

Transcriptional Signature and Memory Retention of Human-Induced Pluripotent Stem Cells

Maria C. N. Marchetto¹, Gene W. Yeo², Osamu Kainohana³, Martin Marsala³, Fred H. Gage¹, Alysso R. Muotri^{4*}

1 The Salk Institute for Biological Studies, Laboratory of Genetics, La Jolla, California, United States of America, **2** University of California San Diego, School of Medicine, Department of Cellular & Molecular Medicine, Stem Cell Program, La Jolla, California, United States of America, **3** University of California, San Diego, School of Medicine, Department of Anesthesiology, La Jolla, California, United States of America, **4** University of California San Diego, School of Medicine, Department of Pediatrics/Rady Children's Hospital San Diego, Department of Cellular & Molecular Medicine, Stem Cell Program, La Jolla, California, United States of America

Abstract

Genetic reprogramming of somatic cells to a pluripotent state (induced pluripotent stem cells or iPSCs) by over-expression of specific genes has been accomplished using mouse and human cells. However, it is still unclear how similar human iPSCs are to human Embryonic Stem Cells (hESCs). Here, we describe the transcriptional profile of human iPSCs generated without viral vectors or genomic insertions, revealing that these cells are in general similar to hESCs but with significant differences. For the generation of human iPSCs without viral vectors or genomic insertions, pluripotent factors Oct4 and Nanog were cloned in episomal vectors and transfected into human fetal neural progenitor cells. The transient expression of these two factors, or from Oct4 alone, resulted in efficient generation of human iPSCs. The reprogramming strategy described here revealed a potential transcriptional signature for human iPSCs yet retaining the gene expression of donor cells in human reprogrammed cells free of viral and transgene interference. Moreover, the episomal reprogramming strategy represents a safe way to generate human iPSCs for clinical purposes and basic research.

Citation: Marchetto MCN, Yeo GW, Kainohana O, Marsala M, Gage FH, et al. (2009) Transcriptional Signature and Memory Retention of Human-Induced Pluripotent Stem Cells. PLoS ONE 4(9): e7076. doi:10.1371/journal.pone.0007076

Editor: Thomas A. Reh, University of Washington, United States of America

Received: June 26, 2009; **Accepted:** August 20, 2009; **Published:** September 18, 2009

Copyright: © 2009 Marchetto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: University of California San Diego Stem Cell Research Program Startup Funds, California Institute for Regenerative Medicine, The Lookout Fund and the Picower Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: muotri@ucsd.edu

Introduction

Genetic reprogramming to a pluripotent state of mouse somatic cells was first achieved by ectopic expression of four factors (Oct4, Sox2, Klf4 and c-Myc) using retroviruses [1]. Such cells were named induced pluripotent stem cells (iPSCs). Subsequently, this method was applied to human cells using the same factors or a different combination in a lentivirus vector (Oct4, Sox2, Lin28 and Nanog) [2–5]. Both mouse and human iPSCs are similar to embryonic stem cells (ESCs) with respect to their morphology, cell behavior, gene expression, epigenetic status and differentiation potential both in culture and *in vivo*. However, to date, a comprehensive transcriptional analysis has not been reported comparing human ESCs and iPSCs. One reason is that the technology used to derive iPSCs is not “footprint-free” and thus, subjected to transcriptional interference.

Viral vectors are known to affect the transcriptional profile from target cells, altering their behavior and sometimes inducing apoptosis [6]. Moreover, the reactivation of the viral transgene was also implicated in tumorigenesis from iPSC-derived chimeric mice [7]. Also, random integration may influence the molecular signatures of iPSCs by interrupting regulatory regions in the human genome. Interestingly, a transcriptional analysis revealed that transgene expression from not completely silenced viral vectors could, in fact, perturb global gene expression in hiPSCs [8].

Several attempts were made to generate a viral-free, integration-free iPSCs. The generation of iPSCs with later excision of reprogramming factors was recently achieved; still, the genome continues to be affected by random solo-LTR insertions from viral vectors [8]. Mouse iPSCs were also generated by multiple transient expression of Oct4, Sox2 and Klf4 from embryonic fibroblasts at very low efficiency [9]. Recently, a two-step seamless factor removal from iPSCs using transposase-stimulated excision was recently reported [10,11]. Although evidence that the system might work in human cells was presented, it needs further validation in more rigorous pluripotent assays [10,11]. A “footprint-free” and highly efficient system of generating human iPSCs would help to determine the molecular mechanism of cellular reprogramming and accelerate the search for efficient compounds that will replace the original factors without side effects.

The timing of the reprogramming and the factors required seem to vary depending on cellular context [12–17]. The susceptibility of a somatic cell to reprogram may depend on how similar its transcriptional profile is to ESCs. Of note, mouse neural stem cells (NSCs) were reprogrammed using only one (Oct4) or two factors (Oct4 and Klf4), due to the endogenously high expression of pluripotent genes, such as Sox2 and c-Myc, as well as several intermediate reprogramming markers [14,17,18]. Fibroblasts that already carry the Oct4 transgene can be reprogrammed with fewer factors, facilitating the study of nuclear reprogramming [19].

Moreover, although reprogramming can be achieved without c-Myc, iPSC generation is more efficient when the gene is present [20,21]. Furthermore, recent data suggest that c-Myc expression primes cells for iPSC conversion, accelerating the initial steps of reprogramming to achieve high efficiency [22]. Such observations prompted us to use human NSCs expressing c-Myc, as a model to facilitate the generation of iPSCs and to study the reprogramming steps.

Results

Oct4 and Nanog can reprogram human neural stem cells

Our starting material was a multipotent, karyotypically normal, c-Myc-immortalized human NSC line derived from a tissue sample of human midbrain (10 weeks of gestation). Our rationale was that the elevated expression of c-Myc and Sox2 in these cells might prompt them to reprogram more easily than reported for other types. First, we examined whether the combination of Oct4 and Nanog would reprogram these cells to a pluripotent state [23,24]. The human NSCs have a typical, undifferentiated neural stem cell morphology when expanding as monolayers on laminin-coated plates (**Fig. 1A**). NSCs were infected once with lentivirus expressing Oct4 and Nanog (ON) and plated onto a layer of irradiated mouse embryonic fibroblasts (MEFs) in human ESCs (hESC) medium [25]. Individual cells positive for alkaline phosphatase (AP), a marker for pluripotent cells, appear as early as 4 days after infection (**Fig. 1A**, inset). Interestingly, the efficiency was around 1–3%, as measured by the number of AP-positive colonies, at 14 days post-infection. Single infection with an empty control virus or Nanog alone did not produce any colonies (**Fig. 1B**). In the first week after infection, hundreds of small colonies grew rapidly and had hESC morphology (**Fig. 1C, D**). Two weeks after infection, iPSC colonies with a mature morphology similar to hESCs were distinguished from the original NSC population (**Fig. 1E–F**). The NSCs-iPSC(ON) colonies were then manually isolated and propagated under feeder-free growth conditions on matrigel-coated dishes. They expressed markers of undifferentiated ESCs, including Lin28, TRA-1-60 and SSEA-4, confirming the genetic reprogramming by the two factors, Oct4 and Nanog (**Fig. 1G**). Several iPSC lineages were established from independent infections and mechanically expanded for at least 20 passages while maintaining a normal karyotype (data not shown).

