Stem Cell Reports, Volume *2*

Supplemental Information

Single-Cell Gene Expression Profiles Define Self-Renewing, Pluripotent, and Lineage Primed States of Human Pluripotent Stem Cells

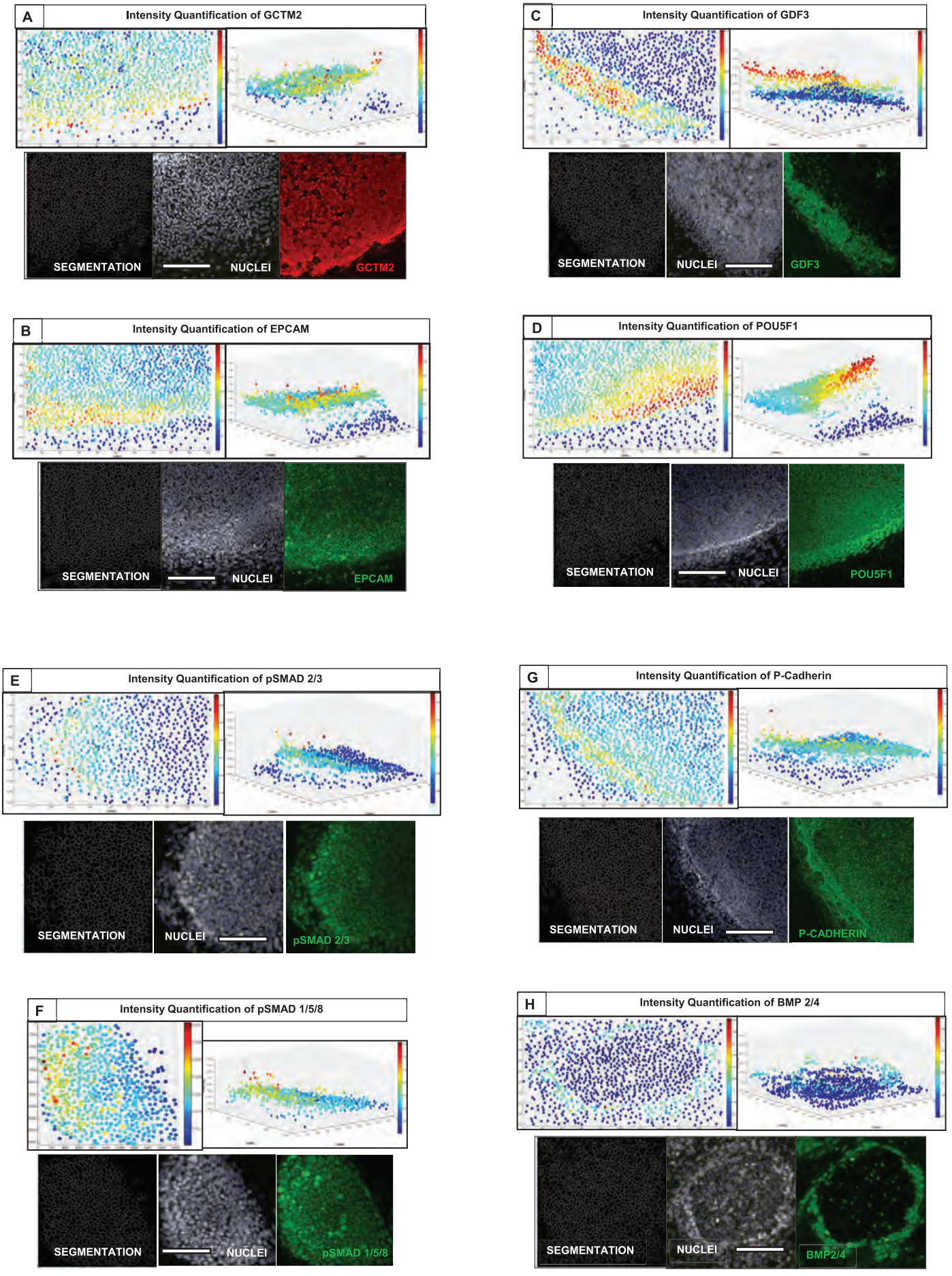
Shelley R. Hough, Matthew Thornton, Elizabeth Mason, Jessica C. Mar,

