

SUPPLEMENTAL METHODS

Cell Culture

hESC lines H7 and H9 were obtained from WiCell (Madison, WI) and the HES2 line was obtained from ES Cell International (Singapore). hiPSC lines are described in Supplemental Table 1. IMR90-derived iPSCs (hiPSC1) were obtained from the WiCell Research Institute and assayed between passages 62-68 (1). hASC-derived iPSCs (hiPSC2 & hiPSC4) were derived by lentiviral transduction (2) or by non-viral minicircle transfection (3) as previously described and assayed at passages 51-55 and 48-52, respectively. A new adult dermal fibroblast-derived iPSC line (hiPSC3) was derived by lentiviral transduction using previously described techniques and assayed at passages 45-58 (4). hiPSC lines were individually isolated and expanded from clonally proliferating cell clusters at day 14-28 post transduction (5). Cells were passaged every 3-5 days using 1 mg/mL collagenase IV (Invitrogen). Passage numbers of the cell lines used in this study were as follows: H7: 78-81; H9: 75-79; HES2: 62-65; hiPSC1: 62-68; hiPSC2: 51-55; hiPSC3: 45-58; hiPSC4: 48-52. Normal karyotypes of the H7, HES2, hiPSC1, and hiPSC3 lines are shown in Supplemental Figure 13. Normal karyotypes of the H9, hiPSC2, and hiPSC4 lines have been shown in previous publications from our laboratory less than six passages prior to their use in this study (2, 3, 6). Southern blotting of genomic DNA isolated from the H7, hiPSC4, and hiPSC4 lines is shown in Supplemental Figure 14.

Single cell qRT-PCR

Single cells were sorted into wells containing 10 μ l of reaction buffer (CellsDirect kit, Invitrogen). Reverse transcription and specific transcript amplification were performed using 1 μ l of SuperScript III Reverse Transcriptase / Platinum Taq Mix (Invitrogen) on the thermocycler

(ABI Veriti) as follows: 50°C for 15 min, 70°C for 2 min, 94°C for 2 min, then 94°C for 15 sec, 60°C for 30 sec, and 68°C for 45 sec for 18 cycles, then 68°C for 7 minutes. We then used a microfluidic platform to conduct single cell qPCR in nanoliter reaction volumes, thereby enabling high throughput processing and enhanced detection sensitivity (16). The amplified cDNA was then loaded into Biomark 48.48 Dynamic Array chips using the Nanoflex IFC controller (Fluidigm). Threshold cycle (C_T) as a measurement of relative fluorescence intensity was extracted by the BioMark Real-Time PCR Analysis software.

Fluorescence activated cell sorting (FACS)

hiPSCs were dissociated as above, then pelleted by centrifugation at 200 g for 5 minutes. The cell pellet was resuspended in 100 μ l of D-PBS, then incubated with 20 μ l of Alexa Fluor 647-conjugated mouse anti-human SSEA-4 and 5 μ l of Alexa Fluor 488-conjugated mouse anti-human Tra-1-60 for 30 minutes on ice. The cell pellet was washed with 600 μ l of D-PBS, then sorted as above.

Teratoma formation assay

H7 and hiPSC1 lines stably expressing the bioluminescence reporter gene construct were derived as previously described (7). hESCs and hiPSCs were harvested by treating cells with cell dissociation buffer (Invitrogen) for 10 minutes. Cells were counted using an automated cell counter (Invitrogen Countess) and suspended in Matrigel® at a concentration of 10^6 cells / 8 μ l Matrigel®:PBS at 1:1 mixture. hESCs and hiPSCs were kept on ice for <45 minutes for optimal viability prior to injection. All animal study protocols were approved by the Stanford Animal Research Committee. Surgical procedures were performed on 8-10 week old female

immunocompromised SCID beige mice (Charles River Laboratories, Inc.) by a single experienced microsurgeon. Briefly, mice were knocked down, then intubated, ventilated, and maintained in anesthesia on 2% inhaled isoflurane. An aseptic laparotomy was performed at the left flank and the left kidney was exposed by making a small incision in the peritoneum. By applying slight pressure to both sides of the incision, the kidney was raised out of the peritoneum. Using a 23-gauge needle, a small nick in the kidney capsule was made. hESCs and hiPSCs resuspended in 30 μ l Matrigel were injected directly under the kidney capsule of recipient mice using a 28.5-gauge insulin syringe. After 6 weeks, animals were sacrificed according to protocols approved by the Stanford Animal Research Committee after the duration of the study. Teratomas were explanted, sectioned, and processed for H&E staining. Slides were interpreted by an expert pathologist.

Optical bioluminescence imaging (BLI) of transplanted animals

BLI was performed on all animals using a Xenogen IVIS system. All animals were imaged on days 0, 2, 7, 14, 21, 28, 35, and 42 following cell transplantation. The reporter probe D-luciferin (375 mg/kg) was administered via intra-peritoneal injection 10 minutes prior to image acquisition, after which animals were imaged for 20 min using 1 second to 5 minute acquisition intervals repetitively for the study period. Regions of interest (ROI) were drawn over the signals using the Igor image analysis software (Wavemetrics, Lake Oswego, OR). BLI signal was standardized for acquisition time and quantified in units of maximum photons per second per square centimeter per steradian (photons/sec/cm²/sr).

Cardiomyocyte differentiation

hESC & hiPSC colonies were dispersed into cell aggregates containing approximately 500 to 1,000 cells using 1 mg/mL collagenase IV (Invitrogen). Cell aggregates were suspension-cultured in ultra-low attachment cell culture dishes in STEMPRO34 medium supplemented with 1% L-Glutamine (Invitrogen) for 14 days with inductive cytokines (R&D Systems) as follows: day 0-1: 10 ng/mL BMP4; day 1-4: 10 ng/mL BMP4, 5 ng/mL bFGF, 3 ng/mL Activin A; day 4-8: 10 ng/mL VEGF, 150 ng/mL DKK1; day 8-14: 10 ng/mL VEGF, 5 ng/mL bFGF. Beating cell aggregates were manually counted in four fields of view and averaged at day 14 (8, 9).

Endothelial cell differentiation

hESC & hiPSC colonies were dispersed into cell aggregates containing approximately 500 to 1,000 cells using 1 mg/mL collagenase IV (Invitrogen). Cell aggregates were suspension-cultured in ultra-low attachment cell culture dishes in differentiation medium containing 400 mL of Knockout DMEM, 100 mL of ES-Qualified FBS, 5 mL of MEM NEAA, 5 ml of 100× anti-anti (Invitrogen) supplemented with inductive cytokines as follows: day 0-7: 20 ng/mL BMP4; day 1-4: 10 ng/mL Activin A; day 2-14: 8 ng/mL FGF-2; day 4-14: 25 ng/mL VEGF-A (R&D Systems). Endothelial progenitor cells were magnetically separated using mouse anti-human CD31 antibody (BD Biosciences) and expanded.

Microarray analysis

RNA was extracted from cell populations using the RNEasy Extraction Kit (Qiagen), hybridized onto Affymetrix HG-U133Plus2 chipss, and expression signals were scanned on an Affymetrix GeneChip Scanner. All data sets were analyzed using GeneSpring GX 11 software (Agilent).

Gene-level estimates were derived from the CEL files. Summarization of gene expression data was performed by implementing the robust multichip averaging algorithm, with subsequent baseline normalization of the log-summarized values for each probe set to that of the median log summarized value for the same probe set in the control group. Expression data were then filtered to remove probe sets for which the signal intensities for all the treatment groups were in the lowest 20th percentile of all intensity values. Hierarchical clustering was performed by Euclidian distance metric and centroid linkage rule incorporating the Benjamini-Hochberg FDR multiple testing correction and visualized as dendrograms.

SUPPLEMENTAL DISCUSSION

Rationale for single cell gene expression profiling of human pluripotent stem cells.

Although mouse iPSCs have been shown to be truly pluripotent through tetraploid blastocyst complementation assays (10-12), the functional equivalence of hiPSCs to hESCs is unclear (13, 14). In the absence of embryo assays to stringently assess pluripotency in humans, reliable molecular surrogates of pluripotency and self-renewal must be devised and implemented (15). Microarray analyses have been useful for understanding gene regulation and expression in different cell types, but substantial variability exists within a population of cells (16, 17). The state of the transcriptional regulatory network that governs cell fate determination is reflected in the abundance of regulatory transcripts, and transitions between states are brought on by intrinsic metastability, stochastic fluctuation, and external signals (17-19). Therefore, the assumption that lineage commitment and the undifferentiated state are binary cell fate choices may not accurately represent the nuances of dynamic gene expression changes. Most currently utilized gene profiling methods are inadequate in their ability to characterize complex cell states due to their inability to assess gene expression levels in single cells. Tagging of gene products with fluorescent proteins has enabled single cell assessment of expression via FACS, and such studies have given tremendous insight into how heterogeneous expression of a single gene of interest can influence stem cell states (19-21). However, assessing multiple gene products in human pluripotent stem cells using this approach is difficult because of the extensive genomic modification required and the limited number of fluorescent proteins with non-overlapping excitation/emission spectra.

Validation of microfluidic platform for assaying gene expression in single pluripotent stem cells. Studies of the transcriptome and epigenome of mammalian ESCs and the inner cell mass of the blastocyst have revealed a network of co-regulated genes that maintain pluripotency and suppress lineage commitment. We selected 28 pluripotency-related transcripts known to be highly expressed across 59 undifferentiated hESC lines but lowly expressed in differentiating cell aggregates (previously known as “embryoid bodies”) as determined by the International Stem Cell Initiative (22). We validated the ability of our microfluidic platform to assay small amounts of mRNA by diluting known quantities of RNA and cDNA to levels comparable to those found in single pluripotent stem cells. We confirmed the expected logarithmic relationship between cDNA concentration and cycle threshold (C_T) value for an array of genes known to characterize the pluripotent and differentiated state (Supplemental Figure 2). Cells were isolated in equal measure from various cell lines (Supplemental Figure 1, Supplemental Table 1). We also assayed in parallel the donor cell types (hASCs and IMR90) that were reprogrammed to yield the tested hiPSCs. In a lognormally distributed population, a cell’s degree of expression is best represented by the geometric mean, because use of the arithmetic mean may overestimate the number of transcripts in a typical cell (23).

Expression of the Dlk1-Dio3 gene cluster in hESCs and hiPSCs. The Dlk1-Dio3 gene cluster is an imprinted region of the genome implicated in maintenance of pluripotency (24). Namely, expression of the genes in the Dlk1-Dio3 gene cluster is repressed in “partially pluripotent” mouse iPSC lines that are unable to generate tetraploid-blastocyst-complemented embryos. We sought to compare the expression level of the Dlk1-Dio3 cluster in our hiPSC and hESCs as a potential explanation for their differing patterns of pluripotency-related gene expression (Figure

1C). We assayed the hESC and hiPSC populations for expression of the DIO3, DLK1, MEG8, RTL1, and miR-541 transcripts and observed that both hESC and hiPSC lines had activated the Dlk1-Dio3 gene cluster to an equivalent extent (Supplemental Figure 6A). We also assayed expression of DLK1, DIO3, and MEG8 in single hESCs and hiPSCs, and observed no statistically significant cell-to-cell variations in their expression levels (Supplemental Figure 6B). Therefore, differential activation of the Dlk1-Dio3 gene cluster is not responsible for heterogeneity in transcript levels among hiPSCs. Rather, the hiPSC lines used in this study meet a putative additional criterion for “full pluripotency” in that they have completely activated expression of the Dlk1-Dio3 gene cluster.