A viral-free, integration-free reprogramming approach

To generate human iPSCs without the use of viral delivery vectors or genomic insertions, the Oct4 and Nanog cDNAs were independently cloned under the CMV promoter into a plasmid (pCEP) with the *trans*-acting Epstein-Barr associated nuclear antigen 1 (EBNA-1) gene and the *cis*-DNA element *oriP*. The combination of EBNA-1 and *oriP* elements allows for a transient extra-chromosomal (episomal) state, avoiding genetic integration in human and non-human primate cells [26–30]. The constructs also contain a mammalian selection marker (the hygromycin resistant gene). Human NSCs were electroporated with equimolar concentrations of the two episomal plasmids (pCEP-Oct4 and pCEP-Nanog) or the EGFP-reporter plasmid and plated on MEFs under hESC conditions (**Fig. 2A**). Previous data in the literature suggested that reprogramming factors should be maintained for up to 12 days during iPSC generation from mouse cells [31,32]. Hygromycin selection was maintained for only a week, but transgene expression from the plasmid carrying the EGFP reporter gene suggested that the plasmid remained in the cells for another week before being eliminated (**Figure S1**). After 10–12 days, small

iPSC colonies were first noted. Colonies were mechanically isolated and propagated under hESC conditions on matrigel. At this point, some colonies seemed unstable, with a strong tendency to spontaneously differentiate and form a heterogeneous population of cells (**Fig. 2B, C**). Undifferentiated cells were manually selected from differentiated cells according to morphology until a homogeneous population of iPSCs was achieved (**Fig. 2D**). The iPSC colonies were morphologically indistinguishable from hESCs, forming tight colonies of cells with a large nucleus to cytoplasm ratio and prominent nucleoli (**Fig. 2E**), and they did not display the NSCs' original cell morphology (**Fig. 2F**). The efficiency was higher (0.1–1%) when compared to fibroblasts reprogrammed with retroviruses. We established several cell lineages from three independent transfection experiments and chose three lines (iPSC1, iPSC2, iPSC3) for further characterization.

These three iPSC colonies expressed several pluripotent markers and were able to form embryoid bodies (EBs) *in vitro* (**Fig. 2G, H**). They were also able to express markers of the three germ layers, suggesting that they re-established pluripotency at the molecular and cellular levels (**Fig. 2I**). PCR DNA fingerprinting confirmed their derivation from NSCs rather than from a contaminating hESC line (**Figure S2**). All iPSC clones could be successfully propagated for more than 30 passages while maintaining a normal karyotype (data not shown). Plasmid transfection may lead to random integration into the genome at low frequency. To test for genomic integration of plasmid DNA, we designed several sets of PCR primers to amplify various parts of the vector and transgenes (**Fig. 3A, B**). Teratomas containing derivatives from all three embryonic germ layers confirmed that the hiPSCs (but not the original NPCs used) were pluripotent and able to differentiate to complex tissues in two different experimental settings (**Fig. 2J** and **Figure S3**). Additionally, southern blot analyses did not detect integration of plasmids in these clones (**Fig. 3C**). DNA from the transfected plasmids was not detected in any established colony using either method, indicating a lack of genomic insertion and suggesting that the episomal vectors had been diluted from the cells over time.

Human iPSCs have similar levels of myc when compared to hESCs

We then analyzed if myc levels from these iPSCs derived from NSCs would change after reprogramming. Interestingly, despite the fact that the NSCs were immortalized with ectopic expression of myc, the transcriptional activity of myc is higher in iPSCs compared to NSCs. Moreover, iPSCs clones have similar myc transcriptional levels to hESCs (**Figure S4**). Together, these observations indicate that the myc expression will likely not interfere with the global transcription profile on the iPSCs.

A transcriptional signature for human iPSCs

Next, we asked if the global molecular signatures of two plasmid-free iPSC lines (iPSC1, iPSC2) resembled those of available hESC lines, namely HUES6 and Cyt25. Gene expression profiles measured using human genome Affymetrix Gene Chip arrays were grouped by hierarchical clustering, and correlation coefficients were computed for all pair-wise comparisons (**Fig. 4A**). We observed that the two iPSCs lines were almost indistinguishable from each other and that the two hESC lines were also highly similar to each other. Clearly, the iPSC and hESC lines were globally more similar to each other than to the NSC line (**Fig. 4A**), and combined with manual inspection of the gene expression of several known pluripotent (Oct4, LIN28, Sox2 and Nanog) and neural stem cell markers (Sox2, Nestin and Musashi2) as measured

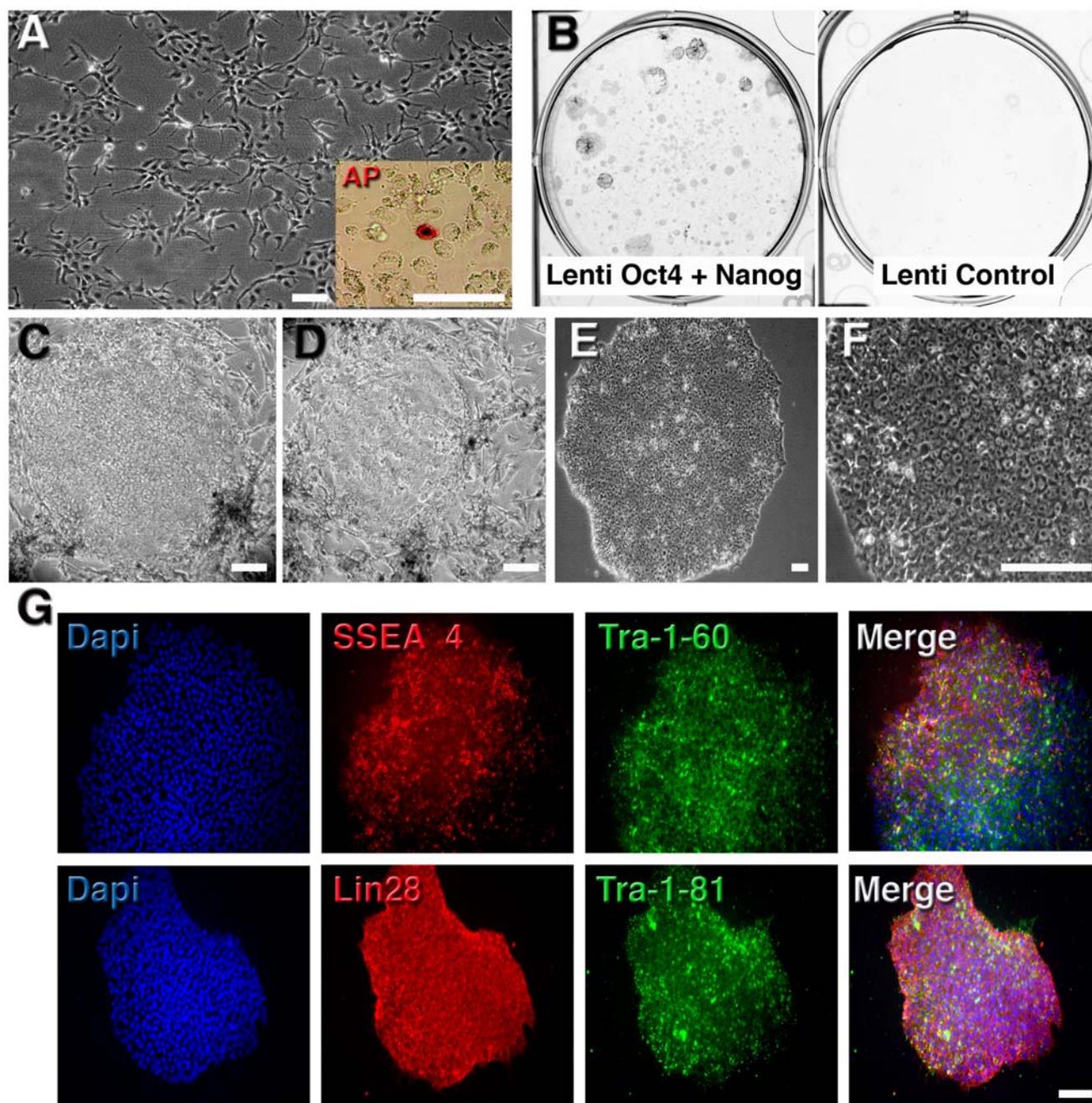


Figure 1. Efficient and rapid generation of iPSCs from human fetal NSCs using two factors. **A**, Morphology of human fetal NSCs before lentiviral infection. Inset: after 3 days post-infection with Lenti-Oct4 and Lenti-Nanog, individual cells expressed alkaline phosphatase (AP). **B**, Example of infected plates stained for AP at 14 days post-infection showing several AP-positive colonies. Control infection did not result in any AP-positive colonies. **C** and **D**, Aspect of colonies 14 days after infection growing in MEFs. **E**, Established human iPSC colonies, with well-defined borders and compact cells, are morphologically similar to hESCs. **F**, Typical image of iPSCs growing in feeder-free conditions. **G**, Representative immunofluorescence analysis of iPSCs growing on matrigel. Clear expression of pluripotent markers is observed. Bar = 150 μ m.
doi:10.1371/journal.pone.0007076.g001

on the arrays, we concluded that the reprogramming was successful (**Fig. 4A**).

