A more efficient method to generate integration-free human iPS cells

Keisuke Okita[1], Yasuko Matsumura[1], Yoshiko Sato[1], Aki Okada[1], Asuka Morizane[1,2], Satoshi Okamoto[3], Hyenjong Hong[1], Masato Nakagawa[1], Koji Tanabe[1], Ken-ichi Tezuka[4], Toshiyuki Shibata[5], Takahiro Kunisada[4], Masayo Takahashi[1,3], Jun Takahashi[1,2], Hiroh Saji[6] & Shinya Yamanaka[1,7–9]

We report a simple method, using p53 suppression and nontransforming L-Myc, to generate human induced pluripotent stem cells (iPSCs) with episomal plasmid vectors. We generated human iPSCs from multiple donors, including two putative human leukocyte antigen (HLA)-homozygous donors who match ~20% of the Japanese population at major HLA loci; most iPSCs are integrated transgene-free. This method may provide iPSCs suitable for autologous and allologous stem-cell therapy in the future.

Genomic integration of transgenes increases the risk of tumor formation and mortality in chimeric and progeny mice derived from induced pluripotent stem cells (iPSCs)[1]. Integration-free human iPSCs have been generated using several methods, including adenovirus[2], Sendai virus[3], the piggyBac system[4], minicircle vector[5], episomal vectors[6], direct protein delivery[7] and synthesized mRNA[8] (**Supplementary Table 1**). However, reprogramming efficiency using integration-free methods is impractically low in most cases. Direct delivery of proteins or RNA is labor-intensive, requiring repeated delivery of the reprogramming factors. Modifying Sendai virus vectors or preparing synthesized RNA are technically demanding.

In the original report describing episomal plasmid vectors for reprogramming, the authors used seven factors, including *POU5F1* (also known as *OCT3/4*), *SOX2*, *KLF4*, *MYC* (also known as *c-MYC*), *NANOG*, *LIN28A* (also known as *LIN28*) and SV40 large T antigen (*SV40LT*), in three different vector combinations[6] (T1–T3 combinations; **Fig. 1a** and **Supplementary Table 2**). In this study, we used two findings from our laboratory to enhance efficiency of reprogramming by episomal plasmids: iPSC generation is markedly enhanced by p53 suppression[9], and L-Myc is more potent and specific than c-Myc during human iPSC generation[10].

We prepared four vector combinations (**Fig. 1a** and **Supplementary Table 2**). The Y1 combination had six factors (*OCT3/4*, *SOX2*, *KLF4*, *c-MYC*, *LIN28* and *NANOG*) in three episomal plasmids. The Y2 combination contained an additional *TP53* (also known as *p53*) shRNA in one of the three plasmids. We replaced *c-MYC* and *NANOG* with *MYCL1* (also known as *L-MYC*) in the Y1 and Y2 combinations, respectively, to yield the Y3 and Y4 combinations (**Fig. 1b**).

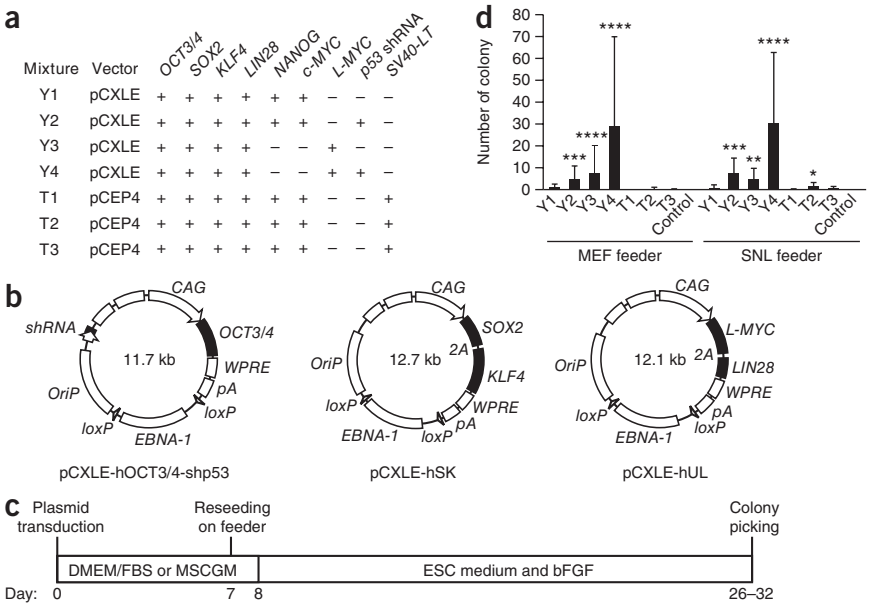
We electroporated these seven combinations of episomal vectors (Y1–Y4 or T1–T3) into three human dermal fibroblast (HDF) lines and two dental pulp cell lines on day 0 (**Fig. 1c**). We trypsinized transfected cells on day 7 and reseeded them onto feeder layers. We maintained the cells in embryonic stem cell (ESC) medium, and small cell colonies became visible ~2 weeks after transfection. We counted the number of colonies with a flat human ESC–like morphology and non-ESC-like colonies around day 30 (**Supplementary Fig. 1**). The Y4 combination resulted in significantly ($P < 0.05$) more iPSC colonies than did any of T1–T3 combinations (**Fig. 1d**). In addition to these five parental cell lines, we obtained iPSC colonies from seven additional HDF lines with the Y4 combination of factors (**Supplementary Table 3**).

