Supporting Information

Koyanagi-Aoi et al. 10.1073/pnas.1319061110

SI Materials and Methods

Cell Culture. The human embryonic stem cells (hESCs) and human induced pluripotent stem cells (hiPSCs) were maintained in Primate ES cell medium (ReproCELL) supplemented with 4 ng/mL of human recombinant basic fibroblast growth factor (bFGF) (Wako) on SNL feeders (1–3). Human dermal fibroblasts (HDFs) were obtained from the Japanese Collection of Research Bioresources or were purchased from Cell Applications, Inc. Dental pulp (DP) cells were kindly provided by Dr. Ken-ichi Tezuka (Gifu University Graduate School of Medicine, Gifu Prefecture, Japan). HDFs and human embryonic carcinoma cell (hECC) lines were maintained in DMEM (Nacalai Tesque) containing 10% (vol/vol) FBS (Thermo) and 0.5% penicillin/streptomycin (Life Technologies). DP cells were cultured in MSBGM medium (Lonza). $CD34^+$ cord blood cells were obtained from the Stem Cell Resource Network in Japan (banks at Miyagi, Tokyo, Kanagawa, Aichi, and Hyogo) through the RIKEN BioResource Center (Tsukuba, Ibaraki, Japan). The peripheral blood was harvested from healthy donors whose written informed consent was obtained in accordance with the ethical committee of the Department of Medicine and Graduate School of medicine, Kyoto University. The mononuclear cells were then isolated by density gradient centrifugation with Ficoll-Paque PLUS (GE Healthcare).

Generation of Human iPSCs. The generation of hiPSCs from HDFs, DP cells, and blood samples using a retroviral system or episomal vectors was performed as described previously (1, 4–6). TKCBV4-2, 5-6, and TKCB7-2 iPS cells were kindly provided by Drs. Koji Eto and Naoya Takayama (Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan) (7).

During the generation of hiPSCs from blood using Sendai viral vectors, vectors encoding POU class 5 homeobox 1 (OCT3/4), SRY (sex determining region Y)-box 2 (SOX2), Kruppel-like factor 4 (gut) (KLF4), and v-myc avian myelocytomatosis viral oncogene homolog (MYC, also known as c-MYC) (CytoTune-iPS; DNAVEC) were infected into $CD34^+$ cells at a multiplicity of infection of 3 or 10 in αMEM medium supplemented with 10% (vol/vol) FBS, 50 ng/mL IL-6, 50 ng/mL sIL-6R, 50 ng/mL stem-cell factor, 10 ng/mL thrombopoietin, 20 ng/mL Flt3 ligand, and 20 ng/mL IL-3. The next day, the infected cells were centrifuged to remove residual virus, plated onto six-well plates covered with MEF feeder cells, and cultured in Primate ES cell medium supplemented with 4 ng/mL of bFGF until colonies were formed. Sendai virus infection and the generation of iPSCs from αβT cells were carried out as described previously (8).

RNA Extraction. We lysed the cells at subconfluent density using TRIzol reagent (Life Technologies) and total RNA was purified by a standard protocol. The RNA concentration and purity were determined through measurement of the A260/280 ratios with a Nanodrop instrument (Thermo Scientific). For microarrays, confirmation of the RNA quality was performed using a 2100 Bioanalyzer (Agilent Technologies).

mRNA Expression Analysis. The gene expression profiling was carried out using the SurePrint G3 human GE microarray (Agilent Technologies) according to the manufacturer's protocol. The data were analyzed using the GeneSpring GX 11.5.1 software program (Agilent Technologies). The data processing was performed as follows: (*i*) Threshold raw signals were set to 1.0, (*ii*) log base 2 transformation was performed, and (*iii*) 75th percentile normalization was chosen as the normalized algorithm (http://genespring-support.com/faq/normalization). The flag setting was performed as follows: Feature is not positive and significant (not detected), not uniform (compromised), not above background (not detected), saturated (compromised), or is a population outlier (compromised). Control probes were removed and only the "detected" probes that were present in at least one sample in all hES/hiPS cell samples were used for the further analysis. The number of probes used in the analysis was 36,757 (Fig. 1*A*) and 36,083 (Fig. 3*A*).

MicroRNA Microarray Analysis. The miRNA expression profiling was carried out using the Agilent Technologies human miRNA microarray release 12.0 according to the manufacturer's protocol. The data were analyzed using the GeneSpring GX 11.5.1 software program (Agilent Technologies), and data processing was performed in the same way as for the mRNA expression analysis, except that 90th percentile normalization was chosen as the normalized algorithm. The number of probes used in the analysis was 476 (Fig. 1*B*).

Genomic DNA Extraction and Bisulfite Treatment. Genomic DNA extraction and purification from cultured cells was carried out using a Gentra Puregene kit (Qiagen). Extracted DNA was quantitated by using the Nanodrop instrument, and the quality was assessed by gel electrophoresis. A total of 500 ng of genomic DNA was treated with bisulfite using the EZ DNA Methylation-Gold Kit (Zymo Research Corp.) according to the manufacturer's protocol.

