bar: +/-s.e.m). D) Images showing quarters of Tra-1-60 stained reprogramming plates upon infection of a non-targeting control and two lincRNA-ST8SIA3 targeting shRNAs. Arrowheads mark Tra-1-60+ iPSC colonies. E) Structure of the lincRNA-ST8SIA3 locus. Green, Blue: Demarcation of the H3K4me-H3K36me domain in ESCs. Red: Structure of lincRNA-ST8SIA3 RNA; asterisk marks position of OCT4/SOX2/NANOG binding (see Figure 3A). Right: Northern hybridization of lincRNA-ST8SIA3 detects a 2.6kb transcript in hFib2-iSP5, but not dH1f (full-length blot see Supplementary Figure 12). F) qRT-PCR verifies lincRNA-ST8SIA3 overexpression from a retroviral vector (pBabe-lincRNA-ST8SIA3) compared with pBabe-puro and pBabe-puro-GFP vectors in dH1f relative to the levels in H9 ESCs and hFib2-iPS5 (n=2, error bars: +/-s.e.m). G) Quantification of Tra-1-60^{+} iPSC colonies upon overexpression of lincRNA-ST8SIA3 compared to pBabe and pBabe-GFP controls (n=5, error bar: +/-s.e.m.). H) Quantification of cell numbers on day 6-7 in lincRNA-ST8SIA3 overexpressing cells and controls. Cell numbers are relative to the pBabe control (day 28 +/-2; n=5, error bar: +/-s.e.m.). I) Image of quarter-plates of Tra-1-60 stained colonies (arrowheads) in pBabe, pBabe-GFP, and pBabe-lincRNA-ST8SIA3 infected samples. Statistical analysis was performed using Student's t-test.

Supplementary Figure Legends

Supplementary Figure 1

Characterization of iPS cell lines used in this study. A) Semi-quantitative RT-PCR (30 cycles) to document silencing of retroviral transgenes in iPSC lines. BG01, MSC: ESC and fibroblast controls, respectively, for detection of pluripotency markers and absence of transgenes. B) qRT-PCR documents the total mRNA levels of reprogramming factor mRNAs of OCT4, SOX2, KLF4, and MYC in iPSC lines in comparison with the levels in ESC lines (H1, H9, BG01). mRNA levels are depicted relative to the levels in H1 ESCs. C) Fluorescence immunostaining shows upregulation of SSEA4 and Tra-1-60 surface markers, and absence of transgene expression as reflected by GFP in iPSCs. DNA was counterstained with Hoechst 33342.

Supplementary Figure 2

Differentiation potential of fibroblast-derived iPSC lines. A) qRT-PCR to document the downregulation of pluripotency markers *OCT4* and *NANOG*, as well as upregulation of differentiation markers of all three germ layers on day 6 of EB differentiation relative to undifferentiated controls. B) Hematoxyolin and Eosin (H+E) stained examples of teratomas showing tissues derived from all three germ layers obtained from iPSC injected into the femoral muscle of Rag2γc mice.

Supplementary Figure 3

Gene Set Enrichment Analysis (GSEA[1]) shows that all iPS cell lines used in this study were significantly enriched for human ESC signature genes and depleted of genes normally underexpressed in ESCs. Gene sets were taken from [2] and [3], and GSEA was performed using the Gene Set Enrichment Analysis Software (Broad Institute, Cambridge, USA), using 2500 permutations of gene sets (A: Genes specifically expressed in ESCs; B) Genes overexpressed in ESCs relative to other tissues; C) Genes underexpressed in ESCs relative to other cell types; D) Genes characteristic for ESCs taken from [3].

Supplementary Figure 4

Expression analysis of lincRNAs in $CD34^+$-derived iPSCs. A) Unsupervised clustering of lincRNA expression segregates $CD34^+$ cells (red) from ESCs and $CD34^+$-derived iPSCs (blue). B) Supervised clustering analysis identifies 152 lincRNAs that were differentially expressed between ESCs/iPSCs and $CD34^+$ cells (>2fold, FWER<0.05). Expression values are represented in shades of red and blue relative to being above (red) or below (blue) the median expression value across all samples (log scale 2, from -3 to +3).

Supplementary Figure 5

Heatmaps from Figure 2A showing lincRNAs with elevated expression in fibroblast-derived iPSCs (left) and $CD34^+$-derived iPSCs relative to hESCs (>2fold, FWER<0.05). The chromosomal coordinates of differentially expressed lincRNA segments are

indicated, and segments corresponding to the 10 overlapping iPSC-elevated lincRNAs are highlighted in green. Expression values are represented in shades of red and blue relative to being above (red) or below (blue) the median expression value across all samples (log scale 2, from -3 to +3).

Supplementary Figure 6

ChIP analysis of iPSC-enriched lincRNA loci in BJ1-iPS2 cells. Top: lincRNA loci demarcated by histone H3K4me3 and H3K36me3 domains (green and blue, respectively). Bottom: ChIP in BJ1-iPS2 cells followed by qPCR analysis detects binding of OCT4, SOX2, and NANOG within lincRNA-SFMBT2, lincRNA-VLDLR, and lincRNA-ST8SIA3 regions. ChIP enrichment values are displayed normalized to a control region (chr12: 7,839,777- 7,839,966; hg18).

Supplementary Figure 7

Validation of siRNA mediated knock-down of OCT4 in hFib2-iPS5 cells and analysis of lincRNA expression upon differentiation of H9 cells. A) Western Blot showing reduction of OCT4 protein levels upon siRNA transfection. Lane 1: hFib2-iPS5 cells, no siRNA transfected; lane 2: hFib2-iPS5 cells, transfected with GAPDH control siRNA; lane 3: hFib2-iPS5 cells, transfected with OCT4 siRNA; lane 4: MRC5 fibroblasts as negative control. B) Cell morphology changes upon knock-down of OCT4 (right) compared to a control siRNA transfected sample (left). C) Changes in iPSC-enriched lincRNA levels upon siRNA-mediated knock-down of OCT4 in H9 ESCs. Left: qRT-PCR of *OCT4* and *NANOG* transcript levels upon depletion of OCT4. Right: qRT-PCR of iPSC-enriched lincRNA levels upon depletion of OCT4. Transcript levels are displayed relative to non-targeting control siRNAs (ctrl siRNA) (n=3, error bars +/-s.e.m). C) iPSC-enriched lincRNA expression during EB differentiation of H9 ESCs. Left: qRT-PCR analysis monitoring transcript levels of pluripotency markers (*OCT4* and *NANOG*) and the differentiation marker *LAMIN A/C* over a ten day differentiation time-course; right: qRT-PCR analysis of iPSC-enriched lincRNAs. RNA levels are depicted relative to undifferentiated cells (n=2, error bars +/-s.e.m).

Supplementary Figure 8

Effect of lincRNA knock-down on reprogramming. A) qRT-PCR to document the efficiency of lincRNA-SFMBT2 knock-down following infection with shRNA-expressing lentivirus relative to a non-targeting control (n=4, error bar +/- s.e.m). B) qRT-PCR shows that lincRNA-SFMBT2 and lincRNA-ST8SIA3 are not expressed in dH1f fibroblasts prior to reprogramming (n=2, error bar +/- s.e.m.). Expression values are depicted relative to the levels in H9 cells. C) Quantification of iPSC colonies derived from dH1f cells upon knock-down of lincRNA-SFMBT2 relative to a non-targeting control (n=2, error bar +/- stdv).

Supplementary Figure 9

Characterization of iPSCs derived from lincRNA-ST8SIA3 knock-down and overexpression reprogramming experiments. A) qRT-PCR to document the expression of reprogramming factor mRNAs; blue: total (transgenic and endogenous) levels of OCT4, SOX2, KLF4, and MYC mRNAs, red: endogenous levels of OCT4, SOX2, KLF4, and MYC mRNAs in H9 ESCs, dH1f fibroblasts, and fibroblast-derived iPSC lines derived from the non-targeting control-sh knock-down (ctrl-sh iPS1), the lincRNA-ST8SIA3 targeting knock-down (linc-sh iPS2), the control experiment with the GFP-expressing pBabe vector (pBabe-GFP iPS1), and the lincRNA-ST8SIA3 overexpression experiment (pBabe-lincRNA-ST8SIA3 iPS3). Expression levels are depicted relative to the levels in H9 ESCs (log scale). B) Top: Bright field images of iPSC colonies. Bottom: live-cell imaging of the same fibroblast-derived iPSC lines to document activation of the SSEA4 and Tra-1-60 cell surface markers. Hoechst 33342 was used as a DNA counter-stain. C) qRT-PCR to document downregulation of OCT4 and NANOG mRNAs, as well as upregulation of differentiation markers GATA4 and AFP (endoderm), RUNX1 and GATA2 (mesoderm), and NCAM (ectoderm) in day 6 EBs derived from ctrl-sh iPS1, linc-sh iPS2, pBabe-GFP iPS1, and pBabe-lincRNA-ST8SIA3 iPS3 cells. Expression values are represented relative to undifferentiated controls.

Supplementary Figure 10

Knock-down or overexpression of lincRNA-ST8SIA3 in dH1f fibroblasts does not affect cell growth. A) Cumulative cell numbers measured over a 14 day time-course after infection of dH1f with a non-targeting control-shRNA expressing virus (ctrl-sh) and a lincRNA-ST8SIA3 targeting virus (linc-sh1). No statistically significant difference in cell growth could be observed (n=4, error bars +/-s.e.m.; P=0.92). B) Cumulative cell numbers measured over a 14 day time-course after infection of dH1f with empty pBabe virus, pBabe-GFP, or lincRNA-ST8SIA3 overexpression virus (pBabe-lincRNA-ST8SIA3). No statistically significant difference in cell growth could be observed (n=2, error bars +/-s.e.m.; P=0.79 (compared with pBabe) and 0.81 (compared with pBabe-GFP)). Statistics were performed using Student's t-test.