Expression of cell cycle regulatory genes in hESCs and hiPSCs. The cell cycle regulatory transcripts CDKN1A (encoding p21/Cip1), CDKN2A (encoding p16/Ink4a), and MYBL2 (also known as B-MYB) have been previously shown to affect the expression of important pluripotency-maintenance genes in hESCs and hiPSCs (25-28). In particular, MYBL2 is a transcription factor involved in cell cycle progression that is highly abundant in ESCs. Downregulation of MYBL2 precedes decreases in POU5F1, NANOG, and SOX2 protein abundance during ESC differentiation. Given the important role of cell cycle progression in transcript abundance, we sought to determine whether the expression pattern of cell cycle regulatory genes differed significantly in our hESC and hiPSC lines. When levels of such transcripts were measured at the population level, however, there were no statistically significant differences in the expression CDKN1A, CDKN2A, and MYBL2 between hESCs and hiPSCs (Supplemental Figure 7A). Gene expression levels of CDKN1A, CDKN2A, and MYBL2 also did not differ significantly between single hESCs and single hiPSCs (Supplemental Figure 7B).

Therefore, the observed heterogeneity in hiPSC transcript abundance is not attributable to differences in cell cycle progression between hiPSCs and hESCs.

Principal components analysis of immunophenotypic gene expression profiling data. In order to better characterize the data obtained from hESCs/hiPSCs isolated by SSEA-4⁺ / Tra-1-60⁺ immunophenotype, we used principal components analysis (PCA). PCA takes data points in a multidimensional space and defines new axes (components) to cut across the space such that the first component captures as much of the variance in the data as possible. The second component, orthogonal to the first, captures as much as possible of the remaining variance, and so on. We graphed data points as individual cells and defined principal components for a 28-D space (representing 28 genes). Applying this analysis to the expression data from the Tra-1-60⁺ / SSEA-4⁺ hESCs and hiPSCs, we observed that hESCs occupied a discrete focus with a small percentage of outliers within principal component space, while hiPSCs spanned a large area of the space with no discrete concentration (Supplemental Figure 10), demonstrating that even Tra-1-60⁺ / SSEA-4⁺ hiPSCs exhibit considerable cell-to-cell variability in gene expression levels.

Comparison of hESC- versus hiPSC-derived teratoma growth kinetics. hESC-derived teratomas grew exponentially in the initial weeks after transplantation and stabilized between days 28-42. Bioluminescence signal intensity reached a plateau due to centralized necrosis in the larger tumor volumes. hESC-injected animals were unable to further support the tumor burden at day 42 and needed to be euthanized. hiPSC-derived teratomas, however, grew exponentially throughout the 42 day period, albeit at a slower rate than hESC-derived teratomas (Figure 3A-B; Supplemental Figure 15). Interestingly, undifferentiated hESCs and hiPSCs displayed equivalent

expression levels of cell cycle regulatory genes at the population and single cell level (Supplemental Figure 7). Therefore, differences in teratoma growth rate are probably not attributable to differences in the proliferative rates of the undifferentiated hiPSCs and hESCs. However, a greater proportion of undifferentiated hiPSCs are “primed” for exit of the pluripotent compartment, as indicated by their reversible downregulation of a number of key pluripotency-maintenance transcripts (Figure 1C). In addition, hiPSC-derived cells produced during *in vitro* directed differentiation protocols can manifest reduced proliferation rates and increased senescence marker expression relative to their hESC-derived counterparts (Figure 3F) (13, 14). Therefore, the retarded growth kinetics of hiPSC-derived teratomas is more likely attributable to the altered differentiation capacity of hiPSCs and deficits in the proliferative capacity of the resulting hiPSC-derived cells, not deficits in the proliferative capacity of the undifferentiated hiPSCs.

Metastability of hiPSCs. Heterogeneous transcript expression among seemingly homogeneous cell populations has been previously shown to underlie important cell fate determinations (16-19, 29-31). Insightful work by Stewart et al. has demonstrated that undifferentiated hESC cultures consist of a heterogeneous mixture of SSEA-3⁺ and SSEA-3⁻ cells that differ in their clonogenic capacity and Oct4/Nanog expression profile (31). Fluctuations in the levels of pluripotency-related transcripts such as Nanog have been previously shown to underlie a propensity towards differentiation and lineage commitment in murine ESCs (19, 20). A metastable subset of ESCs preparing for exit of the pluripotent compartment can be identified by reversible downregulation of pluripotency-related genes (18, 31, 32). We show here that hiPSCs display increased heterogeneity in the expression of key transcription factors that regulate the pluripotency

maintenance network such as OCT4, SOX2, and NANOG. hiPSCs also displayed increased variability in expression of cancer-related proliferative genes such as MYC and TDGF1 (Cripto), which may have significant ramifications for the safety and ultimate clinical applicability of these cell types (33, 34). This variability in expression of key factors reflects an inherent instability in the state of undifferentiated hiPSCs, and may be a residual consequence of the poorly understood, stochastically-driven reprogramming process (35-37). The higher proportion of metastable hiPSCs is reflected in their retarded teratoma growth kinetics and a more limited ability to give rise to therapeutic cell populations such as cardiac and endothelial cells. The retarded growth kinetics of hiPSC-derived teratomas could not be attributed to deficits in the proliferative rate of undifferentiated hiPSCs. Consistent with the results of others, our results indicate that hiPSC-derived cells manifest altered differentiation potential relative to other pluripotent cell types (13, 14, 38, 39). Functional differences in differentiation capacity may be partially attributable to recently delineated “epigenetic memory” distinctions between murine and human iPSCs and their ESC counterparts (37-39). However, the differential gene expression patterns attributed to epigenetic memory are primarily limited to lineage-associated genes (e.g., *Cxcr4*, *Itgb1*, *Gr-1*, *Lysozyme*, *Mab21L1*, *Atbf1*, *Hand1*), not the core pluripotency maintenance network genes discussed here (e.g., *Nanog*, *Sox2*, *Oct4*) (38, 39). The altered differentiation capacity of hiPSCs relative to hESCs can therefore be attributed to two separate contributing factors: 1) cell-to-cell heterogeneity in the expression of pluripotency maintenance genes and 2) residual expression of lineage-associated genes due to epigenetic memory of the tissue of origin. Although epigenetic memory can be mitigated by continual passaging (38, 40), late passage bona fide hiPSCs with full activation of the *Dlk1-Dio3* cluster retained significant cell-to-cell heterogeneity and unreliable differentiation potential. These data highlight our limited current

understanding of transcriptional and epigenetic heterogeneity in cultured pluripotent stem cells.
In their current form, even stringently defined hiPSCs may be of limited clinical usefulness.

SUPPLEMENTAL FIGURE LEGENDS

Supplemental Figure 1. Characterization of hESC & hiPSC populations. (a) Microarray analysis of hESC and hiPSC populations. Clustering analysis shows that hESC types (H7, H9, HES2) bear considerable similarities to one another in terms of global gene expression profile. hiPSCs from different starting cell sources derived using different methods (hiPSC1-4) all have a similar global gene expression profile to hESCs. Pluripotent stem cell types (hESCs and hiPSCs) bear considerable similarity to one another in comparison to the differentiated somatic cell types from which the hiPSCs were derived (hASCs and IMR90). (b) Top panel shows RT-qPCR results for pluripotency genes conducted on RNA extracted from hiPSC populations in comparison to hASCs and H7 hESCs. Note that ΔC_T or $\Delta\Delta C_T$ is calculated such that higher expression results in more negative values (see Methods). Graph indicates $\Delta\Delta C_T$ values (normalized to 18S endogenous control and hASCs) \pm SD. Bottom panel shows RT-qPCR performed on RNA extracted from hiPSC2 and hiPSC3 populations with primers specific for the endogenous transcript (i.e., annealing to the untranslated region) or the total amount of transcript (i.e., annealing to the coding region). No significant difference in total vs. endogenous expression levels is observed at this level of resolution, indicating successful silencing of exogenous lentiviral transgenes. Graph indicates ΔC_T values (normalized to 18S endogenous control) \pm SD. (c) Immunostaining of hiPSC colonies for antigens as labeled shows high expression levels of pluripotency markers Oct4, Nanog, Sox2, SSEA-3, Tra-1-81, and Tra-1-60.

Supplemental Figure 2. Validation of logarithmic relationship between input mRNA levels and output cycle threshold (C_t) values using microfluidic RT-qPCR assay. (a) Reference RNA (Stratagene) was diluted within the picogram range without compromising the expected

logarithmic relationship between mRNA amount and C_t value using Taqman® assays specific for the genes indicated. (b) The mRNA of a hESC colony containing ~100 cells was diluted up to 10,000 fold without compromising the logarithmic relationship between input mRNA and C_t value. (c) The mRNA of a hiPSC colony containing ~100 cells was identically diluted up to 10,000 fold without compromising the logarithmic relationship between input mRNA and C_t value. As expected, when mRNA was diluted to concentrations less than expected in single cells (i.e., 1,000 to 10,000 fold), the standard deviation between repeat measurements increased. Graphs show arithmetic mean \pm SD for $n = 6$ replicates across three separate microfluidic chips.

Supplemental Figure 3. Comparison of single cell gene expression between hESCs and hiPSCs. Each point represents expression value of a single cell for hESCs (in blue) and hiPSCs (in red) assayed in equal measure from the various cell lines as described. Horizontal axes represent C_t value on a \log_2 scale. Significant variation in C_t values is observed amongst hiPSCs for pluripotency-related genes Oct4, Nanog, Sox2, DNMT3B, Myc, TDGF1, GDF3, and NR6A1. Variation amongst hESCs is less pronounced in comparison. No difference was observed between hESCs and hiPSCs in expression level of the control 18S gene (bottom right panel).

Supplemental Figure 4. Variance amongst 282 hESCs and hiPSCs was computed and graphed on the vertical axis for the genes indicated on the horizontal axis. Variance is significantly greater amongst hiPSCs (open bars) than amongst hESCs (solid bars) for the comparisons marked with an asterisk at the $P < 0.004$ level by the Ansari-Bradley Homogeneity of Variances Test. Data used to derive the graph is shown in Supplemental Table S2.

Supplemental Figure 5. Principal component analysis of single hESCs and hiPSCs. Each dot represents a single hESC (H7, H9, and HES2) or hiPSC (hiPSC1-4) in principal component space, labeled according to cell line. No cell line occupies a discrete area in principal component space that is separate from the other pluripotent stem cell lines because the variance in the data set is spread amongst all lines. Principal component axes captured the variance across the 20 pluripotency-related genes (from Figure 1) as follows: PC1 (x-axis) = 64.75%, PC2 (y-axis) = 13.85%, PC3 (z-axis) = 11.75%.

Supplemental Figure 6. Activation of the Dlk1-Dio3 imprinted region in hiPSCs and hESCs. (a) RNA was extracted from hESC and hiPSC populations followed by RT-qPCR for the DIO3, DLK1, MEG8, RTL1, and miR-541 genes. Equivalent ΔC_T values were obtained for all tested hESC and hiPSC lines, indicating equivalent activation of the Dlk1-Dio3 cluster and full pluripotency in the tested lines. (b) Single hESCs (blue squares) and single hiPSCs (red squares) were assayed for expression of the DLK1, DIO3, and MEG8 genes, with no significant differences observed between expression levels or variation in ΔC_T values.