We expanded ESC-like colonies derived with the Y4 combination for additional experiments. The majority of the colonies were expandable and exhibited a cellular morphology similar to that of human ESCs, characterized by large nuclei and scant cytoplasm (**Fig. 2a**,**b**). We termed these episomal plasmid vector-derived iPSCs 'pla-iPSCs'. Ten of eleven clones we analyzed were karyotypically normal (**Supplementary Fig. 2** and **Supplementary Table 4**). Short tandem repeat analyses confirmed that pla-iPSC clones were derived from HDFs and dental pulp cells (**Supplementary Table 5**). Reverse transcription–PCR (RT-PCR) analyses revealed that pla-iPSC clones expressed pluripotent stem cell markers, such as *OCT3/4*, *SOX2*, *NANOG* and *DPPA5*, at levels comparable to those in ESCs and retrovirus-derived iPSC clones (**Fig. 2c** and **Supplementary Figs. 3a**, **4** and **5**). Global gene expression profiles also showed that pla-iPSC clones were similar to ESC and retro-iPSC clones (**Supplementary Fig. 6** and **Supplementary Table 6**). The DNA methylation levels of CpG sites in the promoter region of *NANOG* were high in parental HDFs and dental pulp cells but were low in pla-iPSCs and ESCs (**Fig. 2d**).

To examine whether episomal vectors persisted in pla-iPSCs, first we transfected an episomal vector encoding enhanced GFP

[1]Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan. [2]Department of Biological Repair, Institute for Frontier Medical Sciences, Kyoto University, Kyoto, Japan. [3]Laboratory for Retinal Regeneration, Center for Developmental Biology, RIKEN, Kobe, Japan. [4]Department of Tissue and Organ Development, Gifu University Graduate School of Medicine, Gifu, Japan. [5]Department of Oral and Maxillofacial Science, Gifu University Graduate School of Medicine, Gifu, Japan. [6]Human Leukocyte Antigen (HLA) Laboratory, Kyoto, Japan. [7]Institute for Integrated Cell-Material Sciences, Kyoto University, Kyoto, Japan. [8]Yamanaka Induced Pluripotent Stem Cell Project, Japan Science and Technology Agency, Kawaguchi, Japan. [9]Gladstone Institute of Cardiovascular Disease, San Francisco, California, USA. Correspondence should be addressed to K.O. (okita@cira.kyoto-u.ac.jp) or S.Y. (yamanaka@cira.kyoto-u.ac.jp).

RECEIVED 29 OCTOBER 2010; ACCEPTED 8 FEBRUARY 2011; PUBLISHED ONLINE 3 APRIL 2011; DOI:10.1038/NMETH.1591



(EGFP) into fibroblasts and monitored fluorescence. Sixty-eight percent of the cells were fluorescent 1 week after transfection (**Supplementary Fig. 7**). However, the signal quickly decreased thereafter, and only 2.4% of cells were fluorescent 4 weeks after electroporation. Then we estimated the copy numbers of the episomal vectors in established pla-iPSC clones. We designed a PCR primer pair for *EBNA-1* sequence derived from Epstein-Barr virus to calculate the copy numbers of the episomal vectors and another primer pair for the endogenous *FBXO15* locus to estimate the cell number. We detected ~200 copies of the episomal vectors per cell 6 d after transfection (**Fig. 2e** and **Supplementary Fig. 3b**). In contrast, we detected no *EBNA-1* DNA in five of seven clones tested at passages 11–20 (~80–120 d after transfection). The remaining two clones contained ~0.001 and 2 copies, respectively. The latter clone likely had integrated the plasmid into a chromosome.

