

Generation of human induced pluripotent stem cells from dermal fibroblasts

W. E. Lowry^{††§¶}, L. Richter^{*}, R. Yachechko^{||}, A. D. Pyle^{††§**}, J. Tchieu^{†||}, R. Sridharan^{||}, A. T. Clark^{**§}, and K. Plath^{††§¶||}

Departments of ^{*}Molecular Cell and Developmental Biology, ^{||}Biological Chemistry, and ^{**}Microbiology, Immunology, and Molecular Genetics, [†]Molecular Biology Institute, [‡]Jonsson Comprehensive Cancer Center, and [§]Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, 615 Charles E. Young Drive South, University of California, Los Angeles, CA 90095

Communicated by Owen N. Witte, University of California, Los Angeles, CA, December 19, 2007 (received for review December 5, 2007)

The generation of patient-specific pluripotent stem cells has the potential to accelerate the implementation of stem cells for clinical treatment of degenerative diseases. Technologies including somatic cell nuclear transfer and cell fusion might generate such cells but are hindered by issues that might prevent them from being used clinically. Here, we describe methods to use dermal fibroblasts easily obtained from an individual human to generate human induced pluripotent stem (iPS) cells by ectopic expression of the defined transcription factors *KLF4*, *OCT4*, *SOX2*, and *C-MYC*. The resultant cell lines are morphologically indistinguishable from human embryonic stem cells (HESC) generated from the inner cell mass of a human preimplantation embryo. Consistent with these observations, human iPS cells share a nearly identical gene-expression profile with two established HESC lines. Importantly, DNA fingerprinting indicates that the human iPS cells were derived from the donor material and are not a result of contamination. Karyotypic analyses demonstrate that reprogramming of human cells by defined factors does not induce, or require, chromosomal abnormalities. Finally, we provide evidence that human iPS cells can be induced to differentiate along lineages representative of the three embryonic germ layers indicating the pluripotency of these cells. Our findings are an important step toward manipulating somatic human cells to generate an unlimited supply of patient-specific pluripotent stem cells. In the future, the use of defined factors to change cell fate may be the key to routine nuclear reprogramming of human somatic cells.

reprogramming | stem cell | *OCT4* | *SOX2*

The therapeutic use of stem cells depends on the availability of pluripotent cells that are not limited by technical, ethical, or immunological considerations. Recent work showing that primate ES cells can be derived by somatic cell nuclear transfer (SCNT) from somatic cells opens the door to the possibility that SCNT of human cells will soon allow for the generation of “patient-specific” ES cells (1). An approach toward the same end was recently described, in which murine fibroblasts were reprogrammed by ectopically expressing factors known to be highly expressed in murine ES cells (2). Specifically, transduction of a set of four genes encoding the transcription factors *Oct4*, *Sox2*, *C-Myc*, and *Klf4* globally reset the epigenetic and transcriptional state of fibroblasts into that of pluripotent cells, designated induced pluripotent stem (iPS) cells, that were functionally indistinguishable from murine ES cells (2–5). Application of this approach in human cells would have enormous potential and generate patient-specific pluripotent stem cells to study and potentially ameliorate human disease. Thus, we asked here whether the defined factor approach recently described for murine reprogramming (2) could be applied to induce human fibroblasts to become pluripotent, ES-like cells.

An analysis of published expression datasets indicated that the factors required for murine cell reprogramming are also highly expressed in HESC (data not shown). We therefore reasoned that expression of the same set of four genes with the addition of the *NANOG* transcription factor, which has been shown to facilitate murine cell reprogramming in cell-fusion experiments (6), might

induce reprogramming of human fibroblasts. Although murine reprogrammed cells were first obtained by applying a drug selection scheme for clones that express endogenous ES cell-specific genes (2–5), we and others have recently shown that such a drug-based selection approach is not required to obtain iPS cells (5, 7). Thus, we attempted to isolate human reprogrammed cells simply by overexpressing defined factors in fibroblasts and selection of ES-like colonies appearing in the culture. While our work was completed, two laboratories published elegant work in agreement with our study, demonstrating that human somatic cell reprogramming by overexpression of defined factors is a feasible method to induce pluripotency (8, 9).

Results

cDNAs coding for the human *OCT4* (*POU5F1*, *A isoform*), *SOX2*, *KLF4*, *C-MYC*, and *NANOG* genes were cloned into the pMX retroviral vector, and virus was generated in Phoenix-A cells. Human fibroblasts donated from a single neonatal foreskin (normal human dermal fibroblasts, NHDF1) were infected twice over 3 days at passage 6 with the same volume of each viral supernatant and replated 4 days later onto a feeder layer of irradiated murine fibroblasts. In all experiments, a GFP-expressing pMX virus was added to monitor infection efficiency. Control cells infected with empty pMX virus and the GFP-bearing virus in a 5:1 ratio did not change the morphology of the cells, which continued to grow as a monolayer (Fig. 1A). In contrast, in those fibroblast cultures that were infected with viruses carrying the five defined factors and GFP, colonies formed 14 days after infection (Fig. 1A' and A''). These “early” colonies were highly proliferative and adopted a morphology distinct from fibroblasts. However, further characterization indicated that these clones were consistently infected with only the *OCT4* and *C-MYC* retroviruses and in some cases with the *NANOG* and GFP virus (Fig. 2A', referred to as *OCT4/C-MYC* clones). The *OCT4/C-MYC* colonies did not induce expression of HESC signature genes (Figs. 2B and Table 1) nor HESC-specific cell-surface antigens (data not shown), suggesting that combined *OCT4* and *C-MYC* overexpression in fibroblasts is not sufficient to induce an ES-like expression pattern in fibroblasts. At 21 days after infection, new colonies emerged in the infected fibroblast cultures that adopted a tightly packed morphology and were strongly immunoreactive for the HESC surface antigens TRA-1–81, TRA-1–61, and SSEA-4 (Fig. 1B', B'', and data not shown). An obvious feature of these colonies was their refractive edges and three-

Author contributions: L.R. and R.Y. contributed equally to this work; W.E.L., A.D.P., A.T.C., and K.P. designed research; W.E.L., L.R., R.Y., A.D.P., J.T., R.S., and K.P. performed research; W.E.L., A.D.P., and K.P. contributed new reagents/analytic tools; W.E.L., J.T., R.S., A.T.C., and K.P. analyzed data; and W.E.L. and K.P. wrote the paper.

The authors declare no conflict of interest.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE9865).

[¶]To whom correspondence may be addressed. E-mail: blowry@ucla.edu or klplath@mednet.ucla.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0711983105/DC1.