Despite the global similarity between iPSCs and hESCs, the profiles were not completely indistinguishable, which led us to study what the molecular differences were. Four independent (A versus B) group-wise comparisons were performed to identify differentially expressed genes: (i) iPSC versus hESC (1,952 Refseq-annotated genes were significantly enriched in iPSCs versus

hESCs; 1,072 genes were enriched in hESCs versus iPSCs at $P < 0.01$ after correcting for multiple hypotheses testing); (ii) iPSC versus NSC (3,347 genes were significantly enriched in iPSCs versus NSCs; 2,959 genes were enriched in NSCs versus iPSCs); (iii) hESC versus NSC (2,376 genes were significantly enriched in hESCs versus NSCs; 2,541 genes were enriched in NSCs versus hESCs); (iv) iPSC and hESC versus NSC (3,730 genes were significantly enriched in iPSCs and hESCs, versus NSCs and 3,638

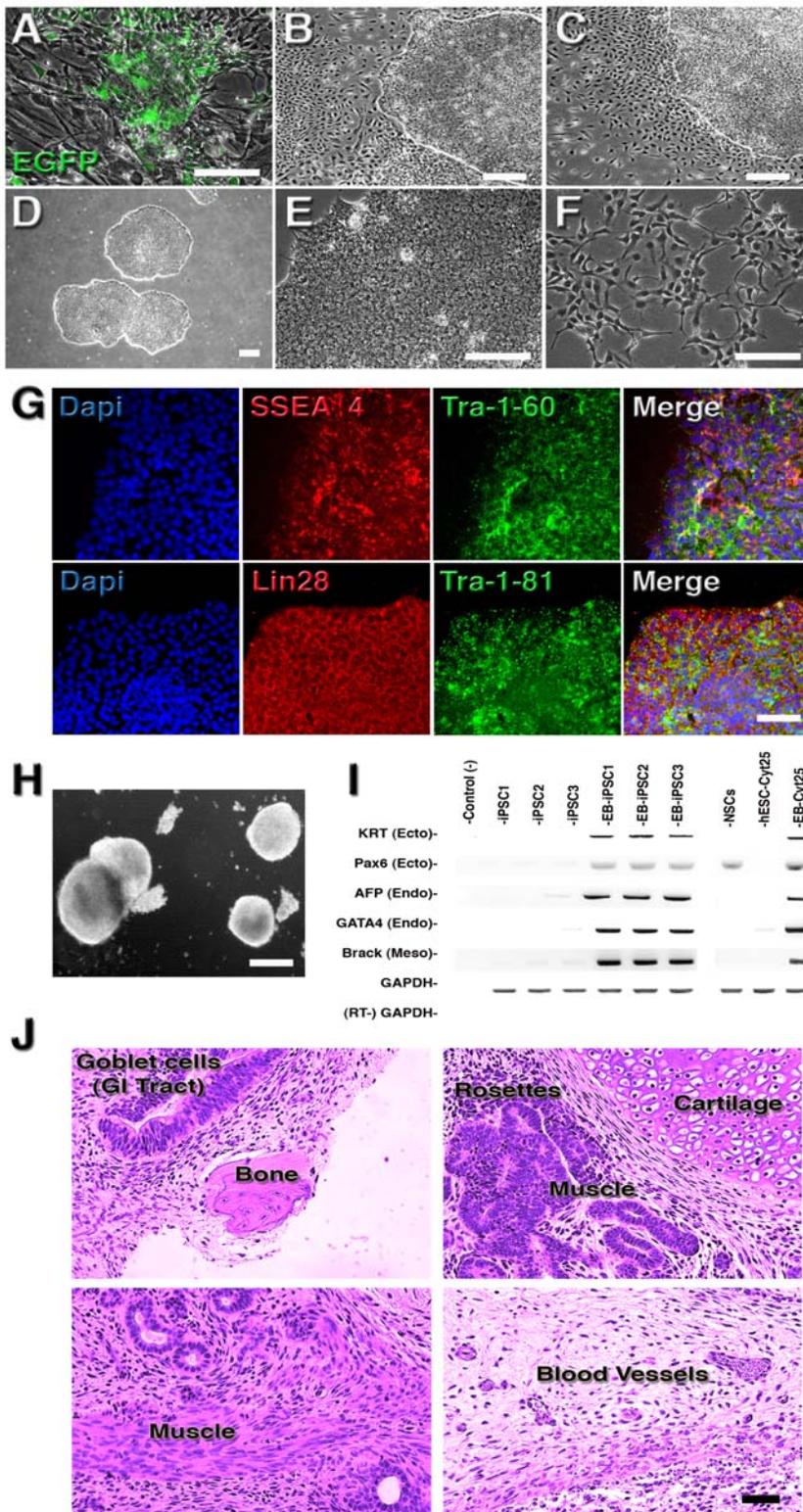


Figure 2. Generation of virus-free, integration-free human iPSCs. **A**, Aspect of human NSCs after plasmid electroporation and plating on MEFs. **B** and **C**, Some selected colonies display a strong differentiation tendency in feeder-free conditions. **D**, Established iPSC lines are morphologically similar to hESCs. **E**, iPSCs have a large nucleus-to-cytoplasm ratio and prominent nucleoli when compared to original NSCs (**F**). **G**, Immunofluorescence analysis of iPSCs growing on matrigel showed clear expression of typical ESC markers. **H**, *In vitro* differentiation of iPSCs into EBs. **I**, RT-PCR from undifferentiated and EB-derived iPSCs showing expression of markers for all three primary germ cell layers. The hESCs Cyt25 was used as a benchmark. **J**, Hematoxylin and eosin staining of teratoma sections generated from integration-free iPSC lines showing differentiation in three germ layers: goblet cells in gastrointestinal (GI) tract (endoderm); neural rosettes (ectoderm) and blood vessels, muscle and cartilage/bone (mesoderm). Bar = 150 μ m.
doi:10.1371/journal.pone.0007076.g002

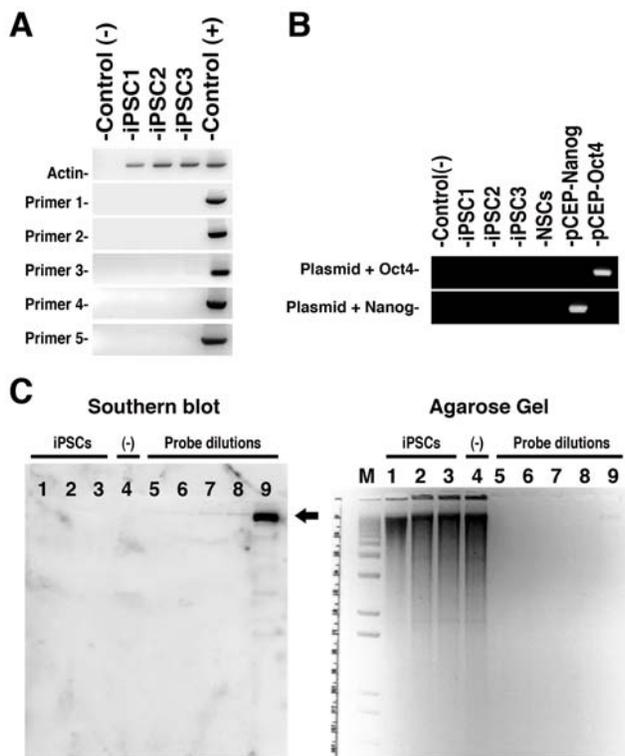


Figure 3. Absence of plasmid integration on virus-free iPSCs. **A** and **B**, PCR analyses for plasmid integration in genomic DNA from the iPSC clones. Controls: (-) water; (+) pCEP4 plasmid. Primers were designed to specifically amplify plasmid backbone (**A**) or transgenes (**B**) (see Methods). **C**, Southern blot (left) membrane hybridization of 10 μ g of *Bam*HI-digested genomic DNA (see corresponding agarose gel on right) using a DNA probe from the pCEP backbone. Plasmid DNAs of pCEP-Oct4 and PCEP-Nanog, diluted to the equivalent of 0.5 integration per genome, were used as controls for probe dilution. Lanes: M, DNA molecular marker; 1- iPSC1; 2- iPSC2; 3- iPSC3; 4- NSCs (negative control); 5- probe 25 μ g; 6- probe 50 μ g; 7- 100 μ g; 8- 200 μ g and 9- 50 μ g. Arrow indicates expected probe size.
doi:10.1371/journal.pone.0007076.g003