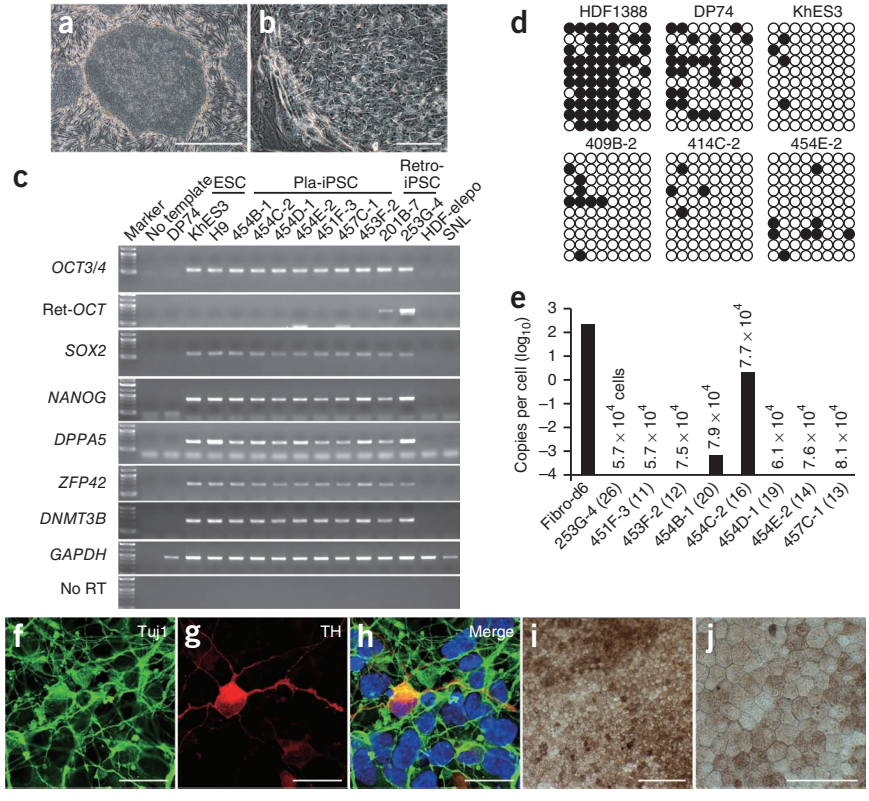


Figure 1 | Establishment of human iPSCs. (**a**) Combinations of reprogramming factors and episomal vectors used in this study. (**b**) Episomal expression vectors in the Y4 combination. *CAG*, CAG promoter; *WPRE*, woodchuck hepatitis post-transcriptional regulatory element; and *pA*, polyadenylation signal. (**c**) Schematic of the pla-iPSC induction protocol. DMEM, Dulbecco's modified Eagle medium; FBS, fetal bovine serum; MSCGM, mesenchymal stem cell growth medium; bFGF, basic fibroblast growth factor. (**d**) Numbers of colonies per 1.0×10^5 cells obtained with different combinations of reprogramming factors. Control, cells transduced with episomal vector encoding EGFP; MEF, mouse embryonic fibroblasts; SNL, mouse embryonic fibroblast cell line. Data are means ± s.d. of numbers of ESC-like colonies obtained from 15 independent induction experiments using five cell lines. $^{****}P < 0.05$ against T1, T2, T3 and control; $^{***}P < 0.05$ against T1, T3 and control; $^{**}P < 0.05$ against T1 and control; $^{*}P < 0.05$ against control.

