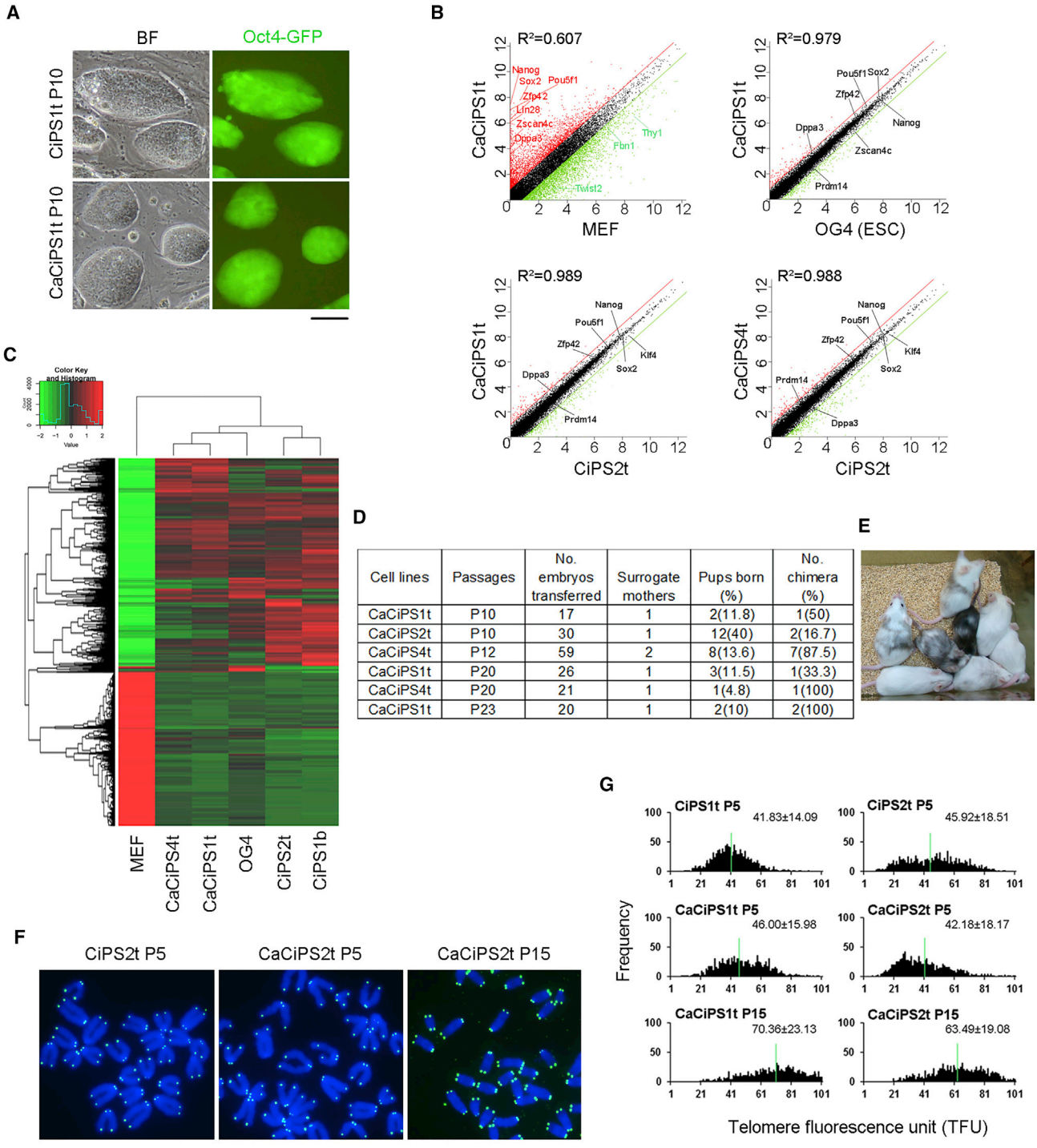
(C) Fisher's exact test between relatively upregulated differential genes at early embryo developmental stages from single-cell RNA-seq data (Fan et al., 2015) and all 196 differential genes from our RNA-seq data (p adjusted < 0.01).

(D) Heatmap highlighting highly expressed 2C genes found in CA-treated cells compared with controls.

(E) ChIP-qPCR analysis of H3K9me3 and HP1α abundance at *Zscan4*, *Chr13_subtel* (sub-telomere region of no. 13 chromosome), and *ERVK10c* loci following treatment with crotonic acid for 12 days during stage II of induction. β-*actin* served as negative control. Data represent mean ± SD from three independent experiments. *p < 0.05.

(F) Schematic diagram illustrating the process of CiPSC induction following addition of crotonic acid, which activates 2C genes including Zscan4, reduces telomere and DNA damage, and maintains telomeres, promoting CiPSC formation.

See also Figure S6.



Cell lines	Passages	No. embryos transferred	Surrogate mothers	Pups born (%)	No. chimera (%)
CaCiPS1t	P10	17	1	2(11.8)	1(50)
CaCiPS2t	P10	30	1	12(40)	2(16.7)
CaCiPS4t	P12	59	2	8(13.6)	7(87.5)
CaCiPS1t	P20	26	1	3(11.5)	1(33.3)
CaCiPS4t	P20	21	1	1(4.8)	1(100)
CaCiPS1t	P23	20	1	2(10)	2(100)

Figure 7. Generation and Characterization of CaCiPSCs Induced by Addition of Crotonic Acid at Stage II

(A) Morphology under bright field with phase contrast optics and Oct4-GFP fluorescence of CiPSCs generated without CA and CaCiPSCs with CA. Scale bar represents 100 μm.

(B) Scatterplots comparing transcriptome profile of CaCiPS1t from CA-treated cells, CiPS2t without CA treatment, OG4 ESCs, and MEFs. Parallel diagonal lines indicate 2-fold threshold.

(C) Hierarchical clustering of transcriptome profile of MEFs, OG4 ESCs, CiPS1b, CiPS2t, CaCiPS1t, and CaCiPS4t.

(legend continued on next page)

activation indicates achievement of sufficient reprogramming and is essential for achieving a naive pluripotency network (Bao et al., 2009; De Los Angeles et al., 2015; Gafni et al., 2013). Notably, telomerase activates earlier than do pluripotency genes *Oct4* and *Sox2*. Early activation of telomerase might result from CHIR99021 used at high concentration (seven times more than routinely used), which activates Wnt/β-catenin. Wnt/β-catenin signaling in turn can directly activate telomerase (Hoffmeyer et al., 2012). Our data suggest that *Oct4* could be expressed in only a few cells, revealed by the distal Oct4-GFP fluorescence during the late stage of chemical induction, such that the protein level in a cell population is minimal by western blot. Thus, timing to activate *Oct4* becomes essential for further enhancing CiPSC induction.

Also, *Nanog* is not activated well during chemical induction until CiPSC clonal formation and passages (Figure 6F). Nanog mediates acquisition of both embryonic and induced pluripotency (Silva et al., 2009). Moreover, biallelic expression of *Nanog* is required in the timely maturation of the ICM into a fully reprogrammed pluripotent epiblast (Miyanari and Torres-Padilla, 2012). Inhibiting Tgf-β signaling promotes the completion of reprogramming through induction of *Nanog* (Ichida et al., 2009), and *Nanog* overexpression can accelerate reprogramming (Hanna et al., 2009). Presumably, chemicals that can activate *Nanog* earlier likely increase CiPSC induction.

Unexpectedly, telomeres shorten, coincident with telomere damage, and incidence of apoptosis also is increased during the late stage of lengthy induction by pure chemicals (Figure 6F). Fewer cells are likely to be selected and survive during reprogramming induction. This is in contrast to conventional iPSCs induced by Yamanaka factors, by which telomeres continue to elongate or at least maintain their length without shortening during the induction period (Wang et al., 2012). Thus, chemicals that can alleviate telomere damage and shortening during the late stage of induction would enhance CiPSC generation.

Activation of *Zscan4* during induction indeed reduces telomere damage and improves iPSC induction and quality (Hirata et al., 2012; Jiang et al., 2013). We find that crotonylation facilitated by crotonic acid can activate 2C genes and increase T-SCE during induction and that increased expression of ZSCAN4 reduces telomere damage and improves telomere maintenance, increasing CiPSC induction efficiency (Figure 6F).

Our data suggest that crotonic acid activates *Zscan4* and 2C genes, likely by reducing heterochromatic histones (e.g., H3K9me3 and HP1α) at telomeres/subtelomeres. This is consistent with the findings that histone crotonylation accumulates with HP1γ (Tan et al., 2011) and reduces HP1α at heterochromatin (Wei et al., 2017b), and reduction of H3K9me3 leads to activation of 2C genes at subtelomeres; e.g., *Zscan4* and *Tcstv1* (Dan et al., 2014, 2015). More experiments are needed to understand whether crotonylation directly or indirectly regulates the processes. Additionally, histone crotonylation has critical and broad function as histone acetylation in transcription (Wei et al., 2017a). Moreover, crotonylation improves CiPSC clone formation at stage II during middle induction, implying that XEN state is primed for action by crotonic acid. Continuous treatment with crotonic acid is not needed after *Zscan4* activation. Likewise, forced expression of *Zscan4* is required only for the first few days of conventional iPSC formation (Hirata et al., 2012).

Telomere maintenance during induction is one of the key elements for successful reprogramming to pluripotency. Histone crotonylation by crotonic acid is one example that protects telomeres by activating 2C genes and *Zscan4* and increasing T-SCE-based ALT-like activity, as shown here. Human iPSCs and ESCs generated thus far have not been demonstrated to exhibit ALT-like mechanisms, unlike mouse ESCs/iPSCs. It is unclear whether this is related to the pluripotency status of ESCs/iPSCs, as presumably conventional human iPSCs/ESCs are more like mouse epiblast stem cells or primed states than naive mouse iPSCs/ESCs (Liu, 2017).

We anticipate that fine-tuning of treatment time and concentration of crotonic acid in combination with other chemicals that can facilitate early activation of *Oct4* distal enhancer and promoter and of *Nanog* may accelerate the reprogramming and improve CiPSC induction.

EXPERIMENTAL PROCEDURES

Mice and Cell Culture

Mice were housed and cared for in the College Animal Facility and the use of mice was approved by the Institutional Animal Care and Use Committee at Nankai University. All the animal experiments were performed following the ethical guidelines approved by the Tianjin Animal Management Committee. ESCs were cultured on mitomycin C-inactivated MEF feeder cells in the ESC medium,

(D and E) Summary table showing efficiency of chimera generation from CaCiPSCs at mid and late passages following injection into four- to eight-cell albino embryos (D). Chimeras were initially identified by coat color (E) and confirmed by microsatellite genotyping.
(F) Representative telomere FISH images. Blue, chromosomes stained with DAPI; green dots, telomeres.
(G) Histogram displaying distribution of relative telomere length shown as TFU by telomere Q-FISH. Medium telomere length (green lines) is shown as mean ± SD above each panel.
See also Figure S7.

which was changed daily, and cells routinely passaged every 2 days. For culture of CiPSCs, 2i (3 μM CHIR99021 and 1 μM PD0325901) were added to the medium. The ESC culture medium consisted of knockout DMEM (Invitrogen) with 20% FBS (Hyclone), 1000 U/mL mouse leukemia inhibitory factor (LIF; ESG1107; Millipore), 0.1 mM non-essential amino acids, 0.1 mM β-mercaptoethanol, 1 mM L-glutamine, penicillin (100 U/ml), and streptomycin (100 μg/mL).

CiPSC Induction from MEFs

MEFs were isolated from OG2 mice. Isolated OG-MEF cells at early passages (up to P3) were used for chemical induction following the method described (Zhao et al., 2015). The details are described in Supplemental Experimental Procedures.

