



Published in final edited form as:

Science. 2009 May 8; 324(5928): 797–801. doi:10.1126/science.1172482.

Human Induced Pluripotent Stem Cells Free of Vector and Transgene Sequences

Junying Yu^{1,2,3,*}, Kejin Hu³, Kim Smuga-Otto^{1,2,3}, Shulan Tian^{1,2}, Ron Stewart^{1,2}, Igor I. Slukvin^{3,4}, and James A. Thomson^{1,2,3,5,*}

¹ Morgridge Institute for Research, Madison, WI, 53707–7365, USA

² Genome Center of Wisconsin, Madison, WI 53706–1580, USA

³ Wisconsin National Primate Research Center, University of Wisconsin-Madison, Madison, WI 53715–1299, USA

⁴ Department of Pathology and Laboratory Medicine, University of Wisconsin-Madison, Madison, WI 53706, USA

⁵ Department of Anatomy, University of Wisconsin-Madison, Madison, WI 53706–1509, USA

Abstract

Reprogramming differentiated human cells to induced pluripotent stem (iPS) cells has applications in basic biology, drug development, and transplantation. Human iPS cell derivation previously required vectors that integrate into the genome, which can create mutations and limit the utility of the cells in both research and clinical applications. Here we describe the derivation of human iPS cells using non-integrating episomal vectors. After removal of the episome, iPS cells completely free of vector and transgene sequences are derived that are similar to human embryonic stem (ES) cells in proliferative and developmental potential. These results demonstrate that reprogramming human somatic cells does not require genomic integration or the continued presence of exogenous reprogramming factors, and removes one obstacle to the clinical application of human iPS cells.

The proliferative and developmental potential of both human ES cells and human iPS cells offer unprecedented access to the differentiated cells that make up the human body (1–3). In addition, iPS cells can be derived with a specific desired genetic background, including patient-specific iPS cells for disease models and for transplantation therapies, without the problems associated with immune rejection. Reprogramming of both mouse and human somatic cells into iPS cells has been achieved by expressing combinations of factors such as *OCT4*, *SOX2*, *c-Myc*, *KLF4*, *NANOG*, and *LIN28* (2–4). Initial methods used to derive human iPS cells employed viral vectors, where both the vector backbone and transgenes are permanently integrated into the genome (2,3). Such vectors can produce insertional mutations that interfere with the normal function of iPS cell derivatives, and residual transgene expression can influence differentiation into specific lineages (2), or even result in tumorigenesis (5). Vector integration-free mouse iPS cells have been derived from liver cells with adenoviral vectors (6), and from embryonic fibroblasts with repeated plasmid transfections (7), but the low frequencies obtained make it unclear how practical these approaches will be for human cells, which generally require longer exposure to reprogramming factors (2,3).

*To whom correspondence should be addressed. E-mail: jyyu2008@gmail.com (J.Y.); thomson@primate.wisc.edu (J.A.T.).

While this manuscript was in review, two alternative approaches were described to remove transgenes from mouse or human iPS cells. In one approach, Cre/LoxP recombination was used to excise integrated transgenes (8,9). This approach successfully removes transgene sequences, but leaves behind residual vector sequences, which can still create insertional mutations. A second approach used seamless excision of piggyBac transposons to produce vector and transgene-free mouse iPS cells (10). Although a promising approach, vector removal from human iPS cells produced by this method has not yet been reported, and removing multiple transposons is labor intensive. Here, we report that human iPS cells completely free of vector and transgene sequences can be derived from fibroblasts by a single transfection with oriP/EBNA1 (Epstein-Barr nuclear antigen-1)-based episomal vectors.

Derived from the Epstein-Barr virus, oriP/EBNA1 vectors are well suited for introducing reprogramming factors into human somatic cells, as these plasmids can be transfected without the need for viral packaging, and can be subsequently removed from cells by culturing in the absence of drug selection. The stable extrachromosomal replication of oriP/EBNA1 vectors in mammalian cells requires only a cis-acting oriP element (11) and a trans-acting EBNA1 gene (12). The oriP/EBNA1 vectors replicate only once per cell cycle, and with drug selection can be established as stable episomes in about 1% of the initial transfected cells (13,14). If drug selection is subsequently removed, the episomes are lost at ~5% per cell generation due to defects in plasmid synthesis and partitioning; thus, cells devoid of plasmids can be easily isolated (15).