© 2008 by The National Academy of Sciences of the USA

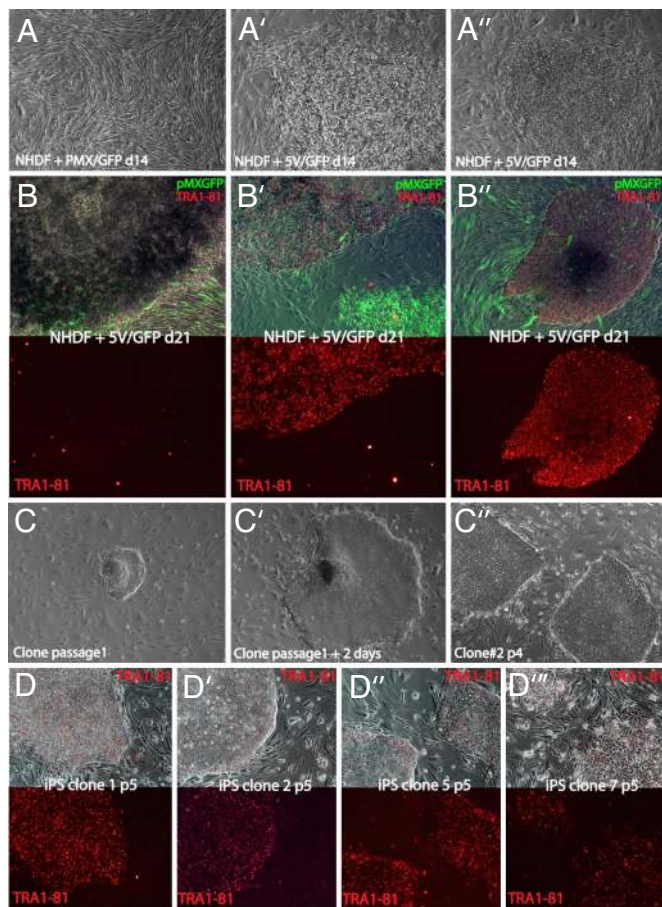


Fig. 1. iPS clones share HESC morphology. (A–A'') Heterogeneous morphology of colonies of NHDF1 infected with empty pMX virus and GFP-containing pMX viruses in a 5:1 ratio (pMX/GFP) or a combination of six pMX viruses each carrying one of the five defined transcription factors or GFP (5V/GFP), at day 14 after infection in phase contrast. (B–B'') Phase-contrast images of different colonies from 5V/GFP transduced cultures merged with live TRA-1-81 staining (red) and GFP fluorescence derived from the pMX-GFP virus (green) (Upper) and the TRA-1-81 channel separately (Lower) from cultures transduced with 5V/GFP. Note that only a minor proportion of colonies are TRA-1-81-positive as seen in B and B'. The staining in TRA-1-81-positive colonies was indistinguishable from that obtained with HESC (data not shown). (C–C'') Phase-contrast images of iPS clones at different passages. (D–D'') "Live" Tra-1-81 staining and merge with phase-contrast appearance of indicated iPS clones at passage 5.

dimensional growth highly reminiscent of HESC colonies. Thus, the HESC-like morphology in combination with HESC surface antigen stainings suggested that these "late" appearing colonies could have been reprogrammed to an ES-like state.

Staining unfixed plates of colonies after 28 days for the TRA-1-81 antigen proved to be an invaluable method to distinguish and isolate these putative accurately reprogrammed colonies. Colonies that stained homogeneously positive for TRA-1-81 were picked from the plate and passaged. Upon replating, these colonies immediately appeared almost identical to HSF1 and H9, two previously established HESC lines that are maintained locally (Fig. 1 C–D'' and data not shown). All analyzed TRA-1-81-positive clones were found to be infected with the viruses bearing *SOX2*, *C-MYC*, *OCT4*, and *KLF4* (Fig. 2A). Similar to GFP, the integration of the *NANOG*-encoding virus was variable between clones, suggesting that *NANOG* is dispensable for the generation of TRA-1-81-positive colonies (Fig. 2A). The clones maintained their morphology and TRA-1-81 expression through five passages (Fig. 1 C–D''') and continue to after at least 4 months with passing

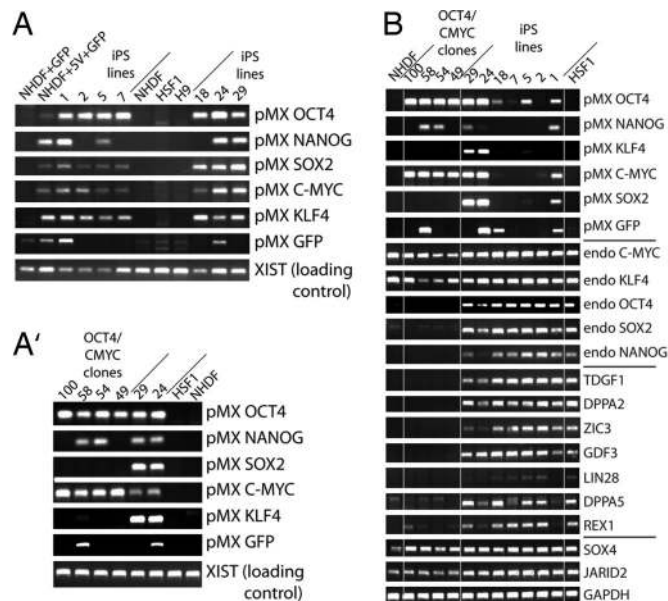


Fig. 2. iPS clones express key HESC markers. (A and A') PCR for retroviral integration events on genomic DNA derived from iPS and "early" *OCT4/C-MYC* clones, control NHDF1, NHDF1 cells infected with control (GFP) or defined factor viruses (5V + GFP), and HSF1 or H9 HESC, with primers that specifically recognize each of the integrated viruses. Loading control: PCR for a genomic region on the X chromosome within the *XIST* locus. iPS clones 24 and 29 are included in A' as a positive control for the PCR conditions. (B) RT-PCR for pMX retroviral transcription and expression of endogenous counterparts of the defined factors as well as additional HESC-specific genes (*TDGF1* through *REX1*) in iPS clones, NHDF1 and the HSF1 HESC, and in *OCT4/CMYC* clones. Note that iPS24 and 29 and the *OCT4/CMYC* clones, respectively, largely failed to suppress expression from the viruses they received.

twice a week. The colonies were cultured in HESC media (with knockout serum replacer and basic FGF) on irradiated feeders and were propagated with standard protocols using collagenase (10, 11). Because these clones appeared to have been reprogrammed to an HESC-like state based on their surface-antigen expression, they were termed human iPS cells and were subsequently analyzed in more detail to understand whether faithful reprogramming to the ES-cell state had indeed occurred. Although we initially isolated 30 human iPS clones, we focused an in-depth characterization around seven of these clones (clones 1, 2, 5, 7, 18, 24, and 29).