Figure 2 | Characterization of pla-iPSC clones. (**a**,**b**) Phase contrast images of an established pla-iPSC line. Scale bars, 1 mm (**a**) and 100 µm (**b**). (**c**) RT-PCR analyses for pluripotent cell markers. Total RNA was isolated from pla-iPSC clones established with the Y1 (clone 454B-1), Y2 (454C-2), Y3 (454D-1) or Y4 (454E-2, 451F-3, 457C-1 and 453F-2) combinations, from retrovirus-derived iPSC clones (retro-iPSC) and from ESC lines. In the lanes labeled *OCT3/4* and *SOX2*, PCR primers only detected endogenous gene expression; in the Ret-*OCT* lane, PCR primers specifically amplified the retroviral *OCT3/4* transgene. *GAPDH* was used as a loading control. As a negative control, *GAPDH* amplification was also performed without reverse transcription (no RT). Fibroblasts 4 d after electroporation of the Y4 mixture (HDF-elepo) and mouse embryonic fibroblast cell line (SNL) were used as other negative controls. (**d**) DNA methylation status of the *NANOG* promoter region in the indicated cell lines. Open and closed circles indicate unmethylated and methylated CpG dinucleotides, respectively. (**e**) Copy numbers of episomal vectors in pla-iPSC clones. Numbers in parentheses indicate passage number. Also shown are the estimated numbers of cells analyzed for each clone. Fibroblasts 6 d after electroporation of the Y4 combination were analyzed (fibro-d6) as a positive control. (**f**–**h**) Differentiation of pla-iPSC clone (454E-2) into dopaminergic neurons. Micrographs are immunostained for Tuj1 (**f**) and tyrosine hydroxylase (TH) (**g**). A merged image with nuclear staining using DAPI (**h**) is shown. Scale bars, 20 µm. (**i**,**j**) Differentiation of pla-iPS clone (454E-2) into retinal pigment epithelial cells. Scale bars, 100 µm (**i**) and 50 µm (**j**).

These data demonstrated that the episomal vectors were spontaneously lost in the majority of pla-iPSC clones.

We examined the differentiation potential of pla-iPSCs *in vivo*. Injection of pla-iPSCs into the testes of immunodeficient mice yielded tumors within 3 months. Histological examination confirmed that these tumors were teratomas and contained tissues of all three germ layers, including neural epithelium, cartilage and gut-like epithelium (**Supplementary Fig. 8**).

We carried out directed differentiation of the pla-iPSCs into dopaminergic neurons *in vitro* (Online Methods). RT-PCR detected upregulation of *SOX1*, a marker of immature neural cells, and downregulation of *OCT3/4* 12 d after induction (**Supplementary Fig. 9a**). Immunostaining showed that the majority of cells expressed Nestin after 29 d, with some cells still proliferating and expressed Ki67 (**Supplementary Fig. 9b–e**). Clusters of Nestin-expressing cells expressed PAX6, and more mature cell clusters expressed tyrosine hydroxylase, a marker of dopaminergic neurons (**Supplementary Fig. 9f,g**). Tyrosine hydroxylase–expressing cells localized with the neural markers Tuj1 and MAP2ab, and the vesicular monoamine transporter VMAT2 (**Fig. 2f–h** and **Supplementary Fig. 9h–l**). Therefore, pla-iPSCs have the potential to differentiate into dopaminergic neurons.

We also examined whether pla-iPSC clones differentiated into retinal pigment epithelial cells using a modified stromal cell–derived inducing activity method (Online Methods). Five of six pla-iPSC clones developed pigmented cell clusters after 30 d in conditioning medium of mouse PA6 stromal cells. The clusters grew and exhibited a squamous and hexagonal morphology, characteristic of retinal pigment epithelial cells (**Fig. 2i,j**).