Library Preparation and RNA Sequencing

A total amount of 3 μg of RNA per sample was used as input material for the RNA sample preparations. Sequencing libraries were generated using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, United States) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. More details are described in the Supplemental Experimental Procedures.

Differential Gene Expression Analysis

For DESeq with biological replicates, differential expression analysis of two groups (two biological replicates per condition) was performed using the DESeq R package (1.18.0). More details are described in the Supplemental Experimental Procedures.

ChIP-qPCR Analysis

ChIP-qPCR analysis was performed as described previously (Dan et al., 2014, 2015).

More methods are detailed in the Supplemental Experimental Procedures.

ACCESSION NUMBERS

The accession number for the RNA-seq data used in this study is GEO: GSE113921, and the number for whole-exome sequencing data is NCBI Sequence Read Archive: SRP144105.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and five tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2018.05.003.

AUTHOR CONTRIBUTIONS

H.F. conducted most experiments, analyzed the data, and prepared the manuscript. C.-l.T., X.Y., X.S., H.W., and Y.L. conducted part of the experiments or provided reagents. L.L. conceived the project, designed experiments, and wrote and revised the manuscript.

ACKNOWLEDGMENTS

We thank Yuan Long and Xin Xie for kind assistance with chemical reprogramming experiments, Weiyu Zhang and Xinyi Lu for help with analysis of RNA-seq data, and Feng Wang and Jinghua Yuan for C-circle experiments and critical reading of the manuscript. This work was supported by the National Natural Science Foundation of China (31430052, 31571546) and PCSIRT (no. IRT13023).

Received: October 23, 2017
Revised: May 2, 2018
Accepted: May 3, 2018
Published: May 31, 2018

REFERENCES

Agarwal, S., Loh, Y.H., McLoughlin, E.M., Huang, J., Park, I.H., Miller, J.D., Huo, H., Okuka, M., Dos Reis, R.M., Loewer, S., et al. (2010). Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. Nature *464*, 292–296.

Bailey, S.M., Brenneman, M.A., and Goodwin, E.H. (2004). Frequent recombination in telomeric DNA may extend the proliferative life of telomerase-negative cells. Nucleic Acids Res. *32*, 3743–3751.

Bao, S., Tang, F., Li, X., Hayashi, K., Gillich, A., Lao, K., and Surani, M.A. (2009). Epigenetic reversion of post-implantation epiblast to pluripotent embryonic stem cells. Nature *461*, 1292–1295.

Bechter, O.E., Zou, Y., Walker, W., Wright, W.E., and Shay, J.W. (2004). Telomeric recombination in mismatch repair deficient human colon cancer cells after telomerase inhibition. Cancer Res. *64*, 3444–3451.

Blackburn, E.H. (2001). Switching and signaling at the telomere. Cell *106*, 661–673.

Chen, J., Liu, J., Yang, J., Chen, Y., Chen, J., Ni, S., Song, H., Zeng, L., Ding, K., and Pei, D. (2011). BMPs functionally replace Klf4 and support efficient reprogramming of mouse fibroblasts by Oct4 alone. Cell Res. *21*, 205–212.

Collado, M., Blasco, M.A., and Serrano, M. (2007). Cellular senescence in cancer and aging. Cell *130*, 223–233.

Dan, J., Liu, Y., Liu, N., Chiourea, M., Okuka, M., Wu, T., Ye, X., Mou, C., Wang, L., Wang, L., et al. (2014). Rif1 maintains telomere length homeostasis of ESCs by mediating heterochromatin silencing. Dev. Cell *29*, 7–19.

Dan, J., Yang, J., Liu, Y., Xiao, A., and Liu, L. (2015). Roles for histone acetylation in regulation of telomere elongation and two-cell state in mouse ES cells. J. Cell. Physiol. *230*, 2337–2344.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev. *19*, 2100–2110.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. Nature *525*, 469–478.

Fan, X., Zhang, X., Wu, X., Guo, H., Hu, Y., Tang, F., and Huang, Y. (2015). Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. Genome Biol. *16*, 148.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., Kalma, Y., Viukov, S., Maza, I., Zviran, A., et al. (2013). Derivation of novel human ground state naive pluripotent stem cells. Nature *504*, 282–286.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63–67.

Gohring, J., Fulcher, N., Jacak, J., and Riha, K. (2014). TeloTool: a new tool for telomere length measurement from terminal restriction fragment analysis with improved probe intensity correction. Nucleic Acids Res. *42*, e21.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creyghton, M.P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. Nature *462*, 595–601.

Henson, J.D., Cao, Y., Huschtscha, L.I., Chang, A.C., Au, A.Y., Pickett, H.A., and Reddel, R.R. (2009). DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. Nat. Biotechnol. *27*, 1181–1185.

Henson, J.D., Lau, L.M., Koch, S., Martin La Rotta, N., Dagg, R.A., and Reddel, R.R. (2017). The C-circle assay for alternative-lengthening-of-telomeres activity. Methods *114*, 74–84.

Hirata, T., Amano, T., Nakatake, Y., Amano, M., Piao, Y., Hoang, H.G., and Ko, M.S. (2012). Zscan4 transiently reactivates early embryonic genes during the generation of induced pluripotent stem cells. Sci. Rep. *2*, 208.

Hoffmeyer, K., Raggioli, A., Rudloff, S., Anton, R., Hierholzer, A., Del Valle, I., Hein, K., Vogt, R., and Kemler, R. (2012). Wnt/beta-catenin signaling regulates telomerase in stem cells and cancer cells. Science *336*, 1549–1554.

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., Zhao, T., Ye, J., Yang, W., Liu, K., et al. (2013). Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. Science *341*, 651–654.

Huang, J., Wang, F., Okuka, M., Liu, N., Ji, G., Ye, X., Zuo, B., Li, M., Liang, P., Ge, W.W., et al. (2011). Association of telomere length with authentic pluripotency of ES/iPS cells. Cell Res. *21*, 779–792.

Huang, Y., Liang, P., Liu, D., Huang, J., and Songyang, Z. (2014). Telomere regulation in pluripotent stem cells. Protein Cell *5*, 194–202.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., and Melton, D.A. (2008a). Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. Nat. Biotechnol. *26*, 795–797.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., Muhlestein, W., and Melton, D.A. (2008b). Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat. Biotechnol. *26*, 1269–1275.

Ichida, J.K., Blanchard, J., Lam, K., Son, E.Y., Chung, J.E., Egli, D., Loh, K.M., Carter, A.C., Di Giorgio, F.P., Koszka, K., et al. (2009). A small-molecule inhibitor of Tgf-Beta signaling replaces sox2 in reprogramming by inducing nanog. Cell Stem Cell *5*, 491–503.

Jiang, J., Lv, W., Ye, X., Wang, L., Zhang, M., Yang, H., Okuka, M., Zhou, C., Zhang, X., Liu, L., et al. (2013). Zscan4 promotes genomic stability during reprogramming and dramatically improves the quality of iPS cells as demonstrated by tetraploid complementation. Cell Res. *23*, 92–106.

Li, Y., Zhang, Q., Yin, X., Yang, W., Du, Y., Hou, P., Ge, J., Liu, C., Zhang, W., Zhang, X., et al. (2011). Generation of iPSCs from mouse fibroblasts with a single gene, Oct4, and small molecules. Cell Res. *21*, 196–204.

Li, W., Li, K., Wei, W., and Ding, S. (2013). Chemical approaches to stem cell biology and therapeutics. Cell Stem Cell *13*, 270–283.

Liu, L. (2017). Linking telomere regulation to stem cell pluripotency. Trends Genet. *33*, 16–33.

Liu, L., Bailey, S.M., Okuka, M., Munoz, P., Li, C., Zhou, L., Wu, C., Czerwiec, E., Sandler, L., Seyfang, A., et al. (2007). Telomere lengthening early in development. Nat. Cell Biol. *9*, 1436–1441.

Londono-Vallejo, J.A., Der-Sarkissian, H., Cazes, L., Bacchetti, S., and Reddel, R.R. (2004). Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. Cancer Res. *64*, 2324–2327.

Long, Y., Wang, M., Gu, H., and Xie, X. (2015). Bromodeoxyuridine promotes full-chemical induction of mouse pluripotent stem cells. Cell Res. *25*, 1171–1174.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57–63.

Maherali, N., and Hochedlinger, K. (2008). Guidelines and techniques for the generation of induced pluripotent stem cells. Cell Stem Cell *3*, 595–605.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., Stadtfeld, M., Yachechko, R., Tchieu, J., Jaenisch, R., et al. (2007). Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. Cell Stem Cell *1*, 55–70.

Marion, R.M., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., Serrano, M., and Blasco, M.A. (2009). Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells. Cell Stem Cell *4*, 141–154.

Miyanari, Y., and Torres-Padilla, M.E. (2012). Control of ground-state pluripotency by allelic regulation of Nanog. Nature *483*, 470–473.

Nichols, J., and Smith, A. (2009). Naive and primed pluripotent states. Cell Stem Cell *4*, 487–492.

Palm, W., and de Lange, T. (2008). How shelterin protects mammalian telomeres. Annu. Rev. Genet. *42*, 301–334.

Pucci, F., Gardano, L., and Harrington, L. (2013). Short telomeres in ESCs lead to unstable differentiation. Cell Stem Cell *12*, 479–486.

Schneider, R.P., Garrobo, I., Foronda, M., Palacios, J.A., Marion, R.M., Flores, I., Ortega, S., and Blasco, M.A. (2013). TRF1 is a stem cell marker and is essential for the generation of induced pluripotent stem cells. Nat. Commun. *4*, 1946.

Shi, Y., Desponts, C., Do, J.T., Hahm, H.S., Scholer, H.R., and Ding, S. (2008). Induction of pluripotent stem cells from mouse embryonic fibroblasts by Oct4 and Klf4 with small-molecule compounds. Cell Stem Cell *3*, 568–574.

Silva, J., Nichols, J., Theunissen, T.W., Guo, G., van Oosten, A.L., Barrandon, O., Wray, J., Yamanaka, S., Chambers, I., and Smith, A. (2009). Nanog is the gateway to the pluripotent ground state. Cell *138*, 722–737.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663–676.

Takai, H., Smogorzewska, A., and de Lange, T. (2003). DNA damage foci at dysfunctional telomeres. Curr. Biol. *13*, 1549–1556.

Tan, M., Luo, H., Lee, S., Jin, F., Yang, J.S., Montellier, E., Buchou, T., Cheng, Z., Rousseaux, S., Rajagopal, N., et al. (2011). Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell *146*, 1016–1028.

Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M.A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-seq analysis. Cell Stem Cell *6*, 468–478.

Teichroeb, J.H., Kim, J., and Betts, D.H. (2016). The role of telomeres and telomerase reverse transcriptase isoforms in pluripotency induction and maintenance. RNA Biol. *13*, 707–719.

Valamehr, B., Robinson, M., Abujarour, R., Rezner, B., Vranceanu, F., Le, T., Medcalf, A., Lee, T.T., Fitch, M., Robbins, D., et al. (2014). Platform for induction and maintenance of transgene-free hiPSCs resembling ground state pluripotent stem cells. Stem Cell Reports *2*, 366–381.

Wang, F., Yin, Y., Ye, X., Liu, K., Zhu, H., Wang, L., Chiourea, M., Okuka, M., Ji, G., Dan, J., et al. (2012). Molecular insights into the heterogeneity of telomere reprogramming in induced pluripotent stem cells. Cell Res. *22*, 757–768.