Supplementary Figure 11

Live-cell imaging to follow the kinetics of reprogramming. A) Quantification of SSEA4+/Tra-1-60+ cell clusters from day 6 to day 16 of reprogramming in the control (ctrl-sh) and lincRNA-ST8SIA3 knock-down (linc-sh1) sample. Four regions per 6well (each region with an area of 53.21mm^2) were quantified for each control and linc-sh1, and average numbers of cell clusters per area are plotted on the y-axis. Error bars represent +/- stdv. B) Representative subregion of SSEA4 (red) and Tra1-60 (purple) stained reprogramming cells. White circles mark SSEA4+/Tra-1-60+ cell clusters identified on day 16.

Supplementary Figure 12

Northern Blot of lincRNA-ST8SIA3 in dH1f and hFib2-iPS5 cells detects a 2.6kb band. Left: ethidium bromide stained agarose gel to show 28S and 18S ribosomal RNA loading controls, as well as the position of RNA marker bands. Right: Blot of lincRNA-ST8SIA3 using a specific probe (see Supplementary Methods) confirms its 2.6kb length.

Supplementary Figure 13

siRNA-mediated knock-down of lincRNA-ST8SIA3 affects ESC and iPSC growth. A) qRT-PCR validation of three lincRNA-ST8SIA3-targeting siRNAs (Linc-siRNA1, -siRNA2, -siRNA3) compared with two non-targeting controls (ctrl siRNA1 and ctrl

siRNA2) as well as non-transfected cells in H9 ESCs (left) and hFib2-iPS5 cells (right). RNA levels are depicted relative to the levels in cells transfected with ctrl siRNA1 ($n \geq 3$, error bar +/-s.e.m.). B) Cell growth of H9 ESCs (left) and hFib2-iPS5 cells (right) upon knock-down of lincRNA-ST8SIA3 compared to non-targeting controls as measured by total cell numbers 72h post-transfection. Cell numbers are depicted relative to the control (ctrl siRNA1). C) Percentage of early apoptotic cells ($AnnexinV^+/Propidium\ iodide^-$) measured by flow cytometry in non-transfected cells, cells transfected with non-targeting control siRNAs, and lincRNA-ST8SIA3-targeting siRNAs in H9 ESCs (left) and hFib2-iPS5 cells (right). Boxes represent the 25^{th} to 75^{th} percentiles of the data, the central line marks median values. Whiskers extend to the most extreme data points. Statistical analysis was performed using Student's t-test.

Supplementary Figure 14

siRNA-mediated knock-down of p53 partially rescues growth deficiency and elevated cell death of H9 cells caused by lincRNA-ST8SIA3 knock-down. A) qRT-PCR documents the efficiency of lincRNA-ST8SIA3 (top) and p53 (bottom) knock-down 72h after siRNA transfection. Expression values are depicted relative to the non-targeting control (ctrl siRNA 1); n=2, error bars +/-s.e.m. B) Relative cell numbers 72h post-transfection in non-transfected cells, or cells transfected with lincRNA-ST8SIA3 targeting siRNA (linc siRNA 1), p53-targeting siRNA (p53 siRNA), or both compared with control-transfected cells (ctrl siRNA 1); n=2, error bars +/-s.e.m. C) Flow cytometric analysis of early apoptotic (AnnexinV+/PI-) and late apoptotic/necrotic (AnnexinV+/PI+) cells 72h following knock-down of lincRNA-ST8SIA3, p53, or both, compared with the non-targeting control and untransfected cells; n=2, error bars +/-s.e.m.

Online Methods

All primer, siRNA, northern probe, cDNA and cloning sequences are listed in Supplemental table 4.

Microarray Analysis

Total RNA was isolated using RNA-Stat60 (Tel-Test) and DNase treated (DNAfree, Ambion). For protein-coding gene expression analysis, total RNA was hybridized to Affymetrix_U133_Plus_2.0 chips and processed as described[10]. For lincRNA expression analysis, total RNA was amplified using Message Amp II (Ambion), labeled and hybridized to lincRNA arrays as described[11].

Statistical Analysis

Affymetrix gene expression arrays were normalized as described[11]. Differentially expressed genes were identified using Student's t-test (two-tailed, two-sample equal variance). LincRNA microarray probe intensities were quantile normalized, log transformed, and significantly enriched lincRNA regions were identified as described[11]. Differentially expressed lincRNAs (Supplementary Table 1) were identified using Student's t-test and significance was estimated using 1000 permutations of class labels to control for a familywise error rate (FWER<0.05).

Unsupervised and supervised hierarchical clustering of gene and lincRNA expression profiles were performed using GenePattern[30].

To compute the correlation between lincRNAs and neighboring protein-coding genes, we computed a Pearson correlation coefficient for each lincRNA to both its left and right neighboring gene across the full datasets. We then permuted gene locations and computed the same correlation coefficient for each lincRNA against randomized gene neighbors. We performed 1000 permutations and assessed the statistical significance of this interaction by comparing the observed scores to the randomly permuted scores.

Statistical significance of overlapping iPSC-enriched lincRNAs in fibro-iPSCs and $CD34^{+}$-iPSCs: Random simulations were used to calculate the probability of obtaining an overlap as large as we identified, while controlling for the set size of both differentially expressed and upregulated lincRNAs in each iPSC-type.

Statistical analysis of iPSC colony yield: Each sample was normalized to the total number of iPSC colonies within one experiment to weigh out the variations in colony numbers across experiments. The resulting values representing fractions of colony numbers within each experiment were then used for statistical analysis using Student's t-test (two-tailed, two-sample equal variance).

Protein-coding genes deregulated upon lincRNA-ST8SIA3 knock-down. Affymetrix gene expression arrays were normalized as described[6]. PaGE analysis was applied to identify genes that are differentially expressed in lincRNA-ST8SIA3 siRNA knock-down iPSC while comparing with control samples[7] (FDR< 0.2, default parameters). Gene Set Enrichment Analysis was performed using 100 permutations of gene sets and a t-test as a test statistics (FDR < 0.2).

qRT-PCR

cDNA was synthesized with Superscript II (Invitrogen) and qPCR performed using the BrilliantSYBRGreenQPCRmix. Relative expression values were calculated (ΔΔCT method) using GAPDH or β-ACTIN as normalizer.

Immunostaining

Cells were fixed with 4% p-formaldehyde and stained with biotin-anti-Tra-1-60 (eBioscience, #13-8863-82) and streptavidin-HRP (Biolegend, #405210) diluted in PBS/3% FCS/0.3% Triton X-100. Staining was developed with the Vector labs DAB kit (#SK-4100), and iPSC colonies quantified with ImageJ software.

Cell culture, siRNA transfection

iPSCs/hESCs were cultured and EB differentiation in suspension was performed as described[7]. For transfections or infections, iPSCs were dissociated with Accutase and plated in mTeSR media (StemCellTechnologies) with 10μM Y-27632 (Calbiochem) at 35k-50k cells/24well or 100k/6well pre-coated with matrigel (BD Bioscience #345277). 100nM OCT4-, GAPDH-, or non-targeting siRNAs were transfected using DharmaFECT1 (Dharmacon #T-2001-01). See Supplement for siRNA sequences.

ChIP assays

ChIP was performed as described[31]. Chromatin extracts were immunoprecipitated using anti-OCT4 (Santa Cruz, sc-8628), anti-SOX2 (R&D Systems, AF-2018), anti-NANOG (R&D Systems, AF-1997), or anti-GFP (Santa Cruz, sc-9996) control. Fold enrichments were calculated by determining the ratio of immunoprecipitated DNA to input, and normalizing to the levels observed at a control region. See supplement for sequences.

Reprogramming assays

Reprogramming infections were performed as described[7]. For knock-down studies, dH1f were infected with reprogramming virus alongside lentivirus expressing non-targeting control (SHC002V), or shRNAs targeting lincRNAs. shRNAs were designed using the iRNAi software, cloned into pLKO.1-puro (Addgene), and verified by sequencing. See Supplement for sequences. After infection, cells were grown for 6 days in αMEM/10% FCS, dissociated, counted, plated onto MEFs. From 24h later, cells were cultured in hESC media; colonies were scored between days 21-28. For overexpression studies, lincRNA-ST8SIA3 cDNA was cloned vial EcoRI into pBabe-puro (Addgene #1764) and verified by sequencing. dH1f were infected twice with pBabe-puro, pBabe-puro-GFP, and pBabe-lincRNA-ST8SIA3 retrovirus, and 1μg/ml puromycin were added 48h later. Puromycin was removed for 48h prior to reprogramming infections.

Cloning of lincRNA-ST8SIA3

5' and 3' RACE were performed using the First Choice RLM-RACE kit (Ambion). See Supplement for primer sequences. PCR products were cloned into pCR4-TOPO using the ZeroBlunt PCR TOPO cloning kit (Invitrogen). Resulting pCR4-lincRNA-ST8SIA3 clones were sequenced. LincRNA-ST8SIA3 cDNA was isolated from pCR4-lincRNA-ST8SIA3 via EcoRI digest and cloned into pBabe-puro. Resulting pBabe-lincRNA-ST8SIA3 clones were sequenced.

Northern Blot

Total RNA was isolated from dH1f and hFib2-iPS5 cells and genomic DNA removed (see above). 20μg RNA per sample were resolved on a 1% denaturing agarose gel.