Hough et al. Supplemental Information

Inventory of Supplemental Information

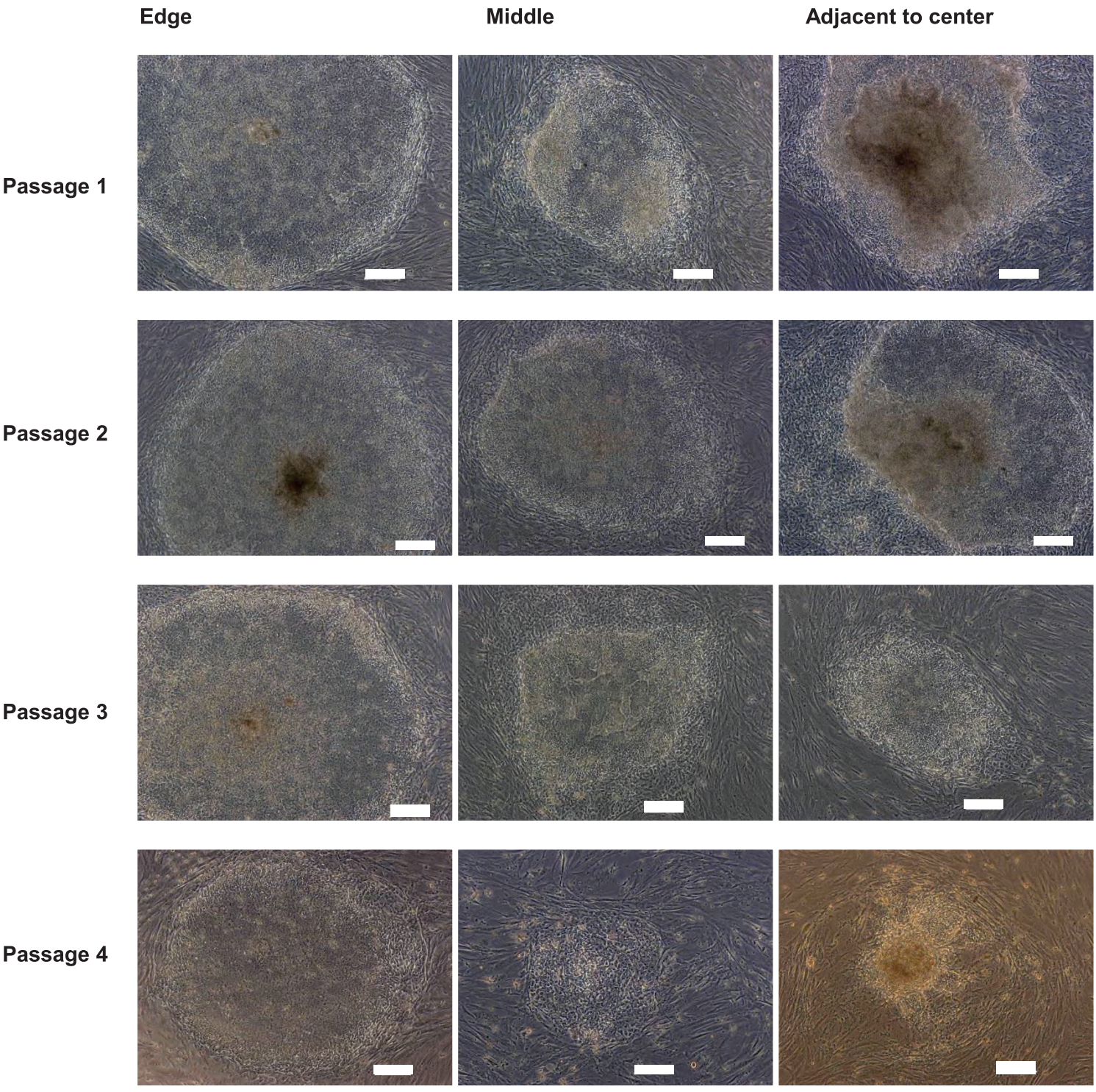
1. Figure S1 Related to Figure 1
2. Figure S2 Related to Figure 5
3. Figure S3 related to Figure 5
4. Figure S4 related to Figure 5
5. Figure S5 related to Figure 5
6. Supplemental Figure Legends
7. Table S1 related to Figures 1 2 and 3
8. Table S2 related to Figure 5
9. Table S3 related to Figures 1 and 5
10. Supplemental Experimental Procedures
11. Supplemental References