Wei, W., Liu, X., Chen, J., Gao, S., Lu, L., Zhang, H., Ding, G., Wang, Z., Chen, Z., Shi, T., et al. (2017a). Class I histone deacetylases are major histone decrotonylases: evidence for critical and broad function of histone crotonylation in transcription. Cell Res. *27*, 898–915.

Wei, W., Mao, A., Tang, B., Zeng, Q., Gao, S., Liu, X., Lu, L., Li, W., Du, J.X., Li, J., et al. (2017b). Large-scale identification of protein crotonylation reveals its role in multiple cellular functions. J. Proteome Res. *16*, 1743–1752.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318–324.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881–894.

Zalzman, M., Falco, G., Sharova, L.V., Nishiyama, A., Thomas, M., Lee, S.L., Stagg, C.A., Hoang, H.G., Yang, H.T., Indig, F.E., et al. (2010). Zscan4 regulates telomere elongation and genomic stability in ES cells. Nature *464*, 858–863.

Zhang, Q., Dan, J., Wang, H., Guo, R., Mao, J., Fu, H., Wei, X., and Liu, L. (2016). Tcstv1 and Tcstv3 elongate telomeres of mouse ES cells. Sci. Rep. *6*, 19852.

Zhao, X.Y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C.L., Ma, Q.W., Wang, L., et al. (2009). iPS cells produce viable mice through tetraploid complementation. Nature *461*, 86–90.

Zhao, Y., Zhao, T., Guan, J., Zhang, X., Fu, Y., Ye, J., Zhu, J., Meng, G., Ge, J., Yang, S., et al. (2015). A XEN-like state bridges somatic cells to pluripotency during chemical reprogramming. Cell *163*, 1678–1691.

Zhu, S., Li, W., Zhou, H., Wei, W., Ambasudhan, R., Lin, T., Kim, J., Zhang, K., and Ding, S. (2010). Reprogramming of human primary somatic cells by OCT4 and chemical compounds. Cell Stem Cell *7*, 651–655.

Stem Cell Reports, Volume *11*

Supplemental Information

Dynamics of Telomere Rejuvenation during Chemical Induction to Pluripotent Stem Cells

Haifeng Fu, Cheng-lei Tian, Xiaoying Ye, Xiaoyan Sheng, Hua Wang, Yifei Liu, and Lin Liu

Supplemental Information

Dynamics of Telomere Rejuvenation During Chemical Induction to Pluripotent Stem Cells

Haifeng Fu, Xiaoying Ye, Cheng-lei Tian, Xiaoyan Sheng, Hua Wang, Yifei Liu and Lin Liu

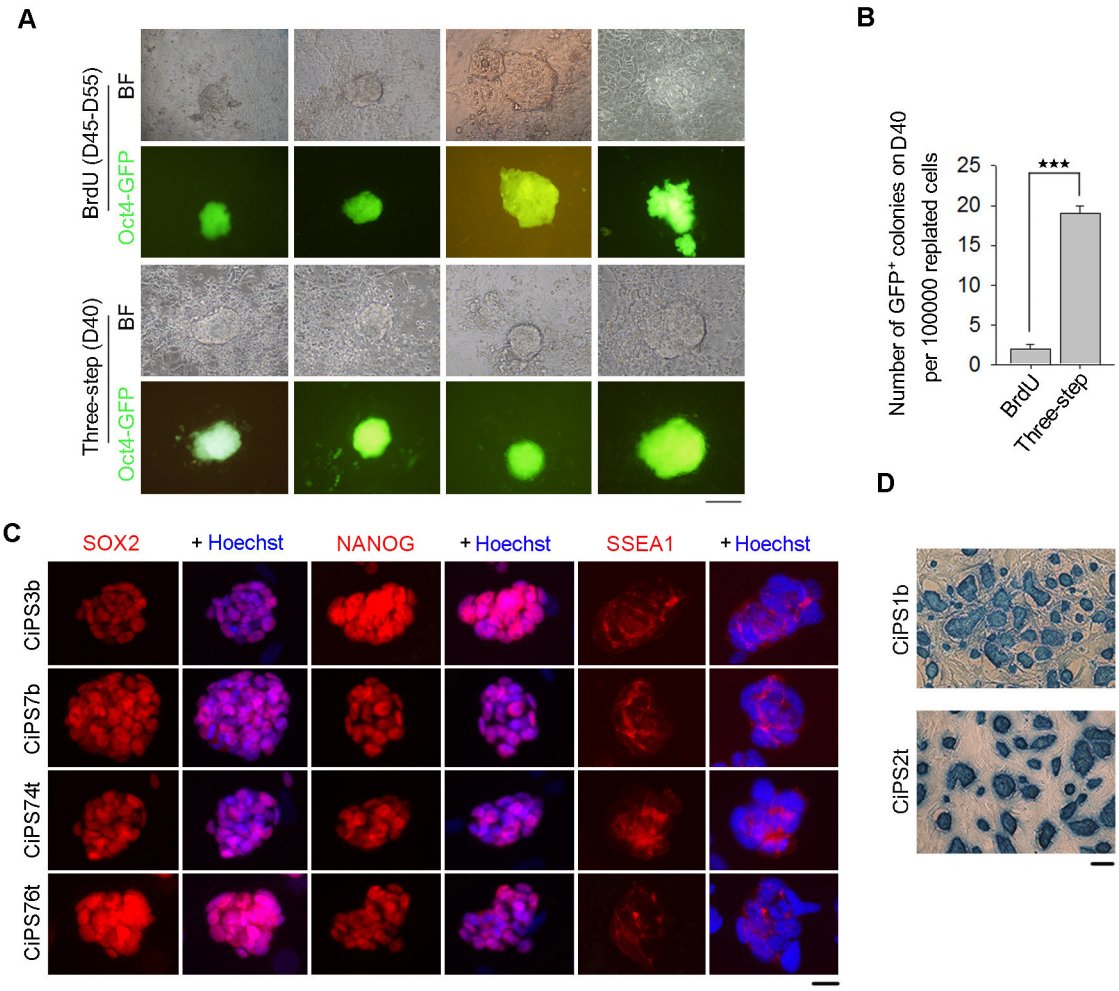


Figure S1. Characterization of CiPSCs. Related to Figure 1.

(A) Morphology of CiPSC primary colonies generated by BrdU and Three-step methods under bright field (BF) with phase contrast optics and Oct4-GFP fluorescence. Oct4-GFP fluorescence were visible in some cells at day 40 by BrdU method, but the colonies with clear boundary appeared later by BrdU method than did Three-step method. Scale bar=100 μm.

(B) Induction efficiency of colony formation based on number of Oct4-GFP fluorescence positive clones. Data represent mean ± SEM from three independent experiments. ***P<0.001.

(C) Immunofluorescence of pluripotent markers SOX2, NANOG and SSEA1 in various CiPSCs (b, BrdU method; t, Three-step method). Scale bar=20 μm.

(D) Alkaline phosphatase (AP) staining of CiPSCs at P5. Scale bar=100 μm.

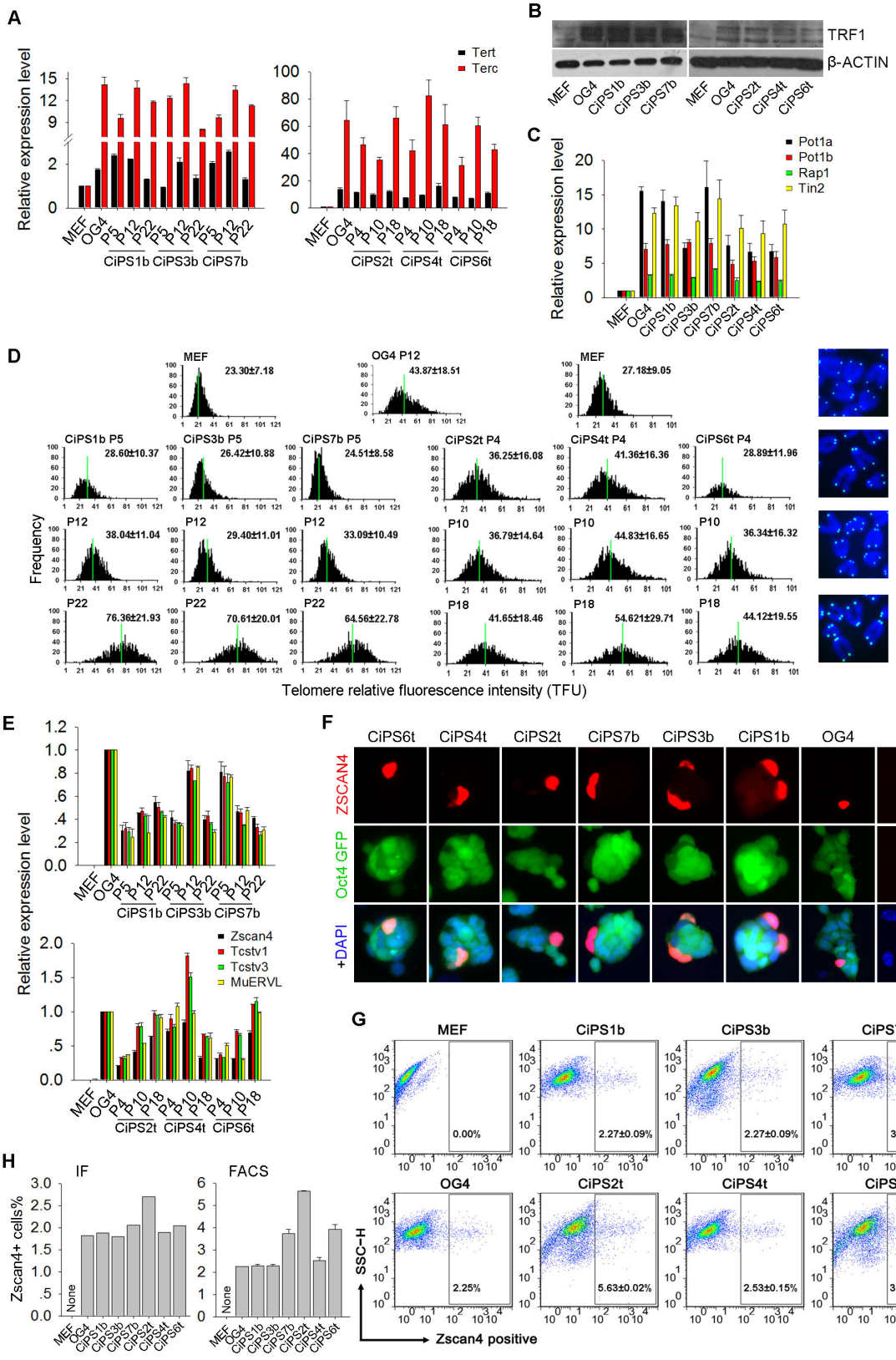

A
Relative expression level
Tert
Terc
MEF
OG4
CiPS1b
CiPS3b
CiPS7b
CiPS2t
CiPS4t
CiPS6t
B
TRF1
β-ACTIN
C
Relative expression level
Pot1a
Pot1b
Rap1
Tin2
D
Frequency
Telomere relative fluorescence intensity (TFU)
E
Relative expression level
Zscan4
Tcstv1
Tcstv3
MuERVL
F
ZSCAN4
Oct4 GFP
+DAPI
G
SSC-H
Zscan4 positive
H
IF
FACS
Zscan4+ cells%
None

Figure S2. Telomere Rejuvenation and Expression of *Zscan4* and Two-Cell Genes in CiPSCs. Related to Figure 2.

(A) Relative expression levels by qPCR of telomerase genes *Tert* and *Terc* in CiPSCs at various passages, compared with MEFs, and ESCs (OG4) also expressing Oct4-GFP fluorescence.

(B and C) Expression of shelterin complex TRF1 protein or genes by Western blot or qPCR analysis.