Following blotting using the Turbo Blotting System (Whatman), the membrane was UV-crosslinked and pre-hybridized for 1h at 42°C inULTRAhyb® Ultrasensitive Hybridization Buffer (Ambion, #AM8669). Radiolabeled probe directed against lincRNA-ST8SIA3 was hybridized overnight. The membrane was washed 3x5min in 2xSSC/0.1% SDS at 42°C, then 3x15min in 0.1xSSC/0.1% SDS. To generate radiolabeled probe, a 231bp fragment of lincRNA-ST8SIA3 was PCR-amplified from pCR4-lincRNA-ST8SIA3. 50ng PCR fragment was used to produce the probe using Ready-To-Go DNA Labelling Beads (GE Healthcare #27-9240-01).

Cells used for microarray expression analysis

Fibroblasts: MRC5 (fetal lung fibroblasts) passage (P) 8; MSC (mesenchymal stem cells derived from bone marrow) P9; hFib2 P13 (adult forearm fibroblasts); BJ1 P7 (neonatal foreskin fibroblasts[1]

$CD34^+$ cells: mobilized peripheral blood CD34+ cells (AllCells, mBP014F). Cells were thawed and cultured as described[2] for two days before RNA isolation.

hESCs: H1 (NIH code WA01) P29; H9 (WA09) P49; BG01 P35

iPSCs: MRC5-iPS7 P13; MRC5-iPS20 P14; MSC-iPS1 P13; MSC-iPS3 P10; hFib2-iPS4 P27; hFib2-iPS5 P21; BJ1-iPS2 P16[1]; CD34-iPS4 P9 ; CD34-iPS8 P9 [2]

Fluorescent Immunostaining and live-cell imaging

Immunostaining: Cells grown in 96well plates (Matrix #4940) coated with hESC-qualified matrigel (BD Biosciences #345277) were fixed for 20-30min with 4% p-formaldeyde / PBS (+/+), washed several times with PBS (+/+), and incubated overnight at 4°C with primary antibody and Hoechst diluted in 3% donkey serum / 3% BSA Fraction VII / 0.01% Triton X-100 / PBS (+/+) (Alexa 647-coupled anti-SSEA4 (1:100), BD Biosciences #560219; Alexa 555-coupled anti-Tra-1-60, BD Biosciences #560121 (1:75); Hoechst, Invitrogen #H3570 (1:20,000)). After several washes with PBS (+/+), images were acquired using a BD Pathway 435 imager equipped with a 10x objective. Areas corresponding to 18.6 mm^2 were imaged. Four images were acquired per frame (Hoechst, GFP, Alexa Fluor 555, Alexa Fluor 647). GFP acquisition settings were optimized for detection of high-level proviral GFP expression.

Live-cell imaging: Cells grown on MEFs in 6well plates were incubated with Alexa 647-coupled anti-TRA-1-60 (BD Biosciences #560122 , 1:75) and Alexa 555-coupled anti-SSEA4 (BD Biosciences #560218 , 1:100) for 2h at 37°C. Where applicable, Hoechst (Invitrogen #H3570, 1:20,000) was added after 1.5h for the remaining 30min. Cells were washed 3x with PBS before 1ml of fresh phenol red-free media were added per well and images were acquired using a BD Pathway 435 imager equipped with a 10x objective. Four areas corresponding to 53.21mm^2 were imaged per 6well. Two or three images were acquired per frame (Alexa Fluor 555, Alexa Fluor 647, with or without Hoechst). Post-acquisition image processing was performed using ImageJ (flatfield-correction, background subtraction; http://rsb.info.nih.gov.ezp-prod1.hul.harvard.edu/ij/) and/or Adobe Photoshop (pseudocoloring, multi-color composites).

Teratoma formation assay

iPSCs grown on matrigel were harvested with dispase (Invitrogen #17105-041, 1mg/ml in DMEM/F12). Cell clumps from one 6well were resuspended in 50µl DMEM/F12, 100µl collagen I (Invitrogen #A1064401), and 150µl hESC-qualified matrigel (BD Biosciences #354277). Cell clumps were then injected into the hind limb femoral muscles (100µl suspension per leg) of Rag2 γ/c mice. After 6-8 weeks, teratomas were harvested, fixed in 4% p-formaldehyde overnight. Samples were then embedded in paraffin, and sections were stained with hematoxylin/eosin (Rodent Histopathology Core, Harvard Medical School, Boston, MA, USA).

Western Blot

Cells were lysed on ice in PBS/1% Triton X-100 containing protease inhibitor cocktail (Roche). 40µg protein was loaded per well of a 10% SDS polyacrylamide gel (BioRad, #345-0011). OCT4 was detected using a monoclonal mouse anti-human OCT4 antibody (Santa Cruz Biotechnology, #sc-5279, 1:1,000), and the GAPDH loading control was detected using a rabbit antibody (Santa Cruz Biotechnology, #sc-255778, 1:1,000). Secondary antibodies used were horseradish peroxidase coupled rabbit or mouse antisera (GE Healthcare, #NA934V or NA931V, 1:5,000). Proteins were detected using the Amersham ECL detection kit as described by the manufacturer.

Flow cytometric analysis

iPSCs were dissociated to single cells with Accutase, washed in PBS and stained with Alexa488-coupled AnnexinV and propidium iodide according to the manufacturer's instructions (Invitrogen Vybrant Apoptosis Assay Kit #2, #V13241). Samples were analyzed using a FACScalibur flow cytometer and data was processed using FloJo Software.

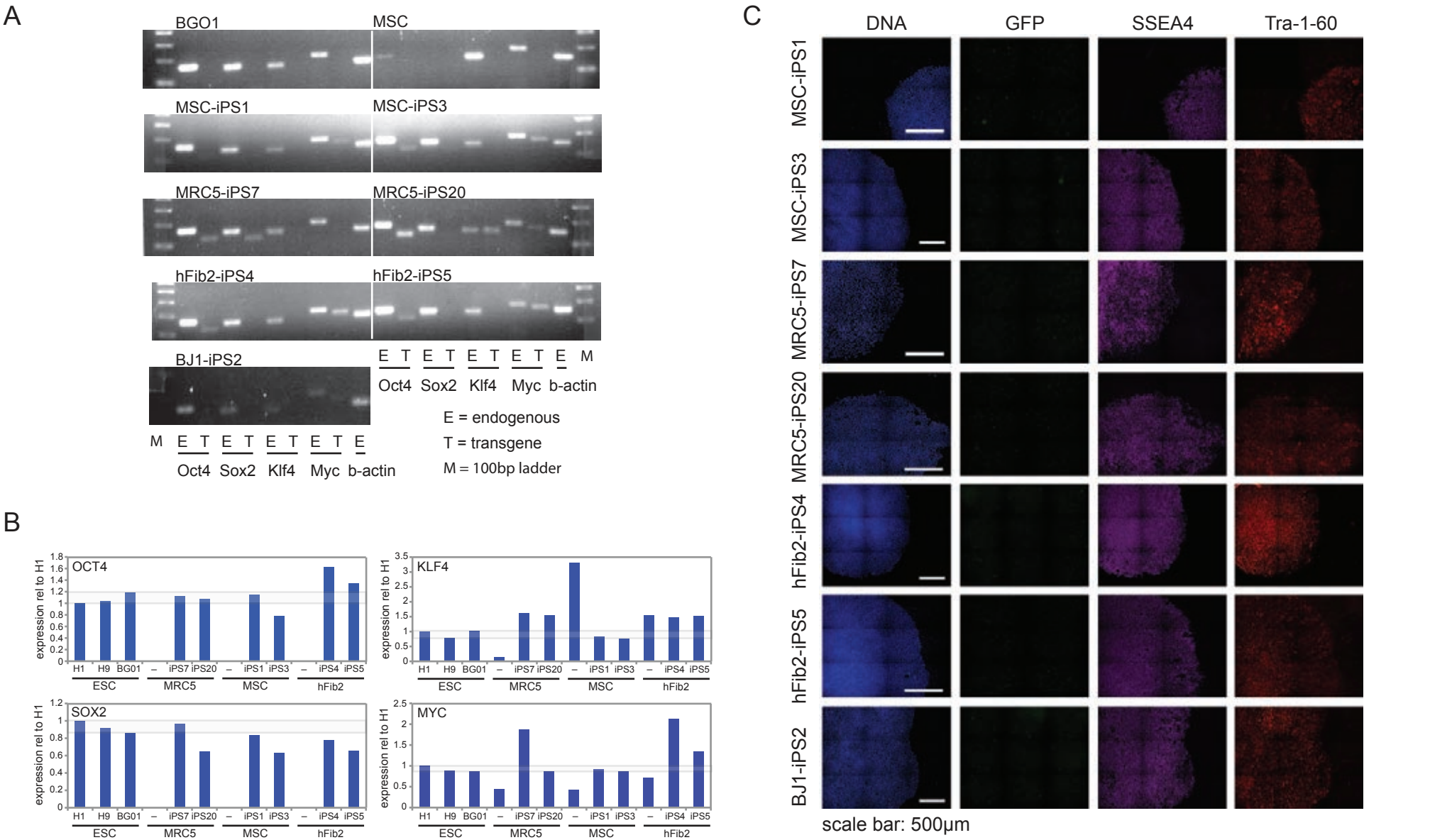
Production of viral supernatants

293T cells were plated at a density of $2.5x10^6$ cells per 10cm dish. The next day, cells were transfected with 2.5µg viral vector, 2µg Gag-Pol vector (pCMV-dR8.2 dvpr Addgene #8455, pUMVC Addgene #8449, or ps-PAX2 #12260), and 0.2µg VSV-G plasmid (pCMV-VSV-G Addgene #8454, or pMD2.G Addgene #12259) using 15µl Fugene 6 (Roche Applied Science #1181509001) in 50µl DMEM per plate. Supernatant was collected 48h and 72h post-transfection and filtered through 45µm pore size filters. For concentration, viral supernatants were centrifuged at 70,000g at 4C for 90min using a Beckman XL-90 ultracentrifuge. Reprogramming viruses were either retroviral[8] (MOI 2.5), or lentiviral[9] (Addgene #21162, 21164; 100µl supernatant).