Figure S1



Scale bars= 200 microns

Figure S2



Scale bars=200 micron

Figure S3

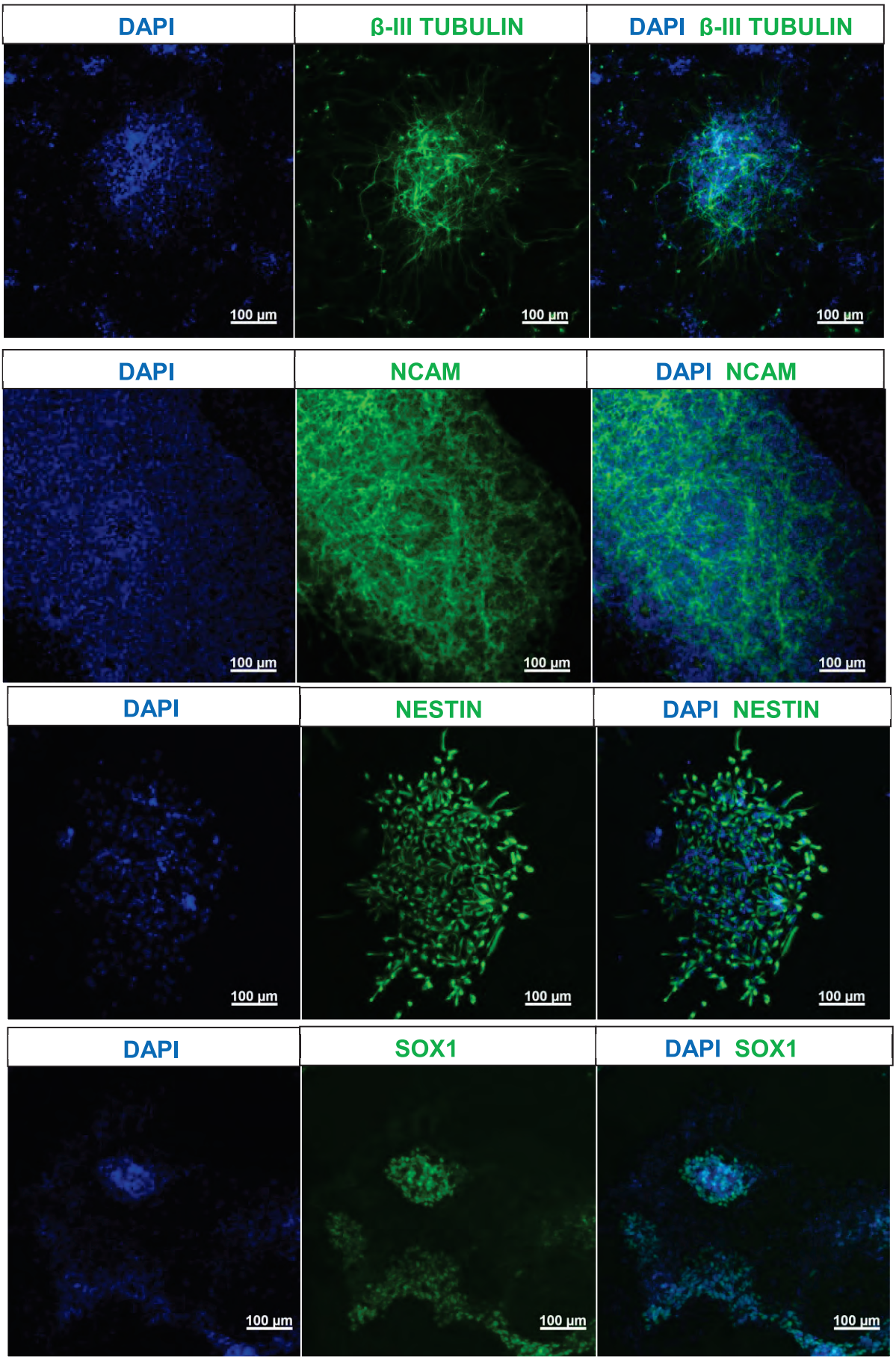

A
DAPI
ß-III TUBULIN
DAPI ß-III TUBULIN
100 µm
B
DAPI
NCAM
DAPI NCAM
100 µm
C
DAPI
NESTIN
DAPI NESTIN
100 µm
D
DAPI
SOX1
DAPI SOX1
100 µm

Figure S4.