genes were enriched in NSCs versus iPSCs and hESCs (**Tables S1 to S8** contain the full list of comparisons). Restricting these differentially expressed genes to ones that changed by at least 4-fold in any comparison, at a stringent p-value cutoff of $P < 0.0001$, we identified three groups of biologically interesting genes. The first group of iPSC-expressed genes was not sufficiently induced to comparable levels as in hESCs and was still at their original levels in NSCs (**Fig. 4B**). This group contained factors that were important in early embryonic fate, such as *Stella*, *ZFP42* (*REX1*), *CLDN10*, *LEFTY1* and *LEFTY2*. It is noteworthy that *ZFP42* has been shown to be dispensable for pluripotency in mouse ES cells [33], which may explain why the factor need not be induced highly. *Lefty1* has been shown to be important for pluripotency as well [34], and it is downstream of Oct4 and Sox2, but perhaps without the use of *Klf4* the *Lefty1* expression is not sufficiently induced. The second group contained iPSC-expressed genes that were not sufficiently repressed, such as *ZIC1*, *OLIG2*, *EN2* and *PTX3*, which were associated with the neuronal lineage (**Fig. 4C**). The third group consisted of genes that were upregulated in iPSCs, which were silenced in both NSCs and hESCs, suggesting that these genes may be downstream factors in the reprogramming step to induce pluripotent cells (**Fig. 5**). Overall, our transcriptome analyses indicated that, whereas the iPSCs are globally similar to

hESCs, they are not indistinguishable, primarily due to the insufficient suppression or induction of NSC-specific or early embryonic-specific genes, respectively, as well as a class of genes that was upregulated during the reprogramming step.

Oct4 alone is able to reprogram human NSCs

Next, we repeated the transient transfection using NSCs derived from the H1 hESC line that contains the EGFP reporter cassette knocked in the endogenous Oct4 gene by homologous recombination [35]. The H1-Oct4-EGFP cell line expressed EGFP, which turned off during differentiation (**Fig. 6A**). NSCs were generated using our previous established protocol and consisted of a cell population with a genetic profile distinct from both human fetal cells and hESCs [36,37]. NSCs derived from the H1-EGFP do not express EGFP (**Fig. 6A**). An EGFP-negative population of NSCs, isolated by FACS, was electroporated with both episomal plasmids carrying Oct4 and Nanog. Several iPSC colonies were observed as early as 10 days after transfection, becoming morphologically indistinguishable from the original H1-Oct4-EGFP cell line (**Fig. 6B, C**). As a control, we electroporated the same cell population with Oct4 only. Interestingly, we detected several colonies when cells were transfected with Oct4 alone. These colonies were positive for pluripotent makers, such as Nanog and *Lin28*, suggesting efficient reprogramming (data not shown). Our findings in human cells recapitulate recent data demonstrating that Oct4 alone is sufficient to reprogram mouse NSCs [17]. The observation that we could reprogram using only Oct4 but not Nanog alone, suggests that Oct4 is likely an upstream factor for cellular reprogramming. In fact, a previous study suggests that Nanog may function to stabilize pluripotency rather than being essential for the pluripotent stage [38].

Discussion

Using a simple methodology (**Fig. 7**), we demonstrated that it is possible to generate human iPSCs at a high frequency without viruses and with no evidence of genomic insertion. Human iPSCs were achieved using transient episomal vectors carrying the cDNAs for Oct4 and Nanog in a cell type that was likely more prone to genetic reprogramming, such as NSCs. Also, we demonstrated here for the first time that a myc-immortalized cell line could be successfully reprogrammed, opening new avenues for the study of several previously characterized immortalized cell lines that are relevant for the biological understanding of several disorders. The Myc-immortalized NSCs represent a reliable, homogenous and commercially available tool to dissect individual factors required for reprogramming. Myc levels after reprogramming is similar to hESCs, thus it may be a better standard model for fundamental reprogramming studies than fibroblasts. Furthermore, our data from hESC-derived NSCs indicate that reprogramming can be achieved without ectopic expression of the tumor-associated genes, *c-Myc* and *Klf4*. Future studies will show whether other sources of primary human NSCs can also be efficiently reprogrammed by such non-viral methodology. Human iPSCs generated by episomal vectors were then used to assess whether human iPSCs and ESCs are really equivalent at the molecular and functional levels, avoiding artifacts that may affect their genetic signature, differentiation behavior or developmental potential. Almost all previous studies have shown that the genetic profile of human iPSCs is comparable, but never identical, to that of hESCs. The slight differences between these two cellular populations could be attributed to viral insertions in the genome or incomplete genetic reprogramming. Our data suggest that, although the global transcriptional profiles of hESCs and iPSCs