Accession Numbers

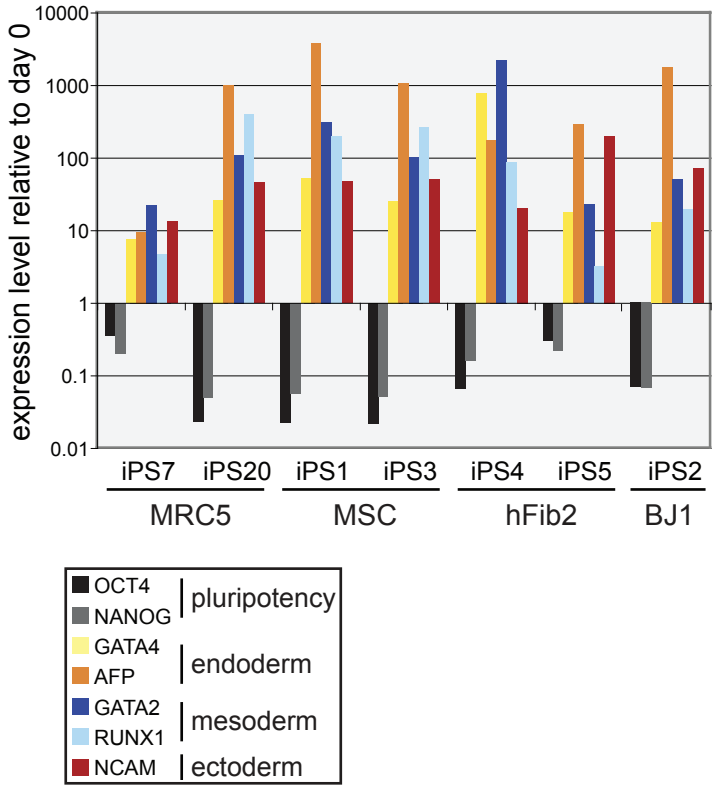
All primary data are available at the Gene Expression Omnibus (GSE24182).

Supplementary Figure 1

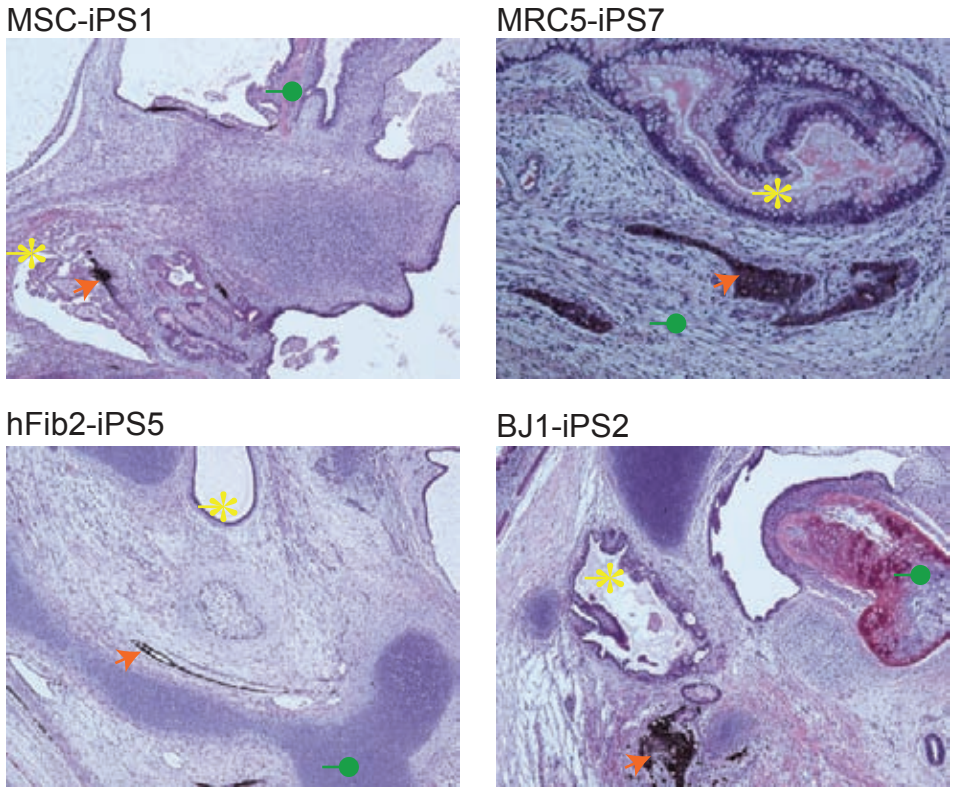


Supplementary Figure 2

A

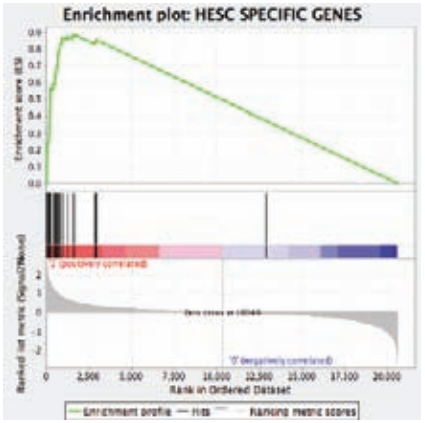


B

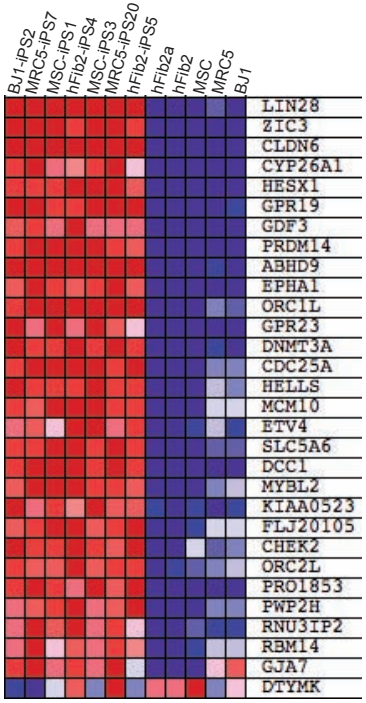


Supplementary Figure 3

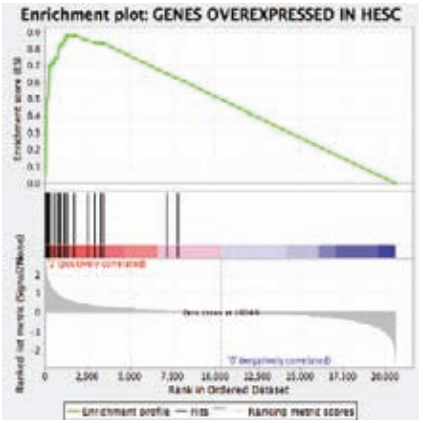
A



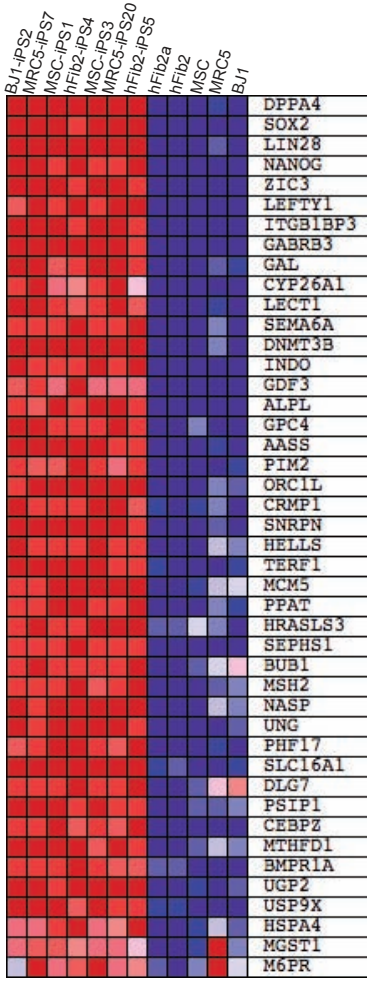
FDR < 0.0025
Normalized Enrichment Score (NES) 2.38



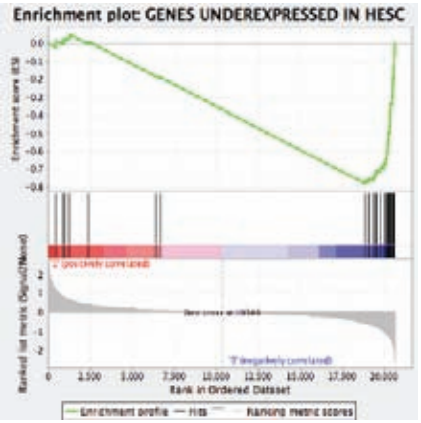
B



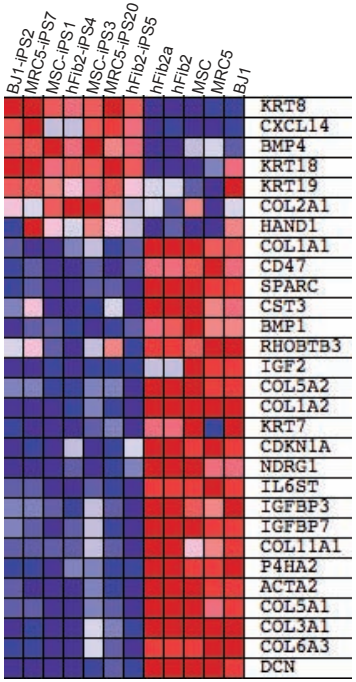
FDR < 0.0025
Normalized Enrichment Score (NES) 2.55