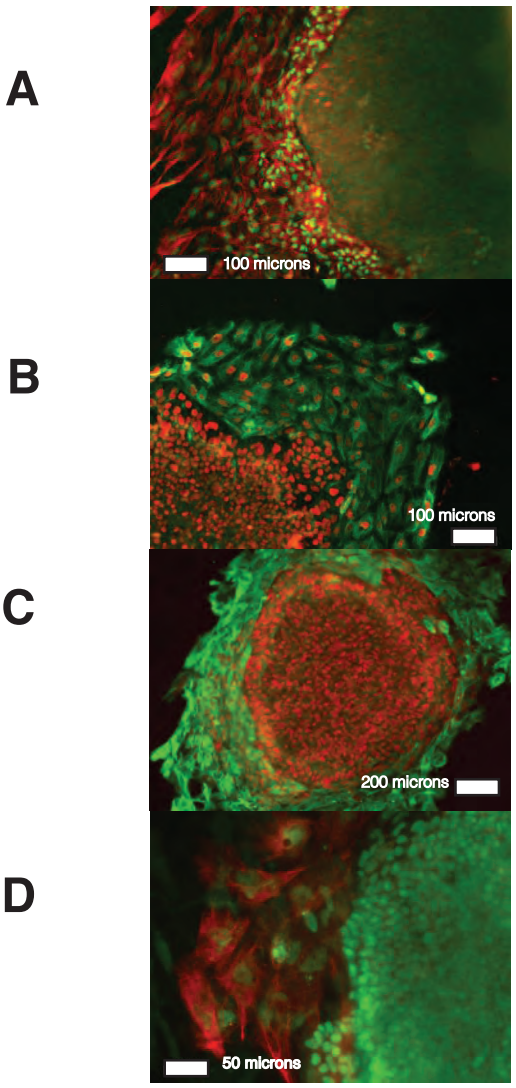
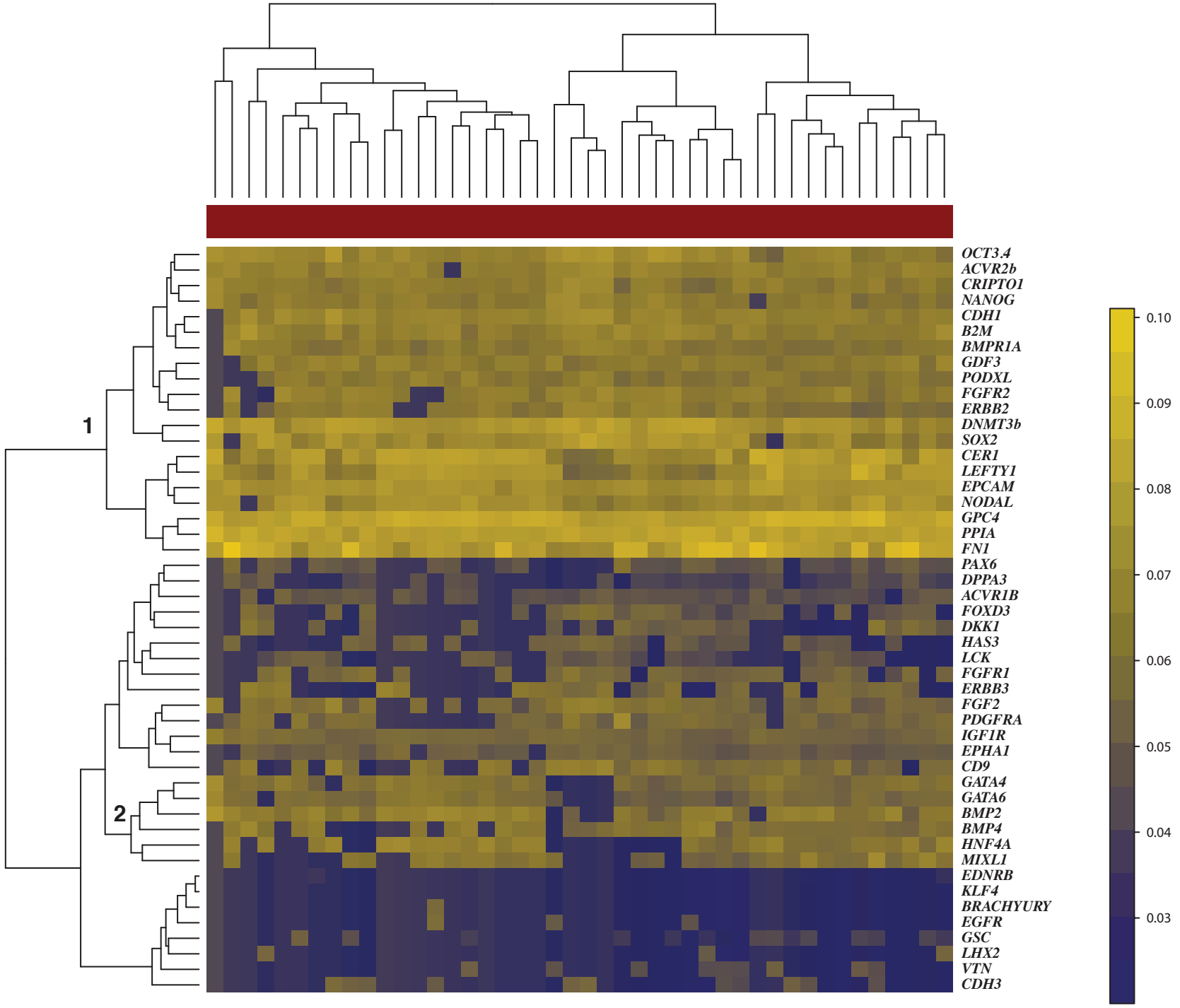