DNA Methylation Analysis with a Beads Array. Genome-wide DNA methylation profiling was performed using the Illumina Infinium Human Methylation27 BeadChip (Illumina). Bisulfite-converted DNA was used, and the remaining assay steps were performed using the reagents supplied by Illumina and their specified conditions. The readout from the array was a β-value, which was defined as the ratio between the fluorescent signal from the methylated allele to the sum of both methylated and unmethylated alleles, and thus correlated with the level of DNA methylation. A β-value of 1.0 corresponds to complete methylation and 0 is equal to no DNA methylation. To exclude potential sources of technical bias, we only used CpG sites with detection P values <0.05 in at least 56 out of 59 samples. Normalization was not performed. The number of probes used in the analysis was 27,445 (Fig. 1*E*).

Generation of Heat Maps. We used Microsoft Excel to visualize the values as heat maps. The color spectrum expands from green (lower value) to magenta (higher value) through black in the gene/miRNA expression analysis, and from white (hypomethylation) to magenta (hypermethylation) in the DNA methylation analysis.

Bioinformatic Analysis. A hierarchical clustering analysis was performed using the GeneSpring GX 11.5.1 software program.

DNA Methylation Analysis with Pyrosequencing and Clonal Sequencing. Pyrosequencing was carried out with primers designed using the Pyromark Assay Design software program, version 2.0 (Qiagen). The primer sequences are shown in Table S1*D*. PCR was performed in a 25-μL reaction mixture containing 25 ng bisulfite-converted DNA, 1× Pyromark PCR Master Mix (Qiagen), 1× Coral Load Concentrate, and 0.2 μM forward and 5′ biotinylated reverse primers. The PCR conditions were 45 cycles of 95 °C for 30 s, 56 °C for 30 s, and 72 °C for 30 s. The PCR product was

bound to streptavidin Sepharose beads (GE Healthcare) and was purified, washed, denatured, and washed again. Then, 16 pmol of the sequencing primer was annealed to the purified PCR product. Pyrosequencing reactions were performed using the PSQ HS 96 Pyrosequencing System. The degree of methylation is shown as the percentage of methylated cytosines divided by the sum of methylated and unmethylated cytosines (percentage of 5 mC). Bisulfite-clonal sequencing was performed as previously described (9). In Fig. 4*F* and Fig. S3*C* we examined the DNA methylation status of LTR7 and its neighboring regions in defective clones with high expression levels of these genes ($n = 6$; TKCBV5-6, TIG118-4f1, TIG108-4f3, 1503-4f1, TIG107-3f1, and 451F3), good clones with low expression levels of these genes ($n = 6$; 703B1, 606B1, 201B7, H1, 253G1, and 454E2), two hECC lines (NTera2 and 2102Ep), and two somatic cell lines (HDF and DP).

Exon Arrays. We performed the exon array of the abhydrolase domain containing 12B (*ABHD12B*) and chromosome 4 open reading frame 51 (*C4orf51*) genes for three defective clones with high expression levels of these genes ($n = 3$, TKCBV5-6, TIG108-4f3, and 451F3), three good clones with low expression levels of these genes ($n = 3$, 201B7, H1, and 253G1), and somatic cells (HDF and DP74). cDNA was generated with the WT expression kit (Ambion) per the manufacturer's instructions. The cDNA was fragmented and end-labeled with a GeneChip WT Terminal labeling kit (Affymetrix). Approximately 5.5 μg of labeled DNA target was hybridized to the Affymetrix GeneChip Human Exon 1.0 ST Array at 45 °C for 17 h, per the manufacturer's recommendations. Hybridized arrays were washed and stained on a GeneChip Fluidics Station 450 and scanned on a GCS3000 Scanner (Affymetrix). An exon array data analysis was performed using the GeneSpring GX11.5.1 software program using ExonPLIER16 with core and extended probe sets. The exon probe sets were accepted if they had a detection above background P value <0.05 in at least one of the samples.

Selection of Microarray Probes Related to LTR7. The RepeatMasker database classifies sequences into subgroups by in silico analysis. We used LTR7 sequences registered in Repeatmasker open 3.3.0 – Repeat Library 20110920, a database of human repetitive sequence (www.repeatmasker.org/species/homSap.html). According to this database, there are 3,523 LTR7 elements, including LTR7A, 7B, and 7C in the human genome. To extract microarray probes that were potentially affected by LTR7 elements, we first selected genes containing LTR7 elements in their gene bodies or in the regions 2 kb upstream from their transcription start sites. We then extracted microarray probes that were located between the each LTR7 and the 3′ end of the corresponding gene body (the direction of LTR7 was not considered). As a result, we extracted 763 probes corresponding to 435 genes from the Agilent Technologies human G3 microarray as LTR7-related probes (Table S1*C*).

Quantitative RT-PCR. To remove any potential contamination by genomic DNA, we treated purified RNA samples with a Turbo DNA free kit (Ambion). After DNase treatment, reverse transcription was performed with the oligo dT_{20} primer using a ReverTra Ace-α-kit (Toyobo). Quantitative RT-PCR was performed with SYBR Premix Ex Taq II (Takara), and samples were analyzed with the StepOne plus real-time PCR system (Applied Biosystems). The primer sequences are shown in Table S1*D*. The relative expression level was calculated by using plasmid DNA containing the PCR product (Fig. 4 and Fig. S5 *C* and *D*).