RT-PCR indicated that the iPS clones silenced expression of exogenous factors to different extents, with clones 2, 5, 7, and 18 silencing most, if not all viral vectors (Fig. 2B). Importantly, all analyzed iPS clones induced expression from the endogenous *OCT4*, *SOX2*, and *NANOG* loci, and of additional HESC signature genes (Fig. 2B), further supporting the conclusion that overexpression of *OCT4*, *C-MYC*, *SOX2*, and *KLF4* in fibroblasts induces an ES-like state.

Semiquantitative RT-PCR and microarray expression analysis was used to compare expression of the defined factors from the retroviral promoter and the endogenous promoter combined (total transcript) and separately (viral or endogenous transcripts) between iPS clones, "early" *OCT4/C-MYC* clones, NHDF1 and the HESC line HSF1 [supporting information (SI) Fig. 6 and data not shown]. The nonreprogrammed *OCT4/C-MYC* clones all failed to shut down expression of the exogenous transcription factors from the retroviral promoter and express much more *C-MYC* than what is normally found in HESC, NHDF1, or iPS clones (Fig. 2B, and SI Fig. 6). In contrast, in most iPS clones, the amount of the defined factors that remain expressed from the retroviral construct does not dramatically change the total transcript levels (SI Fig. 6). Nevertheless, there is some variation in the extent to which different iPS