Figure S5


OCT3.4
ACVR2b
CRIPTO1
NANOG
CDH1
B2M
BMPR1A
GDF3
PODXL
FGFR2
ERBB2
DNMT3b
SOX2
CER1
LEFTY1
EPCAM
NODAL
GPC4
PPIA
FN1
PAX6
DPPA3
ACVR1B
FOXD3
DKK1
HAS3
LCK
FGFR1
ERBB3
FGF2
PDGFRA
IGF1R
EPHA1
CD9
GATA4
GATA6
BMP2
BMP4
HNF4A
MIXL1
EDNRB
KLF4
BRACHYURY
EGFR
GSC
LHX2
VTN
CDH3
1
2
0.10
0.09
0.08
0.07
0.06
0.05
0.04
0.03

Supplemental Figure Legends

Figure S1. Relative quantitation of fluorescence intensity of HES3 cells grown in serum and immunostained for various nuclear and cell surface markers related to pluripotency. Top panels show fluoresence intensity quantitation heatmaps generated using MatLab analysis software applied to images of immunostained colonies; bottom panels show segmentation (left), nuclear stain (middle) and immunofluorescence (right). A, GCTM-2; B, EpCAM; C, GDF3; D, POU5F1; E, phospho-SMAD 2/3; F, phospho-SMAD1/5/8; G, CDH3 (P-cadherin); H; BMP-2/4.

Figure S2. Representative images of colonies during serial passage of cultures originated from ~100 cell fragments from edge, middle and centre of ES cell colonies grown in FCS. Images for all three regions are shown at passages 1-4.

Figure S3. Immunostaining of neural progenitor cells derived from the adjacent centre region of ES cell colonies grown in FCS. Neural cells stained generated β-III tubulin-positive neurons (A) following removal of FGF and EGF and stained positive for progenitor markers NCAM (B), NESTIN (C) and SOX1 (D).

Figure S4. Endodermal cells co-expressing stem cell markers in cultures of ES cells grown in KSR/FGF. Immunofluorescence double label staining for (A), CK8 (red) and HNF4A (green); (B), DNMT3B (red) and CK8 (green); (C), POU5F1 (red) and CK8 (green); (D), CK8 (red) and SOX2 (green).

Figure S5. Single cell gene expression analysis of cells picked from outside perimeter corresponding to CK8 positive areas in A-D of ES colonies grown in KSR/FGF. Gene cluster 1 includes pluripotency markers, positive across >90% of cells; gene cluster 2 includes endodermal markers GATA4, GATA6, BMP2, BMP4, HNF4A and MIXL1. Color scale bar indicates values $1/C_T$ for each cell.

