Supporting Information

Ohnuki et al. 10.1073/pnas.1413299111

SI Materials and Methods

Cell Culture. Human dermal fibroblasts (HDFs) obtained from the Japanese Collection of Research Bioresources were maintained in DMEM (Nacalai tesque) containing 10% FBS (Japan Bioserum) and 0.5% penicillin and streptomycin (Pen/Strep, Life Technologies). PLAT-E (1) and PLAT-GP cells were maintained in the same medium. Embryonic stem cell (ESC) clones obtained from WiCELL (H1 and H9) (2) and Kyoto University (KhES1 and KhES3) (3) were maintained in Primate ESC medium (ReproCELL) supplemented with 4 ng/mL recombinant human basic fibroblast growth factor (bFGF, Wako) and 0.5% Pen/Strep on mitomycin C (MMC)-treated SNL feeder cells, or in mTeSR1 (Veritas) on Matrigel-coated plates (growth factor reduced; BD Biosciences) (4).

Reprogramming. Reprogramming experiments were conducted as described previously (4, 5). To generate retrovirus, we introduced retroviral plasmids into PLAT-E or PLAT-GP cells using the FuGENE 6 transfection reagent (Promega) (1). On the following day, the medium was replaced with an equal volume of fresh medium and cells were incubated overnight. The virus-containing supernatant was collected, filtered through a 0.45-μm pore-size cellulose acetate filter (Whatman) to remove cell debris, and supplemented with 4 μg/mL Polybrene (Nacalai tesque). Appropriate combinations of viruses were mixed and exposed to HDFs expressing the mouse *Slc7a1* gene overnight. This point was designated as day 0. To collect the samples on day 3 posttransduction, we introduced pMXs-internal ribosome entry site-EGFP (IG) encoding SOX2 instead of pMXs-SOX2, along with OKM (OCT3/4, KLF4, and c-MYC), and sorted EGFP (+) cells by flow cytometry. TRA-1-60 (+) intermediate reprogrammed cells were collected by magnetic activated cell sorting as described previously. For reprogramming experiments with shRNA transduction, 1 μg/mL puromycin was added to the medium from days 3–7 to eliminate untransduced cells. Cells were harvested on day 7 and reseeded onto MMC-inactivated SNL feeders to generate induced pluripotent stem cells (iPSCs). On the following day, the medium was replaced with Primate ESC medium supplemented with 4 ng/mL bFGF, and the medium was then changed every other day. The number of iPSC colonies was counted on day 25. We distinguished bona fide iPSC colonies from non-iPSC colonies based on their morphological differences (4).

Neural Differentiation and Flow Cytometry. The single-cell suspensions of pluripotent stem cells were prepared using CTK solution and AccuMAX (Innovative Cell Technology), and then transferred at 9×10^3 cells per well of Lipidure-coated low-binding 96-well plates (NOF Corporation) in 8GMK medium consisting of Glasgow modified eagle medium (Life Technologies), 8% knockout serum replacement (Life Technologies), 1% nonessential amino acids (NEAA; Life Technologies), 1% sodium pyruvate (Sigma), and 100 nM 2-mercaptoethanol (2-ME, Life Technologies) supplemented with 500 nM A-83-01 (Stemgent) and 100 nM LDN193189 (Stemgent). The medium was carefully half-changed on days 5, 8, and 11. On day 14, the aggregates were collected to a 1.5-mL tube and pelletized by centrifugation at $200 \times g$ for 5 min. The pellet was washed with PBS and then dispersed to single cells using AccuMAX, and then cells were used for immunostaining. We used propidium iodide (Life Technologies) or DAPI (Life Technologies) to distinguish dead cells. Samples were analyzed using a FACS Aria II (Becton Dickinson).

Primitive Streak-Like Mesendoderm Differentiation. Primitive streak-like mesendoderm differentiation was performed as described previously (6, 7). In brief, single-cell suspensions of human iPSCs were plated onto fibronectin-coated plates (BD Biosciences) in DMEM/F12 (Life Technologies) supplemented with 1% Insulin-Transferrin-Selenite (Life Technologies), 1% Glutamax (Life Technologies), 1% NEAA, 2% B27 (Life Technologies), 100 nM 2-ME, and 0.5% Pen/Strep. We added 3 μM CHIR99021 (Stemgent) and 50 ng/mL Activin A (Peprotech) on day 1, 3 μM CHIR99021, 25 ng/mL Activin A and 20 ng/mL bFGF on day 2 and 3 μM CHIR99021, 10 ng/mL Activin A, 20 ng/mL bFGF and 40 ng/mL BMP4 (R&D Systems) on day 3.