Supplemental Figure 7. Cell cycle regulatory gene expression in hiPSCs and hESCs. (a) RNA was extracted from hESC and hiPSC populations followed by RT-qPCR for the CDKN1A (encoding p21/Cip1), CDKN2A (encoding p16/Ink4a), and MYBL2 (also known as B-MYB) genes. Equivalent ΔC_T values were obtained for all tested hESC and hiPSC lines, indicating equivalent cell cycle regulation in the tested lines. (b) Single hESCs (blue squares) and single hiPSCs (red squares) were assayed for expression of the CDKN1A, CDKN2A, and MYBL2

genes, with no significant differences observed between expression levels or variation in ΔC_T values.

Supplemental Figure 8. Single hESCs and hiPSCs assayed for expression of differentiation-associated transcripts relating to lineage (ectoderm, endoderm, mesoderm) or major histocompatibility complex (MHC). Ectodermal lineage priming (i.e., co-expression of NES or MAP2 alongside pluripotency markers) was observed amongst both hESCs (solid bars) and hiPSCs (open bars). ANPEP is a marker for differentiated fibroblasts that was used as a secondary control for cell type identification. Vertical axis indicates the percentage of cells with detectable expression of the genes indicated along the horizontal axis.

Supplemental Figure 9. Flow cytometry plots of a representative single cell suspension of hiPSCs immunostained with anti-Tra-1-60 antigen and anti-SSEA-4 antibodies. The viable, Tra-1-60⁺ / SSEA-4⁺ population was selected for single cell sorting and further analysis.

Supplemental Figure 10. Principal component analysis of single hESCs and hiPSCs sorted according to Tra-1-60⁺ / SSEA-4⁺ immunophenotype. Each point represents a single cell in principal component space. Four differentiated IMR90 fibroblasts were included for comparison. hESCs occupy a discrete focus with some variation, while hiPSCs occupy a larger swath of the graph area with large distances between individual cells and no discrete focus. Percent of total variance across 28 genes was captured as follows: PC1 (x-axis) = 65%; PC2 (y-axis) = 35%.

Supplemental Figure 11. Stable transduction of undifferentiated hESCs and hiPSCs. (a) Schematic of the lentiviral expression construct containing firefly luciferase (Fluc) and enhanced green fluorescent protein (eGFP) used to stably transduce hESCs and hiPSCs. (b) Linear correlation between bioluminescence imaging (BLI) signal and number of viable cells is shown for undifferentiated hESCs and hiPSCs.

Supplemental Figure 12. Histological analysis of explanted teratomas. Six weeks after implantation into the kidney capsule, teratomas were explanted, sectioned, and stained for H&E as described. Explanted teratomas displayed histological features of all three germ layers, including (a) columnar epithelium (endoderm), (b) adipocytes (mesoderm), (c) squamous keratinocytes (ectoderm), and (d) mixed elements.

Supplemental Figure 13. Normal karyotypes of the pluripotent stem cell lines used in this study. (a) Normal 46,XX karyotype of hESC line H7. (b) Normal 46,XX karyotype of hESC line HES2. (c) Normal 46,XY karyotype of hiPSC line hiPSC1. (d) Normal 46,XY karyotype of hiPSC line hiPSC3. Normal karyotypes of H9, hiPSC2, and hiPSC4 lines were demonstrated in other recent publications from our lab (2, 3, 6).

Supplemental Figure 14. Southern blotting of genomic DNA isolated from lines hiPSC3, hiPSC4, H7, and fibroblast cells was performed using an Oct4 probe. Three unique integration sites of the hiPSC3 line are revealed, with no weak bands to suggest contaminating subclones. H7, fibroblasts, and hiPSC4 (transgene-free) only contained one endogenous copy of Oct4 as expected.

Supplemental Figure 15. Bioluminescence imaging (BLI) of hESC- and hiPSC-derived teratomas. Data from Figure 3B was plotted for each individual animal. hESC-derived teratomas (blue) grow quickly and reach a plateau by day 28, after which tumor necrosis restricts further substantial increase in the BLI signal intensity. By day 42, hESC-injected animals can no longer support the tumor burden and need to be euthanized. hiPSC-derived teratomas grow more slowly and continue their exponential growth into day 42.

Supplemental Table 1. Cell source and method of derivation for hiPSC lines used in this study. hiPSC1, 2, and 4 are lines that have been well characterized in previous publications as referenced. Note that hiPSC3 is a new line derived for comparative use in this study using well-established techniques (see Methods).

Supplemental Table 2. Statistical comparison of variance in hESC and hiPSC samples. Variance in gene expression measurements amongst hiPSCs is significantly higher than amongst hESCs for the indicated genes with *P*-values for the Ansari-Bradley Two Sample Test of Homogeneity of Variances shown to the right. *P*-values are considered significant at the $P < 0.004$ level.

Supplemental Video 1. Beating cell aggregates derived from hESC (H7) at day 14 of differentiation protocol. hESC-CMs can be derived dependably using established directed differentiation protocols (see Methods) with a beating efficiency of $27.0 \pm 3.6\%$, in contrast to hiPSC-CMs.

References

1. Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318:1917-1920.
2. Sun, N., Panetta, N.J., Gupta, D.M., Wilson, K.D., Lee, A., Jia, F., Hu, S., Cherry, A.M., Robbins, R.C., Longaker, M.T., et al. 2009. Feeder-free derivation of induced pluripotent stem cells from adult human adipose stem cells. *Proc Natl Acad Sci USA* 106:15720-15725.
3. Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., Panetta, N.J., Chen, Z.Y., Robbins, R.C., Kay, M.A., et al. 2010. A nonviral minicircle vector for deriving human iPS cells. *Nat Meth* 7:197-199.
4. Park, I.-H., Lerou, P.H., Zhao, R., Huo, H., and Daley, G.Q. 2008. Generation of human-induced pluripotent stem cells. *Nat. Protocols* 3:1180-1186.
5. Smith, Z.D., Nachman, I., Regev, A., and Meissner, A. 2010. Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nat Biotech* 28:521-526.
6. Xie, X., Hiona, A., Lee, A.S., Cao, F., Huang, M., Li, Z., Cherry, A., Pei, X., and Wu, J.C. 2010. Effects of long-term culture on human embryonic stem cell aging. *Stem Cells Dev* 9:9.
7. Sun, N., Lee, A., and Wu, J.C. 2009. Long term non-invasive imaging of embryonic stem cells using reporter genes. *Nat Protocols* 4:1192-1201.
8. Laflamme, M.A., Chen, K.Y., Naumova, A.V., Muskheli, V., Fugate, J.A., Dupras, S.K., Reinecke, H., Xu, C., Hassanipour, M., Police, S., et al. 2007. Cardiomyocytes derived

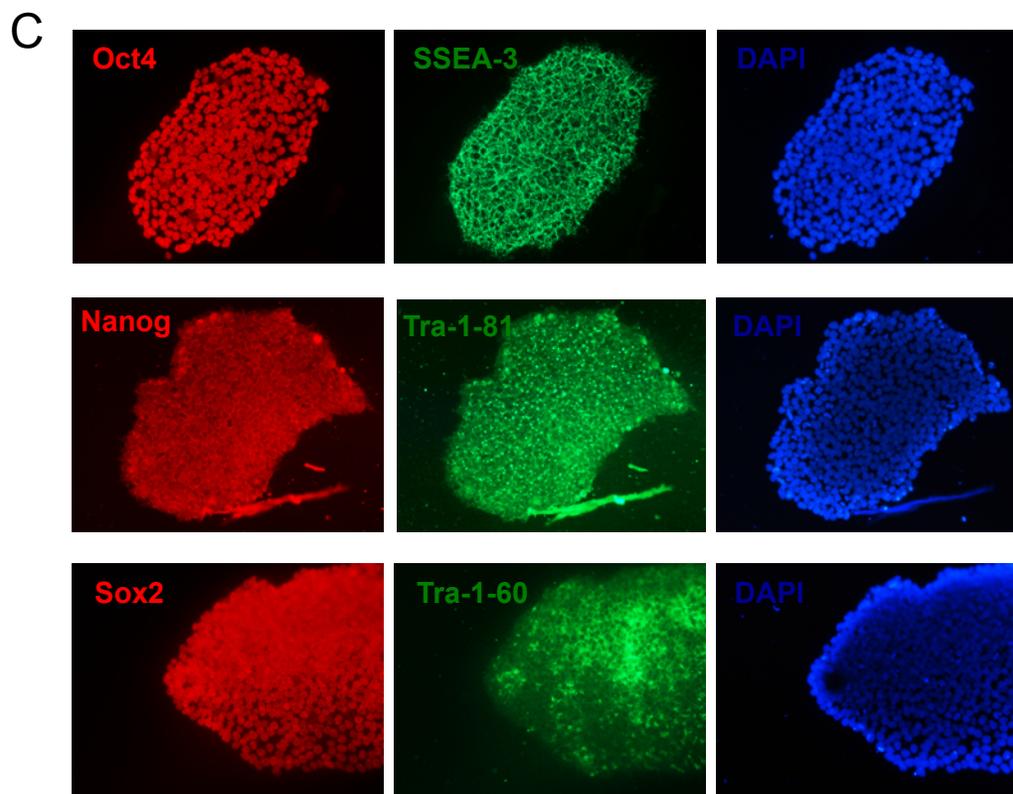
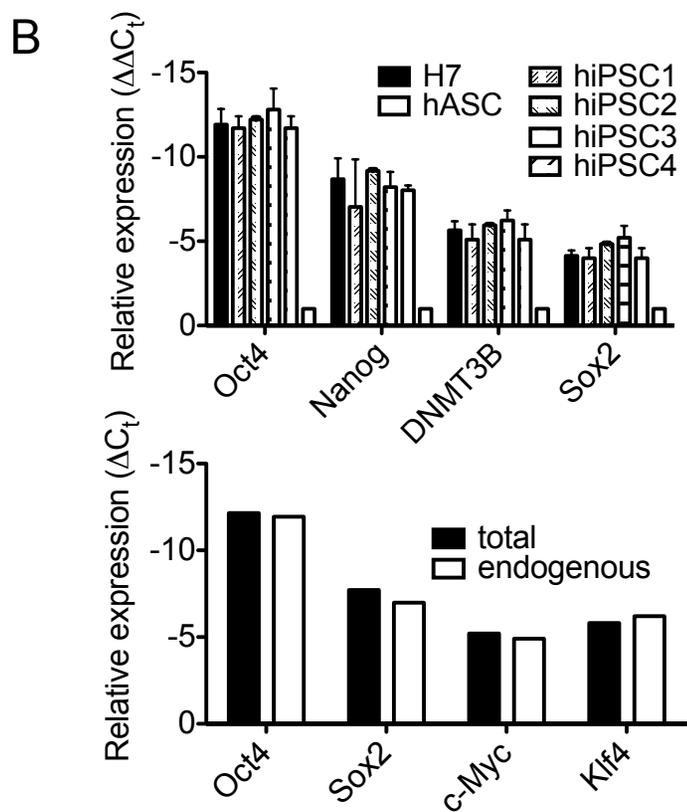
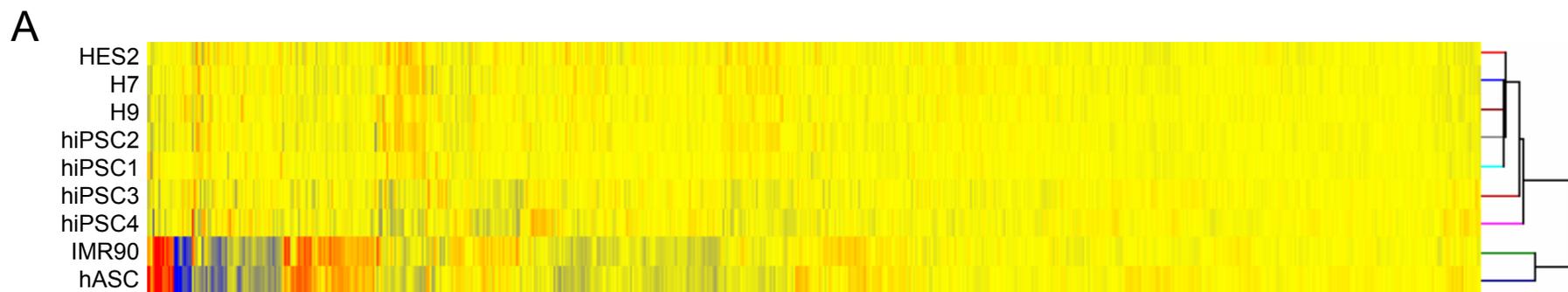
- from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nat Biotech* 25:1015-1024.
9. Cao, F., Wagner, R.A., Wilson, K.D., Xie, X., Fu, J.-D., Drukker, M., Lee, A., Li, R.A., Gambhir, S.S., Weissman, I.L., et al. 2008. Transcriptional and functional profiling of human embryonic stem cell-derived cardiomyocytes. *PLoS One* 3:e3474.
 10. Zhao, X.-y., Li, W., Lv, Z., Liu, L., Tong, M., Hai, T., Hao, J., Guo, C.-l., Ma, Q.-w., Wang, L., et al. 2009. iPS cells produce viable mice through tetraploid complementation. *Nature* 461:86-90.
 11. Boland, M.J., Hazen, J.L., Nazor, K.L., Rodriguez, A.R., Gifford, W., Martin, G., Kupriyanov, S., and Baldwin, K.K. 2009. Adult mice generated from induced pluripotent stem cells. *Nature* 461:91-94.
 12. Okita, K., Ichisaka, T., and Yamanaka, S. 2007. Generation of germline-competent induced pluripotent stem cells. *Nature* 448:313-317.
 13. Hu, B.Y., Weick, J.P., Yu, J., Ma, L.X., Zhang, X.Q., Thomson, J.A., and Zhang, S.C. 2010. Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc Natl Acad Sci USA* 107:4335-4340.
 14. Feng, Q., Lu, S.-J., Klimanskaya, I., Gomes, I., Kim, D., Chung, Y., Honig, G.R., Kim, K.-S., and Lanza, R. 2010. Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells* 28:704-712.
 15. Daley, G.Q., Lensch, M.W., Jaenisch, R., Meissner, A., Plath, K., and Yamanaka, S. 2009. Broader implications of defining standards for the pluripotency of iPSCs. *Cell Stem Cell* 4:200-201.