We examined the human leukocyte antigen (HLA) types of our dental pulp-derived iPSC lines. In a previous study only one HLA type had been detected in two dental pulp lines by a PCR–reverse sequence–specific oligonucleotide probe (rSSOP) protocol[11]: line DP74 had been typed as *HLA-A*24*, –; *HLA-B*52*, –; *HLA-DRB1*15*, –, and line DP94 as *HLA-A*11*, –; *HLA-B*15*, –; *HLA-DRB1*04*, – ('–' means no other allele was detected; **Supplementary Table 7**). We also typed these lines with two additional analyses. A PCR-rSSOP protocol optimized for the Japanese population typed line DP74 and its progeny iPSC lines (454E-2 and 457C-1) as *HLA-A*24:02*, –; *HLA-B*52:01*, –; *HLA-DRB1*15:02*, –, and typed DP94 and its progeny iPSC line (453F-2) as *HLA-A* 11:01*, –; *HLA-B*15:01*, –; *HLA-DRB1*04:06*, –. Sequence-based typing showed that the types of DP74 and DP94 were *HLA-A* 24:02:01*, –; *HLA-B*52:01:01*, –; *HLA-DRB1*15:02:01*, – and *HLA-A*11:01:01*, –; *HLA-B*15:01:01*, –; *HLA-DRB1*04:06:01*, –, respectively. The families of the donors of the two dental pulp lines could not be typed because the lines were established in an anonymous way. Therefore, it is not possible to formally conclude that these donors are homozygous for the HLA haplotypes. Nevertheless, the fact that three independent analyses detected only one type in each donor is indicative of homozygosity.

According to the HLA Laboratory database, frequencies of *HLA-A*24:02*; *HLA-B*52:01*; *HLA-DRB1*15:02* and *HLA-A*11:01*; *HLA-B*15:01*; *HLA-DRB1*04:06* haplotypes in the Japanese population are 8.5% and 1.3%, respectively (http://www.hla.or.jp/hapro_e/top.html; **Supplementary Table 8**). Theoretically, iPSCs established from these two individuals match ~20% of all the combinations of 2,117 haplotypes in Japanese population. Indeed, pla-iPSC lines derived from lines DP74 and DP94 match 32 of 107 donors[11] at the three *HLA* loci (*HLA-A*, *HLA-B* and *HLA-DR*) with the two-digit specification (**Supplementary Table 7**).

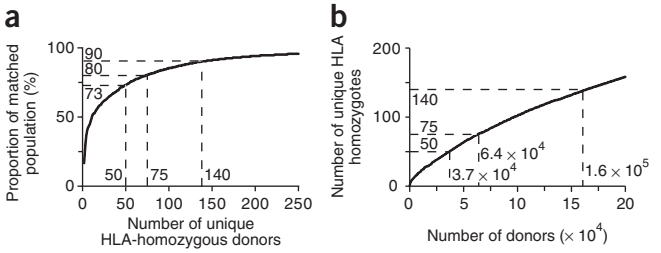


Figure 3 | Estimated coverage of the Japanese population by HLA homozygous donors. (**a**) Estimated cumulative coverage of the Japanese population by theoretical unique *HLA* homozygous donors at *HLA-A*, *HLA-B* and *HLA-DRB1* loci with four-digit specification. (**b**) Estimated numbers of donors required to identify individuals with unique HLA homozygous haplotypes.

Others previously estimated that iPSC lines with 50 unique HLA homozygous haplotypes would match ~90% of the Japanese population at the *HLA-A*, *HLA-B* and *HLA-DRB1* loci with two-digit specification[12]. We performed a similar estimation with four-digit specification using the HLA Laboratory database and found that 50 unique HLA-homozygous donors would cover ~73% of the Japanese population (**Fig. 3a** and **Supplementary Table 8**). Approximately 75 and 140 unique donors would be needed to cover ~80% and 90%, respectively. It would be necessary to type ~37,000, ~64,000 and ~160,000 individuals, respectively, to identify these 50, 75 and 140 donors (**Fig. 3b**).

Allografts using *HLA*-homozygous iPSCs may provide a therapeutic alternative to autologous grafts, for cases in which transplant is likely to be needed soon after injury; furthermore, they allow for the advance selection of safe clones[13]. The beneficial effects of matching at major *HLA* loci are well documented in renal transplantation[14,15], although recipients of allografts derived from *HLA*-homozygous iPSCs would still need immunosuppressants after transplantation because of other HLA antigens, non-HLA antigens and immunity by natural killer cells.

We report a simple, non-integrative method for reprogramming human cells. The increased efficiency and the use of non-transforming Myc should be useful to generate iPSCs from many donors, such as individuals with disease. The approach may also prove beneficial for generating human iPSCs for use in autologous and allologous stem cell therapy.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Accession codes. Addgene: 27076 (pCXLE-hOCT3/4), 27077 (pCXLE-hOCT3/4-shp53-F), 27078 (pCXLE-hSK), 27079 (pCXLE-hMLN), 27080 (pCXLE-hUL), 27081 (pCXLE-Fbx15-cont2) and 27082 (pCXLE-EGFP).