Endoderm Differentiation. Endoderm differentiation was performed as described previously, with slight modification (6, 8). The single-cell suspensions of human pluripotent stem cells were plated onto Matrigel-coated plates in RPMI1640 (Life Technologies) containing 2% B27, 100 ng/mL Activin A, 3 μM CHIR99021, and 0.5% Pen/Strep. We added 0.5 mM sodium butyrate (Sigma) on days 1–3, and then carried out sodium butyrate-free culture until day 7.

Mesoderm Differentiation. The mesodermal differentiation was performed as described previously, with slight modification (6, 9). The single-cell suspensions of human pluripotent stem cells were plated onto collagen I-coated plates (BD Biosciences) in DMEM/F12 containing 2% B27, 100 ng/mL Activin A, 3 μM CHIR99021 and 0.5% Pen/Strep. Forty-eight hours later, the medium was replaced with DMEM/F12 supplemented with 2% B27, 25 ng/mL BMP4, and 0.5% Pen/Strep. The medium was changed every other day until day 8.

Antibodies. The following antibodies were used in the study: Alexa 488-conjugated TRA-1-60 (1:20, 560173, BD Biosciences), anti-polysialylated neuronal cell-adhesion molecule (PSA-NCAM) antibody (1:50; MAB5324, Millipore), APC-labeled anti–C-X-C chemokine receptor type 4 (CXCR4) (1:5; FAB170A, R&D Systems), phycoerythrin (PE)-labeled antiplatelet-derived growth factor receptor-α (PDGFRA) (1:5; 556002, BD Pharmingen), APC-conjugated anti-BRACHYURY (1:5; IC20851A, R&D Systems), Alexa 647-conjugated anti-mouse IgM antibody (1:500; A-21238, Life Technologies), PE-conjugated TRA-1–60 (1:5; 560193, BD Pharmingen), and anti-PE microbeads (130-048-801, Miltenyi Biotec). For Western blot analyses, we used anti-OCT3/4 (1:600; sc-5279, Santa Cruz Biotechnology), anti-SOX2 (1:1,000; ab97959, Abcam), anti-KLF4 (1:500; AF3640, R&D Systems), anti–c-MYC (1:200; sc-42, Santa Cruz), anti–β-actin (1: 5,000; A5441, Sigma), anti-mouse IgG HRP (1:3,000; 7076S, Cell Signaling Technology), anti-rabbit IgG-HRP (1:2,000; 7074S, Cell Signaling), and anti-goat IgG-HRP (1:2,000; sc-2056, Santa Cruz). For ChIP analyses, we used anti-Flag M2 (F1804, Sigma), antitrimethylated lysine 4 of histone H3 antibody (MABI0304, Wako), anti-p300 (sc-585, Santa Cruz), anti-KAP1 (ab10483, Abcam), anti-SET domain bifurcated 1 (sc-66884, Santa Cruz), and antihistone H3Ac (pan acethyl) (39139, Active Motif).

Quantitative RT-PCR. Total RNA was isolated using the Qiazol reagent (Qiagen) and a miRNeasy mini kit (Qiagen). CDNA was generated with a ReverTra Ace-α kit (Toyobo) and oligo dT_{20} primer. For quantitative RT-PCR (qRT-PCR), reaction mixtures were set up in triplicate using SYBR premix Ex-Taq II (Takara) and run using a StepOne instrument (Applied Biosciences). The levels of mRNA were normalized to *G3PDH* expression, and

then relative expression was calculated as the fold-change from the control. Primer sequences for each gene are shown in Dataset S3.

Gene-Expression Analysis by Microarray. Total RNA was purified as described above and the quality was evaluated using a 2100 Bioanalyzer (Agilent Technologies). Total RNA (50 ng) was labeled with cyanine 3-CTP and used for hybridization with a SurePrint G3 Human GE 8 × 60 K array (G4851A, Agilent Technologies) with a one-color protocol. Hybridized arrays were scanned with a Microarray Scanner System (G2565BA, Agilent Technologies), and extracted signals were analyzed using GeneSpring v12.6 software (Agilent Technologies). Gene-expression values were normalized by 75th percentile shifts. Differentially expressed genes between normal and differentiation-defective (DD)-iPSCs were extracted by comparison using *t* tests with Benjamini and Hochberg corrections [FC > 2.0, false-discovery rate (FDR) < 0.05].

CpG Methylation Analysis by Microarray. Genomic DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen). Purified genomic DNA (30 ng) was used for bisulfite CT conversion with an EZ DNA methylation kit (Zymo Research). Bisulfite-treated DNA was whole-genome amplified and hybridized to HumanMethylation450 BeadChip (Illumina), followed by labeling with fluorescence by a single-base extension reaction. Fluorescent signals were read on an Illumina BeadStation GX scanner. Normalization and calculation of β-values, the ratio of methylated to unmethylated cytosine, was performed using GenomeStudio (Illumina). A β-value of 0 indicates nonmethylated cytosine, and that of 1 indicates fully methylated cytosine. The relative distance of the Infinium probe site from a region-of-interest, such as long-terminal repeat 7 (LTR7) and long-interspersed element-1 (LINE-1), was defined as the distance between the center of the region of interest and the probe position.