Neural Induction. We performed the neural differentiation of human pluripotent stem cells with the quick method for serum-free embryoid body formation (SFEBq) as described previously (10). In brief, hESCs and hiPSCs treated with 10 μM Y-27632 were dissociated into single cells and transferred at 9,000 cells per well to 96-well low-cell-adhesion plates (Lipidure-Coat Plate A-U96; NOF Corp.). The cells were cultured for 14 d in DFK5 medium consisting of DMEM/F-12 (Life Technologies), 5% (vol/vol) Knockout Serum Replacement (Life Technologies), 1% (vol/vol) MEM-nonessential amino acids (Life Technologies), 2 mM L-glutamine (Life Technologies), 0.1 mM 2-mercaptoethanol (Life Technologies), and 0.5% penicillin/streptomycin. We used DFK5 medium supplemented with 10 μM Y-27632, 2 μM dorsomorphin (Sigma), and 10 μM SB431542 (Sigma) for the first 4 d.

Adhering neural differentiation of dual SMAD inhibition was performed as described previously (11). Briefly, the cells were plated on matrigel-coated plate and after reaching 90% confluency they were cultured with 100 nM LDN193189 (Stemgent), 10 μM SB431542 in 15% (vol/vol) knockout serum replacement, 2 mM L-glutamine, and 10 μM β-mercaptoethanol–containing DMEM for 12 d.

For dopaminergic differentiation, we first transferred ESCs or iPSCs onto 96-well low-cell-adhesion plates with Y-27632, dorsomorphin, and SB431542 in the same way as indicated for SFEBq. We supplemented the cultures with 100 ng/mL FGF8 (Peprotech) and 20 ng/mL WNT1 (Peprotech) from days 5 to 12, and with 200 ng/mL Sonic Hedgehog (SHH) (R&D) from days 8 to 12. Twelve days after induction, aggregates were transferred onto six-well plates coated with laminin (Becton–Dickinson) and poly-L-ornithine (Sigma) and were cultured with neurobasal medium (Life Technologies) containing 2% (vol/vol) B27 supplement (Life Technologies), 2 mM L-glutamine, and 0.5% penicillin/streptomycin. We added 200 ng/mL SHH to the medium from days 12–15 and 1 ng/mL FGF20 and 12.5 ng/mL bFGF from days 15–22. On day 22, the cells were dissected into clumps and plated on new six-well plates coated with laminin and poly-L-ornithine and were then cultured with neurobasal medium supplemented with 2 ng/mL GDNF (R&D), 20 ng/mL BDNF (R&D), 400 mM dbcAMP (Sigma), and 200 mM ascorbic acid (Sigma) until day 29. TRA-1-60–positive cell labeling and depletion were performed on day 22 using an Anti-TRA-1-60 MicroBead kit and the autoMACS pro device (Miltenyi Biotech).

Flow Cytometric Analysis. Neural aggregates were dissociated and processed for the flow cytometric analysis by a FACS Aria II instrument (Becton–Dickinson). To analyze the proportion of $OCT3/4^+$ cells, the cells were fixed with 3.7% (wt/vol) formaldehyde, permeabilized with 0.2% TritonX-100, and stained with the appropriate antibody. To count the number of $PSA\text{-}NCAM^+$ cells or $TRA\text{-}1\text{-}60^+$ cells, cells were prepared without fixation. To eliminate the number of dead cells from the total cell population, we stained the cells with propidium iodide after labeling them with the anti–TRA-1-60 or anti–PSA-NCAM antibody, or with red fluorescent reactive dye from the LIVE/DEAD Fixable Dead Cell Stain Kits (Invitrogen) before fixing the cell suspension.

Transplantation of ES/iPS Cell-Derived Dopaminergic Neuron Cultures into the Brains of Nonobese Diabetic/Severe Combined Immune-Deficient Mice. To prepare samples for injection, we scraped and mechanically dissected cells by gently pipetting them up and down a few times, suspended them in culture medium (1×10^6 cells/μL), and injected 2 μL of the cell suspension into the right striatum (2 mm lateral, 1 mm rostral to the bregma; depth, 3 mm from the dura) of nonobese diabetic/severe combined immune-deficient mice (6 wk old, female) using a glass micropipette, as described previously (12). The animal experiments were performed in accordance with the Guidelines for Animal Experiments of Kyoto University, the Institutional Animal Care and Use Committee of Kobe Institute in RIKEN, and the Guide for the Care and Use of Laboratory Animals of the Institute of Laboratory Animal Resources (Washington, DC).

Magnetic Resonance Imaging. Graft imaging was performed with an MRmini SA instrument (DS Pharma Biomedical) by using a cylindrical slotted holder with a 20-mm radio frequency coil constructed for mice. T2-weighted images (repetition time, 2,000 ms; echo time, 69 ms) were recorded. The brains were imaged coronally in a single section through the graft center by using an image matrix of 256 × 128, a field of view of 2 × 4 cm^2, and two excitations. Parametric images were generated by using the Sampler XP-NI software program (DS Pharma Biomedical). Graft section surfaces were measured by using the INTAGE Realia Professional imaging software program (Cybernet).

Immunostaining. Anti-NCAM (ERIC1) antibody (Santa Cruz) was used as a primary antibody and anti-Mouse Ig biotin (Dako) was used as a secondary antibody for immunocytechemistry.

Statistical Analysis. ***Gene/miRNA expression.*** We conducted the *t* test (variances assumed equal) for the normalized, filtered data and controlled the false discovery rate (FDR) at 0.05 using the Benjamini–Hochberg method to identify probes that differed significantly between hESCs and hiPSCs, or for neural differentiation good and defective clones.