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank K. Takahashi, T. Aoi and Y. Yoshida for scientific discussion; M. Narita, T. Ichisaka, M. Ohuchi, M. Nishikawa and N. Takizawa for technical assistance; R. Kato, E. Nishikawa, S. Takeshima, Y. Ohtsu and H. Hasaba for administrative assistance; and H. Niwa (RIKEN) and J. Miyazaki (Osaka University) for the *CAG* promoter. This study was supported in part by a grant from the Program

for Promotion of Fundamental Studies in Health Sciences of National Institute of Biomedical Innovation, a grant from the Leading Project of Ministry of Education, Culture, Sports, Science and Technology (MEXT), a grant from Funding Program for World-Leading Innovative Research and Development on Science and Technology (FIRST Program) of Japan Society for the Promotion of Science, Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science and MEXT (to S.Y.) and Senri Life Science Foundation (to K.O.). H.H. is supported by a Japanese government (MEXT) scholarship.

AUTHOR CONTRIBUTIONS

K.O. and S.Y. conceived the project and wrote the manuscript. K.O. constructed the vectors with H.H., M.N. and K. Tanabe, and conducted most of the experiments with Y.M., Y. S. and A.O. A.M. and J.T. carried out the differentiation experiment into dopaminergic neurons. S.O. and M.T. performed differentiation into retinal pigment epithelial cells. K. Tezuka., T.S. and T.K. established dental pulp cell lines. H.S. performed HLA haplotyping in Japanese population and supervised HLA analysis.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at http://www.nature.com/naturemethods/.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Okita, K., Ichisaka, T. & Yamanaka, S. *Nature* **448**, 313–317 (2007).
2. Zhou, W. & Freed, C.R. *Stem Cells* **27**, 2667–2674 (2009).
3. Fusaki, N., Ban, H., Nishiyama, A., Saeki, K. & Hasegawa, M. *Proc. Jpn. Acad. B* **85**, 348–362 (2009).
4. Woltjen, K. *et al. Nature* **458**, 766–770 (2009).
5. Jia, F. *et al. Nat. Methods* **7**, 197–199 (2010).
6. Yu, J. *et al. Science* **324**, 797–801 (2009).
7. Kim, D. *et al. Cell Stem Cell* **4**, 472–476 (2009).
8. Warren, L. *et al. Cell Stem Cell* **7**, 618–630 (2010).
9. Hong, H. *et al. Nature* **460**, 1132–1135 (2009).
10. Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T. & Yamanaka, S. *Proc. Natl. Acad. Sci. USA* **107**, 14152–14157 (2010).
11. Tamaoki, N. *et al. J. Dent. Res.* **89**, 773–778 (2010).
12. Nakatsuji, N., Nakajima, F. & Tokunaga, K. *Nat. Biotechnol.* **26**, 739–740 (2008).
13. Tsuji, O. *et al. Proc. Natl. Acad. Sci. USA* **107**, 12704–12709 (2010).
14. Takemoto, S.K., Terasaki, P.I., Gjertson, D.W. & Cecka, J.M. *N. Engl. J. Med.* **343**, 1078–1084 (2000).
15. Aydingoz, S.E. *et al. Hum. Immunol.* **68**, 491–499 (2007).

ONLINE METHODS

Cell culture. Human fibroblasts HDF1419, HDF1388, HDF1429, HDF1377, HDF1437 and HDF1554 were purchased from Cell Applications, and TIG121, TIG120, TIG114 and TIG107 were obtained from the Japanese Collection of Research Bioresources. Human fibroblasts were cultured in DMEM (Nacalai Tesque) supplemented with 10% FCS (Invitrogen). Human dental pulp (DP) cells were established from human third molars as described previously[11] and were maintained in mesenchymal stem cell growth medium (MSCGM; Lonza). Human ESC lines (KhES-1 and KhES-3) were obtained from Kyoto University. H1 and H9 were from WiCell Research Institute. Mouse embryonic fibroblasts (MEFs) were isolated from embryonic day 13.5 embryos of C57BL/6 mice. All mice used in this study were bred and killed appropriately following code of ethics of animal research committee in Kyoto University. MEF and mouse embryonic fibroblast cell line (SNL) cells[16] were cultured in DMEM supplemented with 7% (v/v) FCS, 2 mM L-glutamine and 50 units and 50 mg ml^{-1} penicillin and streptomycin, respectively. Established iPSCs and ESCs were maintained on mitomycin C–treated SNL cells in primate ESC medium (ReproCELL) containing 4 ng ml^{-1} of bFGF (Wako) as described previously[17].