(D) Histogram displays distribution of relative telomere length shown as TFU by telomere quantitative fluorescence *in situ* hybridization (Q-FISH). The medium telomere length (green bars) is shown as mean ± SD above each panel. On the right next to the cell line are telomere FISH images.

(E) RT-qPCR analysis of two-cell genes *Zscan4*, *Tcstv1*, *Tcstv3* and *MuEVRL* in CiPSCs following passage. MEFs and ESCs served as negative and positive controls.

(F) Immunofluorescence of ZSCAN4 in CiPS cell lines and ESCs OG4 at P15. MEFs as controls do not express ZSCAN4 and Oct4-GFP fluorescence. Scale bar=10 μm.

(G) Analysis of Zscan4^{+} cells in CiPSC and ESC cultures at P15 by flow cytometry.

(H) Quantification and proportion of Zscan4^{+} by immunofluorescence (IF, n=400 cells counted) or by FACS.

Data represent mean ± SEM from three independent experiments.

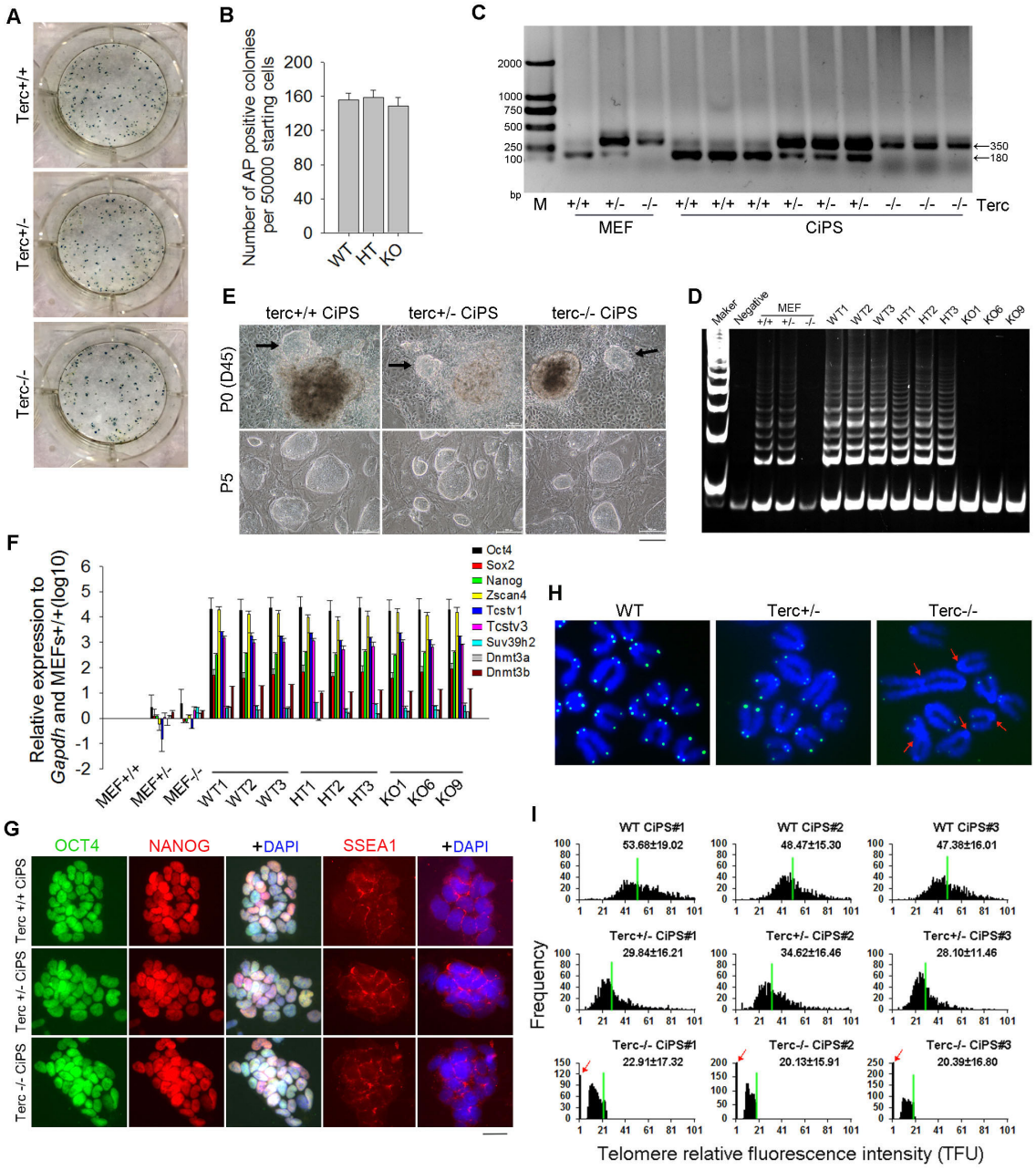


Figure S3. Telomerase Is Dispensable for CiPSC Induction but Required for Telomere Elongation during Passages after CiPSC Formation. Related to Figure 2.

(A) Representative images showing clonal formation based alkaline phosphatase (AP) activity staining. CiPSCs were generated using the three-step method described in Experimental Procedures.

(B) Number of AP positive colonies does not differ among telomerase *Terc* knockout (KO, *Terc*$^{-/-}$), heterozygous (HT, *Terc*$^{+/-}$) and Wild-type (WT, *Terc*$^{+/+}$) cells. Data represent mean ± SEM from three independent experiments. P>0.05.

(C) Genotyping of *Terc* KO (*Terc*$^{-/-}$), heterozygous (HT, *Terc*$^{+/-}$) and Wild-type (*Terc*$^{+/+}$) CiPSC lines and their progenitor MEFs.

(D) Telomerase activity by TRAP assay, confirming that *Terc*$^{-/-}$ cells do not express

telomerase activity.
(E) Typical morphology under phase contrast of primary colonies and formed CiPSCs at passage 5. Scale bar=100 μm.
(F) Relative expression level by qPCR of genes for pluripotency, two-cell genes and methyltransferase genes. Data represent mean ± SEM from three independent experiments.
(G) Immunofluorescence of OCT4, NANOG and SSEA1. Scale bar=20 μm.
(H and I) Telomere Q-FISH images (H) and Histogram displaying distribution of relative telomere length shown as TFU (I). The medium telomere length (green bars) is shown as mean ± SD above each panel. Red arrows indicate chromosome fusion or telomere-signal free ends indicative of telomere loss in *Terc*$^{-/-}$ CiPSCs. CiPSCs at passage 15 were analyzed for telomere lengths estimated by Q-FISH.

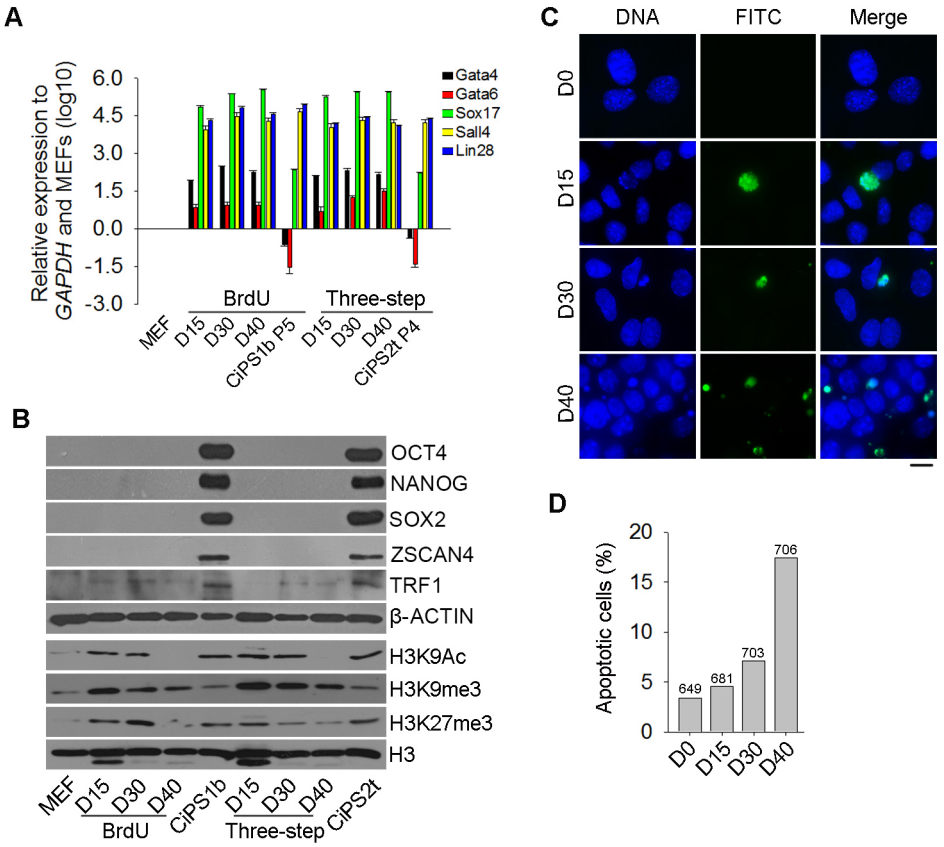


Figure S4. Expression of Genes Related to XEN State and Pluripotency and Apoptosis during Chemical Induction. Related to Figure 4.

(A) RT-qPCR analysis of relative expression levels of *Gata4, Gata6, Sox17, Sall4* and *Lin28* at various days during chemical induction using BrdU or Three-step method. Data represent mean ± SEM from three independent experiments.

(B) Western blot analysis of relative protein expression during induction. H3 and β-actin served as loading control.

(C) Representative images showing apoptotic cells by TUNEL assay. Apoptotic cells are revealed by fragmented nuclei and TUNEL signals (FITC green fluorescence). Scale bar=10 μm.

(D) Percentage of apoptosis cells on various days during chemical induction. Number of cells counted is shown above the bar. Data from three independent experiments.

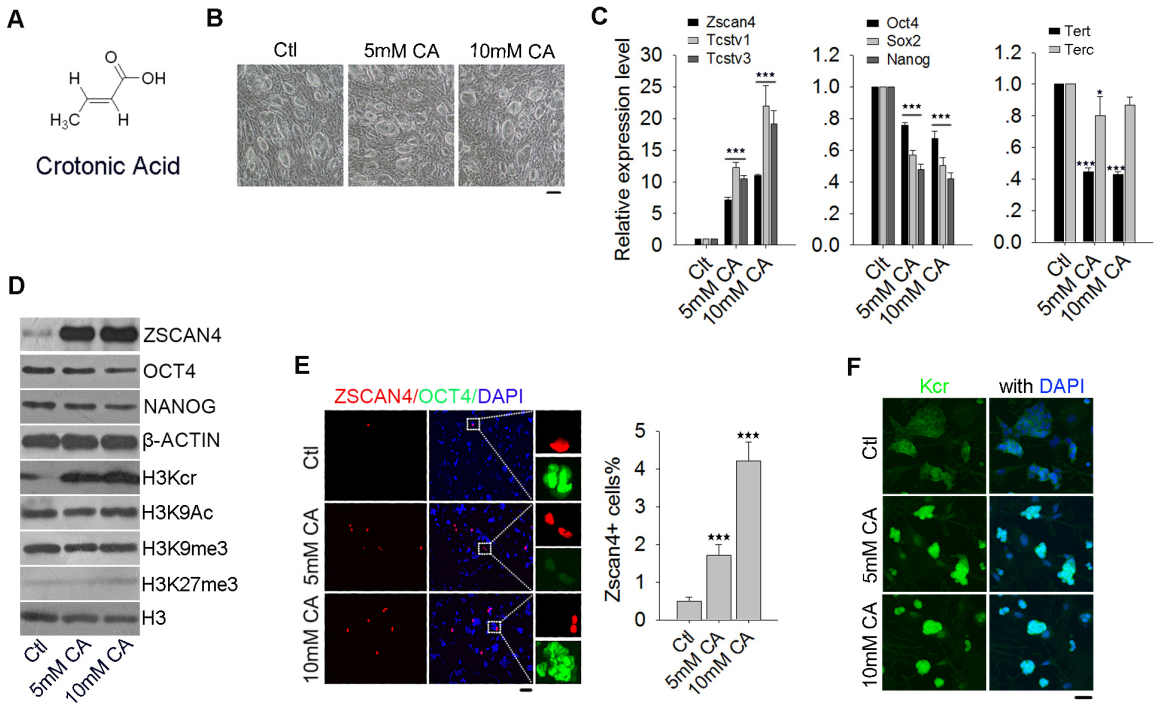


Figure S5. Crotonic Acid Increases *Zscan4* Expression in mESCs. Related to Figure 5.