DNA methylation determined using the beads array. We conducted a Mann–Whitney *U* test on the filtered data controlling the FDR at 0.05 using the Benjamini–Hochberg method to identify probes that differed significantly between hESCs and hiPSCs.

DNA methylation as determined by pyrosequencing and clonal sequencing. A Mann–Whitney *U* test was used to compare the quantitative methylation values between hESCs and hiPSCs or defective and good groups. Calculations were carried out with the Statview software program.

Graft size after transplantation. A *t* test was used to compare the graft sizes derived from defective and good clones (Fig. 4*C*). In the case of comparisons between graft sizes between unsorted and depleted cell cultures (Fig. S4C), we performed a paired *t* test.

1. Takahashi K, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5):861–872.
2. Fujioka T, Yasuchika K, Nakamura Y, Nakatsuji N, Suemori H (2004) A simple and efficient cryopreservation method for primate embryonic stem cells. *Int J Dev Biol* 48(10):1149–1154.
3. Ohnuki M, Takahashi K, Yamanaka S (2009) Generation and characterization of human induced pluripotent stem cells. *Current Protocols in Stem Cell Biology* (Wiley, New York), Chap 4, Unit 4A.2.
4. Okita K, et al. (2011) A more efficient method to generate integration-free human iPS cells. *Nat Methods* 8(5):409–412.
5. Tamaoki N, et al. (2010) Dental pulp cells for induced pluripotent stem cell banking. *J Dent Res* 89(8):773–778.
6. Okita K, et al. (2013) An efficient nonviral method to generate integration-free human-induced pluripotent stem cells from cord blood and peripheral blood cells. *Stem Cells* 31(3):458–466.
7. Kajiwara M, et al. (2012) Donor-dependent variations in hepatic differentiation from human-induced pluripotent stem cells. *Proc Natl Acad Sci USA* 109(31):12538–12543.
8. Seki T, et al. (2010) Generation of induced pluripotent stem cells from human terminally differentiated circulating T cells. *Cell Stem Cell* 7(1):11–14.
9. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126(4):663–676.
10. Morizane A, Doi D, Kikuchi T, Nishimura K, Takahashi J (2011) Small-molecule inhibitors of bone morphogenic protein and activin/nodal signals promote highly efficient neural induction from human pluripotent stem cells. *J Neurosci Res* 89(2): 117–126.
11. Kriks S, et al. (2011) Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* 480(7378):547–551.
12. Miura K, et al. (2009) Variation in the safety of induced pluripotent stem cell lines. *Nat Biotechnol* 27(8):743–745.