OCT4, *SOX2*, *NANOG*, and *LIN28* are sufficient to reprogram human embryonic, neonatal, and adult fibroblasts to iPS cells (2,16), but the reprogramming efficiency is low (<0.01% for newborn foreskin fibroblasts) (2). Such low efficiency makes it difficult to reprogram with oriP/EBNA1-based vectors since the stable transfection efficiency is almost two orders of magnitude less than that of our lentiviral vectors (2). Thus, we first improved reprogramming efficiency with lentiviral vectors. By testing different linkers to coexpress *OCT4* and *SOX2*, we found that the internal ribosome entry site 2 (IRES2) supported higher reprogramming efficiency (fig. S1A and B). Since linkers have less effect on reprogramming efficiency when used to coexpress *NANOG* and *LIN28* (fig. S1B), IRES2 was chosen to coexpress reprogramming factors. Using IRES2-mediated expression of *OCT4*, *SOX2*, *NANOG*, and *LIN28*, the reprogramming efficiency for human foreskin fibroblasts improved about tenfold (~0.1%) over what we previously reported (fig. S1C). The addition of c-Myc and *KLF4* further improved the reprogramming efficiency to over 1%, the highest efficiencies we have achieved for these cells (fig. S1C) (17). Thus, we cloned all six reprogramming factors (*OCT4*, *SOX2*, *NANOG*, *LIN28*, *c-Myc*, and *KLF4*) into an oriP/EBNA1 vector using IRES2 for coexpression. Since our previous experience suggested that both the balance between transgenes and their absolute expression levels are critically important to achieving reprogramming, we tested different transgene arrangements to achieve appropriate levels empirically (table S1).

Initial tests with the six reprogramming genes in the episomal vectors failed to yield human iPS cell colonies (table S2). With this combination of genes, substantial cell death was observed during the first week following transfection, possibly due to the toxic effects of high level c-Myc expression (18). In an attempt to counteract the possible toxic effects of c-Myc expression, we included the SV40 large T gene (*SV40LT*) in some of the combinations (19). Three of these combinations, all of which included *OCT4*, *SOX2*, *NANOG*, *LIN28*, *c-Myc*, *KLF4*, and *SV40LT*, were successful in producing iPS cell colonies from human foreskin fibroblasts using oriP/EBNA1-based vectors (Fig. 1A, fig. S2D and table S2). At least two plasmids in each successful combination express *OCT4* and *SOX2*, consistent with the observation that high level expression of these transgenes improves reprogramming. Clones from two of these combinations (19 from experiment 3 and 6 from experiment 4, Table S2) were chosen for expansion and analysis. These iPS cell colonies exhibited typical human ES cell morphology

(e.g., compact colonies, high nucleus-to-cytoplasm ratios, and prominent nucleoli) (Fig. 1B), and exhibited gene expression profiles that were very similar to human ES cell lines, but dissimilar to the parental fibroblasts (Fig. 1C and table S3). Similar to human ES cells, when injected into immunocompromised mice, these iPS cells formed teratomas consisting of differentiated derivatives of all three primary germ layers (Fig. 1D). PCR analysis failed to detect episomal vector integration in the genome, but did detect their persistence in the episomal fraction (Fig. 1E). The persistence of the episomal vectors suggests a requirement for prolonged transgene expression for successful reprogramming.

Since oriP/EBNA1 episomal vectors are gradually lost from proliferating cells in the absence of selection, we performed subcloning to derive iPS cell clones that had spontaneously lost the episomal vectors. We chose one iPS cell clone derived with combination 6 (iPS-DF6-9) and another one with combination 19 (iPS-DF19-9), and isolated 12 subclones from each. Over one third of the subclones lost their episomal vectors (fig. S3A). We expanded two subclones from each vector combination for detailed analysis (iPS-DF6-9 subclone 9T and 12T, iPS-DF19-9 subclone 7T and 11T). RT-PCR analysis with transgene-specific primers failed to detect any residual transgene expression in any of the four iPS cell subclones (Fig. 2A). In contrast to the parental iPS cell clones, PCR analysis demonstrated the absence of the vector and transgene sequences in both the genomic and the episomal fractions of all four iPS cell subclones (Fig. 2B), which was confirmed by Southern blot analysis (Fig. 2C and fig. S3B).