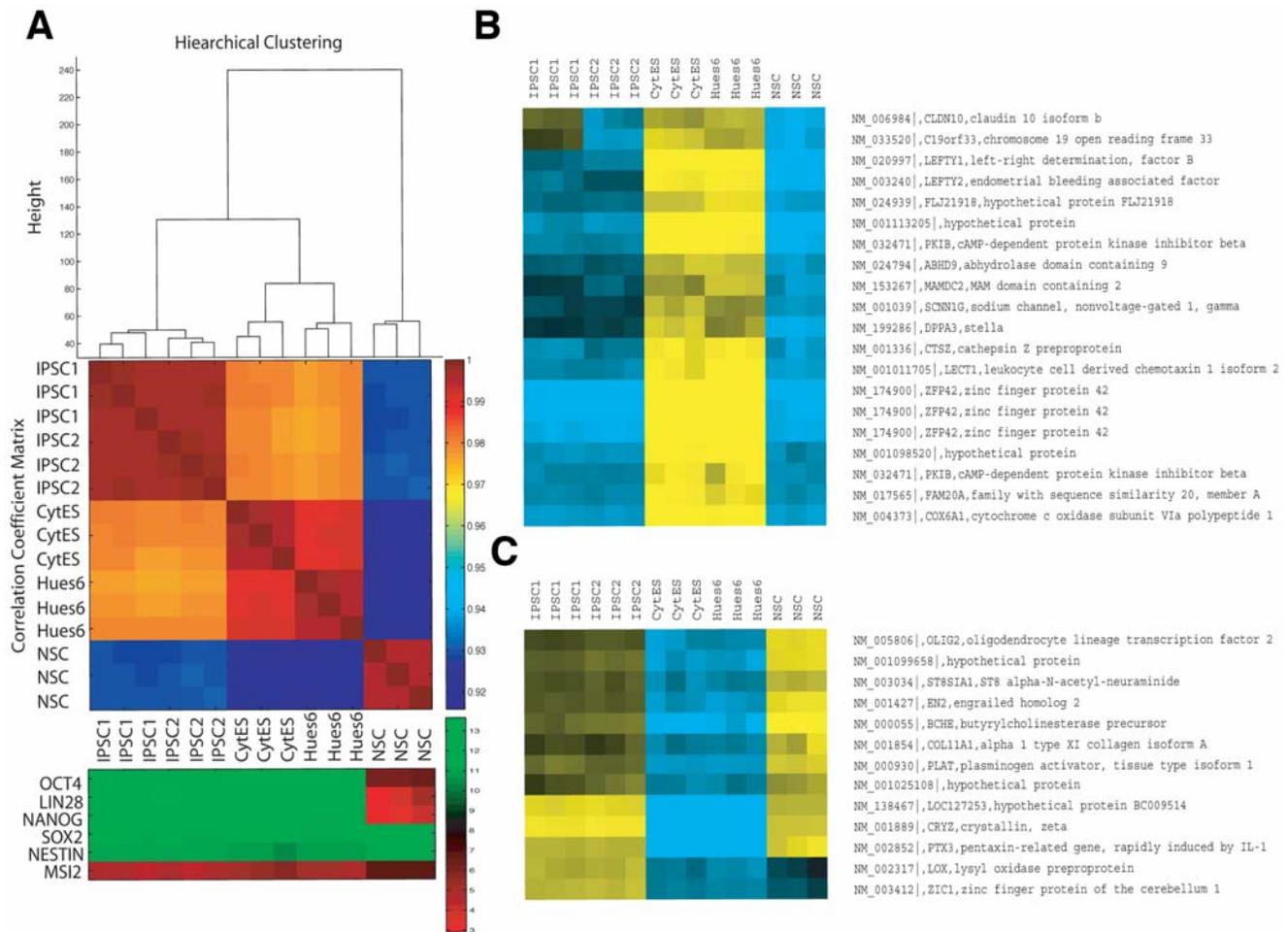


Figure 4. Transcriptional analysis of human integration-free iPSC colonies. **A**, Hierarchical clustering and correlation coefficients of microarray profiles of triplicate iPSC1, iPSC2, CytES (Cyt25 hESC), Hues6 and NSC. Color bar indicates the level of correlation (from 0 to 1). Panel below illustrates marker genes implicated in pluripotency of NSCs, with color bar reporting log₂ normalized expression values (green/red indicates high/low relative expression). **B**, Refseq-annotated genes that were insufficiently induced in iPSCs relative to hESCs (yellow/blue indicates high normalized log₂ expression). **C**, Refseq-annotated genes that were insufficiently silenced in iPSCs relative to hESCs.
doi:10.1371/journal.pone.0007076.g004

were globally similar, small but significant differences indeed exist and cannot be attributed to random viral insertions in the genome. Although we still do not know how relevant these differences are, we anticipate that they may correspond to a unique differentiation potential of iPSCs. Such a potential may also be dependent on the original cell type and may suggest retention of the gene expression memory of the donor cell in iPSCs. We recognize that the use of the oncogene myc to immortalized NSCs might have caused permanent epigenetic changes in the starting cell line that could be carried over into the iPSC stage. The use of only one NSC line further limits the conclusions of this work. Future experiments using iPSCs reprogrammed by the episomal virus-free, transgene-free strategy, in different original primary cell types is needed to validate such hypothesis.

Our results support earlier observations that viral integration is dispensable for genetic reprogramming [15,39]. Our data point to the fact that viral integration does not facilitate iPSC generation, and the efficiency is probably due to the duration and level of the transgenes achieved with episomal plasmids. It has been estimated that each cell contains as many as 50 copies of each episomal plasmid in the nucleus [40]. After a critical amount of time, selection is removed and the episomal vectors are eliminated from

the cells during duplication. Although we never detected episomal plasmids in iPSC established colonies, eventual leftover plasmid will likely be severely methylated when cells reach a pluripotent state, avoiding excess transgene expression after reprogramming [41]. In such a system, the amount and time of gene expression can be easily controlled. We anticipate that different cell types will require a distinct cocktail of pluripotent factors, under specific timing and expression conditions. While this manuscript was in preparation, a similar episomal strategy was used to reprogram human primary fibroblasts using a distinct pluripotent cocktail of factors, validating our methods, but with significant lower efficiency (only 3 to 6 colonies/10⁶ input cells) [42]. Finally, the strategy described here may be a valuable tool for creating safer patient-specific cells and thus could have major implications for future cell therapy.

Materials and Methods

Ethics Statement

All animal work was conducted according to relevant national and international guidelines. Protocols were previously approved by the University of California San Diego Institutional Animal



Figure 5. Refseq-annotated genes that were upregulated in iPSCs relative to both hESCs and NSC. Panel illustrates marker genes implicated in pluripotency of NSCs, with color bar reporting log₂ normalized expression values (green/red indicates high/low relative expression). doi:10.1371/journal.pone.0007076.g005

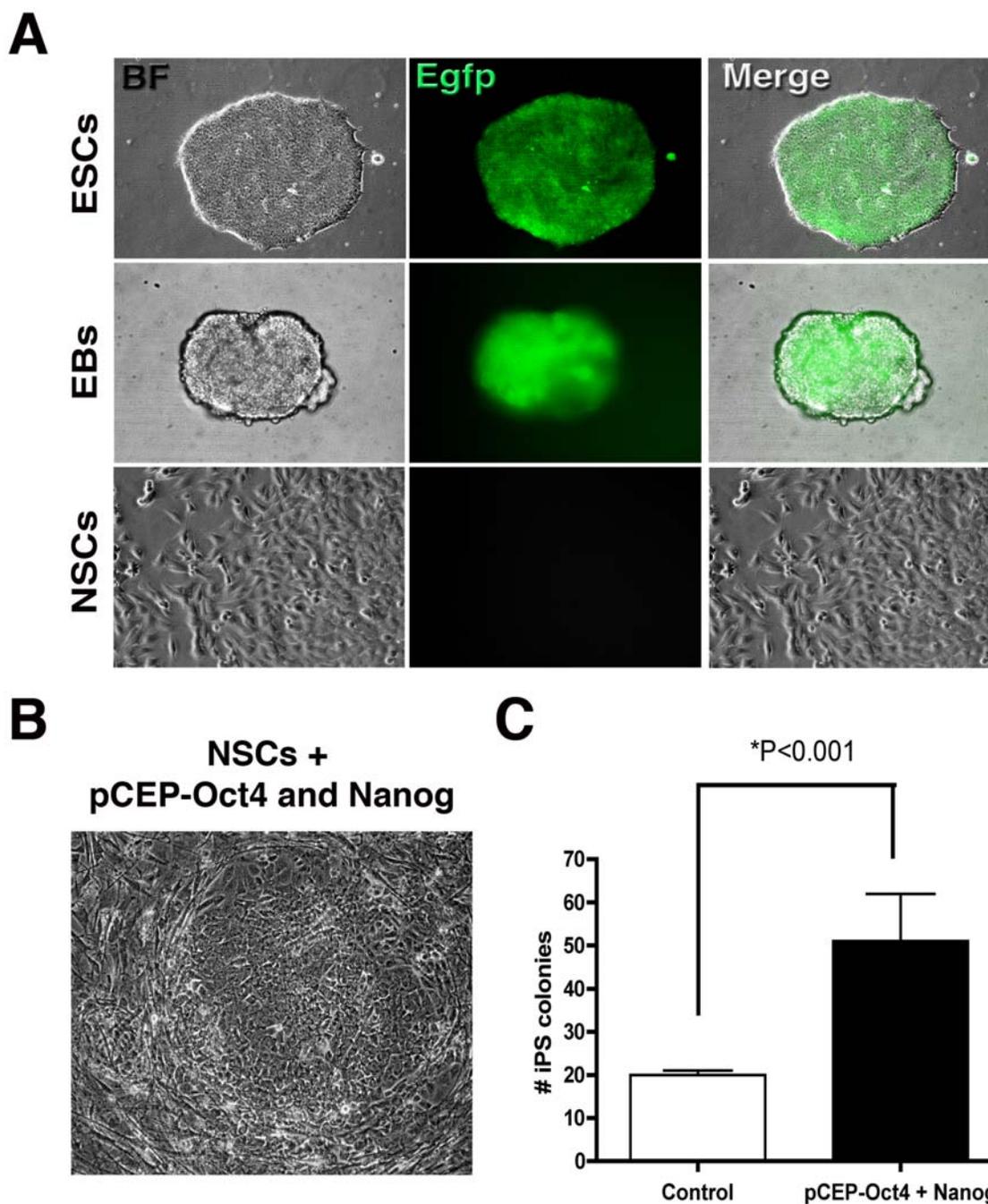


Figure 6. The dynamics of integration-free reprogramming. **A**, Undifferentiated H1 Oct4-EGFP hESC line expresses the EGFP reporter gene that is gradually turned off during NSC differentiation. NSCs are morphologically distinct from hESCs. **B**, Small iPSC colonies can be detected 10 days after transfection with pCEP-Oct4 and pCEP-Nanog. **C**, Typical number of iPSC colonies obtained with electroporation of pCEP-Oct4 and Nanog or with Oct4 alone. Bar = 150 μ m.
doi:10.1371/journal.pone.0007076.g006

Care and Use Committee, the Institutional Review Board and the Embryonic Stem Cell Research Oversight Committee.