(A) Chemical structures of crotonic acid (CA).

(B) Morphology of mouse ESCs (N33) following treatment with 5 or 10 mM crotonic acid (CA) for two passages. Scale bar=100 μm.

(C) Relative expression level by qPCR analysis of two-cell genes, pluripotent genes and telomerase genes of N33 ESCs treated with 5 or 10 mM crotonic acid for two passages. Data represent mean ± SEM from three independent experiments. *P<0.05; ***P<0.001, compared to controls (Ctl).

(D) Protein levels by Western blot of ZSCAN4, OCT4, NANOG, H3Kcr, H3K9me3, H3K9ac, and H3K27me3.

(E) Immunofluorescence of ZSCAN4 in ESCs and proportion of Zscan4^{+} cells following treatment with crotonic acid. n=400 cells counted. ***P<0.001, compared to controls. Scale bar=10 μm. Data represent mean ± SEM from three independent experiments.

(F) Immunofluorescence of lysine crotonylation (Kcr) increases in N33 ESCs treated with 5 or 10 mM crotonic acid. Scale bar=50 μm.

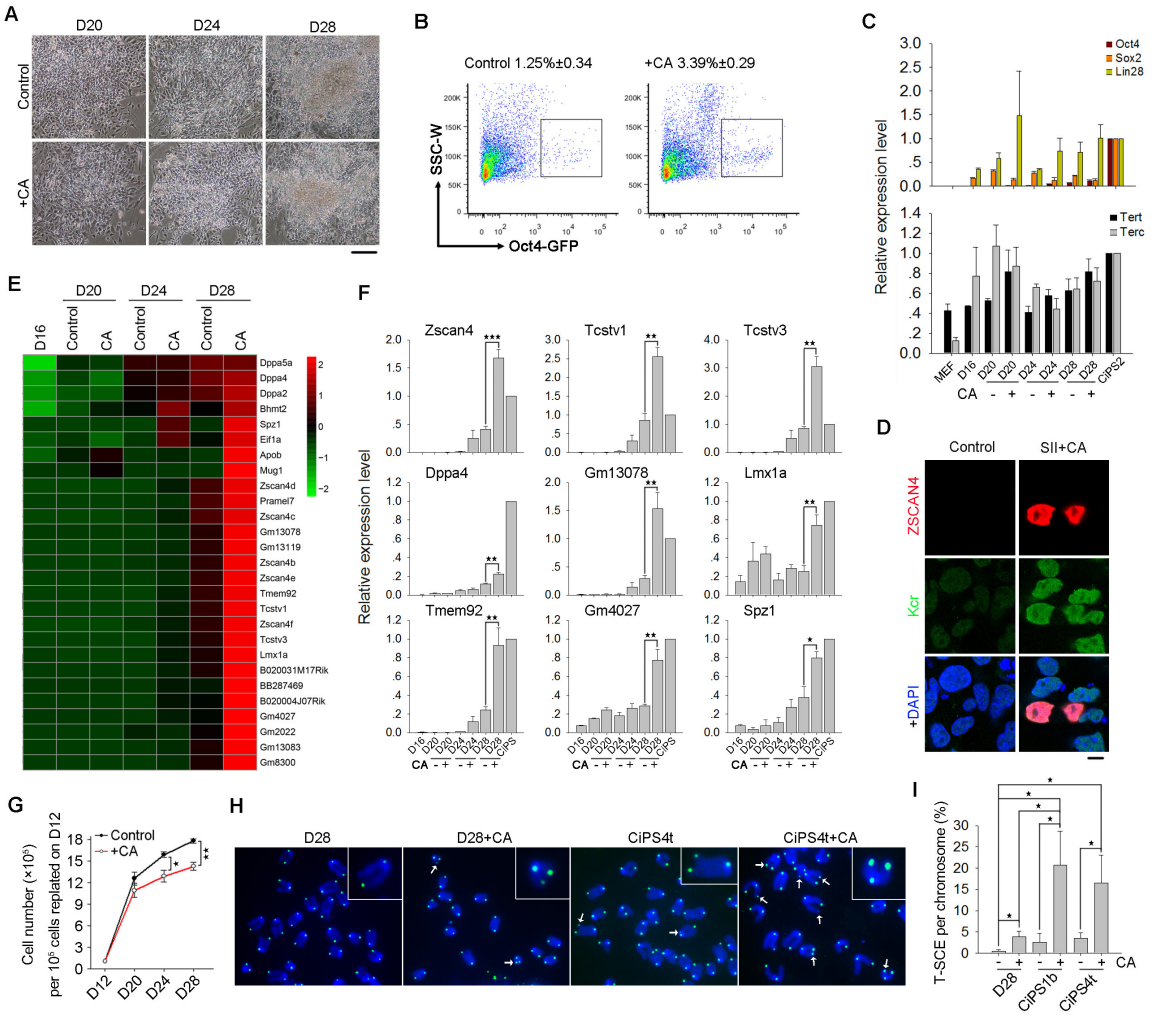


Figure S6. Crotonic Acid Activates Two-Cell Genes and Enhances CiPSC Generation. Related to Figure 6.

(A) Morphology of chemical induction on D20, D24 and D28 following crotonic acid (CA) addition at Stage II, compared with controls without the treatment. Scale bar=100 μm.

(B) Flow cytometry analysis of induction efficiency at D40 based on endogenous Oct4-GFP fluorescence.

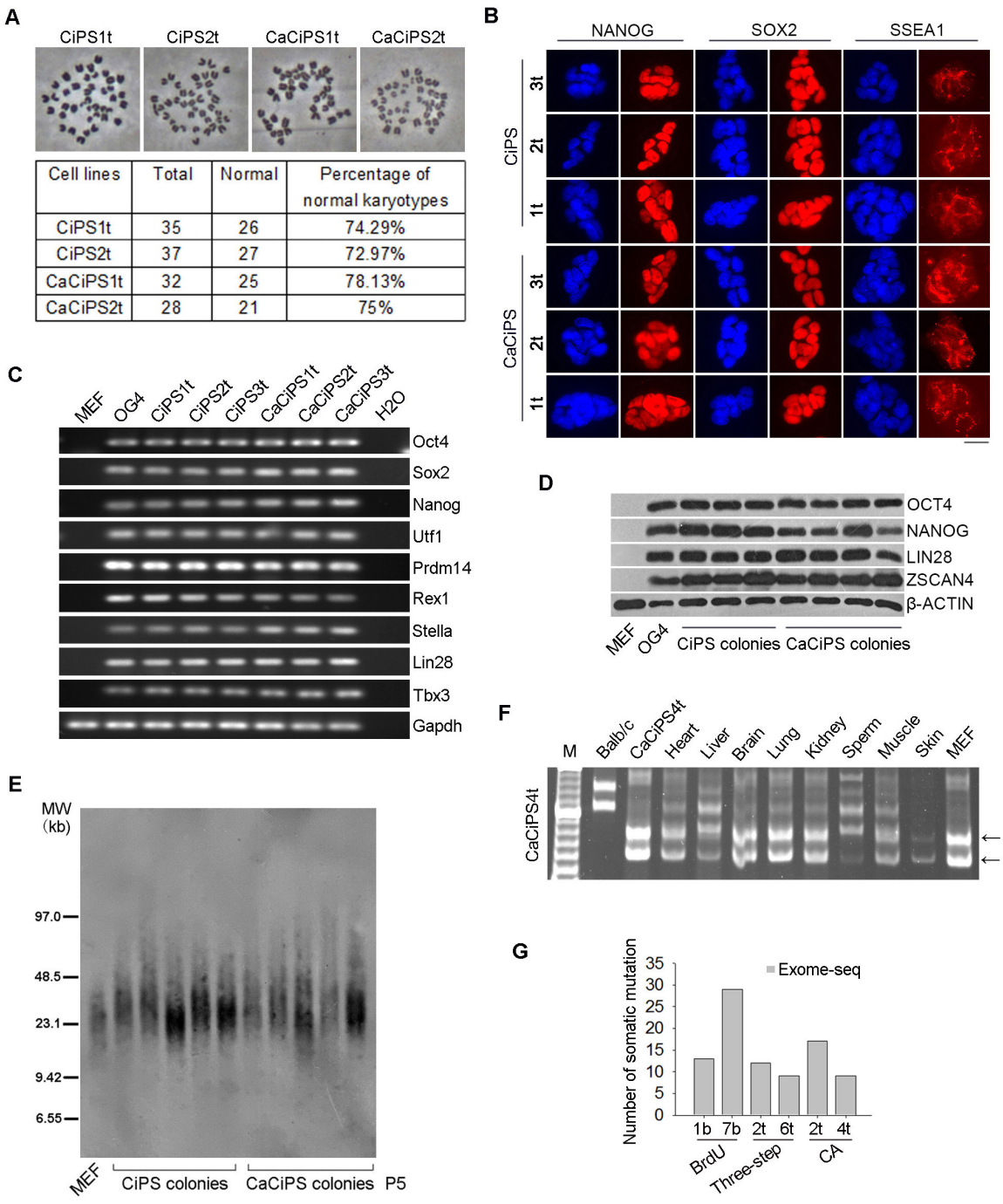
(C) Relative expression levels by qPCR of pluripotent genes *Oct4*, *Sox2*, and *Lin28* and telomerase genes *Tert* and *Terc* on D20, D24 and D28 of cells treated with or without crotonic acid. Data represent mean ± SEM from three independent experiments.

(D) Confocal microscopy showing co-expression of lysine crotonylation (Kcr) and Zscan4 on D28 of induction. Scale bar=10 μm.

(E) Heatmap displaying two-cell genes highly expressed in crotonic acid-treated cells on day 24 and particularly on day 28 by RNA-seq, compared with controls. The two-cell genes were listed in published data (Macfarlan, T.S., *et al.* Nature 2012; 487, 57-63).

(F) qPCR analysis validating the RNA-seq data showing increased expression of two-cell genes by crotonic acid in E. Data represent mean ± SEM from three

independent experiments.
(G) Reduced proliferation of chemical reprogramming cells after CA treatment. 10^5 cells were re-plated on D12, and cells were counted at three time points. Data represent mean ± SEM from three independent experiments.
(H) Representative micrographs showing T-SCE (white arrows) by CO-FISH assay.
(I) T-SCE per chromosome in D28 reprogramming cells and two CiPS cell lines without or with CA treatment. At least 15 metaphase spreads were counted for frequency of T-SCE. Data from two independent experiments. Pairwise comparisons for statistical significance were made by t-tests.
*P<0.05, **P<0.01, ***P<0.001.