The iPS cell subclones were morphologically similar to human ES cells (Fig. 3A); had normal karyotypes (Fig. 3B); expressed human ES cell-specific cell surface markers (Fig. 3D) and genes (Fig. 4A, 4B, fig. S4, and table S4); and differentiated into derivatives of all three germ layers in teratomas (Fig. 4C). Both the *OCT4* promoter and the *NANOG* promoter were demethylated in these iPS cells, similar to human ES cells and in contrast to the parental foreskin fibroblasts (Fig. 3C). As of this writing, combination 19 iPS cells have been in continuous culture for 7 months after the initial fibroblast transfection and have demonstrated no period of replicative crisis. DNA fingerprinting confirmed their origin from foreskin fibroblasts (table S5).

With oriP/EBNA1-based episomal vectors, exogenous DNA is not integrated into the human iPS cell genome, and due to the gradual loss of cellular episomal vectors in the absence of drug selection, vector and transgene-free human iPS cells can be isolated through subcloning without further genetic manipulation. Similar to mouse studies using non-integrating reprogramming methods (6,7), the current reprogramming efficiency of human fibroblasts with oriP/EBNA1 vectors is low (~3 to 6 colonies/ 10^6 input cells). These frequencies are, however, sufficient to recover iPS cells from a reasonable number of starting cells, and fibroblasts are easy to obtain and culture. Since different cell types have different reprogramming frequencies (20), and oriP/EBNA1-based vectors are established as stable episomes at different frequencies in different cell types (13), it might be possible to identify another accessible human cell type more easily reprogrammed with these episomal vectors. The addition of chemical compounds that increase reprogramming efficiency might also facilitate reprogramming by these episomal vectors (21,22). Given the rapid pace of the iPS cell field, it is likely that reprogramming efficiencies will improve significantly, and that it soon will be possible to derive vector and transgene-free human iPS cells by several alternative methods. However, it will be essential to determine which of these methods most consistently produces iPS cells with the fewest genetic or epigenetic abnormalities, because any abnormalities would impact the application of these cells in basic research, drug development, and transplantation therapies much more than the initial reprogramming frequencies. Substantial challenges also remain in cell-specific differentiation and delivery, but the derivation of vector and transgene-free human iPS cells is nonetheless an important advance towards the clinical application of these cells.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

26. Funding was provided by the Charlotte Geyer Foundation and NIH (grants P51 RR000167 and P01 GM081629). We thank A. Griep, P. Powers, B. Sugden, M. Probasco, A. Elwell, C. Glennon, V. Ruotti, and D. J. Faupel for assistance and advice. The authors declare competing financial interests: J.A.T. is a founder, stock owner, consultant, and board member of Cellular Dynamics International (CDI). He also serves as a scientific adviser to and has financial interests in Tactics II Stem Cell Ventures; I.I.S. is a founder, stock owner, and consultant for CDI. A patent on this work is pending. Combination 6 and 19 episomal vectors are deposited in Addgene (Cambridge, MA), and vector-free human iPS cell subclones are deposited in the WiCell International Stem Cell (WISC) Bank (Madison, WI). Microarray data is deposited in the Gene Expression Omnibus (GEO) database (accession number GSE15148).