Pyrosequencing. Genomic DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen). Purified genomic DNA (500 ng) was used for bisulfite CT conversion with an EZ DNA methylation kit (Zymo Research). PCR was performed in a 25-μL reaction mixture containing 25 ng of bisulfite-converted DNA, Pyromark PCR Master Mix (Qiagen), Coral Load Concentrate (Qiagen), and 0.3 μM forward and 5′ biotinylated reverse primers. PCR conditions were 45 cycles at 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s The PCR product was bound to streptavidin Sepharose beads (Amersham Biosciences) and then purified, washed, denatured, and washed again. Then, 0.3 μmol/L pyrosequencing primer was annealed to the purified PCR product. Pyrosequencing reactions were performed in the PSQ HS 96 Pyrosequencing System. The degree of methylation was expressed as the percentage of methylated cytosines divided by the sum of methylated and unmethylated cytosines. To validate the PCR pyrosequencing assay, each CpG dinucleotide position was assayed in triplicate and averages were used in the final analysis. The primer sequences are shown in Dataset S3.

Plasmid Construction. To generate a doxycycline (Dox)-inducible KLF4 expression vector, the KLF4 ORF was cloned into PB/TA/ERN (10) using the Gateway LR reaction system (Life Technologies). Short hairpin RNA sequences were inserted into the AgeI/EcoRI site of pMKO.1-puro (11) (Addgene). To generate a shRNA expression vector for ESCs/iPSCs, a BglII/EcoRI fragment of each shRNA construct was transferred into the BamHI/MfeI site of PB/CAG-GIP/BP. All inserted fragments were verified by sequencing. Target sequences of shRNAs are as follows; LTR7-1 (CCT GTT TGG TGG TCT CTT CAC), LTR7-2 (TGT TTG GTG GTC TCT TCA CAC), long intergenic non-protein coding RNA, regulator of reprogramming (lincRNA-RoR) (AAG CCT GAG AGT TGG CAT GAA), and KLF4 (GAT CAA GCA GGA GGC GGTC TC).

ChIP. For the ChIP assay of exogenous reprogramming factors, such as OCT3/4, SOX2, and KLF4, 3× Flag-tagged factors or nontagged factors were transduced into HDFs using a PLAT-E–produced retroviral system. These cells were harvested at 3 d posttransduction. Cells were fixed with 1% formaldehyde and quenched with 125 mM glycine. Fixed cells were sequentially treated with LB1 [50 mM Hepes-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% (wt/vol) glycerol, 0.5% Nonidet-P 40, and 0.25% TritonX-100], LB2 (10 mM Tris·HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, and 0.5 mM EGTA), LB3 (10 mM Tris·HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, and 0.1% sodium deoxycholate) to obtain nuclear extracts. Chromatin samples were sheared by treatment with a Misonix Astrason S-3000 sonicator (ON, 30 s; OFF, 1 min; Power, 7.0; 15 cycles). A one-tenth volume of 10% (wt/vol) TritonX-100 solution was added to the cell lysates to obtain cleared supernatants after centrifugation. To reduce the nonspecific background, the chromatin samples were incubated with normal IgG (Santa Cruz) linked to Dynabeads (Life Technologies) for 30 min at 4 °C. After removal of the beads, the cleared lysates were incubated with a Dynabead-conjugated antibody overnight at 4 °C with constant rotation. After incubation, the beads were washed twice each with Wash buffer 1 (20 mM Tris·HCl, 8.0, 150 mM NaCl, 2 mM EDTA, 1% TritonX-100, and 0.1% SDS), wash buffer 2 (20 mM Tris·HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% TritonX-100, and 0.1% SDS), wash buffer 3 (10 mM Tris·HCl, pH 8.0, 250 mM lithium chloride, 1 mM EDTA, 1% Nonidet-P 40, and 1% sodium deoxycholate), and TE buffer (10 mM Tris·HCl, pH 8.0 and 1 mM EDTA). The beads were then incubated with elution buffer (25 mM Tris·HCl, pH 7.5, 5 mM EDTA, and 0.5% SDS) supplemented with 0.4 mg/mL protease K (Nacalai tesque) at 42 °C for 1 h and then at 65 °C overnight with constant rotation. Eluates were purified using a Qiaquick PCR purification kit (Qiagen).