Cell lines	Total	Normal	Percentage of normal karyotypes
CiPS1t	35	26	74.29%
CiPS2t	37	27	72.97%
CaCiPS1t	32	25	78.13%
CaCiPS2t	28	21	75%

Figure S7. Characterization of CaCiPSCs Obtained by Addition with Crotonic Acid at Stage II during Induction. Related to Figure 7.

(A) Chromosome spread and karyotype analysis of CiPSCs as controls and CaCiPSCs. Majority of spread exhibit normal ploidy.

(B and C) Expression of pluripotency marker genes by immunofluorescence (B) and conventional RT-PCR (C). Scale bar=20 μm.

(D) Protein levels by Western blot of OCT4, SOX2, LIN28 and ZSCAN4 in CiPSCs and CaCiPSCs compared with MEFs and ESCs (OG4).

(E) Telomere length distribution shown as TRF by Southern blot analysis of CiPSCs and CaCiPSCs at passage 5. MEFs served as control.

(F) Microsatellite genotyping analysis showing contributions of CaCiPSCs to various tissues of the chimera. Four-eight-cell embryos from Balb/c mice were used as recipients. Arrows indicate donor MEFs, the derived CaCiPSCs and contributions in the chimera.
(G) Number of somatic mutations in coding regions in CiPSCs by whole exome sequencing analysis (Data detailed in Table S4 and S5).

Table S1. Primers for RT-qPCR analysis

Genes	Forward	Reverse
Oct4	TTGGGCTAGAGAAGGATGTGGTT	GGAAAAGGGACTGAGTAGAGTGTGG
Sox2	GCACATGAACGGCTGGAGCAACG	TGCTGCGAGTAGGACATGCTGTAGG
Nanog	TTGCTTACAAGGGTCTGCTACT	ACTGGTAGAAGAATCAGGGCT
Tert	ACTGGTGGAGATCATCTTTCTGGG	ACCTGAGGAGTCTGACATATTGGC
Terc	CATTAGCTGTGGGTTCTGGTCT	TCCTGCGCTGACGTTTGTTT
P53	TGTTATGTGCACGTACTCTCCTCC	GTGCTGTGACTTCTTGTAGATGGC
P21	AGACCTGTGAAGACAGGAATGGTC	AGCAGATCACCAGATTAACCCTCC
Sall4	TGGCAGACGAGAAGTTCTTTC	TCCAACATTTATCCGAGCACAG
Zscan4	AAATGCCTTATGTCTGTTCCCTATG	TGTGGTAATTCCTCAGGTGACGAT
Tcstv1	TGAACCCTGATGCCTGCTAAGACT	AGATGGCTGCAAAGACACAACTGC
Tcstv3	AGAAAGGGCTGGAACTTGTGACCT	AAAGCTCTTTGAAGCCATGCCCAG
Rex1	ACGAGTGGCAGTTTCTTCTTGGGA	TATGACTCACTTCCAGGGGGCACT
Pot1b	TCCTCATACGAGGGAAGGTG	GATGCTGGGATCTGGAAAAA
Pot1a	AGCTTCACTCCTCAGGACCA	GGGTTCCATCCCATACCTTT
Rap1	CCGCTACCTCAAGCACCTAC	CACTCCTCCTCAGGCAAGTC
Tin2	GGAGTTTCTGCAGTCCTTGC	TCTGGACTCTGCTGGGAAGT
MuERV-L	CCCATCATGAGCTGGGTACT	CGTGCAGATCCATCAGTAAA
Sox17	CTCGGGGATGTAAAGGTGAA	GCTTCTCTGCCAAGGTCAAC
GATA4	TCTCACTATGGGCACAGCAG	GCGATGTCTGAGTGACAGGA
GATA6	CAAAAGCTTGCTCCGGTAAC	TGAGGTGGTCGCTTGTGTAG
GAPDH	TCAACAGCAACTCCCACTCTTCCA	ACCACCCTGTTGCTGTAGCCGTAT
Lin28a	CCGCAGTTGTAGCACCTGTCT	GAAGAACATGCAGAAGCGAAGA

Primers for RT-PCR analysis

Genes	Forward	Reverse
Oct4	TTGGGCTAGAGAAGGATGTGGTT	GGAAAAGGGACTGAGTAGAGTGTGG
Sox2	GCACATGAACGGCTGGAGCAACG	TGCTGCGAGTAGGACATGCTGTAGG
Nanog	TTGCTTACAAGGGTCTGCTACT	ACTGGTAGAAGAATCAGGGCT
Utf1	TTCGCCGCCGCTCTACT	CAGGGGCAGGTTCGTCATT
Lin28a	CCGCAGTTGTAGCACCTGTCT	GAAGAACATGCAGAAGCGAAGA
Stella	CCCAATGAAGGACCCTGAAAC	AATGGCTCACTGTCCCGTTCA
Rex1	ACGAGTGGCAGTTTCTTCTTGGGA	TATGACTCACTTCCAGGGGGCACT
Tbx3	CCACCCGTTCCTCAATTTGAACAG	CGGAAGCCATTGATGGTAAAGCTG
Prdm14	ACAGCCAAGCAATTTGCACTAC	TTACCTGGCATTTTCATTGCTC
GAPDH	TCAACAGCAACTCCCACTCTTCCA	ACCACCCTGTTGCTGTAGCCGTAT

Primers for ChIP-qPCR analysis

Genes	Forward	Reverse
Actin	CGTGTGACAAAGCTAATGAGGCTG	CTAAGTTCAGTGTGCTGGGAGTCT
Zscan4-1	GCATTATCTGTTCCTCTGGGTC	AACTCCTGTTCCTGGGTGGG
Zscan4-2	TCCCTAGAATACAGTCCTCA	GTAGAATCCTTGATAGTGGG
Ch13_Subtel	GCACACTTGGTGGGCTAAGAAGATG	TTAAATCCTGACCAAAATGCCTGGC
MuERV-L	CAGAGCATTCACACTGGGGA	GTGAGCCTTCCAATTCCGGA
ERVK10c	TGTCAGCTGGCAAAAGAGTAAA	AAGACAGGGGAAGTCAGTTCAG

Table S2. Antibody details

Antibody	Source	Cat. No.
NANOG	Abcam	ab80892
OCT3/4	Santa Cruz Biotechnology	sc-5279
SOX2	Millipore	AB5603
SSEA1	Millipore	MAB4301
ZSCAN4	Millipore	AB4340
TCSTV1/3	Custom-made	
LIN28a	CST	3978S
βIII-TUBULIN	Chemicon	CBL412
TRF1	Alpha Dragrostic	TRF12-S
γH2AX	Millipore	05-636
AFP	Dako	DAK-N1501
α-SMA	Abcam	ab5694-100
NESTIN	Millipore	MAB353
Pan anti-crotonyllysine	PTM Biolabs	PTM-502
β-ACTIN	Abmart	P30002
H3K9me3	Abcam	ab8898
H3K27me3	Millipore	07-449
H3K9ac	Abcam	ab4441
H3	Abcam	ab1791

Table S4. Sequencing Statistics for Mutation Discovery

Cell Line	Quality-Filtered Sequences (bp)	10 x Exome Coverage Sequences (bp)	No. of High-Quality Coding Variants	dbSNP Percentage	No. of Unique Coding Mutations
MEF	5,521,756,147	38,115,851	18,243	87%	—
MEF+BrdU	5,472,472,685	38,382,831	18,387	86%	0
CiPS1b	5,614,716,507	37,886,101	12,942	90%	13
CiPS7b	5,697,170,568	37,890,164	12,277	89%	29
MEF	5,525,093,305	38,415,593	17,981	87%	—
CiPS2t	5,177,923,246	38,037,210	12,667	92%	12
CiPS6t	5,217,052,761	37,706,210	12,771	82%	9
CaCiPS2t	5,059,846,928	36,645,405	17,909	86%	17
CaCiPS4t	7,101,264,336	38,808,513	17,550	87%	9

Table S5. Summary of Somatic Mutations

Cell Line	Mutated Genes
MEF+BrdU	n/a
CiPS1b	*Chd2*, *Sv2b*, *Vmn2r65*, *Xrra1*, *Olfr657*, *Olfr667*, *Olfr480*, *Vmn1r194*, *Shisa9*, *Dgcr8*, *Ufd1*, *Abi3bp*, *Tmem2*
CiPS7b	*Ttll4*, *Olfr1221*, *Olfr1288*,*Eif2ak4*, *Pak6*, *Errfi1*, *Grm3*, *Oasl1*, *Cttnbp2*, *Prss3*, *Lsr*, *Olfr598*, *Syt9*, *Cend1*,*Il12rb1*, *Agt*, *Slc44a2*, *Grik4*, *C1qtnf5_Mfrp*, *Atm*, *Aldh8a1*, *Ddx5*,*Sdk2*, *Aspscr1*, *Fam208b*,*Olfr1368*, *Nutm2*, *Csf2rb2*, *Vwa5b2*
CiPS2t	*Lamc1*, *Pramel7*, *Olfr1154*,*Fpgt*, *Qrich1*, *Usp32*,*Kif2b*, *Akap6*, *Mboat1*, *Arhgap39*, *Ano6*, *Pdzph1*, *Tigd3*
CiPS6t	*Chd6*, *Med8*, *Cyp2b19*, *Ltbp4*, *Map2k7*, *Cdk19*, *Hdac9*, *Calml3*, *Tmem44*
CaCiPS2t	*Mcm3*, *Gli2*, *Olfr1154*, *Gpcpd1*, *Nrd1*, *Capzb*, *Ccl25*, *Aktip*, *Cdh23*, *Mdm2*, *Gpatch8*, *Kif19a*, *Ppp1r13b*,*Bmp1*, *Dgkh*, *Trappc9*, *Atp9b*
CaCiPS4t	*Ttn*, *Abcc9*, *Vmn2r57*, *Olfr834*, *Vmn2r82*, *Gli3*, *Tspan17*, *Adgrv1*, *Krt72*

EXPERIMENTAL PROCEDURES

Mice and Cell Culture

Oct4-GFP (OG2) mice (B6/CBA, JAX stock #004654) that carry Oct4 distal promoter-driven GFP were purchased from Model Animal Research Center of Nanjing University. The transgenic GFP expression of the reporter is under the control of Oct4 promoter and distal enhancer, but the proximal enhancer region is deleted, so the Oct4-ΔPE-GFP transgenic mice have been used for establishing naïve mouse ESCs (Bao et al., 2009; Yeom et al., 1996; Tang et al., 2010). ESC line (OG4) expressing distal Oct4-GFP was derived from OG2 mice and characterized based on the method described (Huang et al. 2008, 2011), but with addition of 2i in the medium. N33 mESC line was derived from C57BL/6 mice (Huang et al. 2011).

Briefly, intact blastocysts were seeded on feeder layers of mitomycin C-treated MEF cells, prepared on 0.1% gelatin-treated four-well culture dishes, in ESC medium consisting of knockout DMEM, 20% knockout serum replacement, supplemented with 1000 units/ml mouse leukemia inhibitory factor, 0.1 mM NEAA, 1 mM L-glutamine, 0.1 mM β-mercaptoethanol, 50 IU/ml penicillin and streptomycin and 2i (3μM CHIR99021 and 1μM PD0325901). Half of the medium was changed daily. Approximately 10 days after seeding, ICM outgrowths were mechanically removed and digested with 0.25% trypsin-ETDA into small clumps, digestion stopped with trypsin inhibitor, and cell suspensions reseeded on fresh feeder cells. Stable ESC lines were routinely obtained after one or two passages. Then, all ESC lines were passaged and cultured in ES medium added with fetal bovine serum and 2i, instead of KSR-based ESC medium.