Table S1. List of primer sets for analysis of gene expression in human ES cells.

	Gene	Name	ABI assay ID
1	ACVR1B	activin A receptor, type IB	Hs00923299_m1
2	ACVR2b	activin A receptor, type II B	Hs00609597_g1
3	B2M	beta-2-microglobulin	Hs99999907_m1
4	BMP2	bone morphogenetic protein 2	Hs00154192_m1
5	BMP4	bone morphogenetic protein 4	Hs00370078_m1
6	BMPR1A	bone morphogenetic protein receptor type I A	Hs01034913_g1
7	CD9	CD9 molecule	Hs01124022_m1
8	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	Hs01023895_m1
9	CDH3	cadherin 3, type 1, P-cadherin (placental)	Hs00354998_m1
10	CER1	cerberus 1, cysteine knot superfamily, homolog (Xenopus laevis)	Hs00193796_m1
11	Cyclophilin (PPIA)	peptidylprolyl isomerase A (cyclophilin A)	Hs99999904_m1
12	DKK1	dickkopf homolog 1 (Xenopus laevis)	Hs00183740_m1
13	DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta	Hs00171876_m1
14	DPPA3	developmental pluripotency associated 3 (STELLA)	Hs01931905_g1
15	EDNRB	endothelin receptor type B	Hs00240747_m1
16	EGFR	epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog, avian)	Hs01076092_m1
17	EPCAM	epithelial cell adhesion molecule	Hs00901885_m1
18	EPHA1	EPH receptor A1	Hs00178313_m1
19	ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)	Hs01001582_m1
20	ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	Hs00951455_m1
21	FGF2	fibroblast growth factor 2	Hs00960934_m1
22	FGFR1	fibroblast growth factor receptor 1	Hs00241111_m1
23	FGFR2	fibroblast growth factor receptor 2	Hs01552926_m1
24	FOXD3	forkhead box D3	Hs01027393_s1
25	FN1	fibronectin 1	Hs00415006_m1
26	GATA4	GATA binding protein 4	Hs00171403_m1
27	GATA6	GATA binding protein 6	Hs00232018_m1
28	GDF3	growth differentiation factor 3	Hs00220998_m1
29	GPC4	glypican 4	Hs00155059_m1
30	GSC	goosecoid homeobox	Hs00418279_m1
31	HAS3	hyaluronan synthase 3	Hs00193436_m1
32	HNF4A	hepatocyte nuclear factor 4, alpha	Hs00230853_m1
33	IGF1R	insulin-like growth factor 1 receptor	Hs99999020_m1
34	KLF4	Kruppel-like factor 4 (gut)	Hs00358836_m1
35	LCK	lymphocyte-specific protein tyrosine kinase	Hs00178427_m1
36	LEFTY1	left-right determination factor 1	Hs00764128_s1
37	LHX2	LIM homebox 2	Hs00180351_m1

38	MIXL1	Mix1 homeobox-like 1	Hs00430824_g1
39	NANOG	Nanog homeobox	Hs02387400_g1
40	NODAL	nodal	Hs01086749_m1
41	OCT 3/4 (POU5F1)	POU class 5 homeobox 1	Hs00999632_g1
42	PAX6	paired box 6	Hs00240871_m1
43	PDGFRA	platelet-derived growth factor receptor, alpha polypeptide	Hs00998018_m1
44	PODXL	podocalyxin	Hs01574644_m1
45	SOX2	SRY (sex determining region Y)-box 2	Hs01053049_s1
46	T (BRACHYURY)	T, brachyury	Hs00610080_m1
47	TDGF1 (CRIPTO1)	teratocarcinoma-derived growth factor 1	Hs02339499_g1
48	VTN	vitronectin	Hs00169863_m1

Table S2. List of antibodies for FACS and immunostaining.