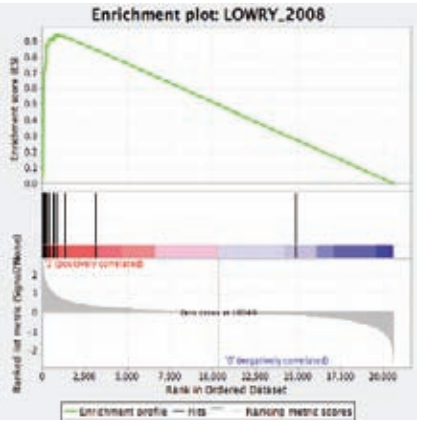
C



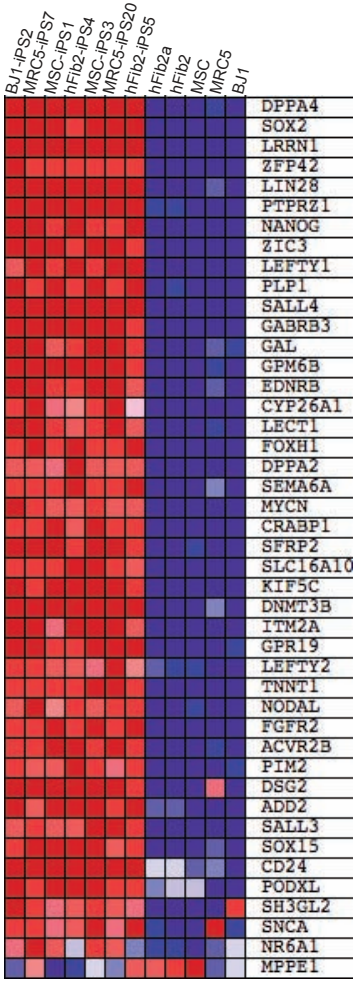
FDR < 0.0025
Normalized Enrichment Score (NES) -2.1