16. Warren, L., Bryder, D., Weissman, I.L., and Quake, S.R. 2006. Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. *Proc Natl Acad Sci USA* 103:17807-17812.
17. Raj, A., and van Oudenaarden, A. 2008. Nature, nurture, or chance: stochastic gene expression and its consequences. *Mol Cell* 135:216-226.
18. Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. 2008. Identification and characterization of subpopulations in undifferentiated ES cell culture. *Development* 135:909-918.
19. Kalmar, T., Lim, C., Hayward, P., Muñoz-Descalzo, S., Nichols, J., Garcia-Ojalvo, J., and Martinez Arias, A. 2009. Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells. *PLoS Biol* 7:e1000149.
20. Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L., and Smith, A. 2007. Nanog safeguards pluripotency and mediates germline development. *Nature* 450:1230-1234.
21. Canham, M.A., Sharov, A.A., Ko, M.S.H., and Brickman, J.M. 2010. Functional heterogeneity of embryonic stem cells revealed through translational amplification of an early endodermal transcript. *PLoS Biol* 8:e1000379.
22. 2007. Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nat Biotech* 25:803-816.
23. Stahlberg, A., Bengtsson, M., Hemberg, M., and Semb, H. 2009. Quantitative transcription factor analysis of undifferentiated single human embryonic stem cells. *Clin Chem* 55:2162-2170.

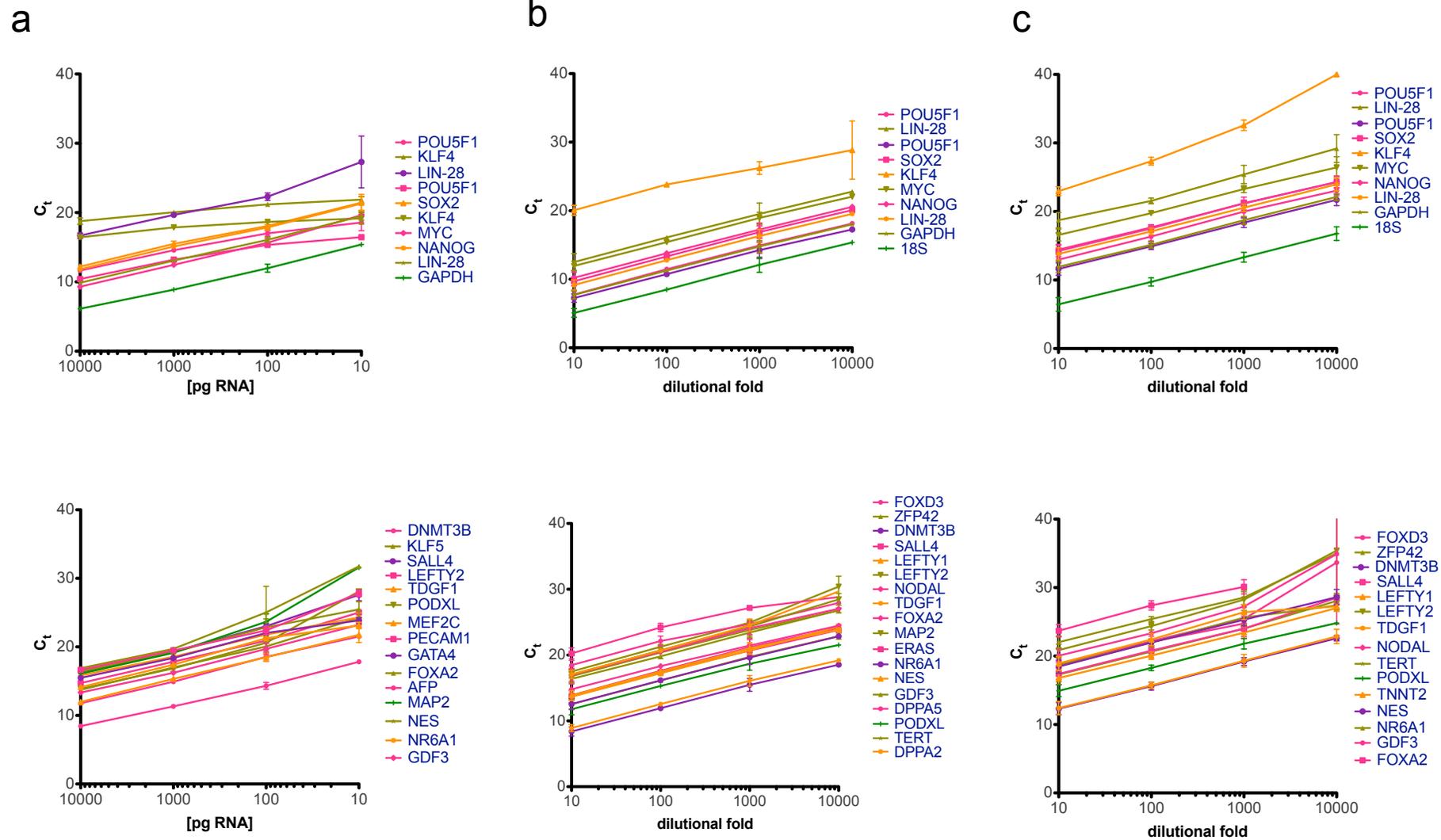
24. Liu, L., Luo, G.-Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., Li, W., Wu, H.-J., Wang, L., Wang, X.-J., et al. 2010. Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *Journal of Biological Chemistry* 285:19483-19490.
25. Tarasov, K.V., Tarasova, Y.S., Tam, W.L., Riordon, D.R., Elliott, S.T., Kania, G., Li, J., Yamanaka, S., Crider, D.G., Testa, G., et al. 2008. B-MYB Is essential for normal cell cycle progression and chromosomal stability of embryonic stem cells. *PLoS One* 3:e2478.
26. Banito, A., Rashid, S.T., Acosta, J.C., Li, S., Pereira, C.F., Geti, I., Pinho, S., Silva, J.C., Azuara, V., Walsh, M., et al. 2009. Senescence impairs successful reprogramming to pluripotent stem cells. *Genes & Development* 23:2134-2139.
27. Kawamura, T., Suzuki, J., Wang, Y.V., Menendez, S., Morera, L.B., Raya, A., Wahl, G.M., and Belmonte, J.C.I. 2009. Linking the p53 tumour suppressor pathway to somatic cell reprogramming. *Nature* 460:1140-1144.
28. Utikal, J., Polo, J.M., Stadtfeld, M., Maherali, N., Kulalert, W., Walsh, R.M., Khalil, A., Rheinwald, J.G., and Hochedlinger, K. 2009. Immortalization eliminates a roadblock during cellular reprogramming into iPS cells. *Nature* 460:1145-1148.
29. Sprinzak, D., and Elowitz, M.B. 2005. Reconstruction of genetic circuits. *Nature* 438:443-448.
30. Franco, C.B., Chen, C.-C., Drukker, M., Weissman, I.L., and Galli, S.J. 2010. Distinguishing mast cell and granulocyte differentiation at the single-cell level. *Cell Stem Cell* 6:361-368.

31. Stewart, M.H., Bosse, M., Chadwick, K., Menendez, P., Bendall, S.C., and Bhatia, M. 2006. Clonal isolation of hESCs reveals heterogeneity within the pluripotent stem cell compartment. *Nat Meth* 3:807-815.
32. Silva, J., and Smith, A. 2008. Capturing pluripotency. *Cell* 132:532-536.
33. Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., Nakagawa, M., Koyanagi, M., Tanabe, K., Ohnuki, M., et al. 2009. Variation in the safety of induced pluripotent stem cell lines. *Nat Biotech* 27:743-745.
34. Levine, A.J., and Brivanlou, A.H. 2006. GDF3, a BMP inhibitor, regulates cell fate in stem cells and early embryos. *Development* 133:209-216.
35. Yamanaka, S. 2009. Elite and stochastic models for induced pluripotent stem cell generation. *Nature* 460:49-52.
36. Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., Bernstein, B.E., Jaenisch, R., Lander, E.S., and Meissner, A. 2008. Dissecting direct reprogramming through integrative genomic analysis. *Nature* 454:49-55.
37. Ghosh, Z., Wilson, K.D., Wu, Y., Hu, S., Quertermous, T., and Wu, J.C. 2010. Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PLoS One* 5:e8975.
38. Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., Apostolou, E., Stadtfeld, M., Li, Y., Shioda, T., et al. 2010. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotech* 28:848-855.