Multiplexed ChIP-seq libraries were prepared from 10 ng of immunoprecipitated DNA fragments using NEBNext ChIP-seq Library Prep Master Mix (New England BioLabs). For sequencing using GAIIx (Illumina), cluster generation was performed using TruSeq SR Cluster Kit v2 (Illumina). Each lane of flow cells contained one sample. Sequencing was performed in single-read run mode with a total 76 cycles, including a 75-bp read and one cycle for phasing. For sequencing using HiSeq2000 (Illumina), cluster generation was performed using TruSeq SR Cluster Kit v3 (Illumina) and six indexed libraries were loaded into each lane of flow cells. Sequencing was performed in multiplexed single-read run mode with total 86 cycles, including 6 cycles for reading the index sequence. Peak call was executed using a model-based analysis of the ChIP-seq algorithm (MACS) v1.4.0rc2 with default parameters.

RNA-Seq Library Preparation and Sequencing. DNase-treated RNA (100 ng) was used to prepare individually indexed and strand-specific RNA-seq libraries using a TruSeq Stranded Total RNA sample prep kit (Illumina). Briefly, ribosomal RNA was depleted using a Ribo-Zero Gold kit (AR Brown), followed by RNA fragmentation, cDNA synthesis, end repair, A-base addition, and ligation with Illumina-indexed adaptors. Cluster generation was performed using a TruSeq PE cluster Kit v3 on an Illumina cBot instrument. Six indexed libraries were loaded into each lane of flow cells. Sequencing was performed on a Hiseq2000 in multiplexed paired-read run mode with a total of 208 cycles, including 6 cycles for reading the index sequence.

Data Analysis of Deep Sequencing. Conversion of Bcl into Fastq format was achieved using the configureBclToFastq.pl program, a part of CASAVA 1.8.2 (Illumina). Sequence reads were then aligned to the human genome (GRCh37/hg19) using the paired-end mode of the BWA v0.5.9rc1 and SAMtools v0.1.17 (ChIP-seq), or TopHat2 software (RNA-seq). Aligned sequence tags were visualized using GenomeJack v2.1 (Mitsubishi Space Software).

Quantitative Expression Analysis of Repeat Sequences. RepeatMasker open-3.3.0 (www.repeatmasker.org), a database of human repetitive sequences, was applied to reference sequences of LTR7 and LINE-1. The LTR7 sequence was from LTR/endogenous retrovirus (ERV)-1 in the database, including LTR7, LTR7A, LTR7B, LTR7C, and LTR7Y, and 3771 LTR7 regions were found. The number of LINE-1 regions (LINE/L1 in the database) was 979,568. For RNA-seq, the number of sequence reads aligned with the LTR7-coding region by TopHat2 was counted using the Partek Genomics Suite. Sample-to-sample variation was normalized as follows: normalized number of sequence tags on LTR7 = number of reads aligned with the LTR7 region/total number of hg19-mappable sequence tags.

Single-Cell Expression Analysis. Single-cell suspensions prepared by treatment with 0.25% trypsin/1 mM EDTA were incubated with PE-conjugated TRA-1-60 and DAPI. TRA-1-60 (+)/DAPI (−) cells were directly sorted into PCR master mix using the FACS Aria II. The template for single-cell PCR was prepared using CellsDirect reagent (Life Technologies). Preamplification of the target sequences was performed for 22 cycles. Subsequent PCR and data collection were conducted with a Biomark system (Fluidigm). Each sample was validated based on expression of *G3PDH* or *ACTB*, with a Ct value < 15. Ct values > 26 were taken to indicate no expression and were unified as 100 for calculation of ΔCt. We used following Taqman assays; abhydrolase domain containing 12B (*ABHD12B*) (Hs00997975_g1), HERV-H LTR-associating 1 (*HHLA1*) (Hs00903176_g1) and chromosome 4 open reading frame 51 (*C4ORF51*) (Hs03037752_m1).

Southern Blotting. Genomic DNA (3–5 μg) was digested with restriction enzymes overnight. Digested DNA were separated on a 0.8% agarose gel and transferred to a nylon membrane (Amersham Biosciences). The membrane was incubated with digoxigenin (DIG)-labeled DNA probe in DIG Easy Hyb buffer (Roche) at 42 °C overnight with constant rotation. After washing, alkaline phosphatase-conjugated anti-DIG antibody (1:10,000; Roche) was added to the membrane. Signals were produced with CDP-star reagent (Roche) and detected using a LAS3000 imaging system (Fuji Film). Primer sequences for generation of probes are shown in Dataset S3.