Table 1. Expression of consensus HESC signature genes in human iPS clones

consensus HESC gene	HSF1 vs NHDF				H9 vs NHDF				iPS2 vs NHDF				iPS5 vs NHDF				Clone54 vs NHDF		Clone100 vs NHDF		NHDF+5V vs NHDF		NHDF
	signal		fold		signal		fold		signal		fold		signal		fold		signal	fold	signal	fold	signal	fold	
	P	I	P	I	P	I	P	I	P	I	P	I	P	I	P	I							
HESRG ^a	8441	P 12.5	I	7041.5	P 12.3	I	6296.5	P 11.1	I	6032.4	P 10.9	I	6	A 2.1	NC	1.6	A 0.1	NC	13.4	A 3.1	NC	1.4	A
DPPA4 ^a	8551	P 11.6	I	6622.6	P 11.1	I	4124.8	P 10.2	I	3433.2	P 10	I	1.7	A -2.8	NC	5	A -0.8	NC	23.1	A 2.7	NC	2.3	A
LRRN1 ^a	2295.4	P 11.1	I	2449.3	P 11.2	I	980	P 9.5	I	943	P 9.7	I	6.3	A 2.4	NC	5	A 0.8	NC	14.5	A 3.4	NC	0.9	A
SOX2 ^{a,b,c,d}	5232.6	P 9.7	I	5120.3	P 9.6	I	2605.2	P 8.8	I	2045.3	P 8.6	I	6.4	A 1.8	NC	1.1	A -0.8	NC	16.6	P 1.1	NC	3.7	A
NANOG ^{a,c,d}	3963.9	P 9.7	I	3276.2	P 10.3	I	1611.5	P 9.8	I	1410.4	P 9.7	I	38.8	A 1.5	NC	10.5	A -0.7	NC	26.6	A 3.3	NC	1.8	A
LIN28 ^{b,c}	8236	P 9.7	I	6788.3	P 9.4	I	6887	P 9.5	I	7215	P 9.5	I	1.3	A 0.5	NC	1.6	A 0.3	NC	15.7	A 0.4	NC	9.8	A
LEFTY2 ^{b,c}	2003.4	P 9.7	I	1774.3	P 9.6	I	2124.9	P 10.1	I	3286.8	P 10.3	I	21.4	A 2.3	NC	84.5	P 1.9	I	24.2	A 1.6	I	2.1	A
ZIC3 ^b	1249.2	P 9.4	I	2427.4	P 10.4	I	1353.9	P 9.4	I	1428.5	P 9.5	I	1.7	A 0.7	NC	4.8	A 1.9	NC	3.8	A 1.6	NC	0.4	A
DPPA2 ^a	668.5	P 9.2	I	338.1	P 8.2	I	177.9	P 7.4	I	193	P 7.3	I	1.6	A -1.5	NC	5.6	A -0.8	NC	25.1	A 4.1	NC	0.8	A
MYCN ^d	1718.1	P 9.2	I	1790.6	P 9.3	I	626.9	P 8.4	I	616.7	P 8.3	I	16.2	A 3.7	NC	0.8	A -0.1	NC	2.1	A 0.8	NC	1	A
PODXL ^c	5752.7	P 9	I	6021	P 9.2	I	6197.9	P 9.1	I	6361.2	P 9.5	I	149.1	P 0.5	NC	159.1	P 0.5	NC	40.4	A 2.9	NC	5.5	A
FGFR2 ^b	1312.7	P 9	I	633.6	P 7.7	I	360.8	P 6.8	I	393.4	P 6.5	I	351	P 7.4	I	530.2	P 7.6	I	1.5	A 0.2	NC	1.2	A
CYP26A1 ^a	1204.2	P 8.7	I	1684.6	P 9.2	I	4273.3	P 10.2	I	5435	P 10.6	I	18.9	A 1.6	NC	2.9	A -0.8	NC	17.2	A 0.7	NC	2.8	A
SH3GL2 ^a	200.2	P 8.6	I	113.2	P 7.1	I	42.2	P 5.2	I	25.1	P 4	I	0.5	A -2.7	NC	15.6	A 1.7	NC	1.3	A 0.6	NC	0.6	A
TDFG1/TDFG3 ^{a,c}	8291.9	P 8.5	I	7825.9	P 8.5	I	7322.2	P 8.4	I	7005.9	P 8.5	I	2.8	A -1.8	NC	23.2	A 0.4	NC	33.7	A 0.6	NC	21	A
SFRP2 ^{a,c}	1799.1	P 8.5	I	1275.3	P 8	I	579.9	P 6.9	I	366.6	P 6.2	I	284.6	P 7.6	I	2.7	A 0.4	NC	12.8	A 0.3	NC	6	A
GPM1 ^a	368.5	P 8.3	I	305.8	P 8.2	I	149.3	P 7.1	I	142	P 7	I	6.3	A 0.9	NC	5.7	A 0.6	NC	7.5	A 1	NC	0.7	A
ZFP42 ^{a,c}	3022.1	P 8.3	I	3294.3	P 8.5	I	1299.1	P 7.4	I	1426.1	P 7.8	I	5.9	A 0.6	NC	117.5	P 4	NC	20.8	P 1.4	NC	5.7	A
SEMA6A ^b	1919.7	P 8.2	I	1540.9	P 7.8	I	1844.8	P 8.2	I	1774.7	P 8.3	I	3.2	A 0.9	NC	3.4	A 0.2	NC	5.8	A 1.1	NC	3.5	A
LEFTY1 ^{a,b,c}	3020.4	P 8	I	3350	P 8.1	I	3232.3	P 7.5	I	3274.1	P 8	I	8.3	A 0.3	NC	19.4	A 1.7	NC	45	M 2.1	NC	11	A
POU5F1 ^{a,b,c,d}	4575.1	P 9	I	4237.8	P 9	I	2967.1	P 8.2	I	1970.1	P 7.8	I	3328.7	P 8.8	I	4800	P 9.2	I	48	P 4	I	21	A
CD24 ^a	10462	P 8	I	9450.4	P 7.8	I	7720.5	P 7.4	I	6714.4	P 7.2	I	56.7	P 0.2	NC	16.3	A -1.5	D	121.6	P 1.3	NC	39	P
PTPRZ1 ^b	3667.8	P 7.9	I	2923.8	P 7.6	I	1575.4	P 6.7	I	1414.5	P 7	I	5.3	A 0	NC	0.3	A -3.2	NC	19.3	P 0.8	NC	12	M
DSG2 ^a	2307.4	P 7.8	I	2495.4	P 7.8	I	1862.3	P 7.5	I	2039	P 7.5	I	7.4	A 2	NC	48.7	P 5.2	I	15.1	A 0.3	NC	11	A
SLC16A10 ^a	234	P 7.7	I	300.7	P 8.3	I	103.8	P 6.5	I	116.8	A 6.1	NC	2.4	A 0.1	NC	3.8	A 0.3	NC	2.5	A 0.7	NC	1.4	A
DNMT3B ^{a,c}	8102.7	P 7.7	I	7437.3	P 7.6	I	3925.4	P 6.7	I	2886.9	P 6.5	I	168.9	P 1.7	I	78.8	P 0.5	NC	40.8	A 0.1	NC	50	A
FOXH1 ^a	1484.9	P 7.3	I	742	P 6.7	I	597.9	P 6.2	I	531.1	P 6.3	I	15.4	A 1	NC	8.5	A 0.5	NC	8.3	A 0.6	NC	4.4	A
SALL3 ^a	975.6	P 7.2	I	342.4	P 5.6	I	424.4	P 5.8	I	373.7	P 5.7	I	18	P 2.4	NC	13.4	A 2.3	NC	15.7	A 0.9	NC	7.6	A
GPM6B ^a	1226.4	P 7.1	I	2427.4	P 8.3	I	365.5	P 5.8	I	361.3	P 5.8	I	12.9	A 2	NC	16	P 2.7	NC	453.7	P 6	I	6.4	A
GAL ^{a,c}	1912.5	P 6.7	I	3510.5	P 7.2	I	4851.1	P 7.6	I	4523	P 7.8	I	1471.3	P 4	I	1089	P 3.5	I	318.9	P 3.7	I	23	A
SNCA ^a	172.1	P 6.6	I	146.9	P 6.1	I	106.7	P 5.4	I	163.8	P 6.2	I	51.7	P 3.3	I	2655.6	P 10.2	I	44	A 4.8	NC	0.9	A
TNNT1 ^a	373.7	P 6.3	I	248	P 5.7	I	427.1	P 6.4	I	386.7	P 6.3	I	49.1	P 1.2	I	21.1	A -0.7	NC	5.9	A 0.1	NC	5.6	A
ITM2A ^b	612.4	P 6.2	I	718.2	P 6.6	I	362.7	P 5.7	I	441.2	P 5.4	I	2.6	A 2.4	I	4.2	P 4.5	NC	130.7	P 3.3	I	6	A
PIM2 ^a	806.7	P 6.2	I	738.5	P 6.1	I	177.8	P 4	I	140.7	P 3.7	I	49.4	P 0.8	NC	117.2	P 1.5	I	18.8	A 1.6	NC	8.2	A
LECT1 ^a	803.2	P 6.1	I	1071.4	P 6.6	I	500.8	P 4.8	I	316.8	P 4.4	I	3.2	A 1	NC	73.1	A 4.7	NC	18.5	A 0.3	NC	10	A
NR6A1 ^c	423.3	P 6	I	415.7	P 5.9	I	413.6	P 5.8	I	391.9	P 5.7	I	3.2	A 0.3	NC	4.7	A 0	NC	6.4	A 0.8	NC	5.6	A
KIF5C ^a	1035.9	P 6	I	794.2	P 5.6	I	736.1	P 5.4	I	791.6	P 5.5	I	31.9	A 2.4	NC	7.3	A 0.4	NC	37.8	A 0.8	NC	20	A
CRABP1 ^a	1485.8	P 6	I	2264.6	P 6.7	I	835.9	P 5.3	I	540.2	P 4.6	I	33.7	P 1.4	NC	112.8	P 3.6	I	126.8	P 2.1	NC	25	A
SALL4 ^{a,b}	2088.5	P 6	I	2053	P 6	I	1236.7	P 5.2	I	1334.3	P 5.4	I	396.9	P 5.7	I	99	A 3.9	I	138.5	A 1.2	NC	30	A
PLP1 ^a	878	P 5.9	I	712.3	P 5.6	I	318.1	P 4.4	I	208.1	P 3.8	I	11	A 0.2	NC	2	A -2.3	NC	19.1	A 0.9	NC	14	A
ADD2 ^b	595.7	P 5.8	I	1102.3	P 6.4	I	291.8	P 4.6	I	204.8	P 4.4	I	39.7	A 4.9	I	1.9	A -0.1	NC	7.4	A 0.2	NC	6.4	A
GABRB3 ^c	1123.5	P 5.7	I	1256.9	P 5.9	I	939.8	P 5.3	I	989.3	P 5.6	I	0.9	A 0.3	NC	1.1	A -0.4	NC	37.3	P 0.6	NC	18	P
GPM6B ^a	1041.7	P 5.6	I	2050.9	P 6.4	I	359.9	P 4.1	I	378.8	P 4.2	I	35.2	M 1	NC	26.6	P 1.1	NC	663.1	P 5.1	I	19	P
MPPE1 ^a	214.8	P 5.4	I	90.9	P 5.1	I	163.9	P 6	I	170.1	P 6	I	8.5	A 3.1	NC	1.8	A 0.8	NC	2.4	A 0.1	NC	2.1	A
EDNRB ^c	113.3	P 5.4	I	215.8	P 6.3	I	399.2	P 7	I	459.4	P 7.5	I	0.3	A 0.4	NC	82.6	P 8.8	I	16	P 2.4	NC	2.6	A
NODAL ^c	441.9	P 5.3	I	286.3	P 5	I	314.8	P 4.7	I	438	P 5.1	I	12.3	A 1.6	NC	3.5	A -0.5	NC	27.4	A 1.6	NC	7.1	A
ACVR2B ^a	2300.2	P 5.3	I	1546.3	P 5.1	I	1076	P 4.6	I	1060.5	P 4.4	I	149.2	P 2.3	I	39.4	A -0.4	NC	93.7	A 0.6	NC	48	A
SOX15 ^a	160.1	P 5	I	157.1	P 4.9	I	109.4	P 4.2	I	95.4	P 4.1	I	4.3	A -0.2	NC	1.5	A -1.1	NC	11.8	A 1.1	NC	4.3	A