Vector construction. Efficient transgene expression was achieved by inserting the woodchuck hepatitis post-transcriptional regulatory element (*WPRE*) upstream of the polyadenylation signal of pCX-EGFP[18]. The episomal cassette was transferred from pCEP4 (Invitrogen). The *EBNA-1* sequence (EcoRI and MfeI sites) was flanked by two *loxP* sequences, and the *loxP–EBNA-1–loxP–OriP* cassette was then digested with BamHI and BglII and inserted into the BamHI site of pCX-EGFP containing the *WPRE*. This episomal vector was designated pCXLE-EGFP.

Human cDNAs encoding *OCT3/4*, *SOX2*, *KLF4*, *c-MYC*, *L-MYC*, *NANOG* and *LIN28* were amplified by PCR and cloned into pCR2.1 (Invitrogen). The translation termination codons of *SOX2*, *c-MYC*, *L-MYC* or *LIN28* were replaced with a BamHI site and then were also cloned into pCR2.1. The cDNAs without a translation termination codon were thereafter ligated with *2A* self-cleavage sequences in pBS-2A[19] with appropriate restriction enzymes to generate pBS–cDNA–2A. *c-MYC–2A* was digested with NotI and BspHI and was ligated into the NotI and NcoI sites of pBS-LIN28-2A in the same reading frame to generate the *c-MYC–LIN28–NANOG* cassette. These *cDNA–2A* or *c-MYC–2A–LIN28–2A* constructs were then ligated to another cDNA or *NANOG* in pCR2.1 with the translation termination codon in the same reading frame using appropriate restriction enzymes. These cDNA-*2A*-cDNA-stop or *c-MYC–2A–LIN28–2A–NANOG*–stop constructs were then inserted into the EcoRI site of pCXLE-EGFP. pCXLE-hOCT3/4-shp53 was constructed by inserting an shRNA expression cassette for *p53*, driven by the mouse *U6* promoter, into the BamHI site of pCXLE-hOCT3/4. The pCXLE-Fbx15-cont2 was generated by inserting the *FBXO15* cDNA into pCXLE-EGFP. Episomal vectors described previously[6] were obtained from Addgene (20922–20927).

Generation of iPSCs with episomal vectors. HDF and DP cells were cultured in DMEM supplemented with 10% FBS and mesenchymal stem cell growth medium (MSCGM), respectively. Three micrograms of expression plasmid mixtures were electroporated into 6×10^5 HDF or DP cells with Microporator (Invitrogen) with a 100-µl kit according to the manufacturer's instructions. The plasmid mixtures used in the experiments are shown in **Supplementary Table 2**. Conditions used were 1,650 V, 10 ms, 3 time pulses for HDF, and 1,800 V, 20 ms, 1 time pulse for DP cells. The cells were trypsinized 7 d after transduction, and 1×10^5 cells were re-plated onto 100-mm dishes covered with an SNL or MEF feeder layer. The culture medium was replaced the next day with primate ESC medium containing bFGF. The colonies were counted 26–32 d after plating, and those colonies similar to human ESCs were selected for further cultivation and evaluation. The pla-iPSC clones used in this study are summarized in **Supplementary Table 9**.

Characterization of pla-iPSC clones. Isolation of total RNA, RT-PCR of marker gene expression, DNA microarray, bisulfite genomic sequencing and teratoma formation were performed as previously described[17]. The primer sequences used in this study are shown in **Supplementary Table 10**. The chromosomal G-band analyses were performed at the Nihon Gene Research Laboratories. Short tandem repeat analyses were performed at Bex Co.; briefly, genomic DNAs were amplified by the PowerPlex 16 system (Promega) and were then analyzed with an ABI PRISM 3100 genetic analyzer and the GeneMapper v3.5 software program (Applied Biosystems). Differentiation of pla-iPSC clones into dopaminergic neurons was performed using the serum-free culture of embryoid body–like aggregates (SFEB) method combined with double SMAD inhibition by a BMP antagonist and an Activin/Nodal inhibitor as described elsewhere[20]. *In vitro* directed differentiation into retinal pigment epithelial cells was performed with the modified stromal cell–derived inducing activity method[21,22]. Briefly, pla-iPSCs were collected with trypsin and collagenase IV, and were treated with inhibitors for WNT and Nodal signaling under serum-free conditions. The cells were then maintained in PA6-conditioning medium for maturation.