1. Morita S, Kojima T, Kitamura T (2000) Plat-E: An efficient and stable system for transient packaging of retroviruses. *Gene Ther* 7(12):1063–1066.
2. Thomson JA, et al. (1998) Embryonic stem cell lines derived from human blastocysts. *Science* 282(5391):1145–1147.
3. Suemori H, et al. (2006) Efficient establishment of human embryonic stem cell lines and long-term maintenance with stable karyotype by enzymatic bulk passage. *Biochem Biophys Res Commun* 345(3):926–932.
4. Takahashi K, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131(5):861–872.
5. Tanabe K, Nakamura M, Narita M, Takahashi K, Yamanaka S (2013) Maturation, not initiation, is the major roadblock during reprogramming toward pluripotency from human fibroblasts. *Proc Natl Acad Sci USA* 110(30):12172–12179.
6. Takahashi K, et al. (2014) Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. *Nat Commun* 5:3678.
7. Oldershaw RA, et al. (2010) Directed differentiation of human embryonic stem cells toward chondrocytes. *Nat Biotechnol* 28(11):1187–1194.
8. 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.
9. Mae S, et al. (2013) Monitoring and robust induction of nephrogenic intermediate mesoderm from human pluripotent stem cells. *Nat Commun* 4:1367.
10. Woltjen K, et al. (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature* 458(7239):766–770.
11. Masutomi K, et al. (2003) Telomerase maintains telomere structure in normal human cells. *Cell* 114(2):241–253.

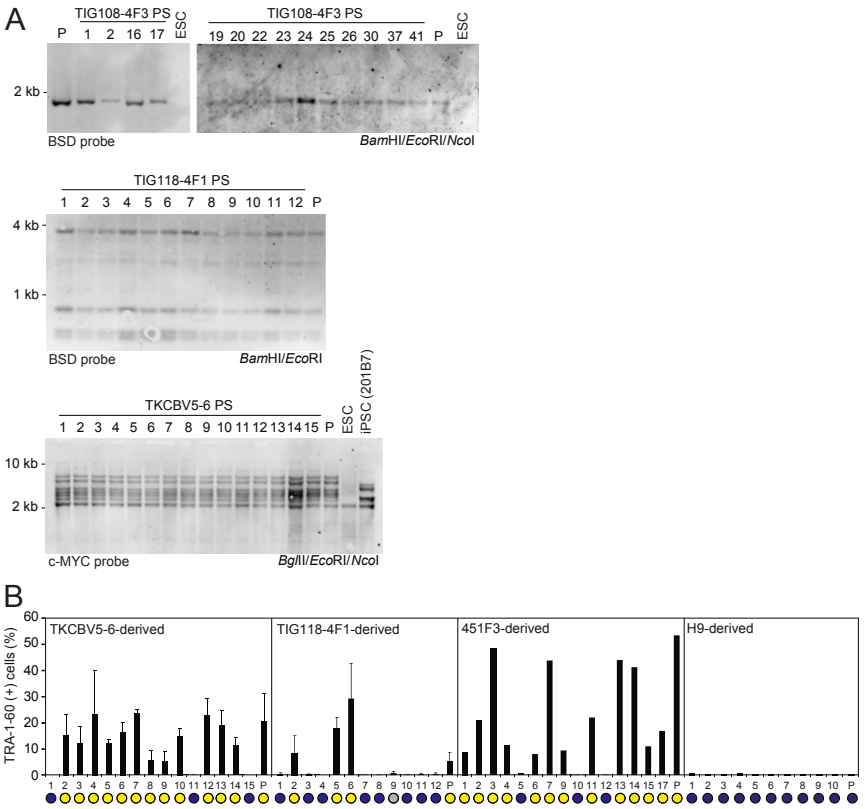

Fig. S1. Subcloning of DD-iPSCs. (*A*) Monoclonal origin of iPSC subclones. The genomic DNA of the subclones derived from TKCBV5-6, TIG108-4F3, and TIG118-4F1 by Southern blotting. Restriction enzymes and probes used for each clone are shown. For subclones derived from the 451F3 integration-free iPSC clone, short tandem-repeat analyses were performed for 16 loci (D3S1358, TH01, D21S11, D18S51, Penta_E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta_D, AMEL, vWA, D8S1179, TPOX, and FGA). (*B*) Percentages of TRA-1-60 (+) cells 14 d after neural induction of each primary subclone and parental cells (P) analyzed by flow cytometry. Blue and yellow circles indicate normal and DD-iPSC subclone, respectively. $n = 2–3$. Error bars are SDs.