iPS clones express most of the HESC signature genes. Genes that are consistently highly expressed between many available HESC lines as determined by the indicated references were analyzed for their expression levels in iPS clones and early OCT4/C-MYC clones 54 and 100 by using microarray data. Note that most HESC genes are induced (denoted with I) in properly reprogrammed iPS clones (2 and 5), whereas these genes in OCT4/C-MYC clones and control cells are often not changed (NC). As determined by GCOS array analysis software: I, increased relative to NHDF; NC, no change relative to NHDF; P, present call; A, absent call; M, marginal call. a, up-regulated in ref. 12; b, up-regulated in ref. 13; c, up-regulated by International Stem Cell Consortium (14); d, defined factors to induce pluripotency; fold is Log₂.

clones repress retroviral transcription in agreement with data shown in Fig. 2B, with iPS clones 24 and 29 exhibiting the least quenching of retroviral transcription (SI Fig. 6). Together, these findings support the notion that reprogramming of fibroblasts to an HESC-like state had indeed occurred upon introduction of the four defined factors OCT4, SOX2, KLF4, and C-MYC and that transgenic expression of these defined genes has ceased, at least in some iPS clones, and given rise to an ES-like endogenous gene-expression pattern.

To exclude the possibility that our iPS clones were simply a contaminant from laboratory HESC cultures, DNA fingerprinting was used to accurately identify the origin of the iPS clones. The data in SI Fig. 7 confirm that all of the iPS clones were derived from NHDF1 and were neither related to HSF1 nor to any other HESC line (genotyping analysis, National Institutes of Health Stem Cell Unit). In addition, although the process of reprogramming remains somewhat of a mystery, karyotypic analysis revealed that gross chromosomal abnormalities were not generated as a result of reprogramming (SI Fig. 7), suggesting that large genomic rearrangements are not required for reprogramming to occur.

To understand how similar the iPS cells generated from NHDF1 were to HESCs, gene-expression profiling was used. Various anal-

yses suggested that iPS clones 2 and 5 were nearly identical in their gene expression profile to two HESC lines (HSF1, H9) (Fig. 3 and Table 1). Scatter-plot analysis of every probe set on a human transcriptome array (HG-U133 + 2, >54,000 probe sets, Affymetrix) emphasized that gene expression levels between iPS cells and HESC closely correlated (Fig. 3A). Similarly, Pearson correlation analysis and hierarchical clustering analysis clearly indicated that iPS cells are significantly more similar to HESCs than to the NHDF1 population from which they were derived (Fig. 3B and C). Examination of the most up- and down-regulated genes in HESCs relative to NHDF1s showed that HESC and iPS cells have nearly identical patterns of the most differentially regulated genes and that there are very few genes expressed by HESC that are not also expressed by iPS cells (Fig. 3D). Table 1 summarizes the expression of 50 genes that are considered to be consensus HESC signature genes (12–14) and further highlights the similarity of gene-expression level between HESC and iPS cells. SI Table 2 depicts the analysis of the top 2,000 up-regulated genes between HESC and NHDF1 extending the observation of similarity. In addition, this table also highlights differences in the gene expression profile between iPS clones. Specifically, iPS clones 24 and 29 induce expression of many fewer HESC-specific genes than iPS clones 2