Isolation of Mouse Embryonic Fibroblasts (MEFs)

MEFs were derived from E13.5 embryos isolated by caesarean section and washed in phosphate-buffered saline (PBS). Heads and visceral tissues were removed, and remaining tissue was washed in fresh PBS, then submerged in 0.05 mM trypsin/1 mM EDTA HBSS solution and incubated at 37°C for 10 min. Tissue was pipetted repeatedly to aid in tissue dissociation, then added to MEF medium containing 10% FBS and plated (passage 0).

CiPSC Induction from Mouse Embryonic Fibroblasts

Mouse embryonic fibroblasts (MEFs) were isolated from OG2 mice. Isolated OG-MEF cells at early passages (up to passage 3) were used for chemical induction following the method described (Zhao et al., 2015). MEFs were seeded at a density of 50000 cells per well on 6 well plate. On the next day (day 0), the MEFs medium was replaced with chemical induction medium. On day 0-12, the induction medium contains 100 ng/ml bFGF, 0.5 mM VPA, 20 μM CHIR99021, 10 μM Repsox, 5 μM Tranylcypromine, 50 μM Forskolin, 0.05 μM AM580 and 5 μM EPZ004777. On day 12, the cells were trypsinized, harvested and then re-plated at 100000 cells per well of a 6 well plate. During day 12-16, concentrations of bFGF, CHIR99021, and

Forskolin were reduced to 25 ng/ml, 10 μM and 10 μM respectively. On day 16-28, the induction medium contains 25 ng/ml bFGF, 0.5 mM VPA, 10 μM CHIR99021, 10 μM Repsox, 5 μM Tranylcypromine, 10 μM Forskolin, 0.05 μM AM580, 0.05 μM DZNep, 0.5 μM 5-aza-dC, and 5 μM SGC0946. On day 28, the culture was transferred into N2B27-2iL medium with 3 μM CHIR99021, 1 μM PD0325901, and 1000 U/ml LIF. After another 8-12 days, Oct4 GFP-positive CiPSC primary colonies emerged and were then picked up for expansion and characterization.

For BrdU method, the induction medium contains the small-molecule cocktails: 0.5 mM VPA, 10 μM CHIR99021, 10 μM Repsox, 5 μM Tranylcypromine, 50 μM Forskolin, 1 μM TTNPB, and 8μM BrdU (Long et al., 2015). DZNep (50 nM) was added to the cell cultures on day 16. On day 28-32, the culture was transferred into N2B27-2iL medium. After additional 16-20 days, GFP-positive CiPSC colonies emerged. The medium was changed every 4 days.

In experiments by treatment with crotonic acid during induction, 6-8 mM crotonic acid (pH=7) was added into induction medium from day 16 to day 28 continuously. Oct4 GFP-positive CiPSC colonies emerged about on day 32.

Embryoid Body Formation Test

Embryoid body (EB) formation *in vitro* was performed as described previously (Mao et al., 2014). CiPSCs were removed off feeder cells twice based on their differences in the adherence to the bottom of dish. The CiPSCs were diluted to 4×10^4 per milliliter. Every 30 μl was pipetted to form a hanging drop on the cover of the 100-mm dish. EBs formed on day 4 and were transferred to 6 well plates for adherent culture. EBs were fixed for immunofluorescence staining using markers of three embryonic germ layers on day 15.

Teratoma Formation Test

1×10^6 CiPSCs were injected subcutaneously into about 6-week-old immunodeficient nude mice. About 4 weeks after injection, the mice were humanely sacrificed, and the teratomas were excised, fixed in 4% paraformaldehyde, dehydrated in gradient ethanol, embedded in paraffin, and sectioned for histological examination by haematoxylin and eosin staining.

Production of Chimeras and Genotyping

Approximately 10-15 CiPSCs were injected into four or eight-cell embryos as hosts using a Piezo injector as described (Huang et al., 2008). Injected embryos were cultured overnight in $KSOM_{AA}$ medium. Blastocysts were transferred into uterine horns of 2.5 dpc surrogate mice. Pregnant females delivered pups naturally at about 19.5 dpc. Pups were identified initially by coat color. The contribution of CiPSCs to various tissues in chimeras was confirmed by standard DNA microsatellite genotyping analysis using D12Mit136 primers: 5′-TT AAT TTT GAG TGG GTT TGGC-3′ and 5′-TTG CTA CAT GTA CAC TGA TCT CCA-3′, about 147 bp.

Immunofluorescence Microscopy

Cells were washed twice in PBS, fixed in freshly prepared 3.7% paraformaldehyde for 30 min at 4°C, washed once in PBS and permeabilized in 0.1% Triton X-100 in blocking solution (3% goat serum plus 0.1% BSA in PBS) for 30 min at room temperature, then washed once in PBS, and left in blocking solution for 2 h. Cells were incubated overnight at 4°C with primary antibodies against OCT4 (sc5279, Santa Cruz), NANOG (ab80892, Abcam), SSEA-1 (MAB4301, Millipore), βIII-tubulin (CBL412, Chemicon), AFP (DAK-N1501, Dako), α-SMA (ab5694-100, Abcam), NESTIN (MAB353, Millipore), ZSCAN4 (AB4340, Millipore), γH2AX (05-636, Millipore), TRF1 (TRF12-S, Alpha Diagnostic). Anti-Pan-Kcr (PTM-502) was permeabilized and blocked in 5% BSA in PBS. Cells were washed three times (each for 15 min) with blocking solution, and incubated for 2 h with secondary antibodies at room temperature. Goat Anti-Mouse IgG (H+L) FITC (115-095-003, Jackson) and Goat Anti-Rabbit IgG (H+L) Alexa Fluor® 594 (111-585-003, Jackson), diluted 1:200 with blocking solution, were used. Samples were washed, and counterstained with 0.5 μg/ml Hoechst 33342 (H1398, MP) or DAPI in Vectashield mounting medium. Fluorescence was detected and imaged using Axio-Imager Z2 fluorescence microscope (Carl Zeiss).

Fluorescence Microscopy of Teratoma Sections

Briefly, after being deparaffinized, rehydration and wash in 0.01 M PBS (pH 7.2-7.4), sections were incubated with 3% H_2O_2 for 10 min at room temperature to block endogenous peroxidase, subjected to high pressure antigen recovery sequentially in 0.01 M citrate buffer (pH 6.0) for 3 min, incubated with blocking solution (5% goat serum and 0.1% BSA in PBS) for 2 h at room temperature, and then with the diluted primary antibodies overnight at 4 °C. The following primary antibodies were used for immunocytochemistry: NESTIN (1:200, MAB353, Millipore), SMA (1:200, ab5694, Abcam) and AFP (1:5, DAK-N150130, Dako). Blocking solution without the primary antibody served as negative control. After washing with PBS, sections were incubated with appropriate secondary antibodies (1:200, goat anti-mouse IgG (H+L) FITC (115-095-003, Jackson) or goat anti-rabbit IgG (H+L) AlexaFluor® 594 (111-585-003, Jackson). Nuclei were stained using Vectashield medium (Vector) added with DAPI and photographed with Zeiss Axio Imager Z2.

Gene Expression Analysis by Quantitative Real-Time PCR

Total RNA was purified using a RNA mini kit (Qiagen), treated with DNase I (Qiagen), and the cDNA was generated from 2μg RNA using Oligo (dT)18 primer (Takara) and M-MLV Reverse Transcriptase (Invitrogen). Real-time quantitative PCR reactions were set up in duplicate with the FS Universal SYBR Green Master (Roche) and carried out on an iCycler MyiQ2 Detection System (BIO-RAD). All reactions were carried out by amplifying target genes and internal control in the same plate. Each sample was repeated three times and normalized using GAPDH as the internal control. The amplification was performed for primary denaturation at 95°C for 10 min, then 40 cycles of denaturation at 95°C for 15 s, annealing and

elongation at 58°C for 1 min, and the last cycle under 55-95°C for dissociation curve. Relative quantitative evaluation of target gene was determined by comparing the threshold cycles. Primers were confirmed their specificity with dissociation curves. Most primers were designed using the IDT DNA website and listed in Table S1.

Western Blot

Cells were washed twice in PBS, collected, and lysed in cell lysis buffer on ice for 30 min and then sonicated for 1 min at 60 of amplitude with 2 s intervals. After centrifugation at 10000g, 4°C for 10 min, supernatant was transferred into new tubes. The concentration of the protein sample was measured by bicinchoninic acid, and then protein samples were boiled in SDS Sample Buffer at 99°C for 10 min. 10μg total protein of each cell extract was resolved by 10% or 12% Acr-Bis SDS-PAGE and transferred to polyvinylidene difluoride membranes (PVDF; Millipore). Nonspecific binding was blocked by incubation in 5% skim milk or 5% BSA in TBST at room temperature for 2 h. Blots were then probed with primary antibodies overnight incubation at 4°C with OCT4 (sc5279, Santa Cruz), NANOG (ab21603, Abcam), SOX2 (AB5603, Millipore), TCSTV1/3 (custom-made), ZSCAN4 (AB4340; Millipore), H3K9me3 (ab8898, Abcam), H3K9ac (04-1003, Millipore), H3K27me3 (07-449, Millipore), Pan-Kcr (PTM-502, PTM biolabs), H3 (ab1791, Abcam), and β-ACTIN (P30002, Abmart). Immunoreactive bands were then probed for 2 h at room temperature with the appropriate horseradish peroxidase (HRP)-conjugated secondary antibodies, anti-Rabbit IgG-HRP (GE Healthcare, NA934V), or goat anti-Mouse IgG (H + L)/HRP (ZSGB-BIO, ZB-2305). Protein bands were detected by Chemiluminescent HRP substrate (Millipore, WBKLS0500). Information for antibodies is listed in Table S2.

Telomerase Activity Assay

Telomerase activity was determined by the Stretch PCR method according to manufacturer's instruction using TeloChaser Telomerase assay kit (T0001, MD Biotechnology). About 2.5×10^4 cells from each sample were lysed. Lysis buffer served as negative controls. PCR products of cell lysate were separated on non-denaturing TBE-based 10% polyacrylamide gel electrophoresis and visualized by ethidium bromide staining.

FACS Analysis

For the FACS analysis of Zscan4 expression, CiPSCs were collected and washed with cold PBS, then fixed in cold 70% ethanol, permeabilized in 0.1% Triton X-100 in blocking solution (3% goat serum in PBS) for 30 min, washed three times and left in blocking solution for 2 h. Cells were incubated overnight at 4 °C with primary antibody against ZSCAN4, washed three times, and incubated for 2h with secondary antibody 594-goat anti-rabbit IgG (H+L) (111-585-003, Jackson) diluted 1:200 with blocking solution. Samples were washed three times with PBS and FACS analysis performed using a FACScan Flow Cytometer (BD Biosciences). For

analysis of endogenous Oct4-GFP, cells were collected and washed with cold PBS, subject to FACS analysis.

Telomere Q-FISH

Telomere length was estimated by telomere Q-FISH as described previously (Herrera et al., 1999; Poon et al., 1999). Telomeres were denatured at 80 ºC for 3 min and hybridized with FITC-labeled $(CCCTAA)_3$ peptide nucleic acid (PNA) probe at 0.5 μg/ml (Panagene, Korea). Chromosomes were stained with 0.5 μg/ml DAPI. Fluorescence from chromosomes and telomeres was digitally imaged on a Zeiss microscope with FITC/DAPI filters, using AxioCam and AxioVision software 4.6. Telomere length shown as telomere fluorescence intensity was integrated using the TFL-TELO program (a gift kindly provided by Peter Lansdorp, Terry Fox Laboratory).