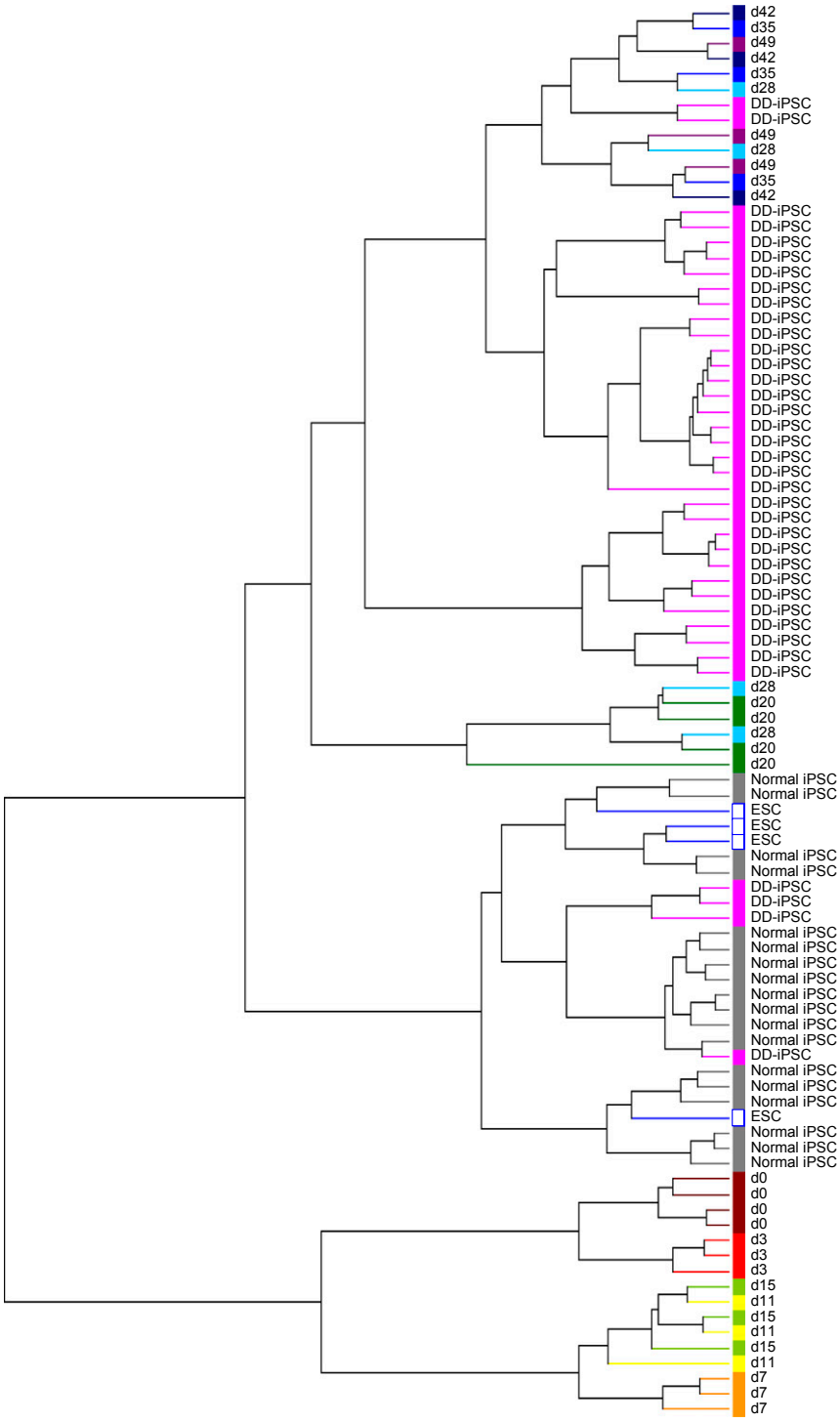

Fig. S2. Reproducibility of DD-iPSC related gene expression. Shown are the hierarchical clustering of 144 DD-iPSC marker expression in parental HDFs [day (d) 0], intermediate reprogrammed cells at the indicated time points (days 3–49), ESCs, normal iPSCs, and DD-iPSCs.

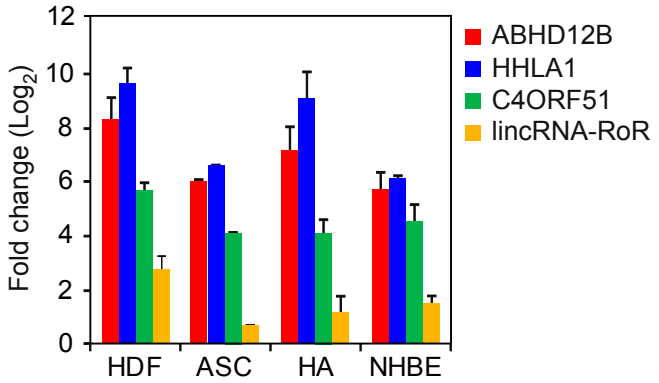

Fig. S3. The expression of DD-iPSC markers in TRA-1-60 (+) cells from various origins. Fold-increases of *ABHD12B*, *HHLA1*, *C4ORF51*, and *lincRNA-RoR* expression in TRA-1–60 (+) intermediate reprogrammed cells on day 28 compared with original somatic cell lines, determined by microarray. ASC, adipose tissue-derived mesenchymal stem cells; HA, human astrocytes; NHBE, normal human bronchial epithelium. $n = 3$. Error bars are SDs.