References and Notes

1. Thomson JA, et al. *Science* 1998;282:1145. [PubMed: 9804556]
2. Yu J, et al. *Science* 2007;318:1917. [PubMed: 18029452]
3. Takahashi K, et al. *Cell* 2007;131:861. [PubMed: 18035408]
4. Takahashi K, Yamanaka S. *Cell* 2006;126:663. [PubMed: 16904174]
5. Okita K, Ichisaka T, Yamanaka S. *Nature* 2007;448:313. [PubMed: 17554338]
6. Stadtfeld M, Nagaya M, Utikal J, Weir G, Hochedlinger K. *Science* 2008;322:945. [PubMed: 18818365]
7. Okita K, Nakagawa M, Hyenjong H, Ichisaka T, Yamanaka S. *Science* 2008;322:949. [PubMed: 18845712]
8. Kaji K, et al. *Nature*. 2009
9. Soldner F, et al. *Cell* 2009;136:964. [PubMed: 19269371]
10. Woltjen K, et al. *Nature*. 2009
11. Yates J, Warren N, Reisman D, Sugden B. *Proc Natl Acad Sci U S A* 1984;81:3806. [PubMed: 6328526]
12. Yates JL, Warren N, Sugden B. *Nature* 1985;313:812. [PubMed: 2983224]
13. Leight ER, Sugden B. *Mol Cell Biol* 2001;21:4149. [PubMed: 11390644]
14. Yates JL, Guan N. *J Virol* 1991;65:483. [PubMed: 1845903]
15. Nanbo A, Sugden A, Sugden B. *Embo J* 2007;26:4252. [PubMed: 17853891]
16. Ebert AD, et al. *Nature* 2009;457:277. [PubMed: 19098894]
17. Liao J, et al. *Cell Res* 2008;18:600. [PubMed: 18414447]
18. Evan GI, et al. *Cell* 1992;69:119. [PubMed: 1555236]
19. Hahn WC, et al. *Nature* 1999;400:464. [PubMed: 10440377]
20. Aasen T, et al. *Nat Biotechnol* 2008;26:1276. [PubMed: 18931654]
21. Shi Y, et al. *Cell Stem Cell* 2008;3:568. [PubMed: 18983970]
22. Huangfu D, et al. *Nat Biotechnol* 2008;26:795. [PubMed: 18568017]