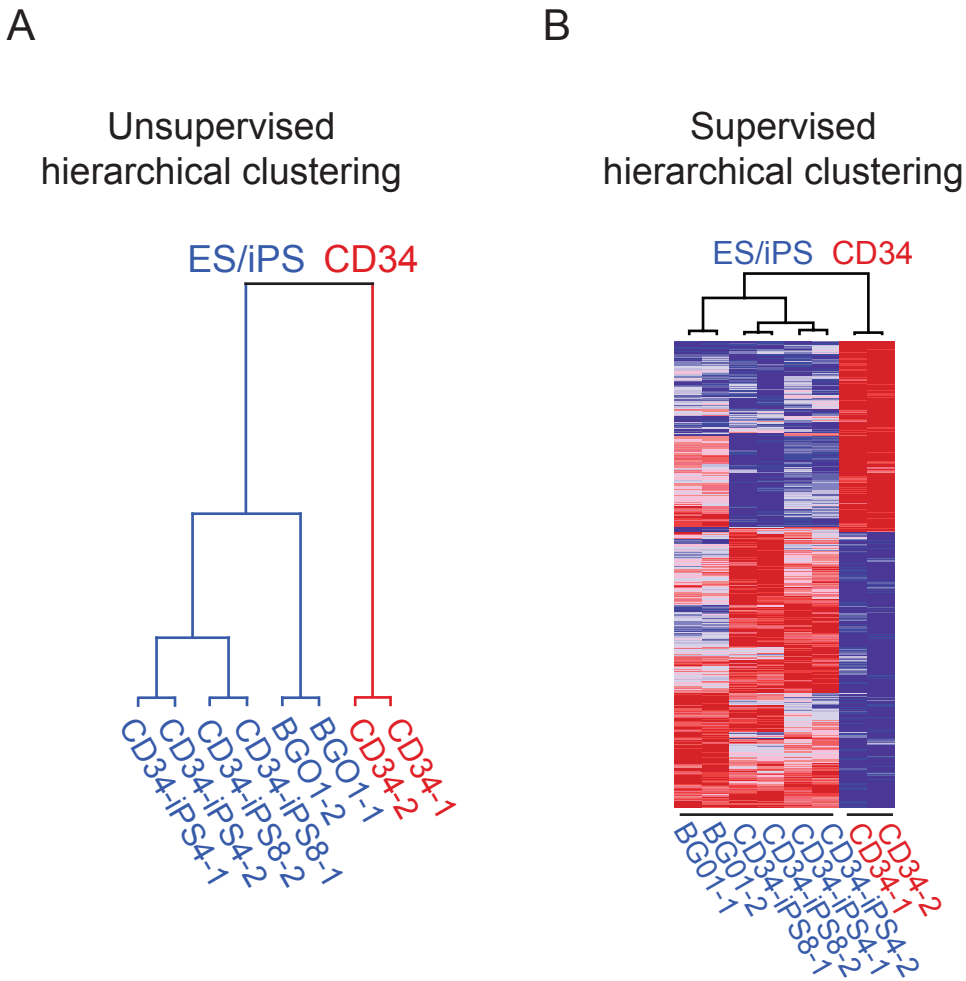
D



FDR < 0.0025
Normalized Enrichment Score (NES) 2.73

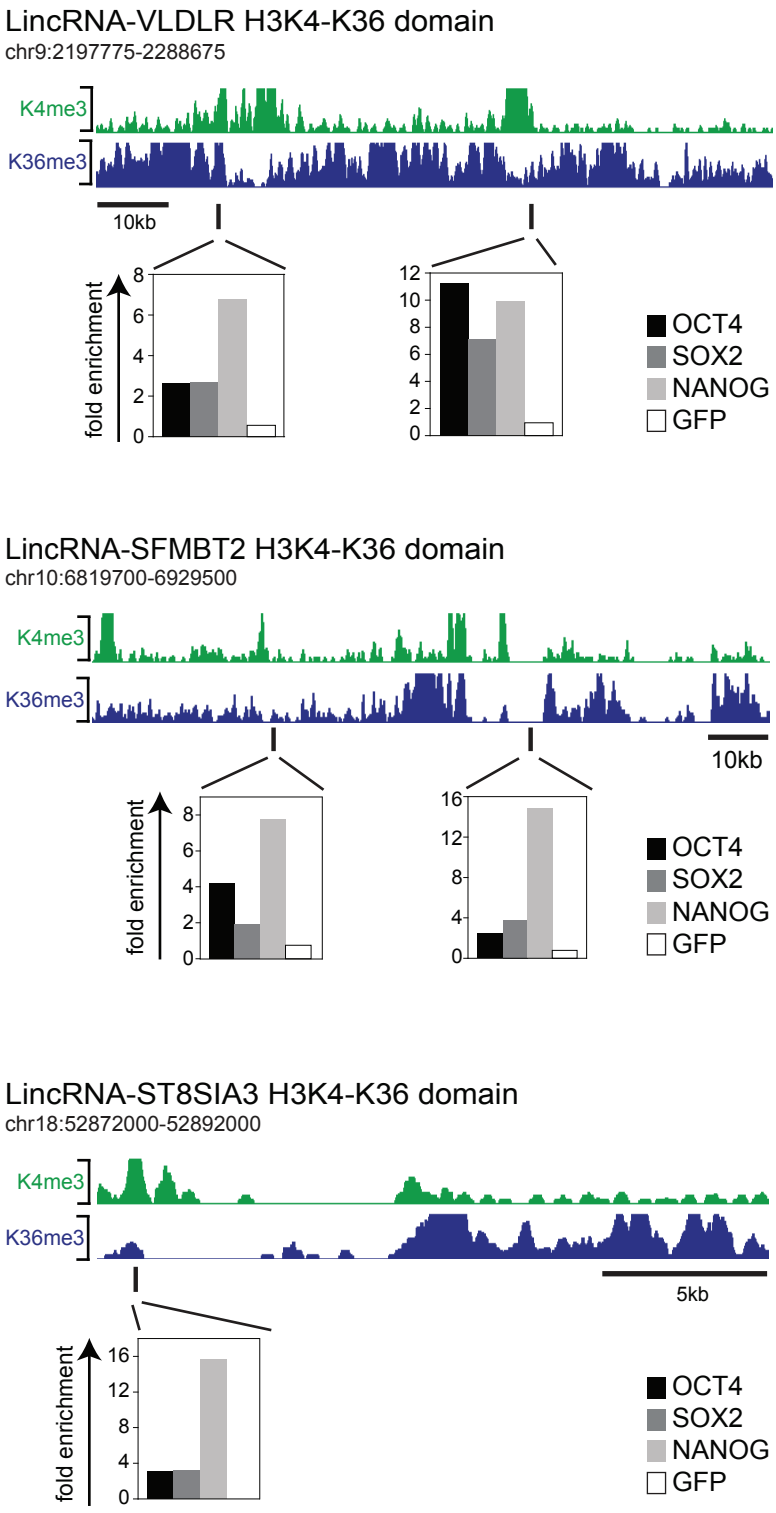


Supplementary Figure 4

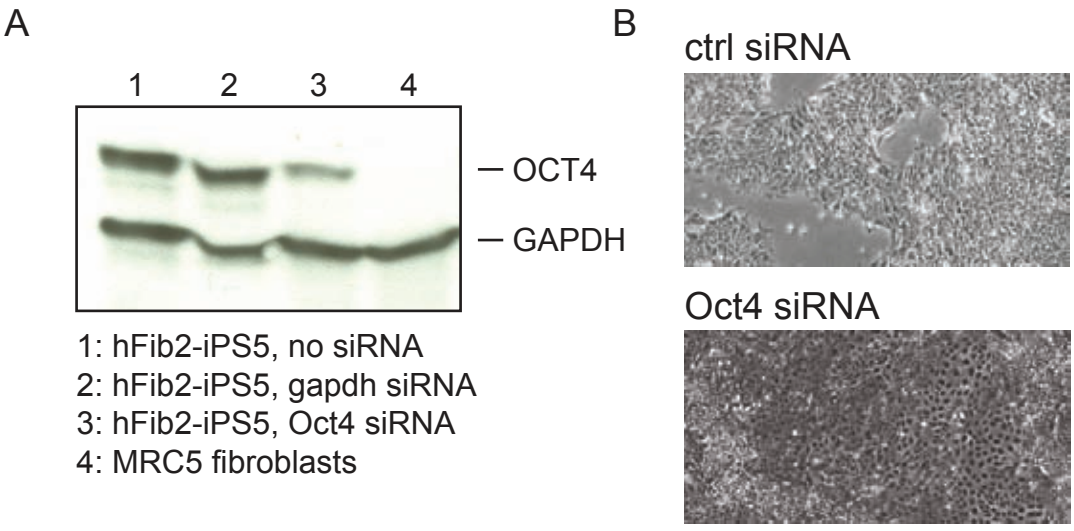


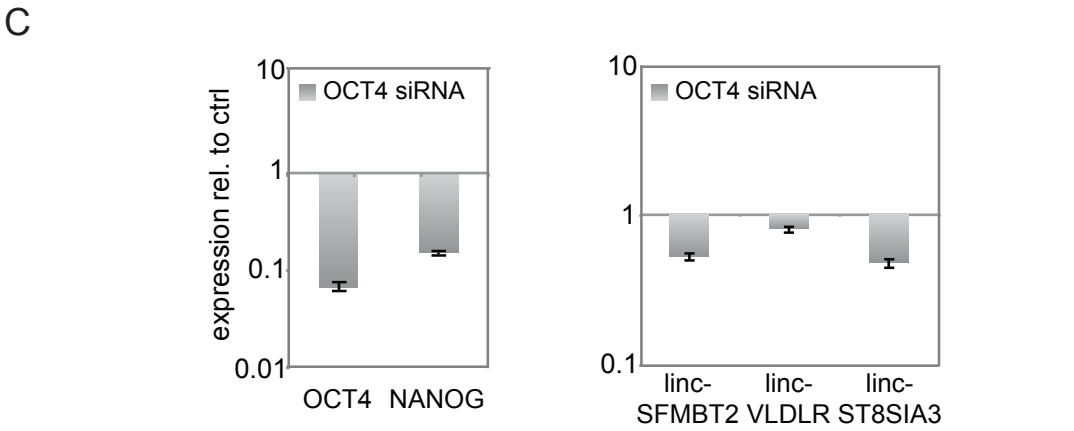
Supplementary Figure 5

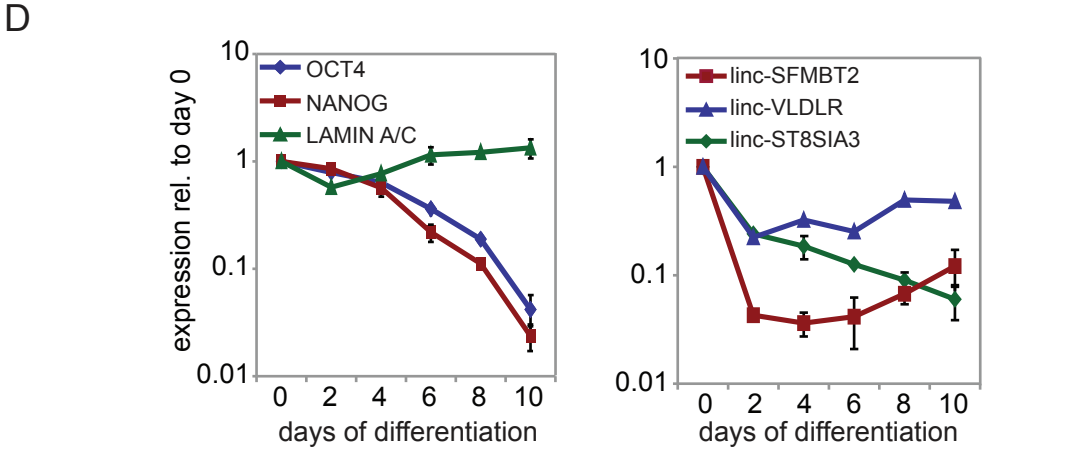
Supplementary Figure 6



Supplementary Figure 7