primary antibody	source	catalog #	clone	host	isotype
AFP	Sigma	A8452	C3	mouse	IgG2a
anti-mouse CD90.2 (Thy1.2) PE	BD Bioscience	BD553006	53-2.1	rat	IgG2a
anti-mouse CD90.2 (Thy1.2) PE-Cy7	eBioscience	25-0902	53-2.1	rat	IgG2a
BMP-2/4	R&D Systems Systems	AF355		goat	IgG
CD9	Pera Lab		TG30	mouse	IgG2a
cerberus1	R&D Systems	mab 1075	162836	mouse	IgG2b
cytokeratin 8	Santa Cruz Biotech	sc-8020	C51	mouse	IgG1
DKK1	R&D Systems	AF1096		goat	pab
DNMT3b	Santa Cruz Biotech	sc-10235	N19	goat	pab
E-Cadherin	R&D Systems	mab 18381	180224	mouse	IgG2b
EPCAM	Santa Cruz Biotech	sc-33687	KS1/4	mouse	IgG2a
EPCAM	DAKO	Ber-EP4 M0804	Ber-EP4	mouse	IgG1
EPCAM-PerCP-Cy5.5	BD Bioscience	BD347199	EBA-1	mouse	IgG1
follistatin	R&D Systems	mab 669	85918	mouse	IgG2a
GATA6	R&D Systems	mab 1700	222228	mouse	IgG1
GCTM2	Pera Lab		GCTM2	mouse	IgM
GDF3	abcam	ab38547		rabbit	pab
HNF4a / 1-6	R&D Systems	PP-K9218-00	K9218	mouse	IgG2a
NCAM	Life Tech	MHCD5600	MEM-188	mouse	IgG2a
Nestin	Millipore	MAB5326	10C2	mouse	IgG1
OCT3/4	Santa Cruz Biotech	sc-5279	C10	mouse	IgG2b
PAX6	Developmental Studies Hybridoma Bank	PAX6		mouse	IgG1
P-Cadherin	R&D Systems	MAB861	104805	mouse	IgG1
phospho-SMAD 1/5/8	cell signaling technology	9511		rabbit	mab
phospho-SMAD3	epitomics	1880-1		rabbit	mab
SOX1	R&D Systems	AF3369		goat	pab
SOX2	Chemicon	AB5603		rabbit	pab
TRA-1-60	abcam	ab16288	TRA-1-60	mouse	IgM
Tubulin, β-III	Millipore	MAB1637	TU-20	mouse	IgG1
αSMA	DAKO	M0851 (N1584)	1A4	mouse	IgG2a

Table S3. List of primer sets for single cell analysis of gene expression in neural precursors isolated by FACS.

	Gene	Name	ABI assay ID
1	CD9	CD9 molecule	Hs01124022_m1
2	CDH1	cadherin 1, type 1, E-cadherin (epithelial)	Hs01023895_m1
3	CDH3	cadherin 3, type 1, P-cadherin (placental)	Hs00354998_m1
4	CER1	cerberus 1, cysteine knot superfamily, homolog (Xenopus laevis)	Hs00193796_m1
5	Chordin	chordin	Hs00415315_m1
6	DKK1	dickkopf homolog 1 (Xenopus laevis)	Hs00183740_m1
7	DNMT3b	DNA (cytosine-5-)-methyltransferase 3 beta	Hs00171876_m1
8	EPCAM	epithelial cell adhesion molecule	Hs00901885_m1
9	FN1	fibronectin 1	Hs00415006_m1
10	Follistatin	follistatin	Hs00246256_m1
11	FOXD3	forkhead box D3	Hs01027393_s1
12	FOXG1	forkhead box G1	Hs01850784_s1
13	GATA4	GATA binding protein 4	Hs00171403_m1
14	GATA6	GATA binding protein 6	Hs00232018_m1
15	GFAP	glial fibrillary acidic protein	Hs00909238_g1
16	GSC	goosecoid homeobox	Hs00418279_m1
17	LEFTY1	left-right determination factor 1	Hs00764128_s1
18	LHX2	LIM homebox 2	Hs00180351_m1
19	MIXL1	Mix1 homeobox-like 1	Hs00430824_g1
20	Musashi-1	musashi homolog 1 (Drosophila)	Hs00159291_m1
21	NANOG	Nanog homeobox	Hs02387400_g1
22	NCAM	neural cell adhesion molecule 1	Hs00941821_m1
23	Nestin	nestin	Hs00707120_s1
24	Noggin	noggin	Hs00271352_s1
25	OCT3/4	POU class 5 homeobox 1	Hs00999632_g1
26	OLIG2	oligodendrocyte lineage transcription factor 2	Hs00377820_m1
27	OTX2	orthodenticle homeobox 2	Hs00222238_m1
28	PAX6	paired box 6	Hs00240871_m1
29	PPIA	peptidylprolyl isomerase A (cyclophilin A)	Hs99999904_m1
30	SHH	sonic hedgehog	Hs00179843_m1
31	SIX3	SIX homeobox 3	Hs00193667_m1
32	SOX1	SRY (sex determining region Y)-box 1	Hs01057642_s1
33	SOX2	SRY (sex determining region Y)-box 2	Hs01053049_s1
34	T	T, brachyury	Hs00610080_m1
35	VTN	vitronectin	Hs00169863_m1