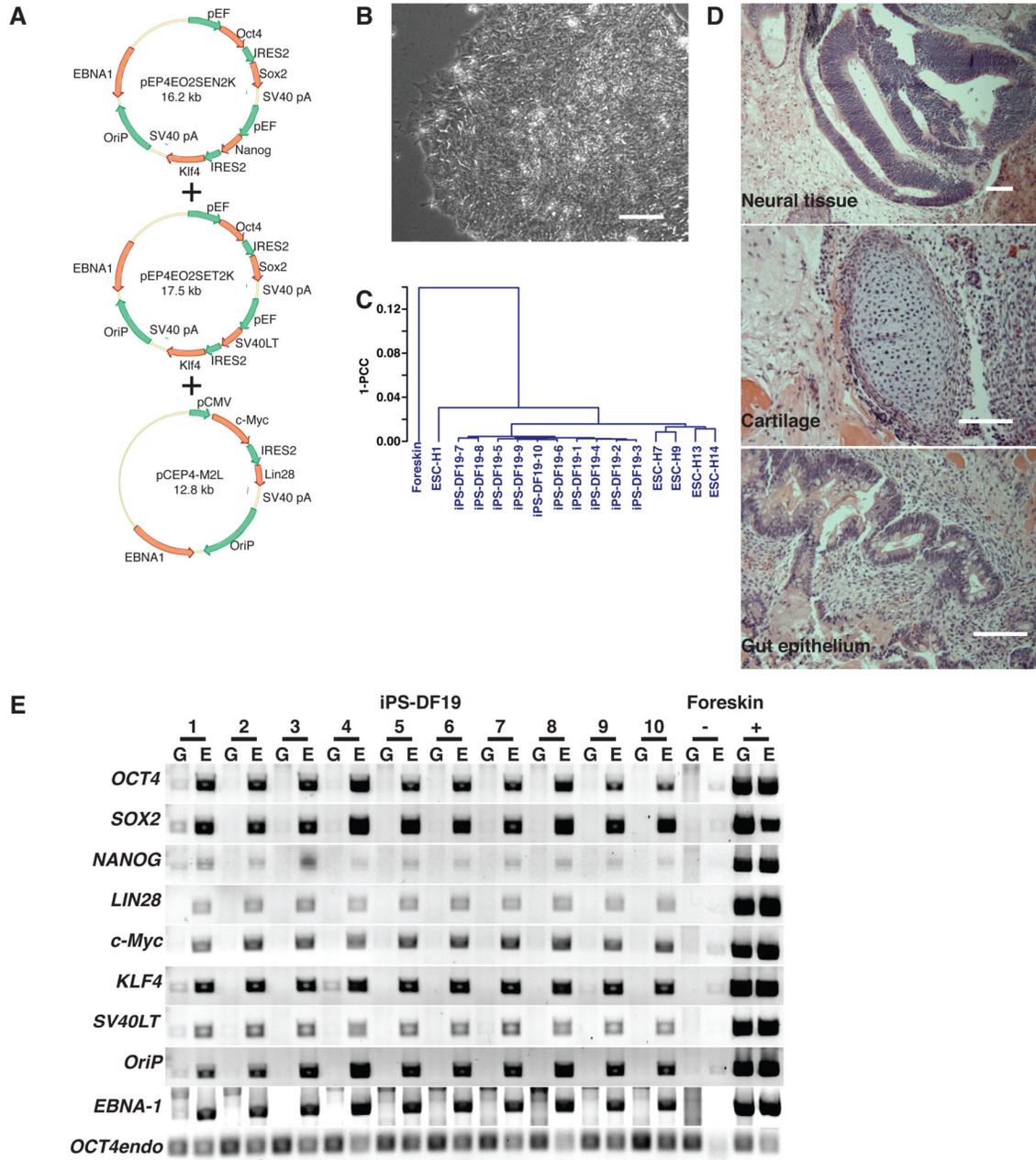


Fig. 1. Reprogramming human foreskin fibroblasts without genomic vector integration. **(A)** Episomal vectors (combination 19 from experiment 3, Table S2). pEF: the eukaryotic elongation factor 1 α promoter; pCMV: the cytomegalovirus immediate-early promoter. Transgenes and other features of vectors are indicated by red and green arrows respectively. **(B)** Bright-field image of iPS cells obtained by transfection of combination 19 episomal vectors (clone DF19-9: “Defined Factor” combination 19, and clone 9). Scale bars, 0.1 mm. **(C)** Pearson correlation analyses of global gene expression (51,337 transcripts) in human fibroblast-derived iPS cell clones (combination 19). 1-PCC: Pearson Correlation Coefficient. **(D)** Hematoxylin and eosin staining of teratoma sections of iPS cell clone DF19-9 (53 days after injection). Teratomas were obtained from all ten iPS-DF19 clones. Scale bars, 0.1 mm. **(E)** PCR analysis of episomal

DNA in iPS-DF19 clone 1 to 10. G: genomic DNA template; E: episomal DNA template. Genomic and episomal DNA from nontransfected and combination 19 episomal vector-transfected (day 17 posttransfection) fibroblasts were used as negative (-) and positive (+) controls respectively. 32 PCR cycles were used for all primer sets.