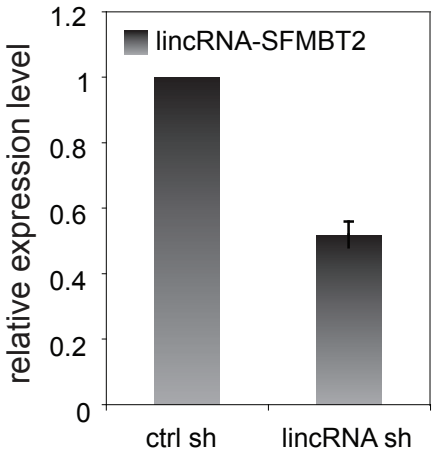




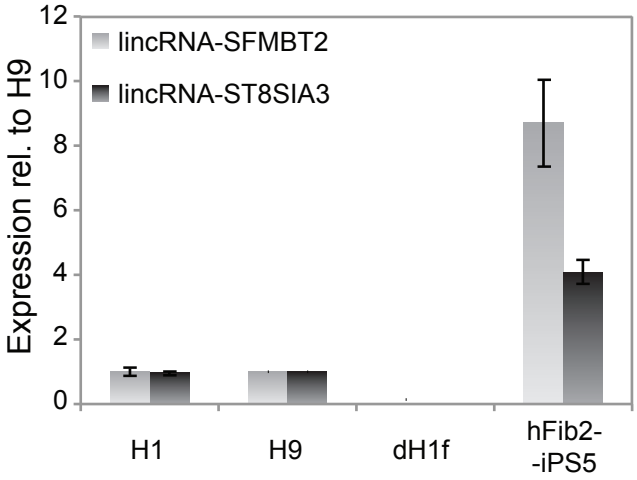


Supplementary Figure 8

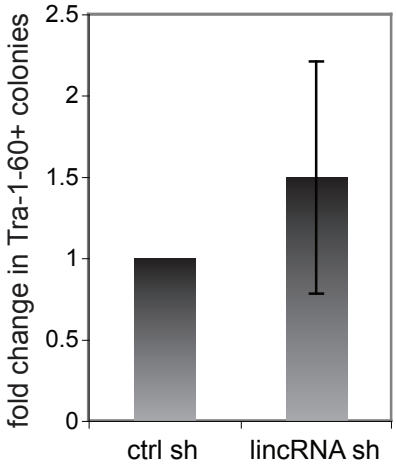
A



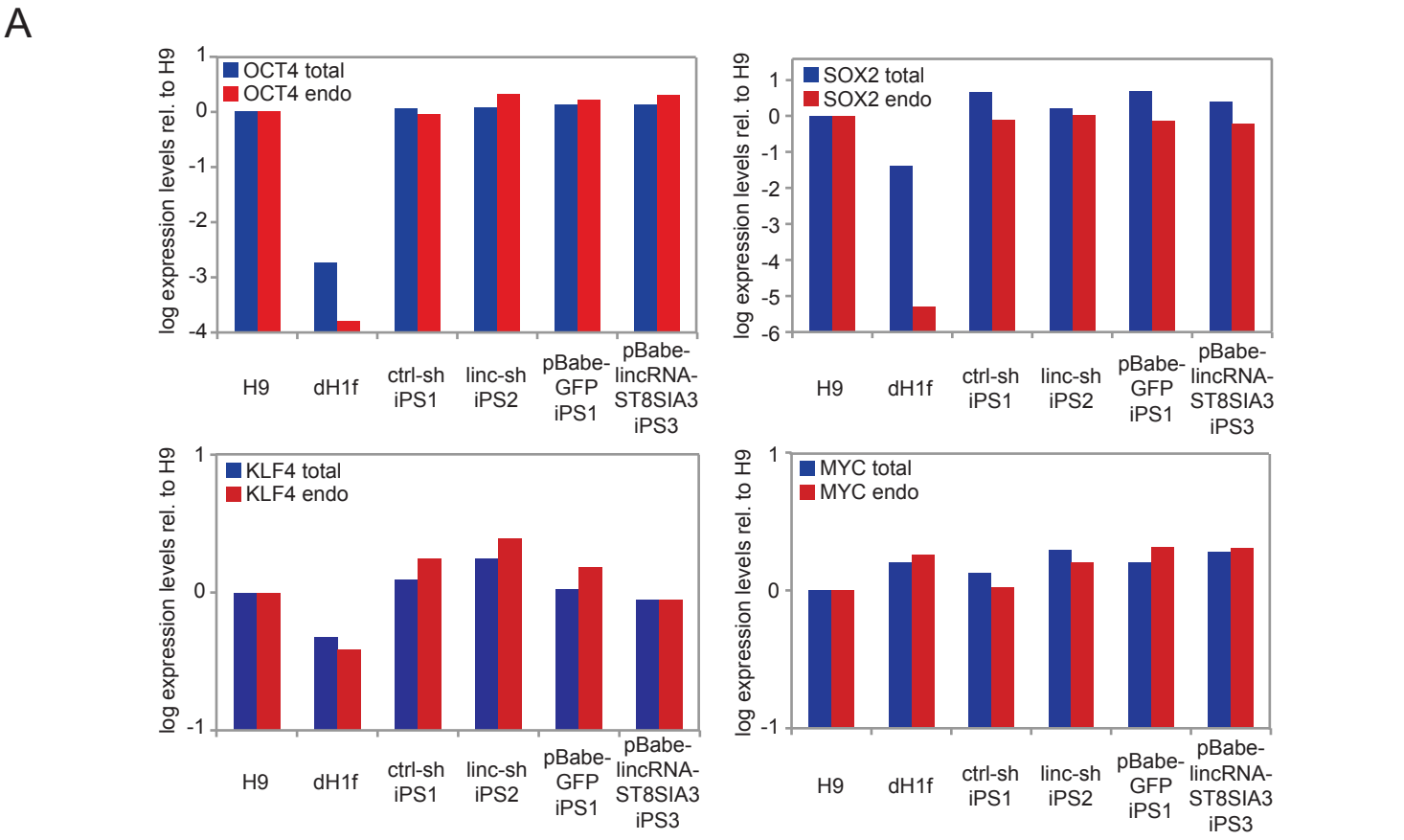
B

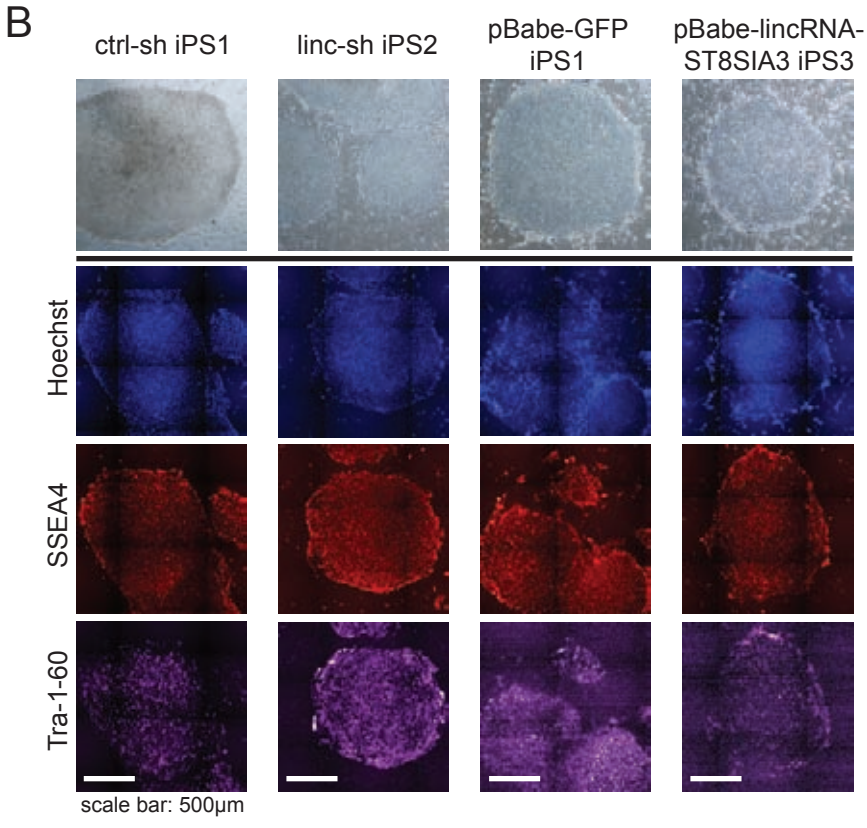


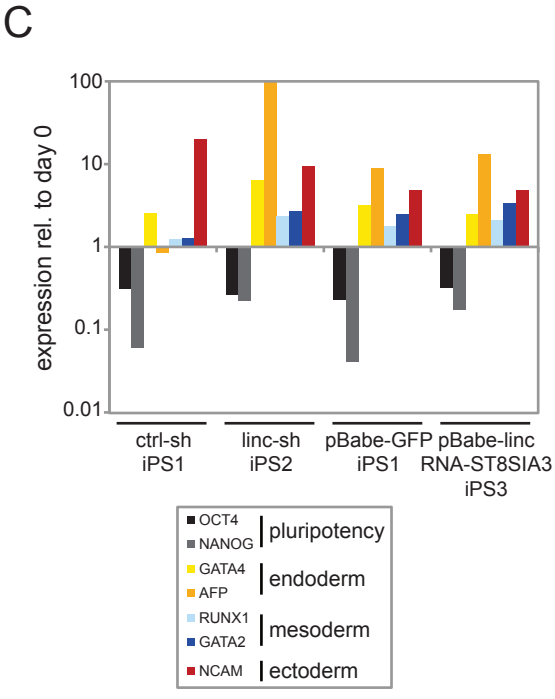
C



Supplementary Figure 9

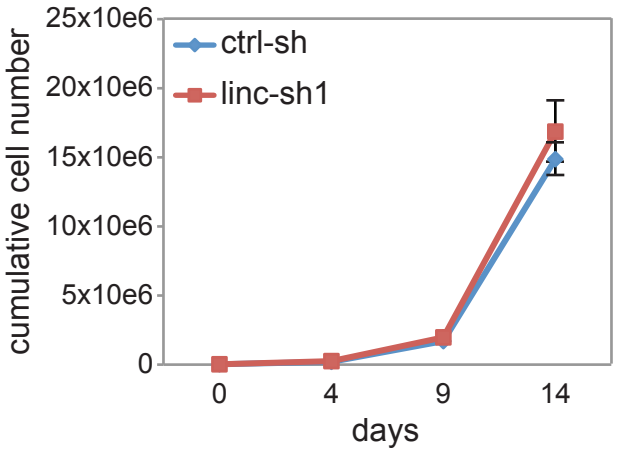