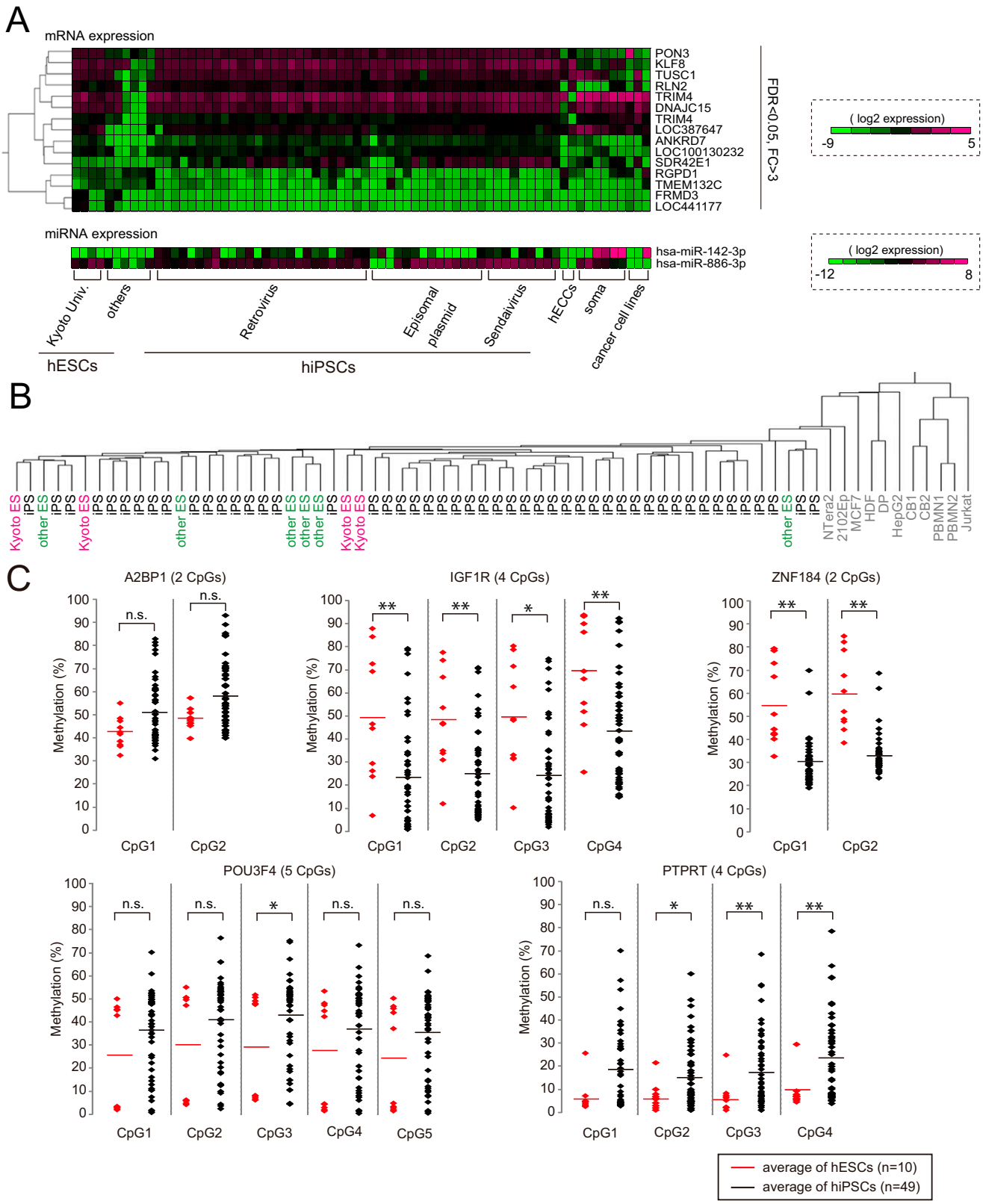

Fig. S1. Gene expression patterns among Kyoto hESCs, other hESCs, and hiPSCs and DNA methylation levels of previously reported ES-iPS differentially methylated regions (DMRs) in our cell lines. (*A*) Heat maps showing the expression levels of 15 probes that were differentially expressed between hESCs and hiPSCs [FDR <0.05 and absolute fold change (FC) >3] and hsa-miR-142-3p and hsa-miR-886-3p in various cell lines. A hierarchical clustering analysis for 15 probes was performed using the Euclidean distance and average linkage algorithm. (*B*) A hierarchical clustering analysis of the global gene expression patterns in various cell lines was performed using the Euclidean distance and average linkage algorithm. (*C*) Previously reported hES-iPS DMRs, A2BP1, IGF1R, ZNF184, POU3F4, and PTPRT, were examined by pyrosequencing in 10 hESCS and 49 hiPSCs in our laboratory. Each CpG dinucleotide position was assayed in triplicate, and average values were plotted. A Mann–Whitney *U* test was used to compare the quantitative methylation values between hESCs and hiPSCs. (n.s., not significant; $*P < 0.05$, $**P < 0.01$).

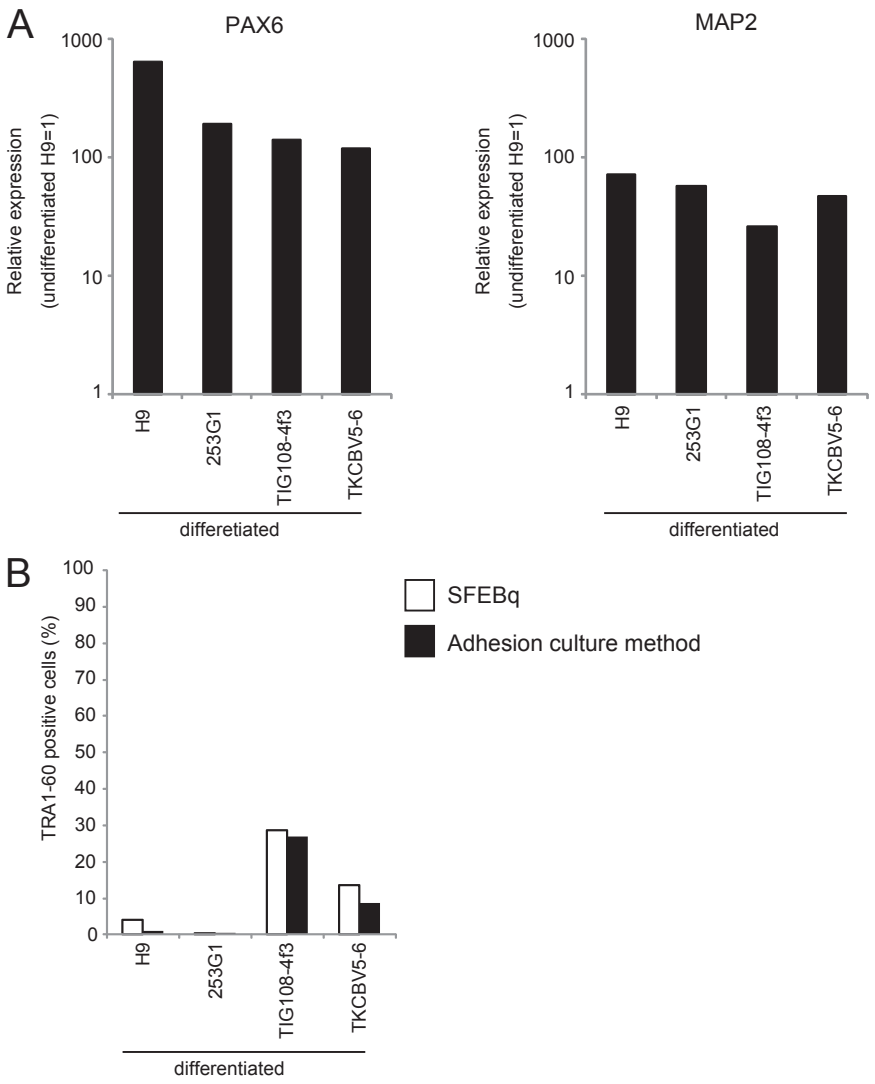

Fig. S2. Most of the cells can differentiate into neural cells although some clones retain undifferentiated cells after neural differentiation. (*A*) The expression levels of PAX6 (*Left*) and MAP2 (*Right*) in neurospheres of differentiated hESC line H9 and hiPSC lines (253G1, TIG108-4f3, and TKCBV5-6) were determined by quantitative RT-PCR. The expression levels in hESC line H9 before differentiation were set to 1, and relative expression levels were presented in log scale. (*B*) A comparison of the proportion of TRA-1-60–positive cells after neural induction between the SFEBq method (white) and the adhesion culture method (black).