Cell culture

Human fetal NSCs (ReNCell VM, Chemicon) were cultured on laminin-coated dishes in ReNcell maintenance medium (Chemicon) in the presence of basic fibroblast growth factor 2 (bFGF2), following the manufacturer's instructions. The hESC Cyt25 (Cythera, San Diego) and HUES6 cell lines were cultured as previously described [25]. Two

days after infection/transfection, cells were plated on mitotically inactivated MEFs (Chemicon), with hESCs medium, in the presence or not of 50 μ g/ml of hygromycin B (Invitrogen). After 2 weeks, iPSC colonies were directly transferred to feeder-free conditions, on matrigel-coated dishes (BD) using mTeSRTM1 (StemCell Technologies). Established iPSC colonies were kept in feeder-free conditions indefinitely and passed using mechanical dissociation. EBs were formed by mechanical dissociation of cell clusters and plating into low-adherence dishes in hESC medium without bFGF2 for 7 days.

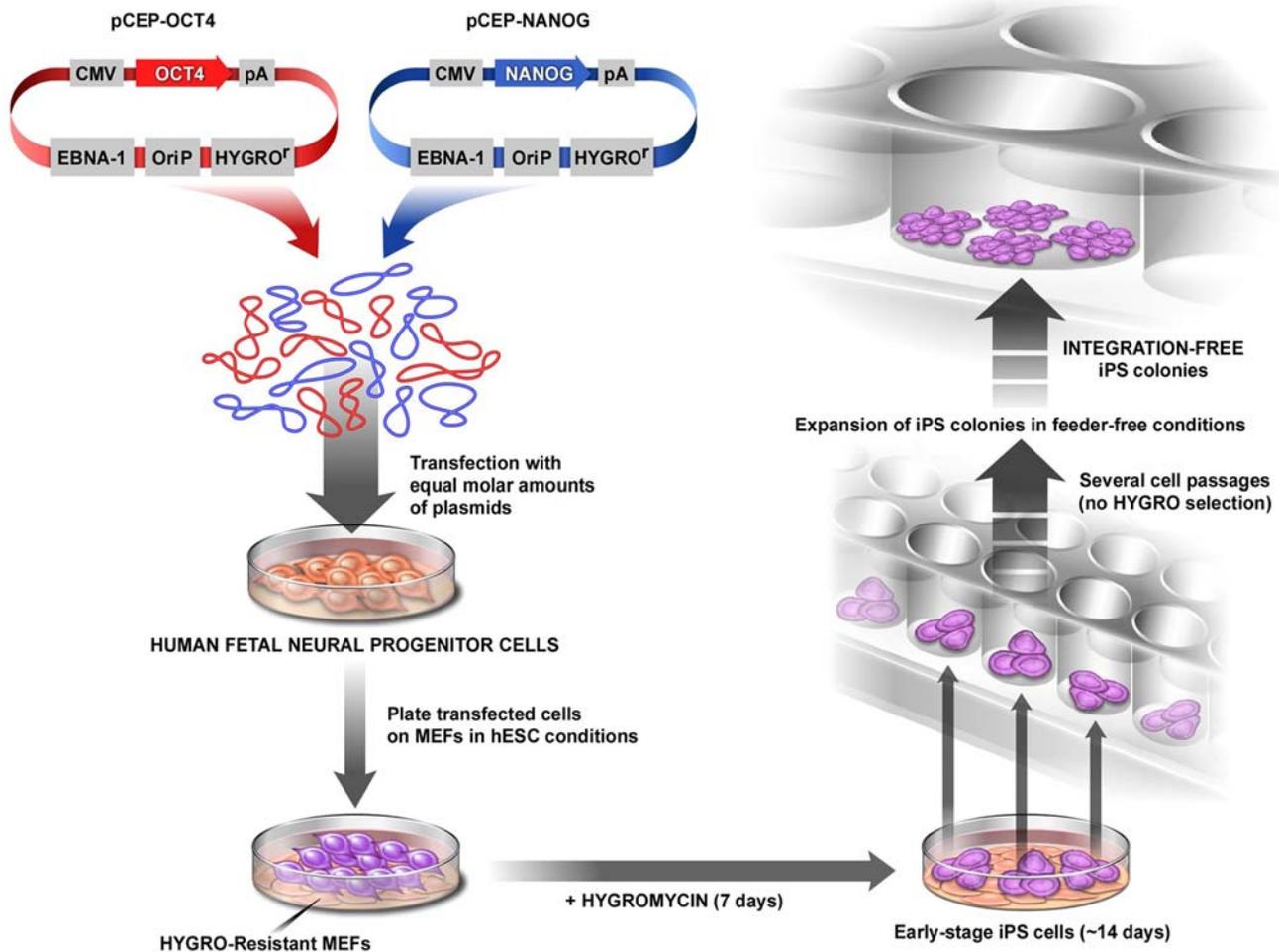


Figure 7. Schematic model of integration-free human iPSC generation from NSCs. Episomal plasmids carrying reprogramming factors are transfected into NSCs and cells are plated on MEFs. On the following day, medium is changed to the hESC condition. Resistant selection is kept for a week. After 14 days, iPSC colonies are visible and can be transferred to a feeder-free condition. Individual colonies are expanded and ready for characterization. At this time, no evidence of plasmid integration is found.
doi:10.1371/journal.pone.0007076.g007

Lentiviral and episomal plasmids

Lentiviral vectors containing the Oct4 and Nanog human cDNAs from Yamanaka's group were obtained from Addgene. The cDNAs were then subcloned into the pCEP4 β episomal plasmid (Invitrogen). Plasmid transfections were done by electroporation of equimolar amounts of pCEP-Oct4 and pCEP-Nanog (5 μ g each) using the nucleofactor for rat NSCs, following the manufacturer's instructions (Lonza/Amaxa Biosystem). Lentiviruses were produced by triple transfection of HEK293T cells followed by ultracentrifugation as previously described elsewhere [25]. Fetal NSCs were infected with both Lenti-Oct4 and Lenti-Nanog at a titer of 0.5×10^{10} gene transfer units/ml overnight, followed by a 2-day recovery period before being plated on mitotically inactive MEFs.

Immunocytochemistry

Cells were briefly fixed in 4% paraformaldehyde and then permeabilized with 0.5% Triton-X in PBS. Cells were blocked in 0.5% Triton-X with 5% donkey serum for 1 hour before incubation with primary antibody overnight at 4°C. After 3 washes in PBS, cells were incubated with secondary antibodies (Jackson ImmunoResearch) for 2 hours at room temperature.

Fluorescent signals were detected using a Zeiss inverted microscope and images were processed with Photoshop CS3 (Adobe Systems). Primary antibodies used in this study are SSEA-4, TRA-1-60, TRA-1-81 (1:100, Chemicon) and Lin28 (1:500 R&D Systems). Alkaline phosphatase activity was detected in live cells using the Vector Red Alkaline Phosphatase substrate kit (Vector Laboratories).

Genomic PCR and Southern blot

Genomic DNA was isolated and prepared using standard molecular techniques. The PCR primers were designed to recognize the pCEP4 episomal vector (Invitrogen). The primer pairs used to amplify the plasmid back bone were: CEP19-F: 5'-tatgatgacacaaacccccgccag -3' and CEP19-R: 5'-aaagcagcagattcttcgccctcc -3'; CEP20-F: 5'-gaaaagcctgaactcaccgc -3' and CEP20-R: 5'-aaagcagcagattcttcgccctcc -3'; CEP21-F: 5'-ggcgaa-gaattctcgtcttc -3' and CEP21-R: 5'-cggtgctgctcatcacagttg -3'; CEP22-F: 5'-cgcaaggaatcggtcaatacactac -3 and CEP22-R: 5'-tccatacaagccaaccacgg -3'; CEP23-F: 5'-ggatttcggctccaacaatgctc -3' and CEP23-R: 5'-tgaacaaacgaccaaacacc -3'. The primers used to amplify the transgene only were: CEP1-F1: 5'-gcgtgga-tagcggttgactc -3'; Oct4R1: 5'-aaatcgaagccaggtgctc -3';

NanogR1: 5'-cagtcggatgcttcaag-3'. Southern blot with 10 µg of genomic DNA, previously digested with *Bam*HI, was performed using standard molecular techniques. The probe used was a fragment of pCEP4 plasmid cut with *Nru*I and *Sal*I enzymes.