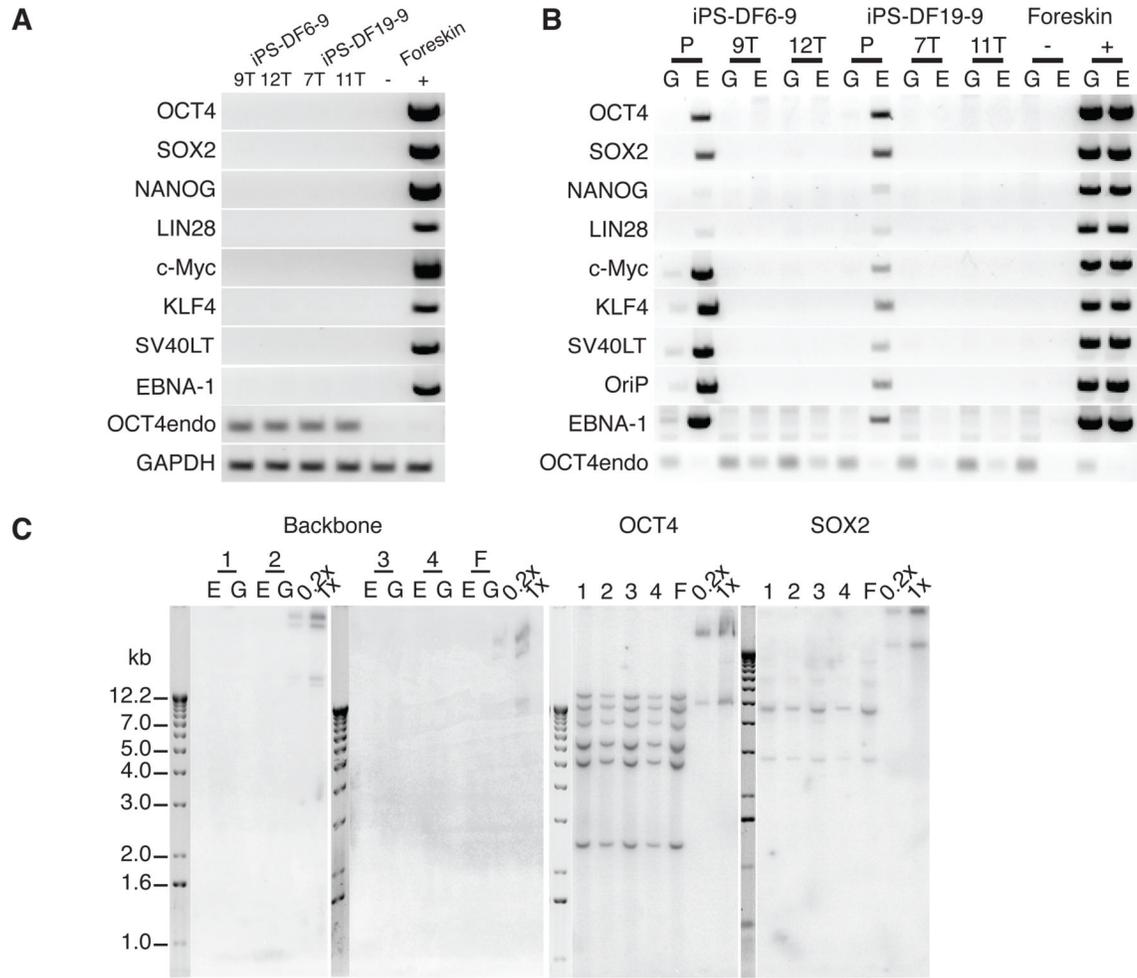


Fig. 2. Human foreskin fibroblast-derived iPS cells free of vectors and transgenes. **(A)** RT-PCR analysis of transgene expression in iPS-DF6-9 subclone 9T and 12T, and iPS-DF-19-9 subclone 7T and 11T. Negative control (-): fibroblasts. Positive control (+): fibroblasts transfected with combination 19 episomal vectors (day 4 posttransfection). 32 PCR cycles were used for all primer sets. **(B)** PCR analysis of episomal DNA in iPS-DF6-9 (P: parental clone), iPS-DF6-9 subclone 9T and 12T, iPS-DF19-9 (P), and iPS-DF19-9 subclones 7T and 11T. G: genomic DNA template; E: episomal DNA template. Negative (-) and positive (+) controls were the same as in Fig. 1E. 32 PCR cycles were used for all primer sets except *OCT4*endo (28 cycles). **(C)** Southern blot analysis of exogenous DNA in iPS-DF6-9 and iPS-DF-19-9 subclones. The pCEP4 vector was used as a probe to detect the presence of vector backbone, and the open reading frames of *OCT4* and *SOX2* were used as probes to examine both the endogenous gene and possible transgenes. 1: iPS-DF6-9-9T; 2: iPS-DF6-9-12T; 3: iPS-DF19-9-7T; 4: iPS-DF19-9-11T; F: foreskin fibroblasts. E: undigested episomal DNA; G: digested genomic DNA. Combination 19 episomal vector DNA diluted to the equivalents of 0.2 and 1 integration per genome was used as positive controls (0.2x and 1x).