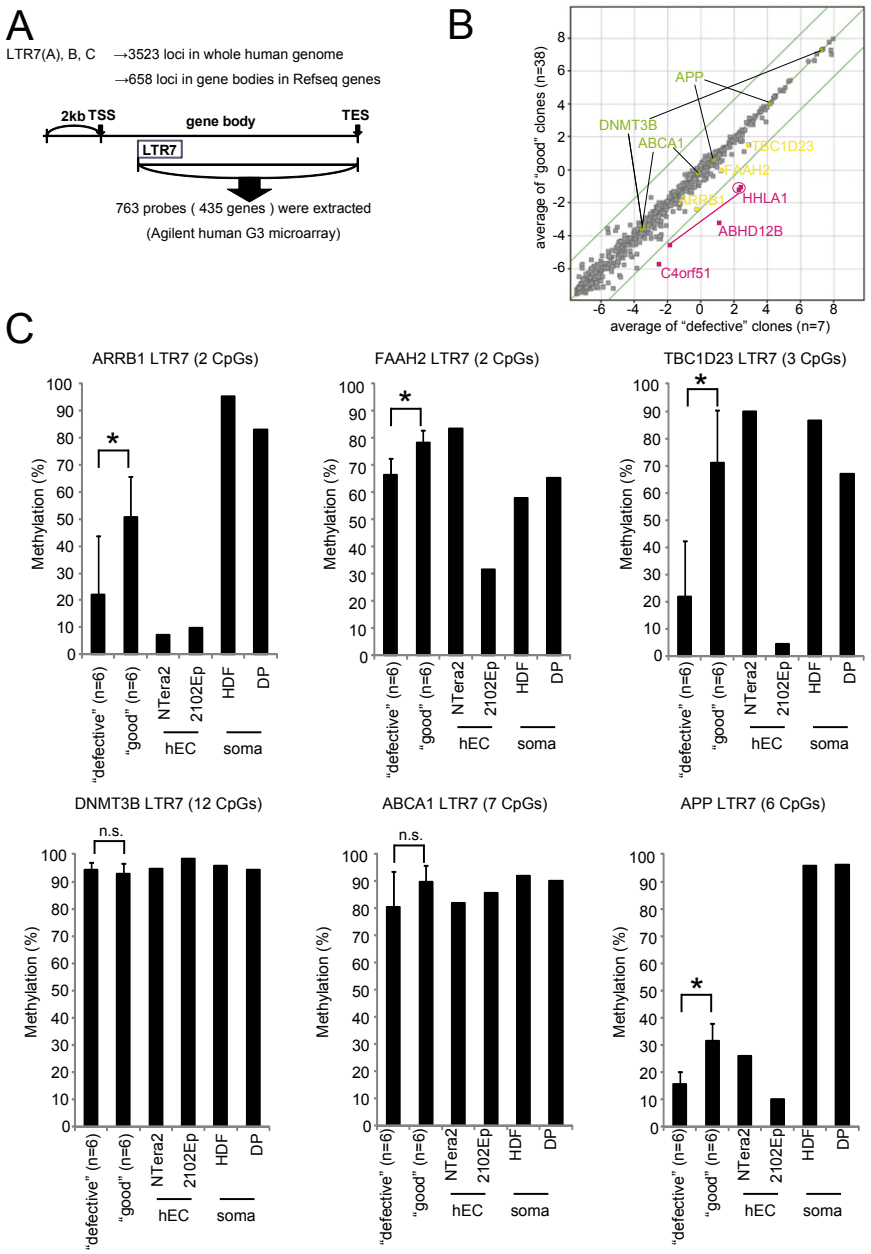

Fig. S3. Activation of LTR7 is not confined to HHLA1, ABHD12B, and C4orf51; DNA hypomethylation exists in some, but not all, LTR7s in defective hiPSC clones. (*A*) The extraction of 763 probes corresponding to 435 genes as LTR7-related probes from the Agilent Technologies human G3 microarray (design ID 028004). (*B*) A comparison of the expression levels of 763 LTR7-related probes between good and defective clones. Magenta-colored genes (ABHD12B, HHLA1, and C4orf51) and yellow-colored genes (ARRB1, FAAH2, and TBC1D23) are differentially expressed between the good and defective clones (FDR <0.05 and FC >5 or FC >2, respectively), and green-colored genes (DNMT3B, ABCA1, and APP) did not show any differences. (*C*) The DNA methylation status of LTR7 and its neighboring regions of ARRB1, FAAH2, and TBC1D23 were examined by pyrosequencing and those of DNMT3B, ABCA1, and APP were examined by clonal sequencing. A Mann–Whitney *U* test was used to compare the quantitative methylation values between defective and good clones (n.s., not significant; $*P < 0.05$).