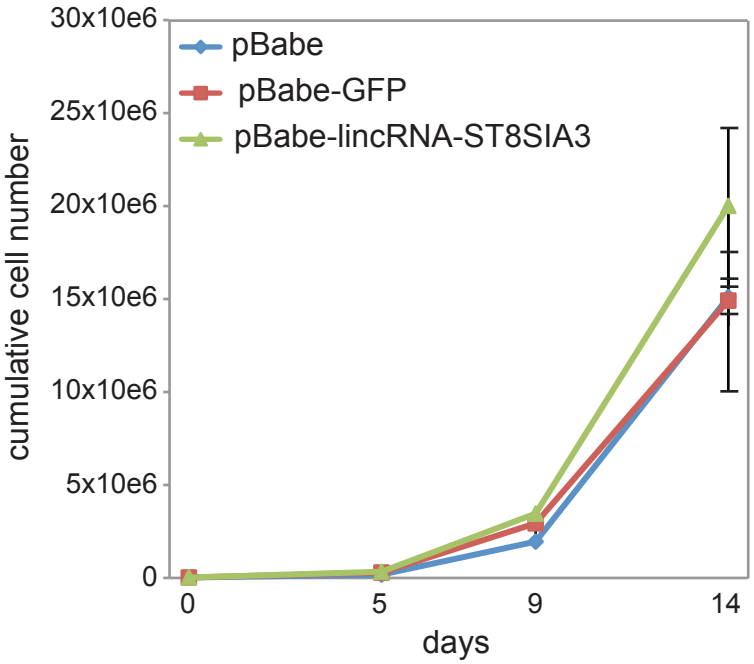


Supplementary Figure 10

A

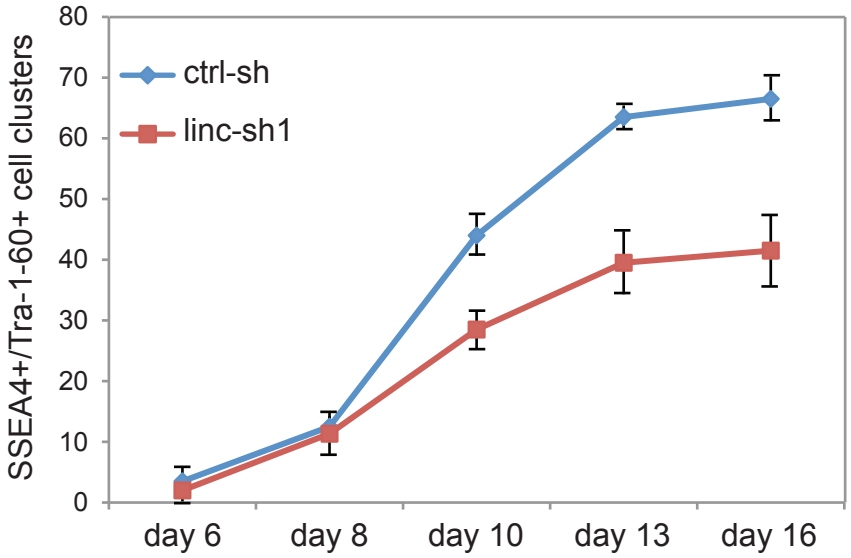


B

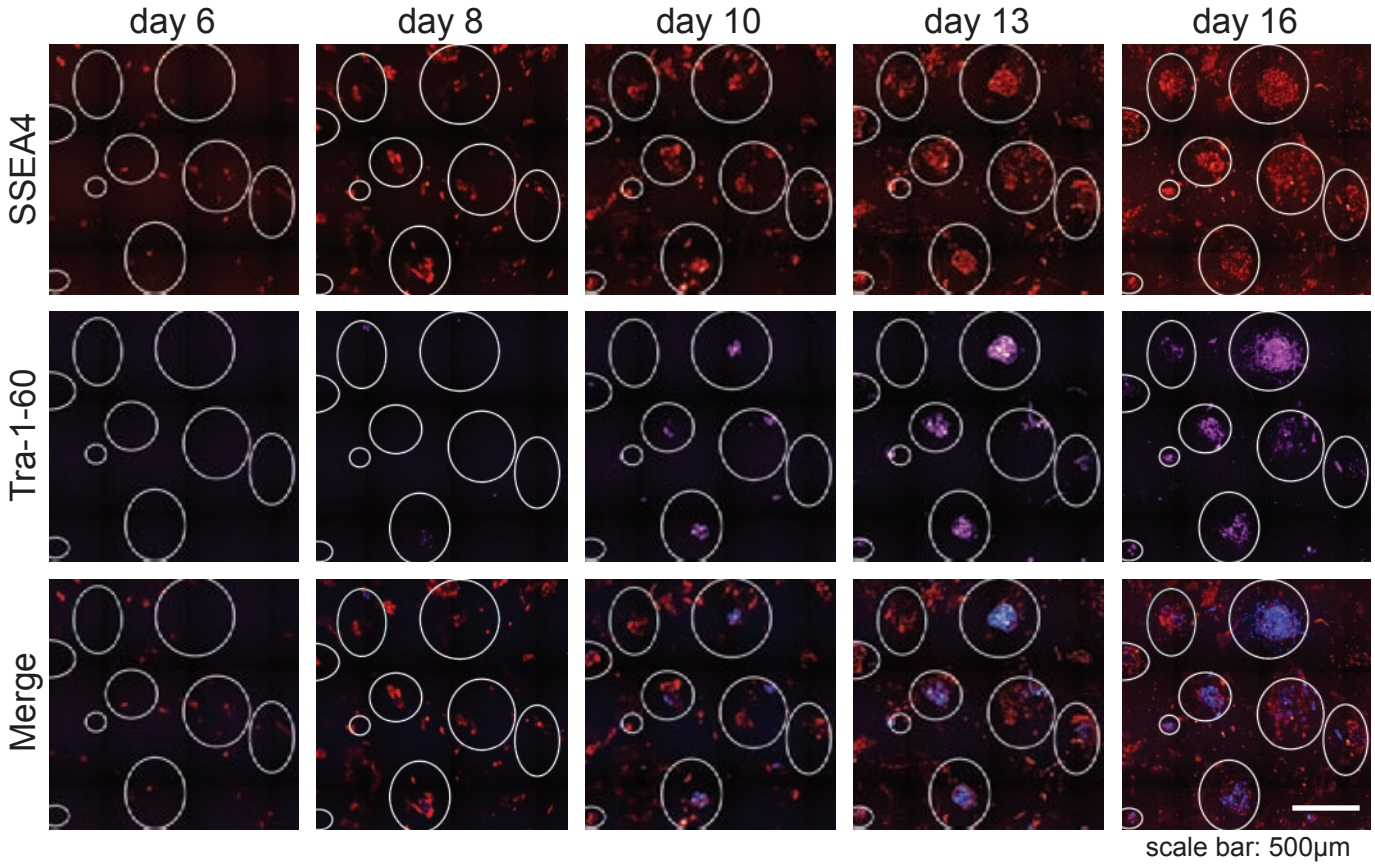


Supplementary Figure 11

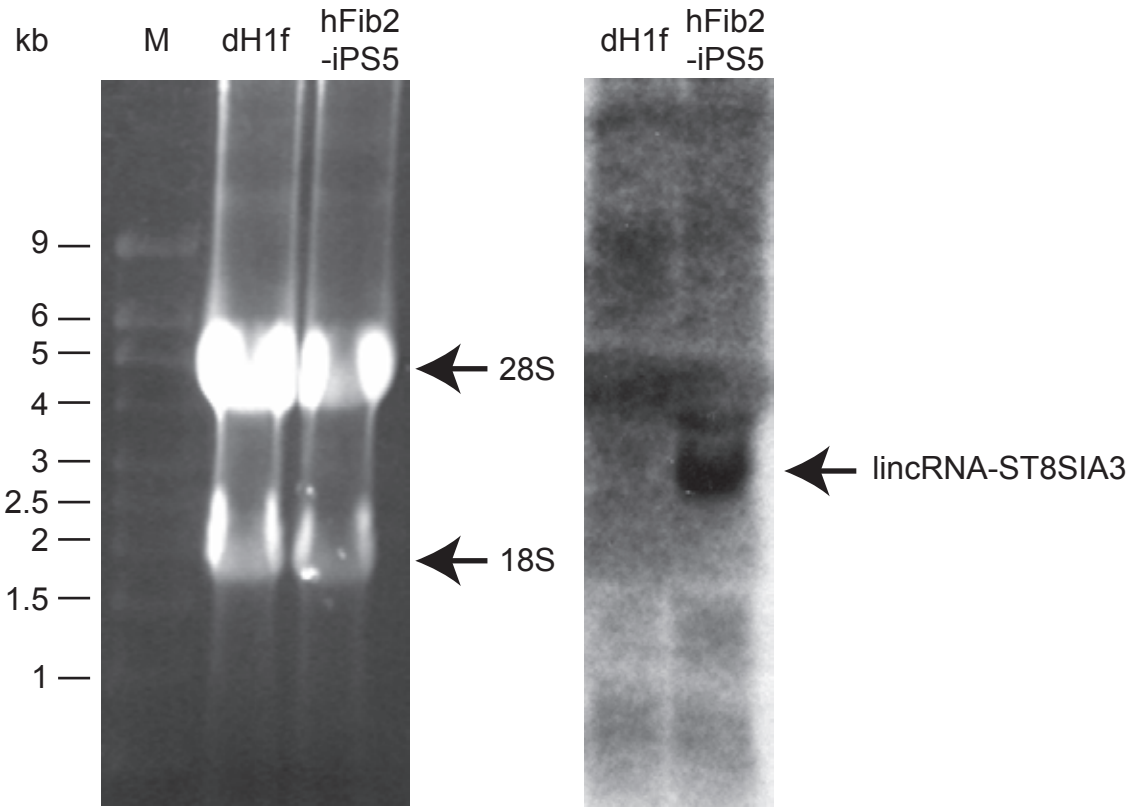
A



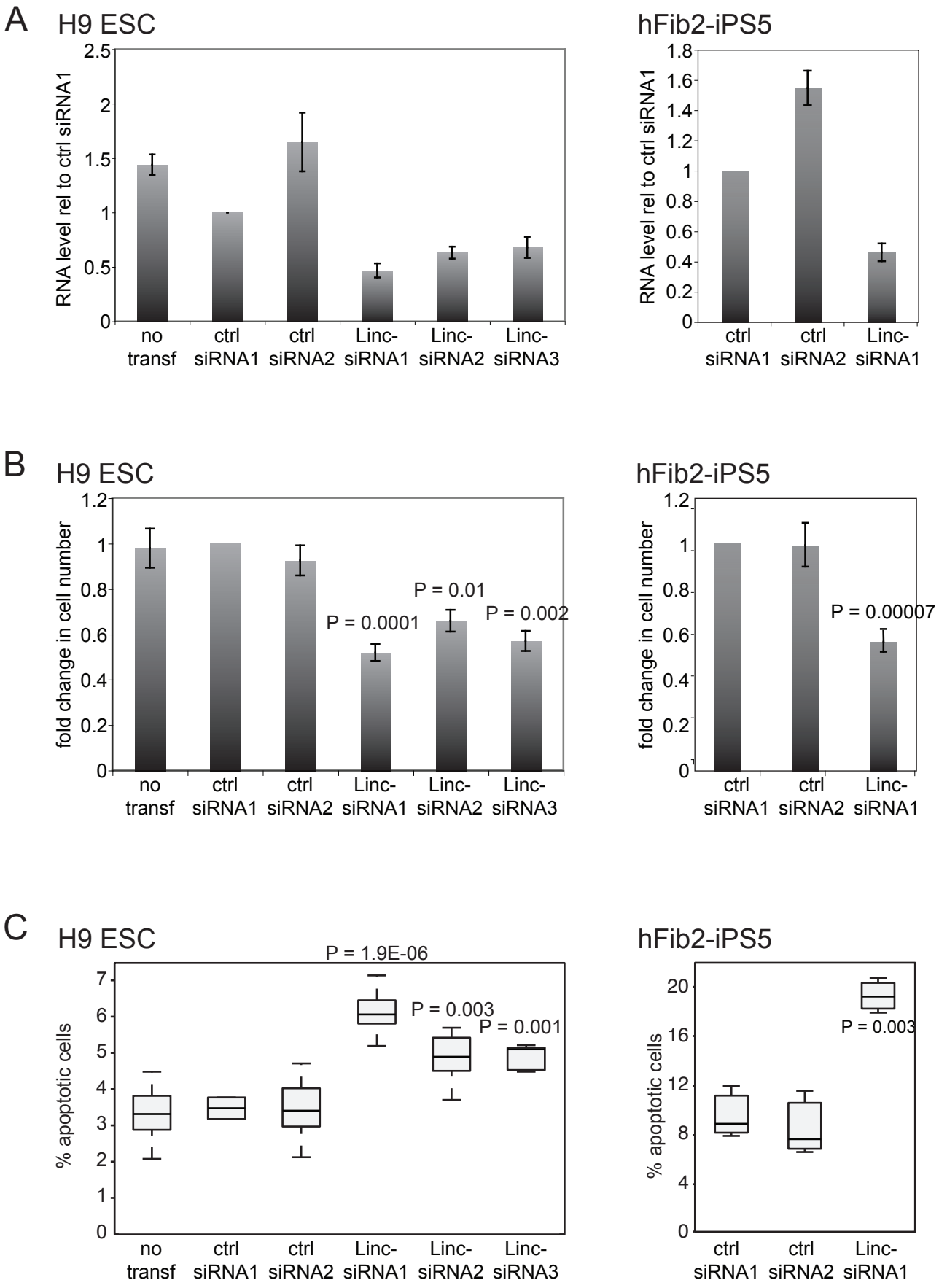
B



Supplementary Figure 12



Supplementary Figure 13



Supplementary Figure 14