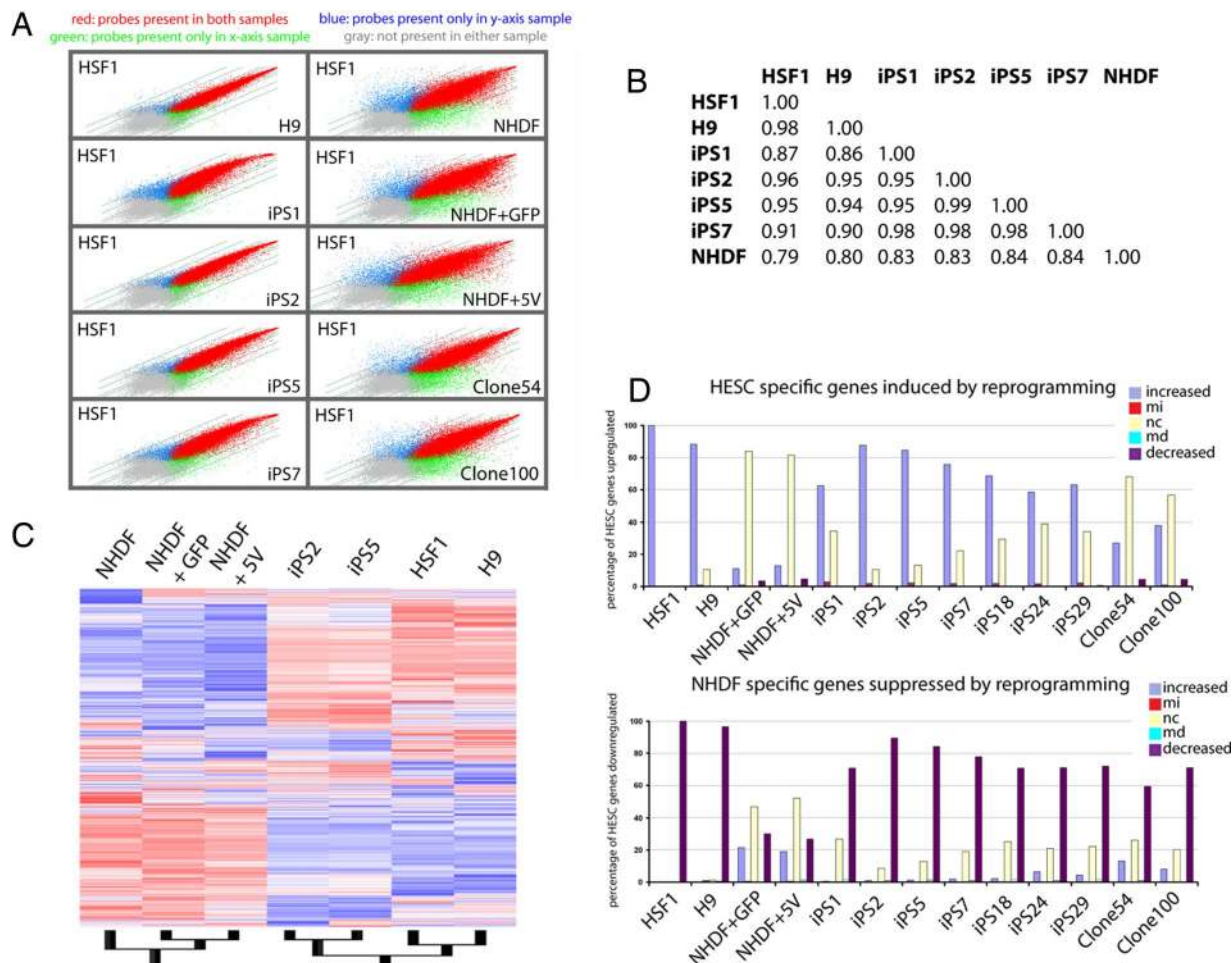


Fig. 3. The transcriptome of iP clones is highly similar to that of HESC. (**A**) Scatter-plot presentation of the expression values for all probe sets derived from genome-wide microarray expression data of indicated cell types. NHDF1 + GFP and NHDF1 + 5V denote a pool of fibroblasts infected with pMX/pMX-GFP control viruses or viruses carrying the five defined factors plus GFP at day 18 after infection. Like the H9 HESC line, iP clones 2 and 5 appear highly similar to the HSF1 HESC, whereas iP lines 1 and 7 appear slightly less similar to HESC. (**B**) Global Pearson correlation of the entire expression data (from Affymetrix microarrays) between indicated cell types. (**C**) Hierarchical clustering of gene-expression data of the indicated cell types. Normalization and expression analysis was performed with DNA-chip analyzer (dChip). A 20% presence call was used to filter genes for clustering, and redundant probe sets were removed. (**D**) The 2,000 most up- and down-regulated genes in HSF1 versus NHDF were determined from genome-wide expression datasets and analyzed for up-regulation, down-regulation, or no change in expression between iP clones or pools of infected NHDF cells and NHDF. MI and MD denote statistically marginal increase or decrease, respectively.

and 5, demonstrating that only partial transcriptional reprogramming has occurred in iP clones 24 and 29. Furthermore, in agreement with RT-PCR data in Fig. 2*B*, *OCT4/C-MYC* clones (clones 54 and 100) induce almost none of the HESC consensus genes (Fig. 3*D* and Table 1). Together, these data demonstrate that expression of four defined transcription factors is sufficient to reset the transcriptional landscape of human fibroblasts to that of HESCs and that efficient silencing of the retroviruses is correlated with more faithful reprogramming to an ES-like expression pattern.

To demonstrate whether iP clones are pluripotent, they were assayed for their ability to differentiate into lineages representative of the three embryonic germ layers. By using standard protocols used for HESC differentiation (15–17), iP clones 2 and 5 were subjected to the embryoid body (EB) formation assay. Fig. 4*A* shows that the iP cells formed canonical EB structures upon collagenase treatment. After growing in suspension for 5 days, the EBs were replated in adherent conditions and driven to differentiate under various conditions. Fig. 4*B* depicts outgrowths from iP cell-derived EBs with distinctive morphologies under different culturing conditions. It should be noted that iP clones 24 and 29 only very inefficiently formed EBs (data not shown).

Upon EB formation and differentiation, iP cells shut down the expression of pluripotency genes *OCT4* and *NANOG* in a similar manner as HSF1 under the same conditions (Fig. 5). Assaying for expression of genes specific for ectoderm, endoderm, and mesoderm, respectively, revealed that iP EBs shared a similar ability to up-regulate different lineage markers as the HESC line HSF1, thus demonstrating pluripotency (Fig. 5). Retinoic acid (RA), known to strongly induce neural differentiation in the EB assay, dramatically induced both *NCAM* and Tyrosine Hydroxylase, a marker of dopaminergic neurons. *Keratin14* (*KRT14*), a marker of epidermis and the ectodermal lineage was also induced under various conditions. Endodermal differentiation was revealed by expression of *PDX1* (pancreas), *SOX7* (parietal endoderm), and *AFP* (liver), whereas mesodermal differentiation was highlighted by *PECAM* (vascular), *SCL* (hematopoietic), and *Desmin* (muscle). The data show that these differentiation markers were essentially absent in undifferentiated HESC and iP cells and strongly induced only after differentiation of the EBs in each condition. Immunostaining further demonstrated that, upon induction of differentiation by EB formation, the iP cells turned on expression of SSEA1, a hallmark of differentiating HESC, and of *NESTIN*, a marker of neural differentiation in response to RA (SI Fig. 8). Only pluripotent cells

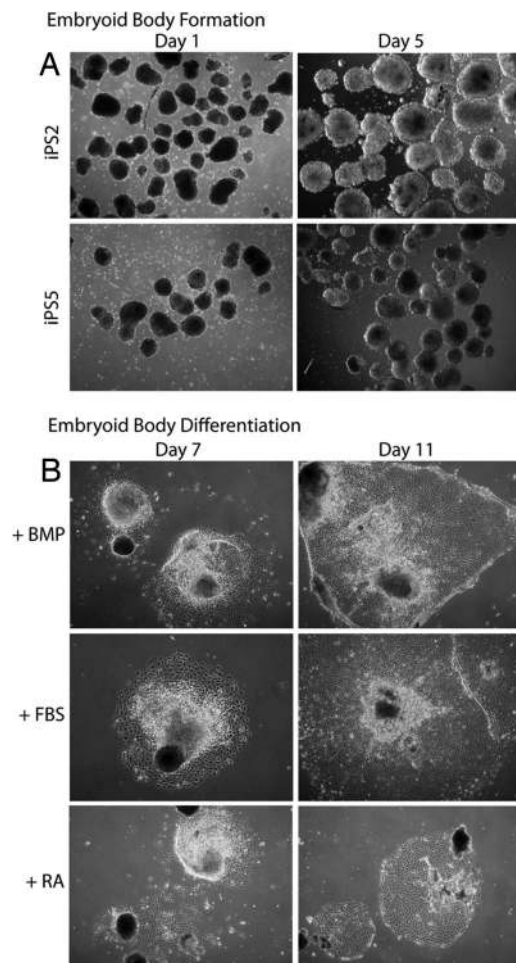


Fig. 4. iPS cells form embryoid bodies similarly to HESCs. (A) Phase-contrast images of EBs generated from iPS clones 2 and 5. (B) Growth of iPS-derived EBs upon plating onto adherent tissue culture dishes under three different media conditions. BMP, bone morphogenetic protein 4.

would have the ability to up-regulate markers of all three embryonic germ layers providing strong evidence for the pluripotency of iPS clones. Whether the iPS clones exhibit pluripotency *in vivo* by teratoma formation and direct transplantation of *in vitro* differentiated cells into appropriate tissues to demonstrate their functional significance remains to be explored.