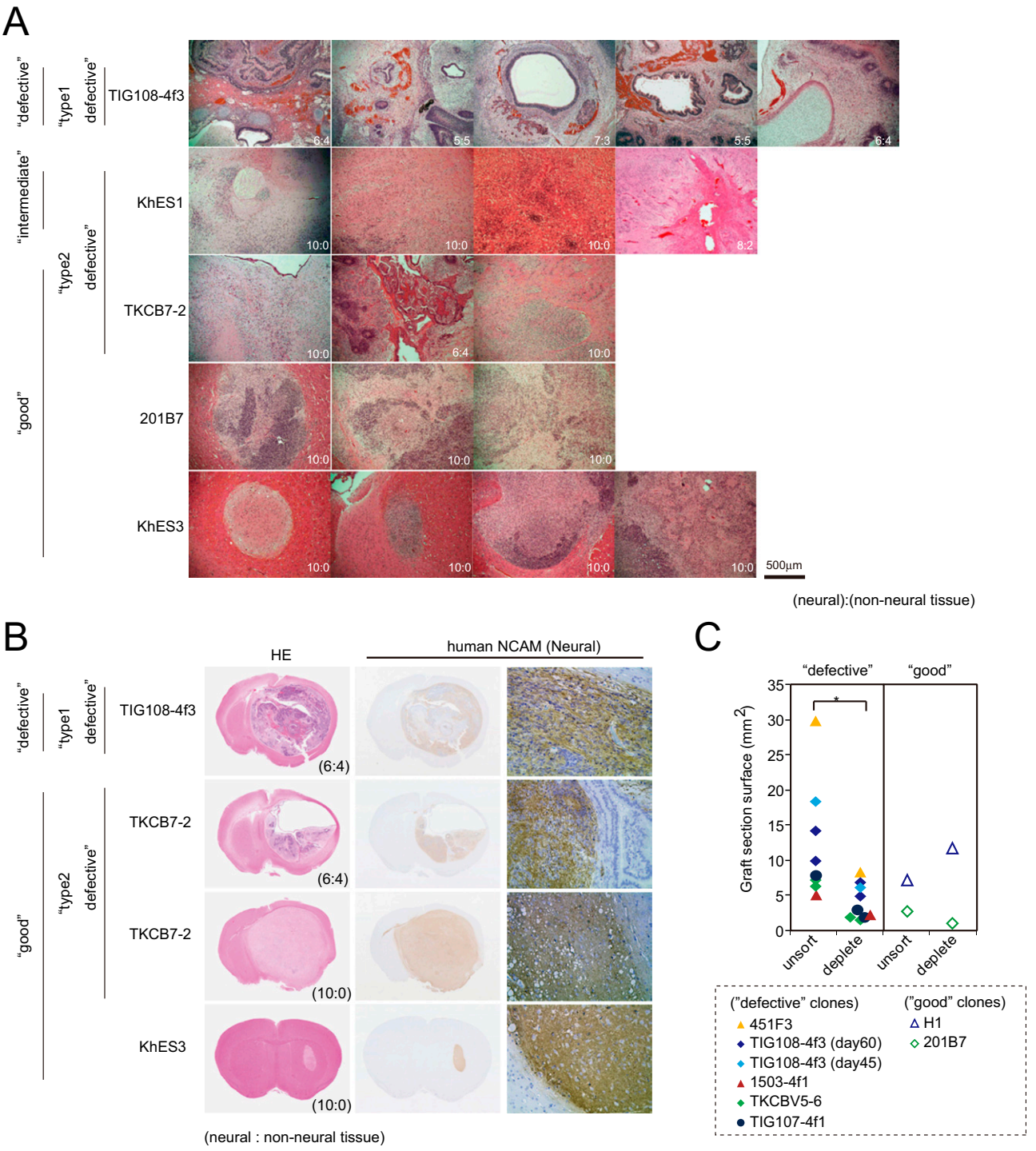

Fig. S4. The histology of grafts derived from hESCs and hiPSCs. (*A*) Transplanted mouse brains were fixed with formaldehyde, embedded in paraffin, sectioned, and stained with H&E. The ratios indicate neural cells and nonneural cells as determined by a microscopic observation. (Scale bar, 500 μm.) (*B*) H&E-stained sections (*Left*) and human NCAM-stained sections (*Center* and *Right*) of mouse brains transplanted with 29-d differentiated hESC/iPSCs. (*C*) The maximum surface size of graft sections 45 or 60 d after transplantation. Transplanted cells were prepared with or without depletion of TRA-1-60^{+} cells 22 d after differentiation (*t test, paired $P < 0.05$).