Telomere Restriction Fragment (TRF) Measurement

The TRF analysis was performed using a commercial kit (TeloTAGGG Telomere Length Assay, catalog no. 12209136001, Roche Life Science), based on the method described previously (Sung et al., 2014). DNA was extracted from cells. A total of 3 μg DNA was digested overnight with MboI at 37°C and electrophoresed through 1% agarose gels in 0.5 × TBE at 14°C for 16 h at 6 V/cm with an initial pulse time of 1 s and end in 12 s using a CHEF Mapper pulsed field electrophoresis system (Bio-Rad). The gel was blotted and probed using reagents in the kit.

Telomere Chromosome Orientation-Fluorescence *In Situ* Hybridization (CO-FISH)

CO-FISH assay was performed as described previously with minor modification (Bailey et al., 2004). Subconfluent ES or CiPS cells were incubated with BrdU (10 μM) for 10-12 h, and MEF or day 28 reprogramming cells incubated with BrdU (10 μM) for 20-24 h. Nocodazole at 0.3 μg/ml was added for 3 h (and 5 h for day 28 reprogramming cells) prior to cell harvest, and metaphase spreads were prepared by a routine method. Chromosome slides were treated with RNase A, fixed with 4% formaldehyde, then stained with Hoechst 33258 (0.5 mg/ml), incubated in 2 × SSC (Invitrogen) for 15 min and exposed to 365 nm UV light (Stratalinker 1800UV irradiator) for 40 min. The BrdU-substituted DNA was digested with Exonuclease III (Takara). The slides were then dehydrated through ethanol series and air-dried. PNA-FISH was performed with FITC-OO-$(CCCTAA)_3$ (Panagene, F1009). Slides were hybridized, washed, dehydrated, mounted, and counter-stained with 1.25 μg/ml DAPI in VectaShield antifade medium. Digital images were captured using CCD camera on Zeiss Imager Z2 microscope. To analyze telomere sister chromatid exchange (T-SCE), one signal at each end of the chromosome was counted as no T-SCE, while two signals at both chromatids on one chromosome end were identified as one T-SCE. At least 15 metaphase spreads were counted for frequency of T-SCE. Pairwise comparisons for statistical significance were made by t-tests.

C-circle Assay

The protocol for c-circle amplification was slightly modified from the method described (Henson et al., 2017). Briefly, 400ng genomic DNA was digested with MboI restriction enzymes at 37°C overnight and purified by phenol-chloroform extraction. Genomic DNA from ALT positive (U2OS) cells and ALT negative (Hela) cells was used as a positive and negative control. DNA was diluted in double distilled water. Samples (10 µl) were combined with 10 µl 1×Φ29 Buffer (NEB, UK) containing BSA, 0.2 mM each dATP, dGTP, dTTP and incubated in the presence or absence of 5U ΦDNA polymerase (NEB, UK) at 30°C for 12 h and then at 65°C for 20 min. Added 200 µl 6×SSC to reaction products and dot-blotted onto a 6×SSC-soaked nylon membrane. DNA was cross-linked onto the membrane and hybridized with a DIG-labeled probe $(CCCTAA)_4$ to detect C-circle amplification products. Blots were washed, and exposed to Tanon 5200.

Apoptosis by TUNEL Assay

Apoptosis was revealed by catalytically incorporating fluorescein-12-dUTP at 3´-OH DNA ends with the Terminal Deoxynucleotidyl Transferase, Recombinant, enzyme (rTdT) using a commercial kit (TB235, Promega). Briefly, after being fixed in 4% paraformaldehyde, washed in PBS, and permeabilized in 0.2% Triton X-100, cells were incubated with rTdT buffer at 37°C for one hour, stained with DAPI for 15 min, washed, mounted in Vectashield, and immediately analyzed under a fluorescence microscope using a standard fluorescein filter set.

Library Preparation and RNA-sequencing

Briefly, mRNA was purified from total RNA using poly-T oligo-attached magnetic beads. Fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5x). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H^-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3’ ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. To select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 µl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37 °C for 15 min followed by 5 min at 95 °C before PCR. PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system.

Clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer’s instructions. After cluster generation, the library preparations

were sequenced on an Illumina Hiseq platform and 125 bp/150 bp paired-end reads were generated.

Differential Gene Expression Analysis

DESeq provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting P-values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate (Hochberg and Benjamini, 1990). Genes with an adjusted P-value <0.05 found by DESeq were assigned as differentially expressed. For DEGSeq without biological replicates: prior to differential gene expression analysis, for each sequenced library, the read counts were adjusted by edgeR program package through one scaling normalized factor. Differential expression analysis of two conditions was performed using the DEG-Seq R package (1.20.0). The P-values were adjusted using the Benjamin & Hochberg method(Hochberg and Benjamini, 1990). Corrected P-value of 0.005 and $\log_2$ (fold change) of 1 were set as the threshold for significantly differential expression.

Clustering and analysis of the early embryo development genes and two-cell genes provided in the heatmap was obtained using RNA-seq data published (Fan et al., 2015; Macfarlan et al., 2012). All 196 differentially expressed genes in our RNA-seq data (Table S3) were compared with genes expressed at various early development stages obtained by single cell RNA-seq data (Fan et al., 2015), using fisher's exact test (P adj$<$0.01). The 196 differentially expressed genes also were compared with the RNA-seq data from ESC cultures (Macfarlan et al., 2012). The calculated z score of selected genes was used for heatmap.

ChIP-qPCR analysis

Cells were harvested and fixed by freshly prepared 1% paraformaldehyde solution for 8 min at room temperature. Glycine was added to quench unreacted formaldehyde. The nuclei were extracted, lysed, and sonicated. DNA fragments were then enriched by immunoprecipitation with H3K9me3 (ab8898; Abcam) and HP1α (05-689; Millipore) antibody. The eluted protein:DNA complex was reverse-crosslinked at 65 °C overnight. DNA was recovered after proteinase and RNase A treatment. ChIP-enriched DNA was analyzed by qPCR using primers for *Zscan4*, *Chr13_subtel* and *ERVK10c* loci (Table S1). *β-actin* served as negative control.

Whole-exome Sequencing

Paired-end DNA library were prepared according to manufacturer's instructions (Agilent). Genomic DNAs (gDNA) from cell samples were sheared into 180~280 bp fragments by Covaris S220 sonicator. Ends of the gDNA fragments were repaired; 3′ ends were adenylated. Both ends of the gDNA fragments were ligated at the 3′ ends with paired-end adaptors (Illumina) with a single 'T' base overhang and purified using AMPure SPRI beads from Agencourt. The adaptor-modified gDNA

fragments were enriched by six cycles of PCR using SureSelect Primer and SureSelect ILM Indexing PreCapture PCR Reverse Primer. The concentration and the size distribution of the libraries were determined on an Agilent Bioanalyzer DNA 1000 chip. Whole exome capture was carried out using Agilent's SureSelect Mouse All Exon V1 Agilent 5190-4642. An amount of 0.5 μg prepared gDNA library was hybridized with capture library of biotinylated RNA baits for 5 min at 95 °C, 24 h at 65 °C. The captured DNA−RNA hybrids were recovered using Dynabeads MyOne Streptavidin T1 from Dynal. DNA was eluted from the beads and desalted using Qiagen MinElute PCR purification columns. The purified capture products were then amplified using the SureSelect ILM Indexing Post Capture Forward PCR Primer and PCR Primer Index 1 through Index 16 (Agilent) for 12 cycles. 50 Mb of DNA sequences of 221,784 exons from 24,306 genes were captured. After DNA quality assessment, captured DNA library were sequenced on Illumina Hiseq 4000 sequencing platform (Illumina) according to manufacturer's instructions for paired-end 150 bp reads (Novogene, Beijing). Libraries were loaded onto paired-end flow cells at concentrations of 14–15 pM to generate cluster densities of 800,000–900,000 per mm2 using Illumina cBot and HiSeq paired-end cluster kit version 1.

Sequence Alignment and Variant Calling

Variant calling was performed as previously described (Gore et al., 2011). Briefly, paired-end clean reads in FastQ format generated by the Illumina pipeline were aligned to the mouse reference genome mm10 by Burrows-Wheeler Aligner (BWA) to obtain the original mapping results stored in BAM format. SAMtools, Picard (http://broadinstitute.github.io/picard/) were used to sort BAM files and do duplicate marking to generate final BAM files for computation of the sequence coverage and depth. BCFtools was used to call variants, and VariantAnnotation was performed to do annotation for VCF (Variant Call Format). Variant positions were obtained at this step. RefSeq was applied for annotation to determine amino acid alternations. Variants obtained from previous steps were compared against SNPs present in the dbSNP to discard known SNPs. The consensus sequences in mouse CiPSC samples were then used to compare with progenitor MEF samples to find candidate novel mutations. Each heterozygous SNP identified in CiPSC lines that was not observed in the progenitor line was considered candidate mutations. We used DAVID to analyze the somatic-coding mutations and check for commonly mutated pathways. The mutations do not seem to be mutated in any common pathways.

Statistical Analysis

Data were analyzed by student's t-test or $\chi 2$ test or fisher's exact test for paired comparison, or by ANOVA and means compared by Fisher's protected least-significant difference (PLSD) for multiple comparisons using the StatView software from SAS Institute Inc. (Cary, NC). Significant differences were defined as *P< 0.05, **P< 0.01, or ***P< 0.001.

REFERENCES

Bailey, S.M., Brenneman, M.A., and Goodwin, E.H. (2004). Frequent recombination in telomeric DNA may extend the proliferative life of telomerase-negative cells. Nucleic Acids Res *32*, 3743-3751.

Fan, X., Zhang, X., Wu, X., Guo, H., Hu, Y., Tang, F., and Huang, Y. (2015). Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. Genome Biol *16*, 148.

Gore, A., Li, Z., Fung, H.L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature *471*, 63-67.

Henson, J.D., Lau, L.M., Koch, S., Martin La Rotta, N., Dagg, R.A., and Reddel, R.R. (2017). The C-Circle Assay for alternative-lengthening-of-telomeres activity. Methods *114*, 74-84.

Herrera, E., Samper, E., Martin-Caballero, J., Flores, J.M., Lee, H.W., and Blasco, M.A. (1999). Disease states associated with telomerase deficiency appear earlier in mice with short telomeres. EMBO J *18*, 2950-2960.

Hochberg, Y., and Benjamini, Y. (1990). More powerful procedures for multiple significance testing. Stat Med *9*, 811-818.

Huang, J., Deng, K., Wu, H., Liu, Z., Chen, Z., Cao, S., Zhou, L., Ye, X., Keefe, D.L., and Liu, L. (2008). Efficient production of mice from embryonic stem cells injected into four- or eight-cell embryos by piezo micromanipulation. Stem Cells *26*, 1883-1890.

Macfarlan, T.S., Gifford, W.D., Driscoll, S., Lettieri, K., Rowe, H.M., Bonanomi, D., Firth, A., Singer, O., Trono, D., and Pfaff, S.L. (2012). Embryonic stem cell potency fluctuates with endogenous retrovirus activity. Nature *487*, 57-63.

Mao, J., Zhang, Q., Ye, X., Liu, K., and Liu, L. (2014). Efficient induction of pluripotent stem cells from granulosa cells by Oct4 and Sox2. Stem Cells Dev *23*, 779-789.

Poon, S.S., Martens, U.M., Ward, R.K., and Lansdorp, P.M. (1999). Telomere length measurements using digital fluorescence microscopy. Cytometry *36*, 267-278.

Sung, L.Y., Chang, W.F., Zhang, Q., Liu, C.C., Liou, J.Y., Chang, C.C., Ou-Yang, H., Guo, R., Fu, H., Cheng, W.T.*, et al.* (2014). Telomere elongation and naive pluripotent stem cells achieved from telomerase haplo-insufficient cells by somatic cell nuclear transfer. Cell Rep *9*, 1603-1609.