Discussion

While our work was completed, two other groups published elegant studies similar to ours demonstrating that a variety of human fibroblasts can be reprogrammed to an embryonic state (8, 9). Although Takahashi *et al.* (9) used the same combination of factors (i.e., *OCT4*, *SOX2*, *KLF4*, and *C-MYC*), Yu *et al.* (8) demonstrated that an alternative combination of defined factors, namely *OCT4*, *SOX2*, *NANOG*, and *LIN28* can induce pluripotency in fibroblasts. The fact that the 7 iPS clones analyzed here always carried at least the *OCT4*-, *SOX2*-, *KLF4*-, and *C-MYC*-bearing retroviruses suggests that this combination of defined factors is sufficient for reprogramming in our studies. Whether expression of *NANOG* improves reprogramming efficiency by *OCT4*, *SOX2*, *KLF4*, and *C-MYC* remains to be determined. Nonetheless, these studies demonstrate the reproducibility and feasibility of the defined factor-reprogramming approach in human somatic cells.

Our data demonstrate that defined-factor transduction into fibroblasts leads preliminarily to the formation of nonreprogrammed colonies that are transduced with the *OCT4* and *C-MYC*

retroviruses, and subsequently, TRA-1-81-positive ES-like colonies are generated. Our analysis shows that some of these iPS clones are only partially reprogrammed to an ES cell state as measured by their gene expression program and their inability to form embryoid bodies (clones 24 and 29), whereas other clones (2 and 5) appear faithfully reprogrammed based on all of the criteria tested. One difference between these two classes of clones is that the partially reprogrammed iPS clones still express all ectopic factors, whereas the properly reprogrammed clones appear to quench expression from the retroviral constructs more efficiently. Although stringent data are lacking, defined factor reprogramming in the murine system suggests that shutdown of the exogenously expressed transcription factors is indicative and required for the establishment of the pluripotent state of murine iPS cells and thus occurred more consistently in murine iPS clones that were faithfully reprogrammed to the pluripotent state ((2–5), and data not shown). Clearly more work is required to understand the correlation between retroviral silencing and faithful reprogramming.

We propose live-TRA-1-81 staining as a method for the selection and isolation of reprogrammed clones. Because the proportion of pluripotent colonies generated by introduction of defined factors is low relative to the total number of colonies, the use of live staining for the TRA-1-81 antigen should facilitate future work. Although we cannot state with certainty that TRA-1-81-negative colonies would not give rise to iPS cells upon extended culturing and passaging, all TRA-1-81 positive colonies we obtained had an ES-like morphology and induced the endogenous ESC gene-expression pattern and were either partially or faithfully reprogrammed to the ES-cell state.

The finding that iPS cells generated by expression of defined factors in human fibroblasts are morphologically and physiologically highly similar to HESCs indicates that the mechanism by which murine fibroblasts are reprogrammed to an ES-like state is conserved across species. Although the generation of iPS cells clearly will have an impact on regenerative medicine, the specific role that each of the four transcription factors plays remains unclear. The elucidation of the mechanism by which reprogramming occurs will likely include genomic, epigenetic, and biochemical regulation and promises to contribute significantly to our understanding of self-renewal, differentiation, and the pathogenesis of cancer.

Future challenges include developing methods to transduce the exogenous transcription factors, to reprogram fibroblasts from patients with diseases, and to show that the reprogrammed cells can be differentiated into cell types that function properly upon transplantation and rescue disease models. These steps might allow for general human cell reprogramming through defined factor expression as a therapeutic approach applicable in a clinical setting.

Materials and Methods

Cell Culture Methods. cDNAs for *OCT4*, *SOX2*, *C-MYC*, *NANOG*, *KLF4*, and GFP were cloned into the retroviral pMX vector and separately transfected into Phoenix Ampho Cells (Orbigen) by using Fugene (Roche). Viral supernatants were harvested 3 days later, combined, and used to infect human neonatal dermal fibroblasts (NHDF1; Lonza) in DMEM with 10% FBS, nonessential amino acids, L-glutamine, and penicillin–streptomycin. A second round of infection was performed at day 3, and the transfection efficiency of each virus as extrapolated from that of GFP in the viral mix was 15–20%, suggesting that nearly 100% of cells received at least one virus. Four days later, cells were passaged onto irradiated murine embryonic fibroblasts (MEFs). Reprogrammed cells and HESC cells were cultured on irradiated MEFs as described (10, 11) in DMEM F12 supplemented with L-glutamine, nonessential amino acids, penicillin–streptomycin, knockout serum replacement (Invitrogen), and 10 ng/ml basic FGF. For early passages, iPS cells were propagated manually, whereas subsequent passaging was performed with collagenase treatment as described (10, 11). TRA-1-81 (Chemicon) detection was done without fixation in HESC media, and images were taken within 1 h after secondary antibody incubation. To initiate EB formation, colonies were detached from the feeder layer with collagenase, media exchanged to HESC media without bFGF, and cell clusters plated in non-tissue-culture-treated plates. After 5 days, EBs were transferred onto adherent, gelatin-coated tissue-culture dishes in media con-