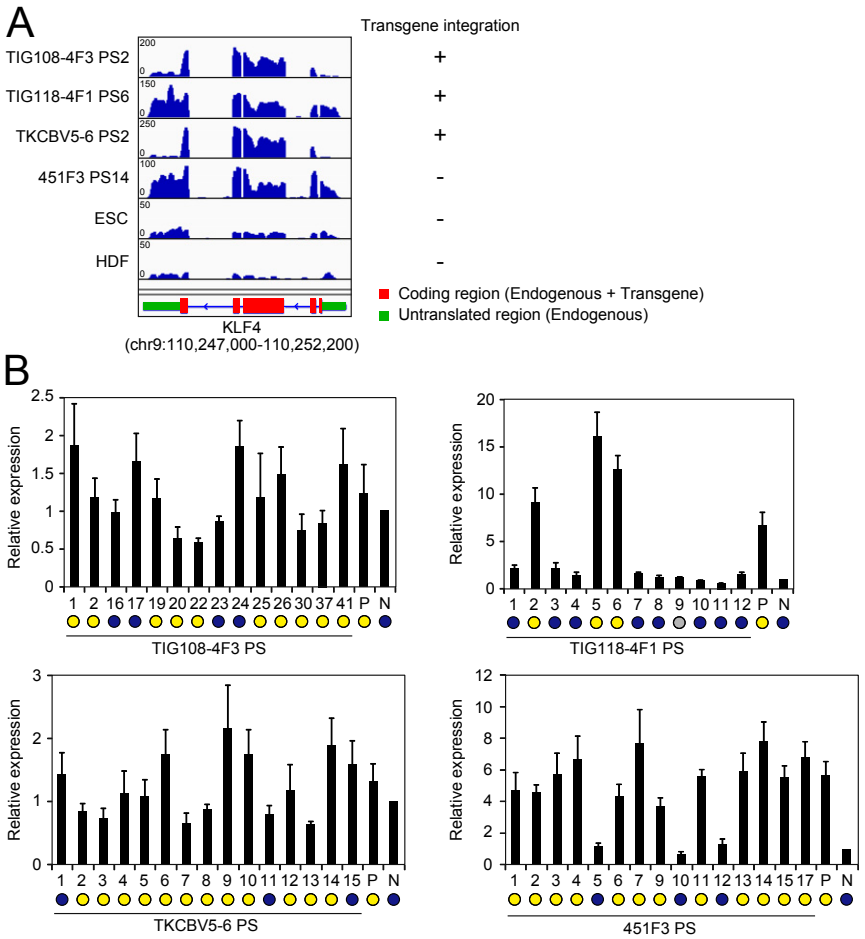

Fig. S4. KLF4 transcripts in DD-iPSCs. (*A*) The histogram shows the representative transcription patterns of *KLF4* in DD-iPSC primary subclones (PS), ESCs, and HDFs analyzed by RNA-seq. Red and green box indicate protein coding region (common in transgene and endogenous locus) and untranslated region (only in endogenous locus), respectively. (*B*) The relative expression of endogenous *KLF4* in DD-iPSCs and their primary subclones (PS) compared with those in normal iPSCs (N) analyzed qRT-PCR. Blue and yellow circles indicate normal and DD-iPSC subclones/parents, respectively. $n = 3$. Error bars are SDs.

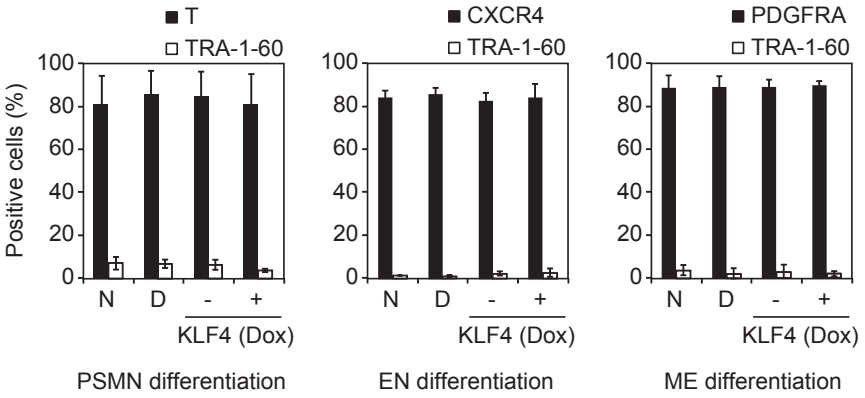

Fig. S5. Differentiation potential of DD-iPSCs and KLF4-overexpressing iPSCs. Normal iPSCs (N), DD-iPSCs (D), and normal iPSCs transduced with Dox-inducible KLF4 were differentiated into primitive streak-like mesendoderm (PSMN), endoderm (EN), or mesoderm (ME) in the conditions with (+) or without (–) Dox. Closed bars in each graph represent the differentiation efficiencies into PSMN, EN, and ME evaluated by immunostaining with T, CXCR4, and PDGFRA, respectively. Open bars indicate the percentages of residual undifferentiated cells labeled by TRA-1-60. $n = 3$. Error bars are SDs.