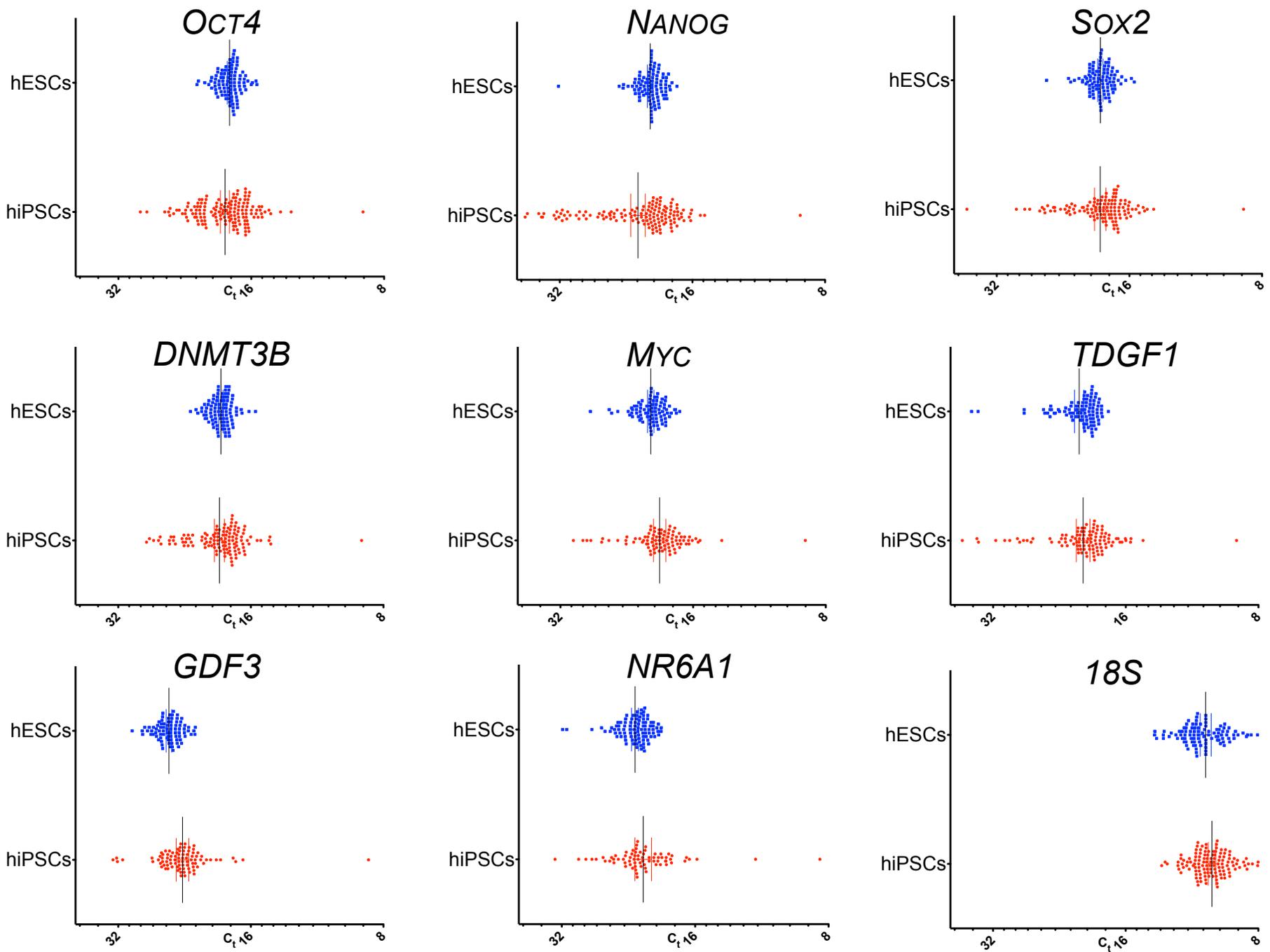
39. Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M.J., Ji, H., Ehrlich, L.I.R., et al. 2010. Epigenetic memory in induced pluripotent stem cells. *Nature* 467:285-290.
40. Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiwu, O., Richter, L., Zhang, J., et al. 2009. Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5:111-123.