RNA extraction and RT-PCR

Total cellular RNA was extracted from $\sim 5 \times 10^6$ cells using the RNeasy Protect Mini kit (Qiagen, Valencia, CA), according to the manufacturer's instructions, and reverse transcribed using the SuperScript III First-Strand Synthesis System RT-PCR from Invitrogen. The cDNA was amplified by PCR using Accuprime Taq DNA polymerase system (Invitrogen). The primer sequences were: hNanog-Fw: 5' cctatgctgtgattgtgg 3' and hNanog-Rv: 5' ctgggacctgtctctctt 3'; hBRACHYURY-F: 5' gcctctctccctccca-cgcacag 3' and hBRACHYURY-R: 5' cggeccgtgtctcacagaccacag-gg 3'; hKRT-18-F: tctgtggagaacgacatcca and KRT-18-R: 5' ct-gtagctcagctctgtga 3'; h-AFP-F: 5' aaaagccactccagcattc 3' and AFP-R: 5' cagacaatccagcattc 3'; GATA-4-F: 5' ctctctcagcagtg-gagagc 3' and GATA-4-R: 5' gagatgcagtgctctgctgc 3'; hGAPDH-Fw: 5' accacagtcctcatccac 3', hGAPDH-Rv: 5' tcaccaccctg-ttctgta 3'. PCR products were separated by electrophoresis on a 2% agarose gel, stained with ethidium bromide and visualized by UV illumination.

Teratoma formation in nude mice

Around $1-3 \times 10^6$ cells were injected subcutaneously into the dorsal flanks of nude mice (CByJ.Cg-Foxn1nu/J) anesthetized with isoflurane. Five to 6 weeks after injection, teratomas were dissected, fixed overnight in 10% buffered formalin phosphate and embedded in paraffin. Sections were stained with haematoxylin and eosin for further analysis.

In vivo spinal iPSCs grafting and identification of teratomas

Adult Sprague-Dawley male rats (320–350 g; n=6) were anesthetized with isoflurane (1.5–2% maintenance; in room air), placed into a spinal unit apparatus (Stoelting, Wood Dale, IL, USA) and a partial Th12–L1 laminectomy performed using a dental drill (exposing the dorsal surface of L2–L5 segments). Using a glass capillary (tip diameter 80–100 µm) connected to a microinjector (Kopf Instruments, Tujunga, CA), rats were injected with 0.5 µl (10,000 cells per injection) of the iPSC (n=3) or proliferating H9 cells in DMEF/F12 media. The duration of each injection was 60 s followed by 30 s pause before capillary withdrawal. The center of the injection was targeted into the dorsal horn (distance from the dorsal surface of the spinal cord at L3 level: 0.5–0.7 mm). Ten injections (500–800 µm rostrocaudally apart) were made on each side of the lumbar spinal cord. After injections, the incision was cleaned with penicillin-streptomycin solution and sutured in two layers. Three or four weeks after cell grafting, rats were deeply anesthetized with pentobarbital and phenytoin and transcardially perfused with 200 ml of heparinized saline followed by 250 ml of 4% paraformaldehyde in PBS. The spinal cords were dissected and postfixed in 4% formaldehyde in PBS overnight at 4°C and then cryoprotected in 30% sucrose PBS until transverse sections (30 µm thick) were cut on a cryostat and mounted on Silane-Prep slides (Sigma). Sections were stained with H&E or immunostained overnight at 4°C with primary human specific (h) or non-specific antibodies made in PBS with 0.2% Triton-X100: mouse anti-nuclear matrix protein/h-nuc (hNUMA; 1:100; Millipore, Temecula, CA, USA); goat anti-doublecortin (DCX; 1:1000; Millipore); mouse anti-Nestin (hNestin; Chemicon). After incubation with primary antibodies, sections were

washed 3× in PBS and incubated with fluorescent-conjugated secondary donkey anti-mouse, or donkey anti-goat antibodies (Alexa 488, 546; 1:250; Invitrogen Corp., Carlsbad, CA, USA) and DAPI for general nuclear staining. Sections were then dried at room temperature, covered with Prolong anti-fade kit (Invitrogen Corp., Carlsbad, CA, USA) and analyzed with confocal microscopy (Olympus, Fluoview 1000).

DNA fingerprinting

DNA fingerprinting analysis was performed by Cell Line Genetics (Madison, WI).

Microarray analysis

The Affymetrix Power Tools (APT) suite of programs and Affymetrix HG-U133 Plus 2.0 library files and annotation were obtained from <http://www.affymetrix.com/support>. Gene-level signal estimates were derived from the CEL files by RMA-sketch normalization as a method in the apt-probeset-summarize program. Hierarchical clustering of the full dataset of 15 (2 hiPSC lines samples, 2 hESC lines, 1 NSC line in triplicate each) by 54,675 probeset values was performed by complete linkage using Euclidean distance as a similarity metric in Matlab. The t-statistic $t_{A,B} = (m_A - m_B) / \sqrt{((n_A - 1)s_A^2 + (n_B - 1)s_B^2)(n_A + n_B) / ((n_A n_B)(n_A + n_B - 2))}$, where n_A and n_B were the number of replicates, m_A and m_B were the mean, and s_A^2 and s_B^2 were the variances of the expression values for the two datasets, was calculated representing the differential enrichment of a gene using gene-level estimates in cell-type(s) A relative to cell-type(s) B. Multiple hypothesis testing was corrected by controlling for the false discovery rate (Benjamini-Hochberg). Four independent (A versus B) comparisons were performed to identify differentially expressed genes: (i) iPSCs versus hESCs; (ii) iPSCs versus NSCs; (iii) hESCs versus NSCs; and (iv) iPSCs and hESCs versus NSCs. A total of 653 probesets were retained at a stringent cutoff of $p < 0.0001$ and fold-change of 4. Probesets were centered by mean expression values, and hierarchical clustering was performed by complete linkage and uncentered correlation as the similarity metric using Cluster 3.0 program. Results were visualized using Java Treeview. Gene ontology analysis was performed as described elsewhere [43].

Supporting Information

Figure S1 Sustained expression using episomal vectors. A, Human fetal NSCs were electroporated with an episomal plasmid carrying the EGFP reporter gene. Transfection efficiency was around 95%. B, Percentage of cells expressing EGFP in the presence or not of hygromycin. Bar = 150 µm. Found at: doi:10.1371/journal.pone.0007076.s001 (1.00 MB TIF)

Figure S2 Integration-free iPSC colonies are genetically identical to the original human fetal NSCs. DNA fingerprinting analysis at 16 independent loci indicated that both iPSCs generated by lentivirus infection (iPSC colony 19) and by transient transfection with episomal vectors (iPSC colony 1) and the original human fetal NSCs (ReNCell VM) shared all alleles investigated and were different from commonly available hESC lines. Found at: doi:10.1371/journal.pone.0007076.s002 (0.81 MB TIF)

Figure S3 Development of teratomas after spinal injections of iPSCs into lumbar gray matter. Lumbar spinal cord sections were stained with H&E at 3 weeks after grafting (A, B). The presence of rosette-like structures (A, yellow arrow) and ectoderm-derived squamous epithelium was identified (B, yellow arrow). Staining with human-specific nestin (green) and DCX (red) antibody show

well organized nestin positive cells in primitive neuronal tube and numerous postmitotic DCX-positive neurons at the periphery of grafts (C, D).

Found at: doi:10.1371/journal.pone.0007076.s003 (9.63 MB TIF)

Figure S4 Myc levels in neural stem cells before and after reprogramming. The myc levels in iPSCs are similar to hESCs.