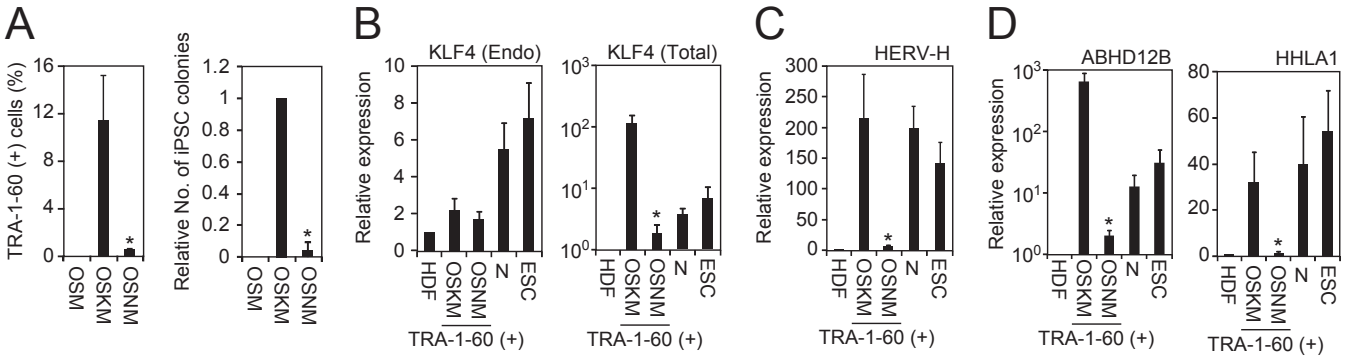

Fig. S6. The effect of exogenous NANOG homeobox (NANOG) on human ERV-H (HERV-H) activity. (*A*) Low reprogramming activity of NANOG. Shown are the proportion of TRA-1-60 (+) cells on day 7 posttransduction of indicated factors (*Left*) and relative numbers of iPSC colonies counted on day 28 compared with those induced by OCT3/4, SOX2, KLF4, and c-MYC (OSKM) (*Right*). Error bars are SD. $n = 3$. *$P < 0.05$ vs. OSKM was calculated by *t* test. (*B*) NANOG does not enhance KLF4 expression. Shown are the relative expression of endogenous (Endo) and total *KLF4* in HDFs, TRA-1-60 (+) cells induced by OSKM or OSNM (OSM with NANOG instead of KLF4) on day 7, normal iPSCs (N), and ESCs compared with those in HDFs, revealed by qRT-PCR. Each value was normalized to that of *G3PDH*. $n = 3$. Error bars are SD. *$P < 0.05$ vs. OSKM was calculated by *t* test. (*C*) Low activity of NANOG for HERV-H induction. Shown are the relative expression of HERV-H in HDFs, TRA-1-60 (+) cells induced by OSKM or OSNM on day 7, normal iPSCs (N), and ESCs compared with those in HDFs, revealed by qRT-PCR. Each value was normalized to that of *G3PDH*, and then the value of Mock was adjusted to be 1. $n = 3$. Error bars are SD. *$P < 0.05$ vs. OSKM was calculated by *t* test. (*D*) NANOG slightly enhanced the expression of LTR7-related genes. Shown are the relative expressions of LTR7-related genes *ABHD12B* and *HHLA1* in HDFs, TRA-1-60 (+) cells induced by OSKM or OSNM on day 7, normal iPSCs (N), and ESCs compared with those in HDFs, revealed by microarray. $n = 3$. Error bars are SD. *FDR < 0.05 vs. OSKM was calculated by *t* test.

Dataset S1. Short tandem-repeat analyses of 451F3 subclones

Dataset S1

Shown are the results of short tandem-repeat analyses for 16 loci (D3S1358, TH01, D21S11, D18S51, Penta_E, D5S818, D13S317, D7S820, D16S539, CSF1PO, Penta_D, AMEL, vWA, D8S1179, TPOX, and FGA) in 451F3 integration-free parental iPSCs (P) and its primary subclones (PS).

Dataset S2. DD-iPSC markers

Dataset S2

Microarray probes, genomic location, and gene symbols of DD-iPSC marker genes, *P* values and fold-change (FC) for comparison between normal and DD-iPSC subclones.

Dataset S3. Primer sequences used in the study

Dataset S3

The purposes, targets and sequences of primers used in this study are listed.