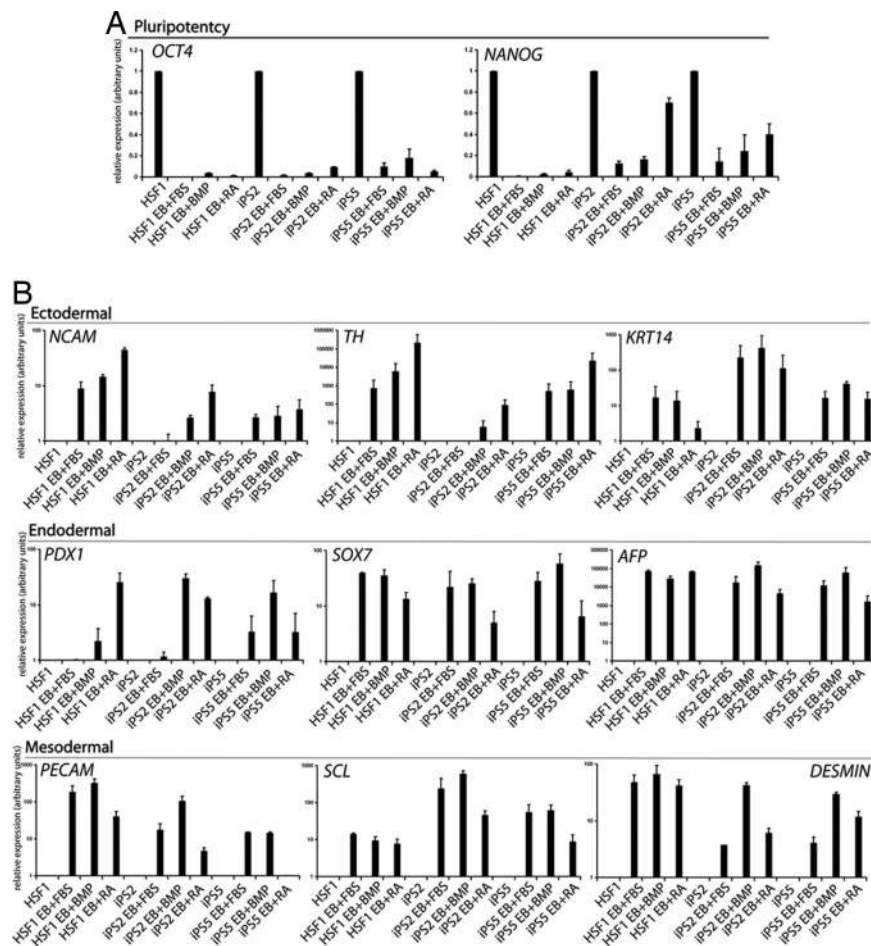


Fig. 5. Pluripotency of iPS cells and up-regulation of ectodermal, endodermal, and mesodermal markers. (A) Real-time RT-PCR analysis of pluripotency gene expression in iPS and control HESC (HSF1) induced to differentiate by EB formation and subsequent plating under indicated conditions [BMP4, FBS, retinoic acid (RA)] relative to *GAPDH* expression. The y axis represents relative fold change upon differentiation. Note that EB differentiation induces down-regulation of pluripotency markers such as *OCT4* and *NANOG*. (B) As in A except that expression of marker genes for different germ layers was analyzed. The specificity of each marker for a given germ layer is indicated. The y axis represents relative fold of induction over undifferentiated cells. Note that although the degree of induction of lineage markers is sometimes variable between HESC and iPS clones, the pattern is consistent.

taining 100 ng/ml BMP4 (R & D Systems), 5% FBS, or 1 μ M *all-trans* retinoic acid (Sigma) and harvested for RNA isolation 9 days later.

Expression Analysis. Total RNA was isolated by using the Absolutely RNA kit (Stratagene) and reverse-transcribed with the SuperScript III First-Strand Synthesis System (Invitrogen) with oligo dT primers. Sequences of all primers are available upon request. In real-time PCR experiments, transcript levels were determined in duplicate reactions and normalized to a *GAPDH* control. Whole-genome expression analysis was performed with the HG-U133 + 2 array (Affymetrix) at the University of California DNA microarray core. Normalization and expression analysis was performed with DNA-chip analyzer [dChip (18)]. Invariant set normalization was used to normalize arrays at the probe level, and the model-based method was used for calculating expression values. A 20% presence call was used to filter genes for clustering, resulting in 20,001 probe sets representing individual genes. Hierarchical clustering analysis was performed to distinguish arrays with similar expression patterns (19). The expression values for a gene across the arrays were standardized by setting the mean signal to 0 and

standard deviation to 1. The expression for each cell type was analyzed by a single microarray experiment, reasoning that biological replicates would be more informative than technical ones.

DNA Analysis. DNA was isolated by using the DNeasy kit (Qiagen) and analyzed for retroviral integration events by PCR with specific primers. DNA fingerprinting and cytogenetic analysis, were performed by Cell Line Genetics.

ACKNOWLEDGMENTS. We thank Kelsey Martin, Gerry Weinmaster, and members of our laboratories for critical reading of the manuscript; Toshio Kitamura (University of Tokyo, Tokyo, Japan) for pMX retroviral constructs; and Shuling Guo (University of California, Los Angeles) for RT-PCR primers. We would especially like to acknowledge Cell Line Genetics for conducting DNA fingerprinting and karyotyping experiments described. K.P. was supported by the Margaret E. Early Trust Foundation, the Jonsson Cancer Center Foundation, and the Kimmel and V Scholar Foundations. R.S. was supported by a California Institute of Regenerative Medicine (CIRM) training grant. W.E.L. was supported by the CIRM and the Jonsson Cancer Center Foundation.

- Byrne JA, et al. (2007) Producing primate embryonic stem cells by somatic cell nuclear transfer. *Nature* 450:497–502.
- Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126:663–676.
- Okita K, Ichisaka T, Yamanaka S (2007) Generation of germline-competent induced pluripotent stem cells. *Nature* 448:313–317.
- Wernig M, et al. (2007) *In vitro* reprogramming of fibroblasts into a pluripotent ES cell-like state. *Nature* 448:318–324.
- Maherali N, et al. (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* 1:55–70.
- Silva J, Chambers I, Pollard S, Smith A (2006) Nanog promotes transfer of pluripotency after cell fusion. *Nature* 441:997–1001.
- Meissner A, Wernig M, Jaenisch R (2007) Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nat Biotechnol* 25:1177–1181.
- Yu J, et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science*.
- Takahashi K, et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131:861–872.
- Amit M, Itskovitz-Eldor J (2006) Derivation and maintenance of human embryonic stem cells. *Methods Mol Biol* 331:43–53.
- Akutsu H, Cowan CA, Melton D (2006) Human embryonic stem cells. *Methods Enzymol* 418:78–92.
- Bhattacharya B, et al. (2004) Gene expression in human embryonic stem cell lines: Unique molecular signature. *Blood* 103:2956–2964.
- Rao RR, et al. (2004) Comparative transcriptional profiling of two human embryonic stem cell lines. *Biotechnol Bioengineer* 88:273–286.
- Adewumi O, et al. (2007) Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotechnol* 25:803–816.
- Itskovitz-Eldor J, et al. (2000) Differentiation of human embryonic stem cells into embryoid bodies comprising the three embryonic germ layers. *Mol Med* 6:88–95.
- Chadwick K, et al. (2003) Cytokines and BMP-4 promote hematopoietic differentiation of human embryonic stem cells. *Blood* 102:906–915.
- Schuldiner M, et al. (2001) Induced neuronal differentiation of human embryonic stem cells. *Brain Res* 913:201–205.
- Schadt EE, Li C, Su C, Wong WH (2000) Analyzing high-density oligonucleotide gene expression array data. *J Cell Biochem* 80:192–202.
- Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci USA* 95:14863–14868.