Found at: doi:10.1371/journal.pone.0007076.s004 (0.27 MB TIF)

Table S1 IPSC-enriched probes in IPSC versus ES. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s005 (5.43 MB DOC)

Table S2 ES-enriched probes in IPSC versus ES. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s006 (2.66 MB DOC)

Table S3 IPSC-enriched probes in IPSC versus NSC. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s007 (8.67 MB DOC)

Table S4 NSC-enriched probes in IPSC versus NSC. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s008 (9.01 MB DOC)

Table S5 ES-enriched probes in ES versus NSC. Probesets enriched in group-wise comparisons: Column headings are

probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s009 (6.06 MB DOC)

Table S6 NSC-enriched probes in ES versus NSC. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s010 (7.45 MB DOC)

Table S7 IPSC, ES-enriched probes in IPSC, ES versus NSC. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s011 (10.18 MB DOC)

Table S8 NSC-enriched probes in IPSC, ES versus NSC. Probesets enriched in group-wise comparisons: Column headings are probeset identifiers, T-statistic, P-value, Fold-Change (log₂), Refseq identifier and Description of the gene. (NA indicates no Refseq annotation).

Found at: doi:10.1371/journal.pone.0007076.s012 (12.08 MB DOC)

Acknowledgments

We would like to thank M. P. Hefferan for his help with confocal microscopy and M.L. Gage for editorial comments.

Author Contributions

Conceived and designed the experiments: MM ARM. Performed the experiments: MM OK MM ARM. Analyzed the data: MM GY ARM. Contributed reagents/materials/analysis tools: GY FHG ARM. Wrote the paper: MM FHG ARM. Financial Support for the project: FHG.

References

- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126: 663–676.
- Takahashi K, Tanabe K, Ohnuki M, Narita M, Ichisaka T, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131: 861–872.
- Yu J, Vodyanik MA, Smuga-Otto K, Antosiewicz-Bourget J, Frane JL, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318: 1917–1920.
- Lowry WE, Richter L, Yachechko R, Pyle AD, Tchieu J, et al. (2008) Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc Natl Acad Sci U S A* 105: 2883–2888.
- Park IH, Lerou PH, Zhao R, Huo H, Daley GQ (2008) Generation of human-induced pluripotent stem cells. *Nat Protoc* 3: 1180–1186.
- Best SM (2008) Viral subversion of apoptotic enzymes: escape from death row. *Annu Rev Microbiol* 62: 171–192.
- Nakagawa M, Koyanagi M, Tanabe K, Takahashi K, Ichisaka T, et al. (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nat Biotechnol* 26: 101–106.
- Soldner F, Hockemeyer D, Beard C, Gao Q, Bell GW, et al. (2009) Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell* 136: 964–977.
- Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S (2008) Generation of mouse induced pluripotent stem cells without viral vectors. *Science* 322: 949–953.
- Kaji K, Norrby K, Paca A, Mileikovsky M, Mohseni P, et al. (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature*.
- Woltjen K, Michael IP, Mohseni P, Desai R, Mileikovsky M, et al. (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature*.
- Aasen T, Raya A, Barrero MJ, Garreta E, Consiglio A, et al. (2008) Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat Biotechnol*.
- Eminli S, Utikal J, Arnold K, Jaenisch R, Hochedlinger K (2008) Reprogramming of neural progenitor cells into induced pluripotent stem cells in the absence of exogenous Sox2 expression. *Stem Cells* 26: 2467–2474.
- Kim JB, Zaehres H, Wu G, Gentile L, Ko K, et al. (2008) Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* 454: 646–650.
- Aoi T, Yae K, Nakagawa M, Ichisaka T, Okita K, et al. (2008) Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* 321: 699–702.
- Hanna J, Markoulaki S, Schorderet P, Carey BW, Beard C, et al. (2008) Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* 133: 250–264.
- Kim JB, Sebastiano V, Wu G, Arauzo-Bravo MJ, Sasse P, et al. (2009) Oct4-induced pluripotency in adult neural stem cells. *Cell* 136: 411–419.
- Silva J, Barrandon O, Nichols J, Kawaguchi J, Theunissen TW, et al. (2008) Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition. *PLoS Biol* 6: e253.
- Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K (2008) Induced pluripotent stem cells generated without viral integration. *Science* 322: 945–949.
- Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* 448: 313–317.
- Wernig M, Meissner A, Cassady JP, Jaenisch R (2008) c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* 2: 10–12.

22. Sridharan R, Tchiew J, Mason MJ, Yachechko R, Kuoy E, et al. (2009) Role of the murine reprogramming factors in the induction of pluripotency. *Cell* 136: 364–377.
23. Breier JM, Radio NM, Mundy WR, Shafer TJ (2008) Development of a high-throughput screening assay for chemical effects on proliferation and viability of immortalized human neural progenitor cells. *Toxicol Sci* 105: 119–133.
24. Donato R, Miljan EA, Hines SJ, Aouabdi S, Pollock K, et al. (2007) Differential development of neuronal physiological responsiveness in two human neural stem cell lines. *BMC Neurosci* 8: 36.
25. Muotri AR, Nakashima K, Toni N, Sandler VM, Gage FH (2005) Development of functional human embryonic stem cell-derived neurons in mouse brain. *Proc Natl Acad Sci U S A* 102: 18644–18648.
26. Margolske RF (1992) Epstein-Barr virus based expression vectors. *Curr Top Microbiol Immunol* 158: 67–95.
27. Van Craenenbroeck K, Vanhoenacker P, Duchau H, Haegeman G (2000) Molecular integrity and usefulness of episomal expression vectors derived from BK and Epstein-Barr virus. *Gene* 253: 293–301.
28. Van Craenenbroeck K, Vanhoenacker P, Haegeman G (2000) Episomal vectors for gene expression in mammalian cells. *Eur J Biochem* 267: 5665–5678.
29. Leight ER, Sugden B (2000) EBNA-1: a protein pivotal to latent infection by Epstein-Barr virus. *Rev Med Virol* 10: 83–100.
30. Leight ER, Sugden B (2001) Establishment of an oriP replicon is dependent upon an infrequent, epigenetic event. *Mol Cell Biol* 21: 4149–4161.
31. Brambrink T, Foreman R, Welstead GG, Lengner CJ, Wernig M, et al. (2008) Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell* 2: 151–159.
32. Stadtfeld M, Maherali N, Breault DT, Hochedlinger K (2008) Defining molecular cornerstones during fibroblast to iPS cell reprogramming in mouse. *Cell Stem Cell* 2: 230–240.
33. Masui S, Ohtsuka S, Yagi R, Takahashi K, Ko MS, et al. (2008) Rex1/Zfp42 is dispensable for pluripotency in mouse ES cells. *BMC Dev Biol* 8: 45.
34. Nakatake Y, Fukui N, Iwamatsu Y, Masui S, Takahashi K, et al. (2006) Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Mol Cell Biol* 26: 7772–7782.
35. Zwaka TP, Thomson JA (2003) Homologous recombination in human embryonic stem cells. *Nat Biotechnol* 21: 319–321.
36. Cezar GG, Quam JA, Smith AM, Rosa GJ, Piekarczyk MS, et al. (2007) Identification of small molecules from human embryonic stem cells using metabolomics. *Stem Cells Dev* 16: 869–882.
37. Yeo GW, Xu X, Liang TY, Muotri AR, Carson CT, et al. (2007) Alternative splicing events identified in human embryonic stem cells and neural progenitors. *PLoS Comput Biol* 3: 1951–1967.
38. Chambers I, Silva J, Colby D, Nichols J, Nijmeijer B, et al. (2007) Nanog safeguards pluripotency and mediates germline development. *Nature* 450: 1230–1234.
39. Varas F, Stadtfeld M, De Andres-Aguayo L, Maherali N, di Tullio A, et al. (2008) Fibroblast derived induced pluripotent stem cells show no common retroviral vector insertions. *Stem Cells*.
40. Belt PB, Groeneveld H, Teubel WJ, van de Putte P, Backendorf C (1989) Construction and properties of an Epstein-Barr-virus-derived cDNA expression vector for human cells. *Gene* 84: 407–417.
41. Kameda T, Smuga-Otto K, Thomson JA (2006) A severe de novo methylation of episomal vectors by human ES cells. *Biochem Biophys Res Commun* 349: 1269–1277.
42. Yu J, Hu K, Smuga-Otto K, Tian S, Stewart R, et al. (2009) Human Induced Pluripotent Stem Cells Free of Vector and Transgene Sequences. *Science*.
43. Yeo GW, Van Nostrand E, Holste D, Poggio T, Burge CB (2005) Identification and analysis of alternative splicing events conserved in human and mouse. *Proc Natl Acad Sci U S A* 102: 2850–2855.