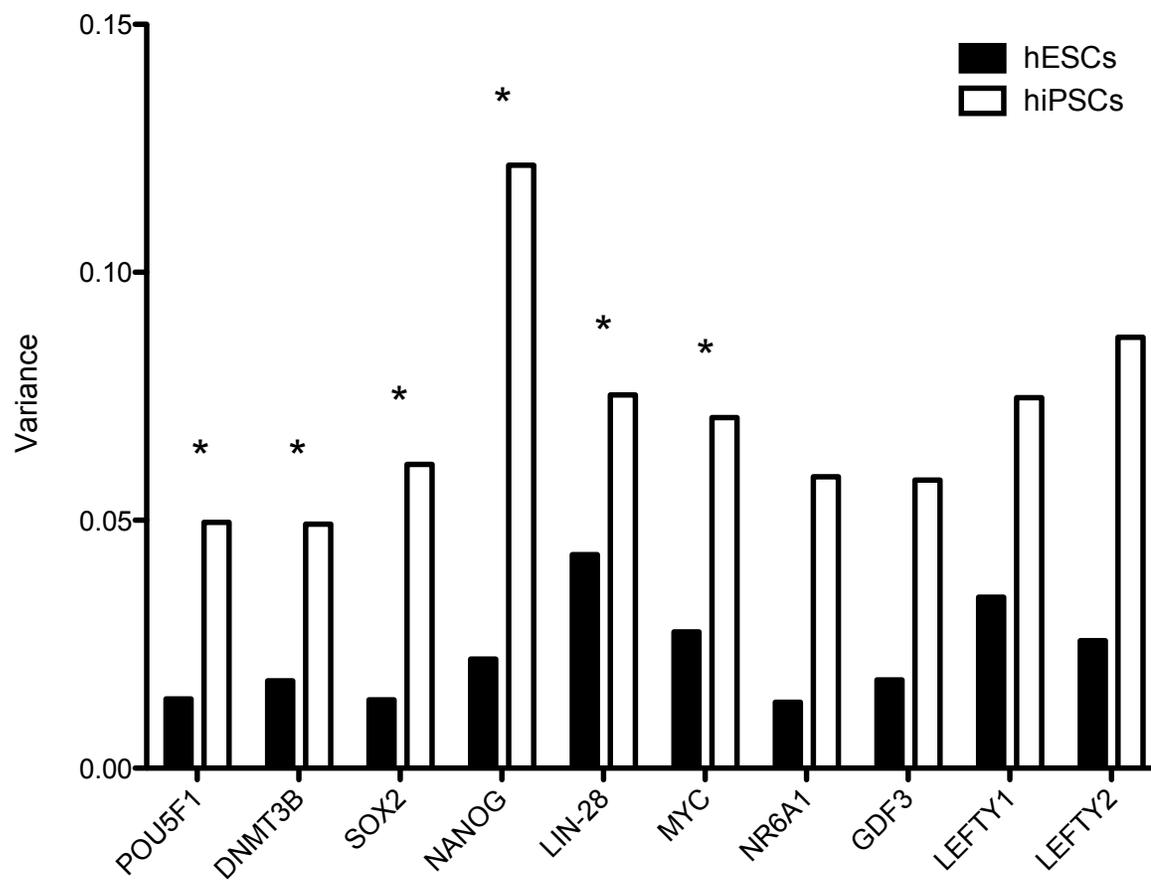
Supplemental Figure 1



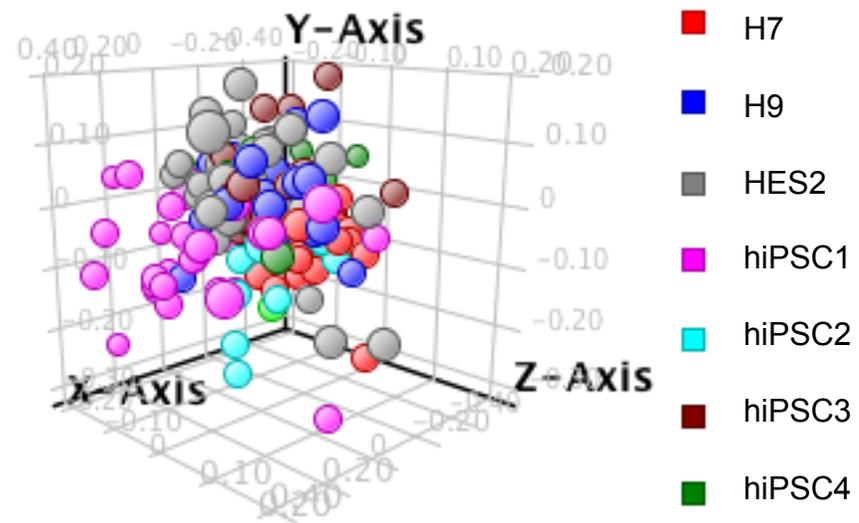
Supplemental Figure 2



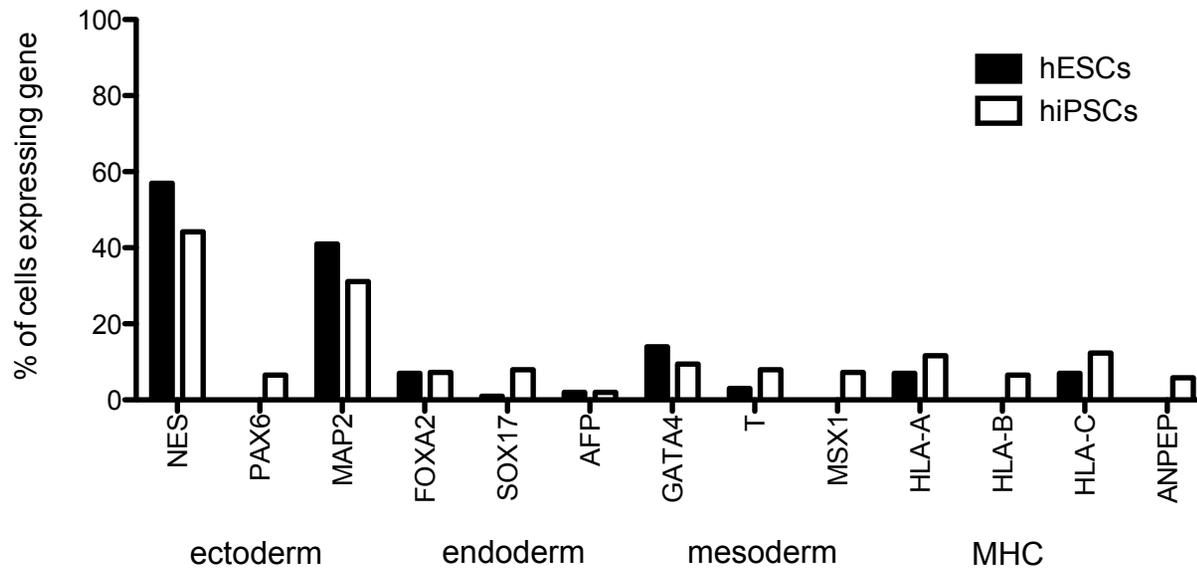
Supplemental Figure 3



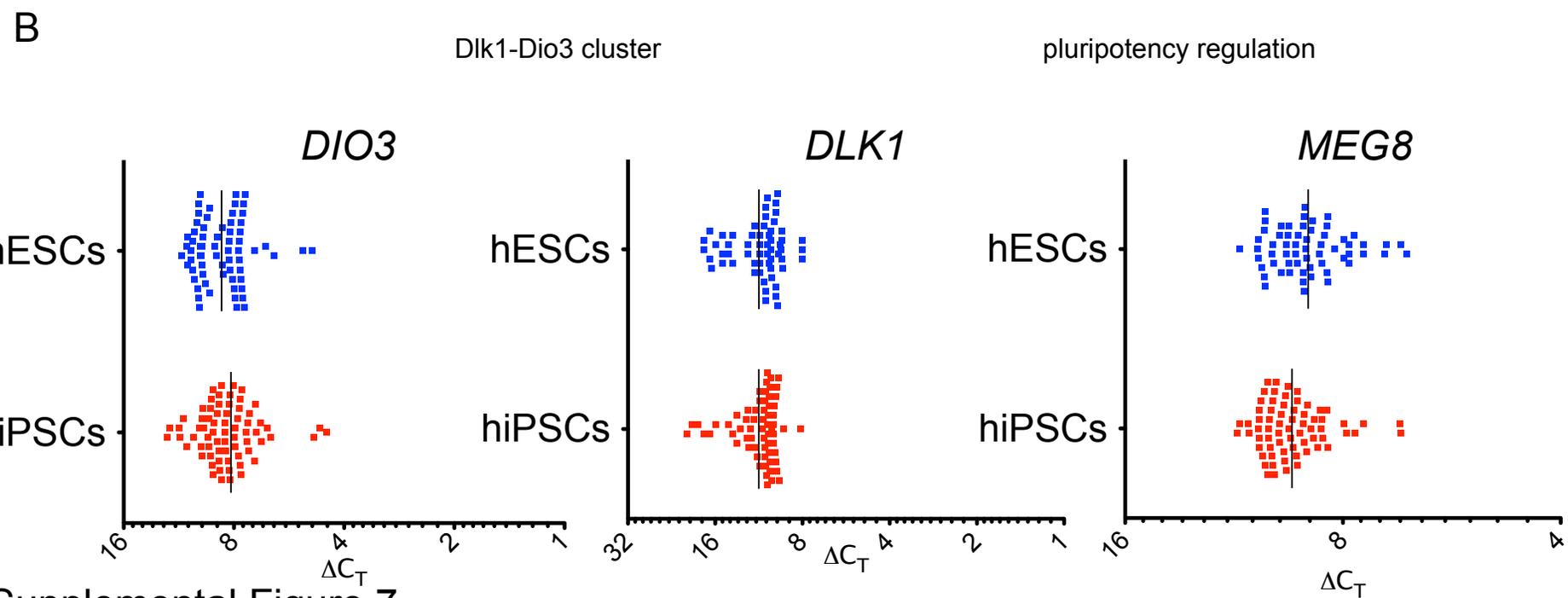
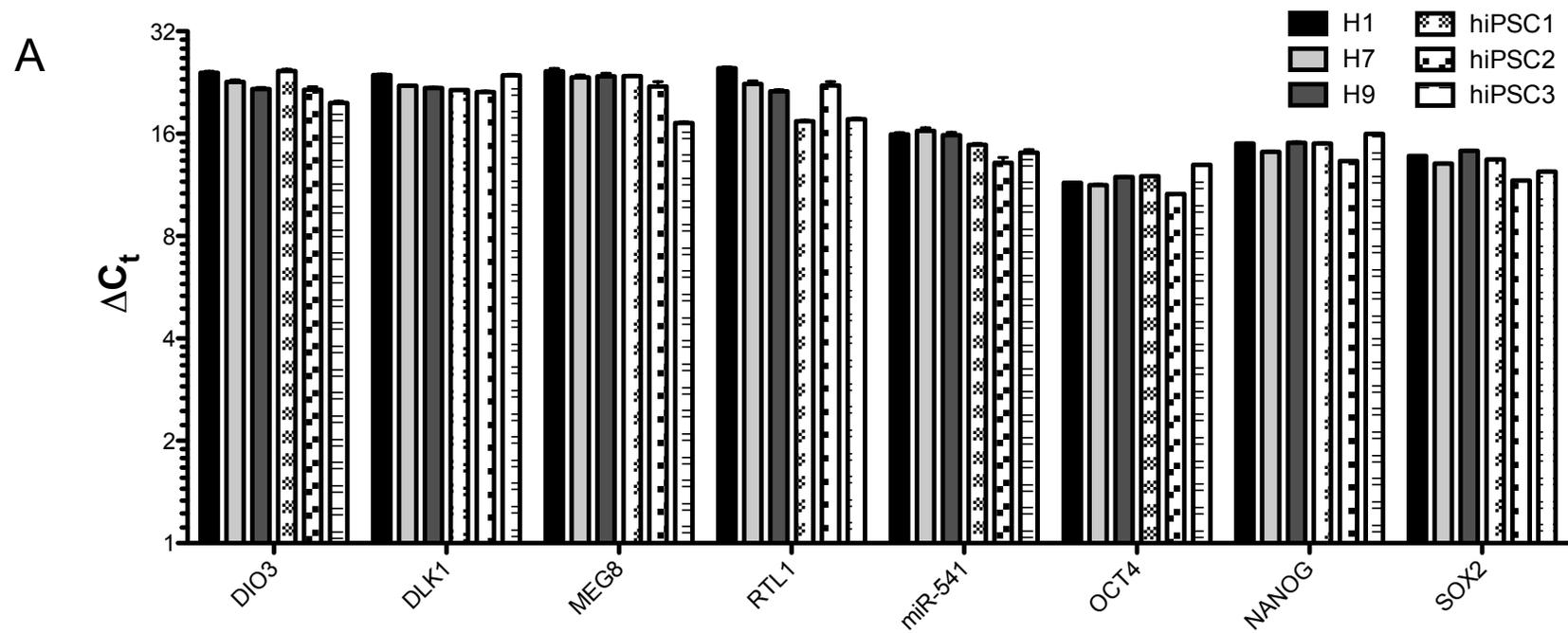
Supplemental Figure 4



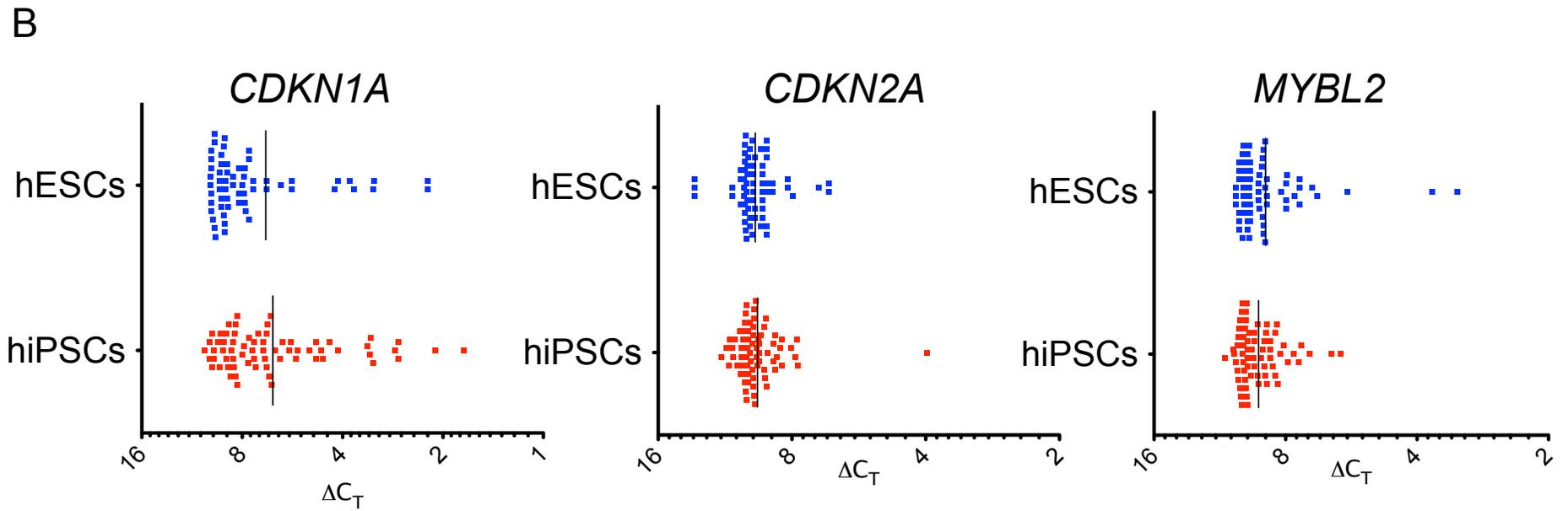
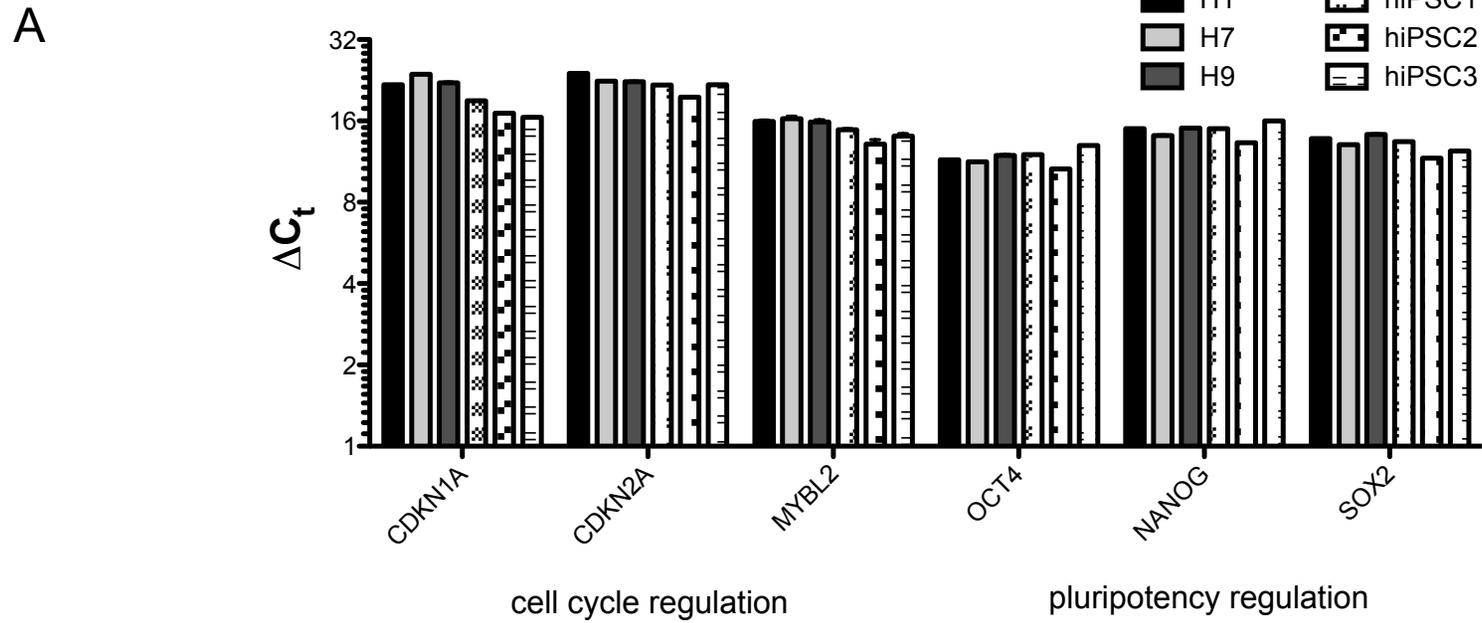
Supplemental Figure 5