Episomal copy-number detection. Cells cultured in 60-mm dishes were collected with a cell scraper after removing feeder cells by treatment with dissociation solution consisting of 0.25% trypsin, 1 mg ml^{-1} collagenase, 1 mM $CaCl_2$ and 20% KSR in PBS. The cells were then placed into tubes and centrifuged, and the cell pellets were lysed with 200 µl of lysis solution, containing 1× Ex Taq buffer (Takara) and 167 µg ml^{-1} proteinase K. The lysates were incubated at 55 °C for 3 h, and proteinase K was inactivated at 95 °C. The lysates were used for quantitative PCR analysis. The pCXLE-hFbx15-cont2 plasmid was used to generate a standard curve to determine the correlation between copy number and threshold cycle (Ct) values for *FBXO15* or *EBNA-1*. Then the copy number of *FBXO15* and *EBNA-1* in each iPSC sample was estimated from the observed Ct values. The cell number in each reaction was estimated by dividing the estimated copy number of *FBXO15* by two since each cell had two *FBXO15* alleles. One reaction included up to 1.2×10^4 cells. The total copy number of *EBNA-1* was measured in ~5×10^4 cells by repeating six or seven reactions.

HLA typing and estimation of coverage. HLA typing of 107 DP cell lines was performed with the PCR-reverse sequence specific oligonucleotide probe (rSSOP) method using LABType SSO

(One lambda) at Repro Cell[23]. Additional HLA typing was performed with PCR-rSSOP using WAKFlow (Wakunaga Pharmaceutical) at HLA Laboratory. We performed pedigree study of 4,743 Japanese families (17,325 members) and identified 2,117 haplotypes, including interlocus recombinant haplotypes, which were detected in family studies. The haplotype frequency was calculated by direct counting on the parents in the families. Sequence-based typing was performed with AlleleSEQR (Atria Genetics) at Mitsubishi Chemical Medience Corporation.

To estimate coverage of Japanese population by HLA homozygous donors, we first calculated the frequencies of all possible combinations of the 2,117 HLA haplotypes shown in **Supplementary Table 8**. Haplotype combinations that can be covered by a given homozygous donor were then identified and their frequencies were added to estimate coverage by the homozygous donor. When one *HLA-A, HLA-B, HLA-DRB1* heterozygous individual was covered by multiple homozygous donors, we counted only once to avoid overestimation.

The expected number (EN) of each homozygous haplotype at a given population size (PS) was first calculated as; EN = (haplotype frequency)^2 × PS. EN (if EN < 1) or 1 was then summed for each homozygous haplotype to estimate the expected numbers of unique HLA haplotype donors at the given PS.

Statistical analyses. Data are shown as the mean ± s.d. Statistical significance among multiple groups was evaluated with the Steel-Dwass test.

16. McMahon, A.P. & Bradley, A. *Cell* **62**, 1073–1085 (1990).
17. Takahashi, K. *et al. Cell* **131**, 861–872 (2007).
18. Niwa, H., Yamamura, K. & Miyazaki, J. *Gene* **108**, 193–199 (1991).
19. Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. & Yamanaka, S. *Science* **322**, 949–953 (2008).
20. Morizane, A., Doi, D., Kikuchi, T., Nishimura, K. & Takahashi, J. *J. Neurosci. Res.* **89**, 117–126 (2010).
21. Kawasaki, H. *et al. Proc. Natl. Acad. Sci. USA* **99**, 1580–1585 (2002).
22. Osakada, F. *et al. J. Cell Sci.* **122**, 3169–3179 (2009).
23. Itoh, Y. *et al. Immunogenetics* **57**, 717–729 (2005).

doi:10.1038/nmeth.1591