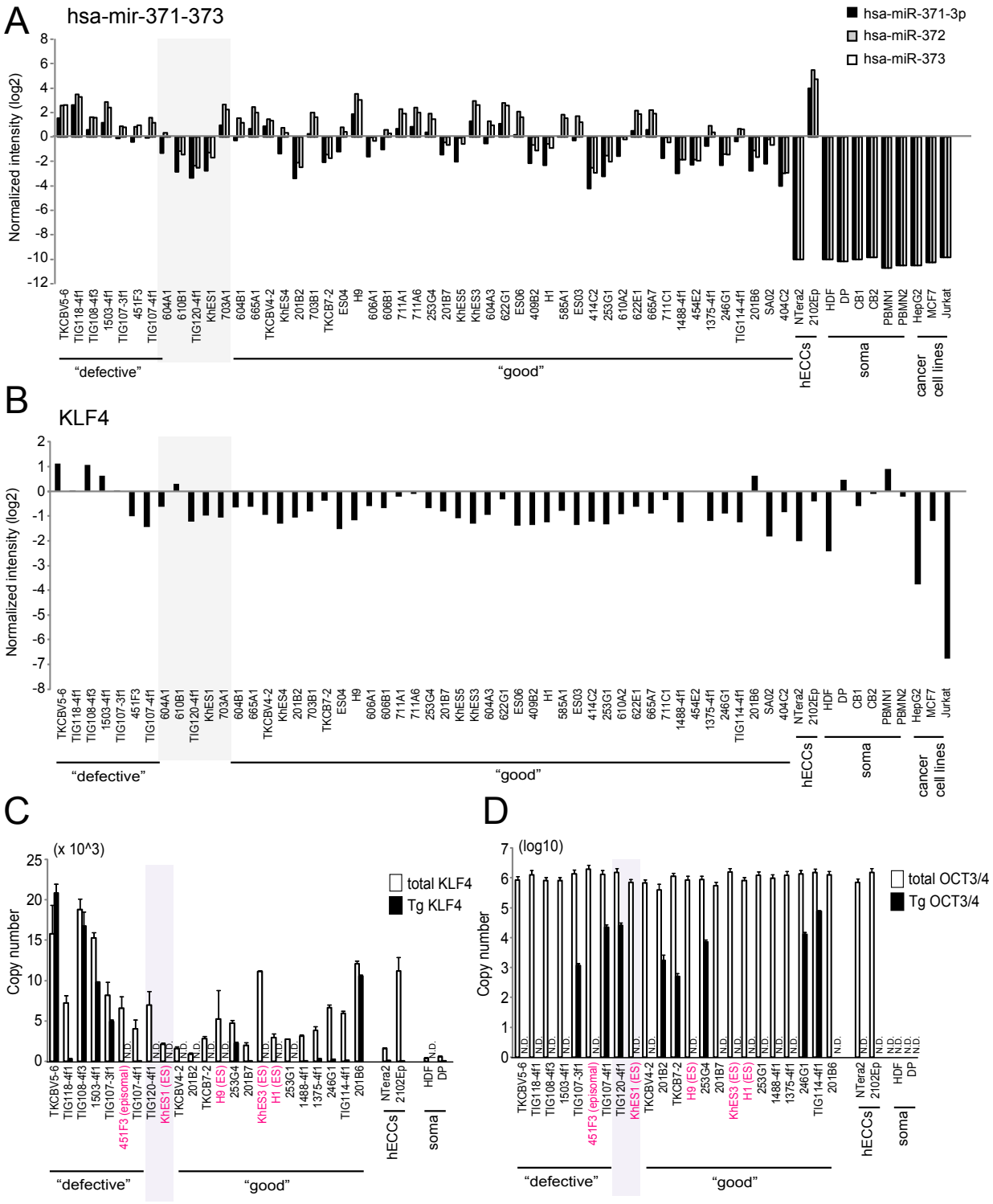

Fig. S5. High expression levels of the hsa-mir-371–373 cluster, KLF4, and transgenes are not absolute markers for type-1 defectiveness. The expression levels of hsa-mir-371–373 (*A*) and KLF4 (*B*) were examined by a microarray analysis in seven defective clones, five intermediate clones, 38 good clones, two hECC lines (NTera2 and 2102Ep), six somatic cell lines (HDF, DP, CB1, CB2, PBMN1, and PBMN2), and three cancer cell lines (HepG2, MCF7, and Jurkat). The total and retroviral transgene expression levels of KLF4 (*C*) and OCT3/4 (*D*) were measured by qPCR in 18 hiPS clones established by a retroviral method, the defective clone 451F3, which was generated using an episomal plasmid vector, four hESCs, two hECCs, and two somatic cell lines (HDF and DP).

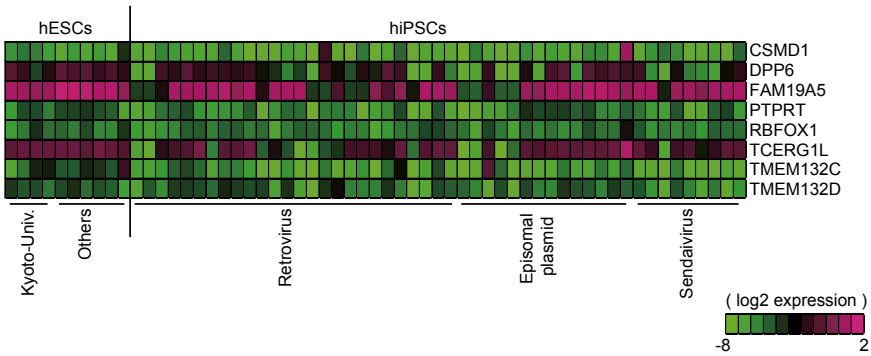

Fig. S6. Previously reported iPS-specific aberrantly methylated genes' expressions. A heat map for 10 hESCs and 49 hiPSCs examined in our laboratory based on the gene expression levels of reported aberrantly methylated genes that can distinguish hiPSCs and hESCs. Of nine previously reported genes, probes for C22ORF34 were not detected in all of the samples in our microarray platform, so we only evaluated the other eight genes (eight probes, detected in at least two clones of our samples).

Other Supporting Information Files

Table S1 (DOCX)
Table S2 (DOCX)
Dataset S1 (XLSX)