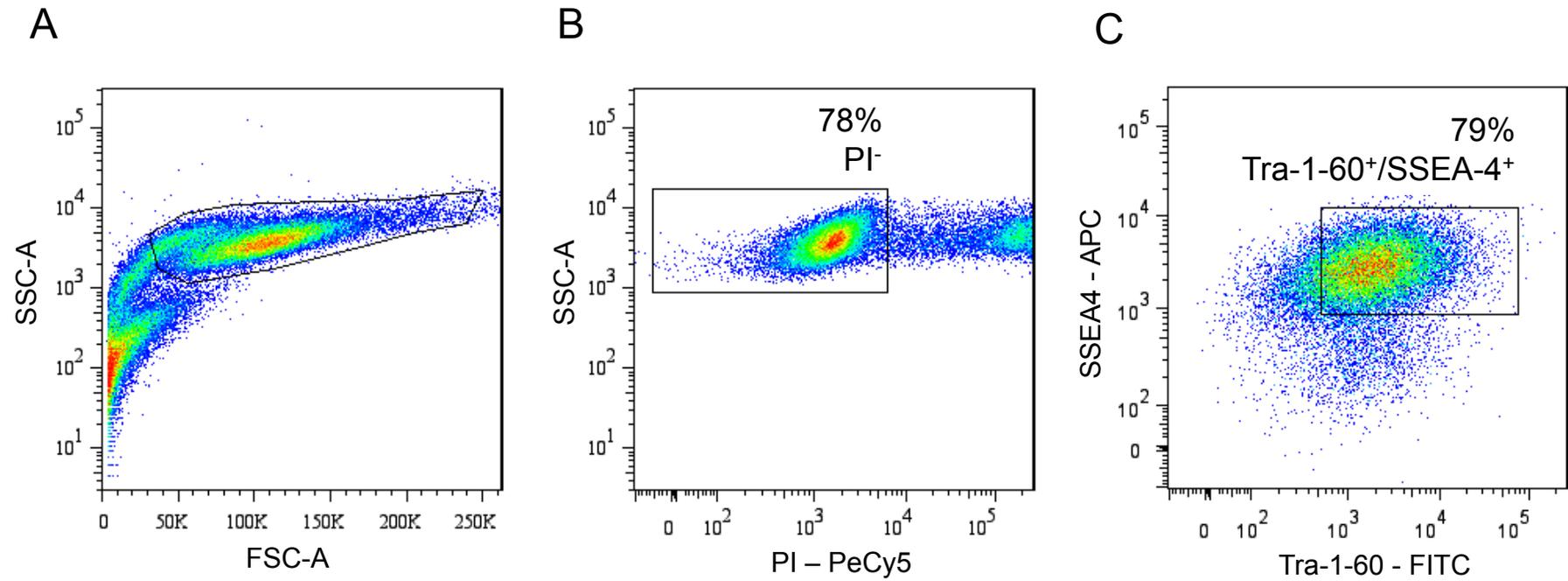
Supplemental Figure 6



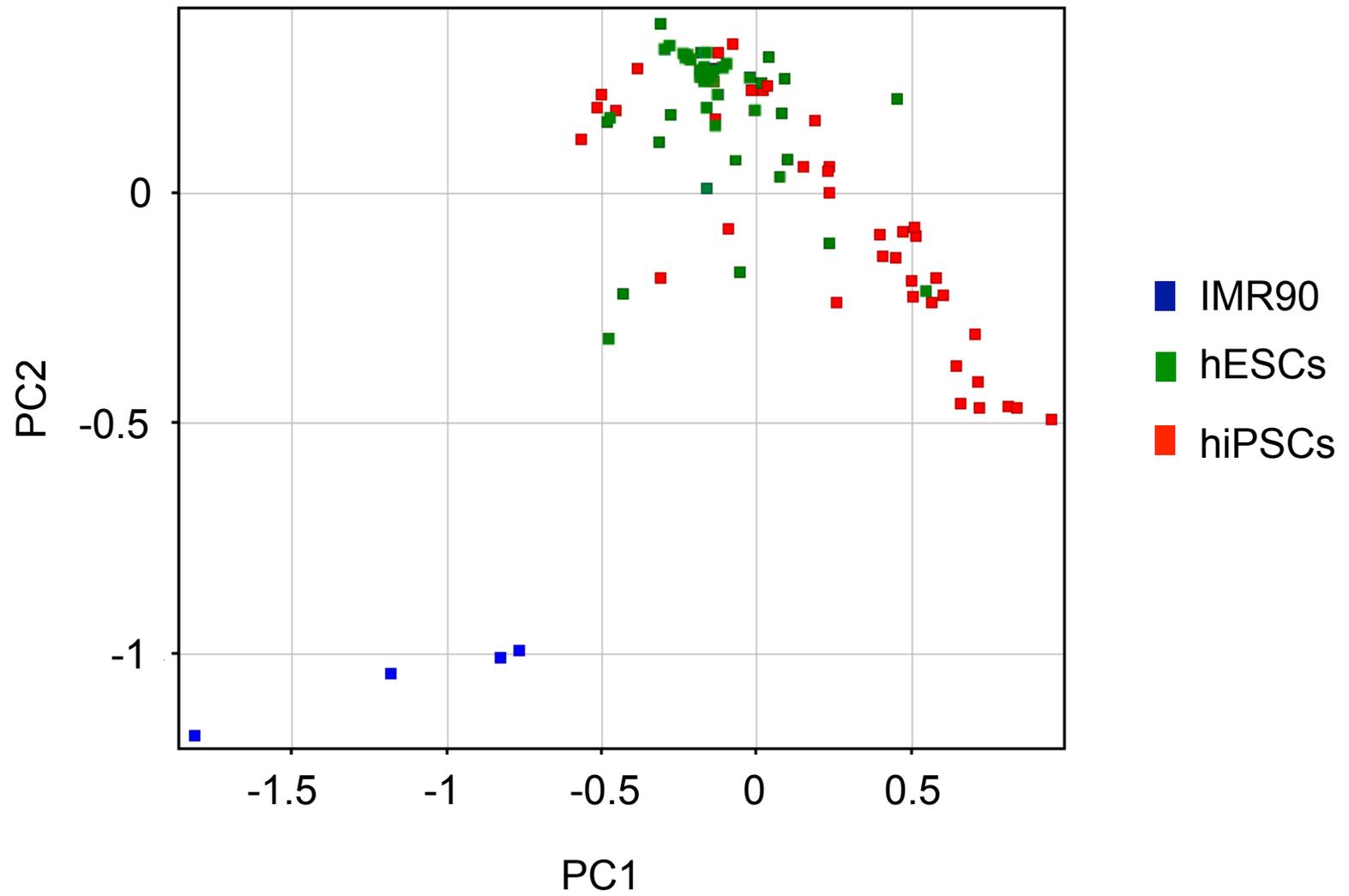
Supplemental Figure 7



Supplemental Figure 8



Supplemental Figure 9

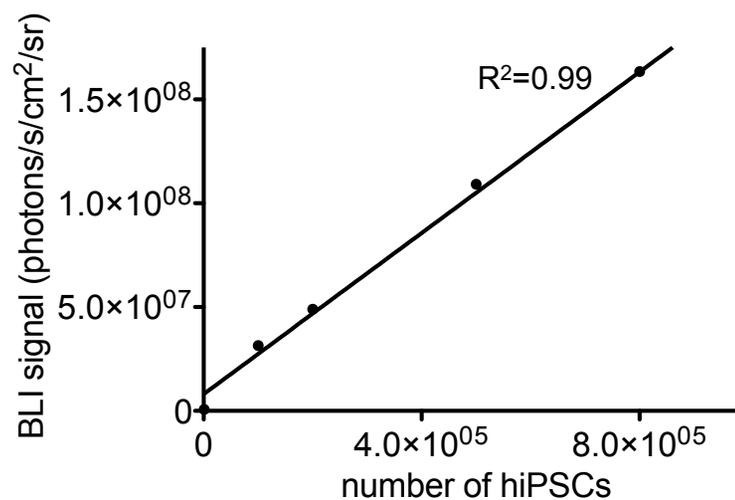
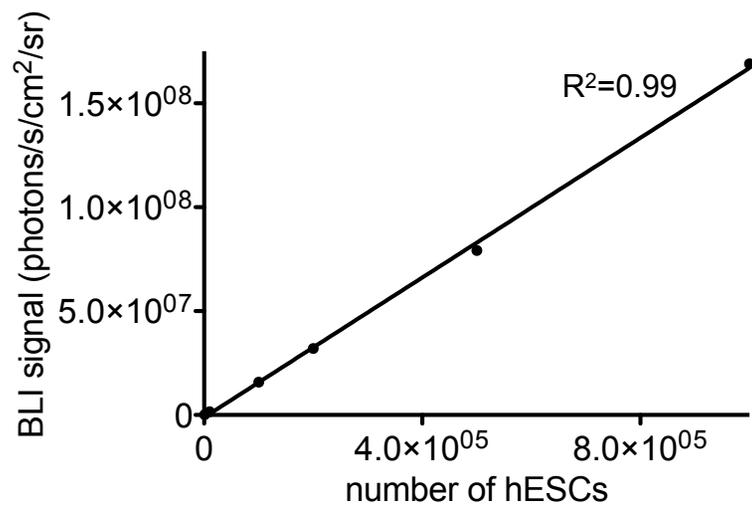