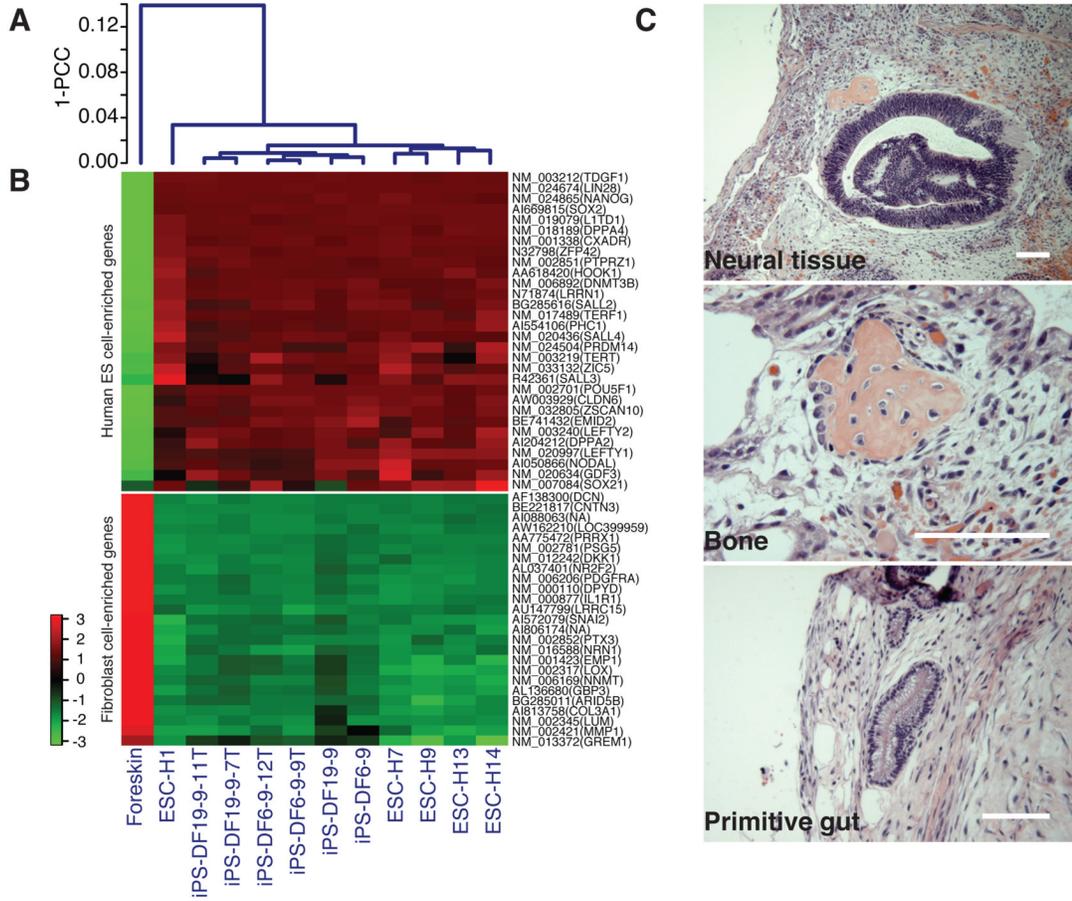


Fig. 4. Characterization of iPS cell subclones. **(A)** Pearson correlation analyses of global gene expression (51,337 transcripts) in iPS cell parental clone DF6-9 and DF19-9; iPS cell subclone DF6-9-9T, DF6-9-12T, DF19-9-7T, and DF19-9-11T; five human ES cell lines; foreskin fibroblasts. 1-PCC: Pearson Correlation Coefficient. **(B)** Expression of genes that are differentially expressed between human ES cells and foreskin fibroblasts. Top panel: 30 well-known human ES cell-enriched genes; bottom panel: top 25 foreskin fibroblast-enriched genes. The color key is shown on the left. **(C)** Hematoxylin and eosin staining of teratoma sections of iPS-DF19-9-11T (7 weeks after injection). Teratomas were obtained from all four iPS cell subclones. Scale bars, 0.1 mm.