Supplemental Figure 10

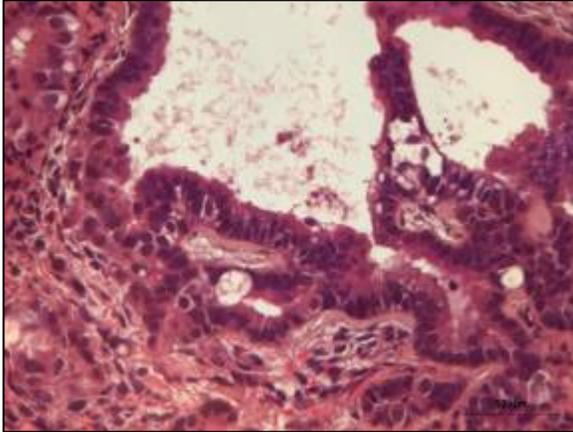
A



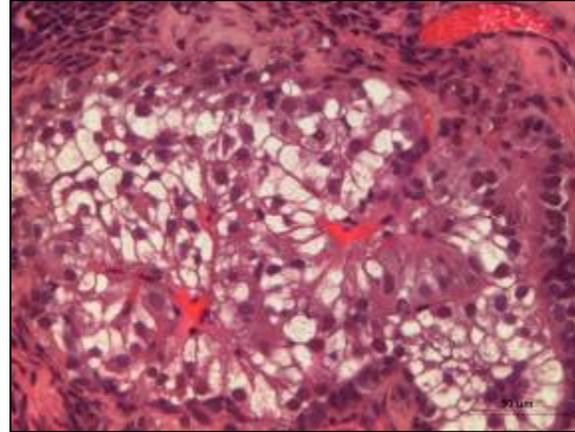
B



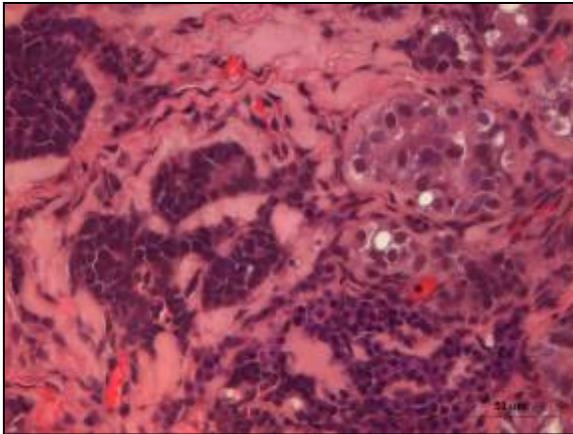
A



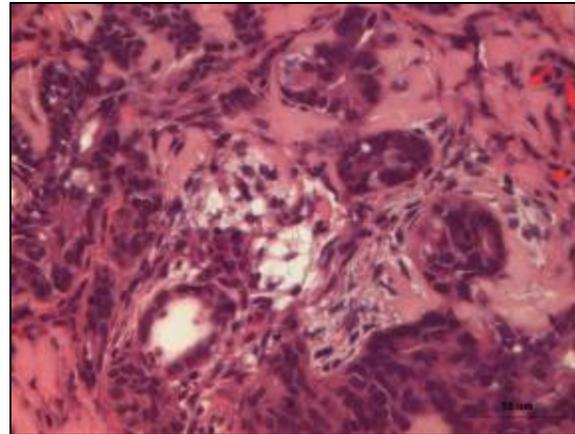
B



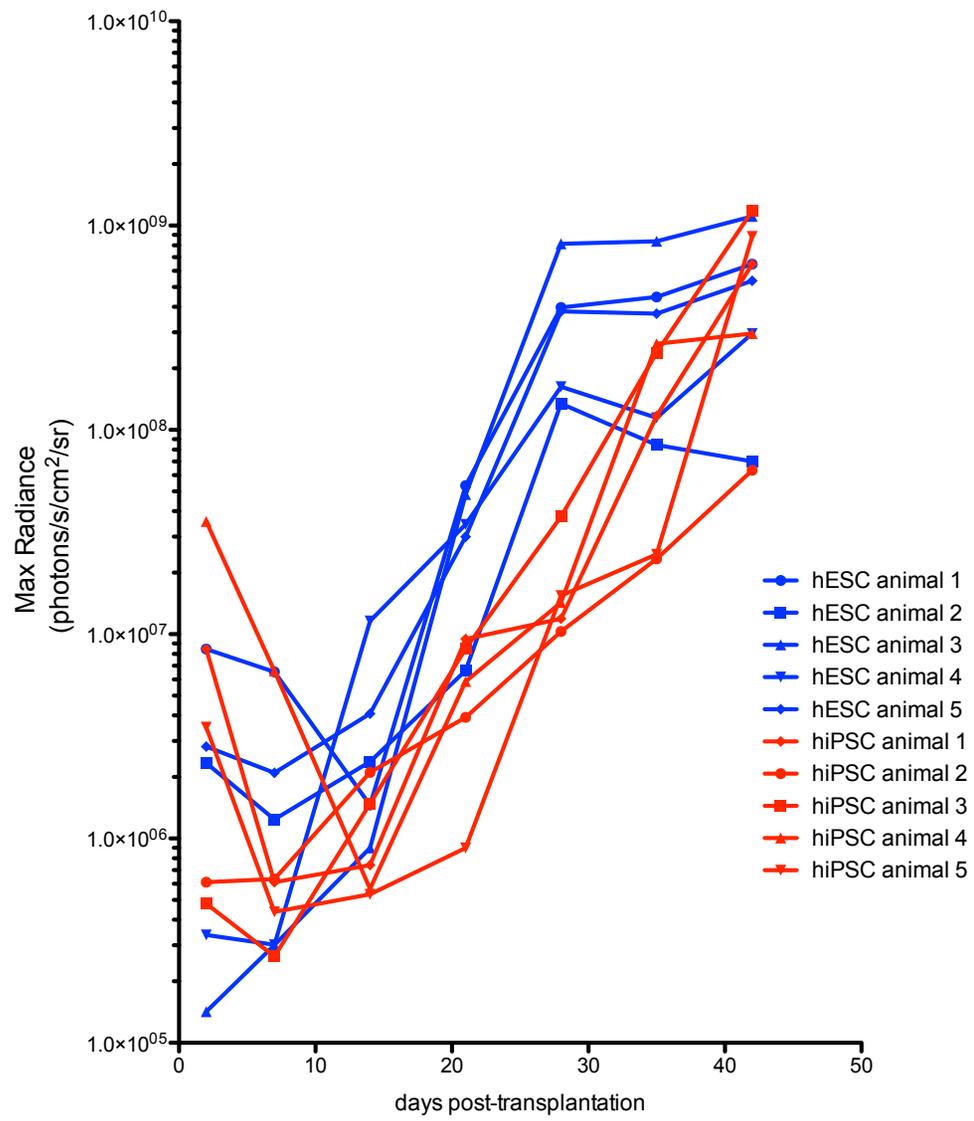
C



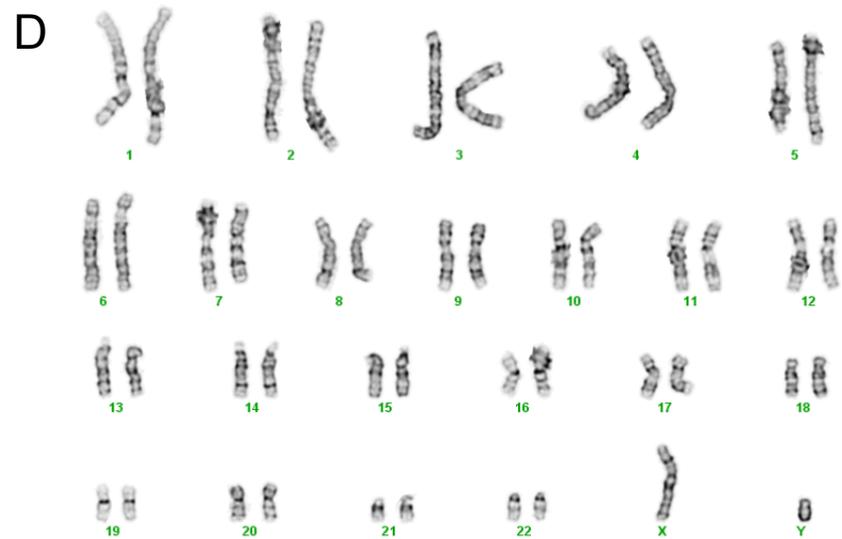
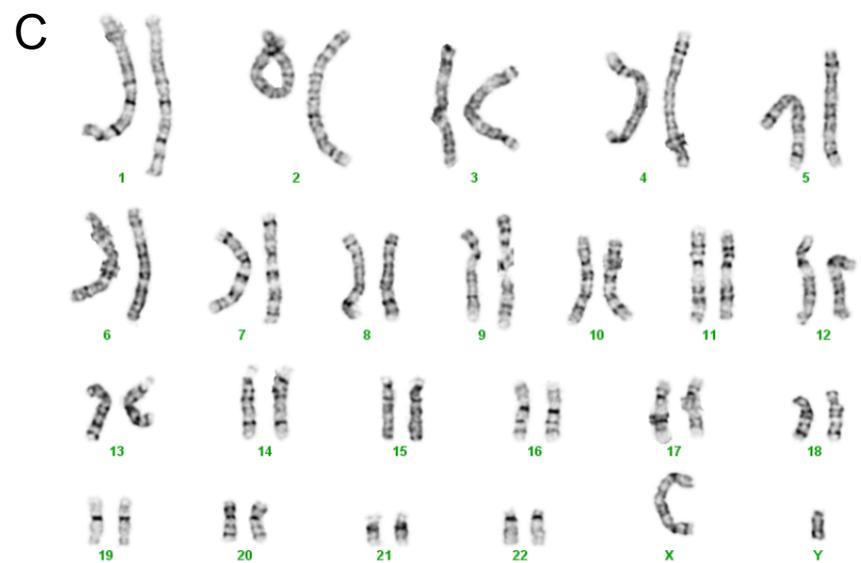
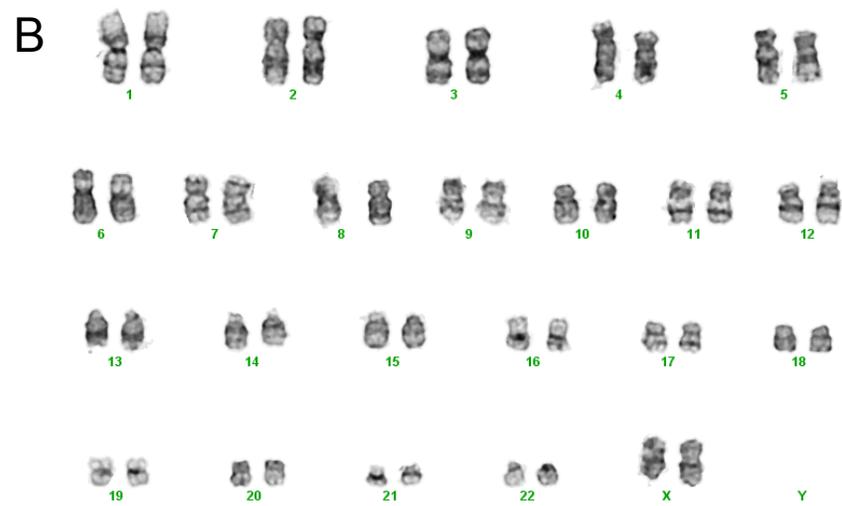
D



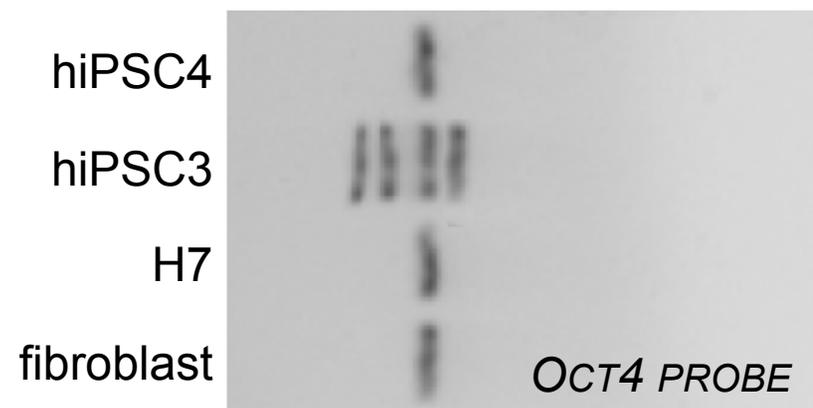
Supplemental Figure 12



Supplemental Figure 13



Supplemental Figure 14



Supplemental Figure 15

| Name | Cell source | Vector | Factors | Assays | | | | | Reference |
|--------|--------------------------|------------|---------------------------|-----------|-----------------------|--------|----|----------|--------------|
| | | | | karyotype | immunohisto-chemistry | RT-PCR | EB | teratoma | |
| hiPSC1 | IMR90 | Lentivirus | Oct4, Sox2, Nanog, Lin-28 | normal | ✓ | ✓ | ✓ | ✓ | Yu et al. |
| hiPSC2 | hASCs | Lentivirus | Oct4, Sox2, c-Myc, Klf4 | normal | ✓ | ✓ | ✓ | ✓ | Sun et al. |
| hiPSC3 | adult dermal fibroblasts | Lentivirus | Oct4, Sox2, c-Myc, Klf4 | normal | ✓ | ✓ | ✓ | ✓ | Present work |
| hiPSC4 | hASCs | Minicircle | Oct4, Sox2, Nanog, Lin-28 | normal | ✓ | ✓ | ✓ | ✓ | Jia et al. |

Supplemental Table 1

| Gene | Variance | | P-value |
|--------|---------------|----------------|---------|
| | Amongst hESCs | Amongst hiPSCs | |
| POU5F1 | 0.0140 | 0.0496 | <0.0001 |
| DNMT3B | 0.0176 | 0.0492 | <0.0001 |
| SOX2 | 0.0138 | 0.0613 | <0.0001 |
| NANOG | 0.0220 | 0.1216 | <0.0001 |
| LIN-28 | 0.0431 | 0.0753 | <0.0003 |
| MYC | 0.0275 | 0.0707 | <0.0001 |
| NR6A1 | 0.0133 | 0.0588 | 0.0007 |
| GDF3 | 0.0178 | 0.0581 | 0.0007 |
| LEFTY1 | 0.0345 | 0.0747 | 0.0067 |
| LEFTY2 | 0.0257 | 0.0869 | 0.0070 |

Supplemental Table 2