Supplemental Experimental Procedures

Routine Cell Culture

Human ES cell culture

Cultures grown in serum-supplemented medium fibroblast feeder cell support (FCS condition): Human embryonic stem cells (HES3 and ENVY lines, (passage 25-45 and 80-94, respectively) were maintained as described previously (Reubinoff et al., 2000). Briefly, hESC colonies were maintained on mitotically inactivated mouse embryonic fibroblasts (MEF) (density of 60,000 cells / cm^2) in media consisting of DMEM (Life Technologies cat. no. 11960-044, L-glutamine 1% v/v (Life Tech cat. no. 25030-081), penicillin / streptomycin 0.5% v/v (Life Tech cat. no. 15070-063), MEM non-essential amino acids 1% v/v (Life Tech cat. no. 11140-050), insulin-transferrin-selenium 1% v/v (Life Tech cat. no. 41400-045), β-mercaptoethanol 0.18% v/v (Life Tech cat. no. 21985-023), with 20% (v/v) FBS. Colonies were passaged weekly using the cut-and-paste method.

Cultures grown in serum replacement plus FGF2 (KSR/FGF condition): HES3, ENVY, USC01, HES2 and H9 lines were maintained on inactivated MEF (density of 15,000 cells / cm^2) in media consisting of DMEM / F12 with L-glutamine (Life Tech cat. no. 11330), 20% (v/v) KNOCKOUT™ serum replacer (Life Tech cat. no. 10828), 4 ng/mL bFGF (Peprotech) and (MEM non-essential amino acids, β-mercaptoethanol, and penicillin / streptomycin (as above). Cells were enzymatically passaged twice weekly at a split ratio of 1:3 using CTK solution consisting of 1 mg/mL collagenase IV (Life Tech cat. no 17104019), 0.25% trypsin (Life Tech cat. no. 15090) 20% KSR (Life Tech cat. no. 10828), 1 mM calcium chloride in PBS.

Cultures grown in defined feeder-free conditions (mTeSR conditions): For feeder-free culture conditions, ENVY and H9 lines were maintained in mTESR1™ medium (Stem Cells) on Matrigel and passaged using dispase according to the manufacturers protocol.

Neural progenitor cultures

To assess the neurosphere-generating capacity, HES3 colonies grown in FCS were washed twice with neural stem cell growth medium (below) and small pieces (~100 μM^2) were excised using a fine needle from the edge, middle, adjacent to center, or center colony regions with the aid of a dissecting microscope. Single pieces were transferred to individual wells of 96-well ultra-low attachment culture plates (Corning Costar 3474) containing neural stem cell growth medium consisting of neurobasal medium (cat. no. 12349), B-27 supplement (2% v/v, cat. no. 17504-044), insulin-transferrin-selenium-A (1% v/v, cat. no. 51300-044), N2 supplement (1%, v/v, cat. no. 17502-048), 2 mM L-glutamine (cat. no. 25030), penicillin / streptomycin (0.5% v/v) (cat. no. 15070), (all from Life Tech) with 20 ng/ml each FGF2 and EGF (Peprotech). Medium was changed every other day. Following expansion for 1 week, resulting neurospheres were transferred to monolayer culture format for immunostaining or were processed for gene expression analysis by QRT-PCR. For monolayer culturing, neurospheres were dissociated by gentle trituration with a 200 μL pipette tip. Dissociated cells were replated onto poly-L-ornithine (Sigma P3655), laminin (5 μg/ml, Sigma L2020) coated 48-well culture plates in NSC growth medium containing FGF2 and EGF (each at 20 ng/ml). Medium was

changed every 48 hr. Confluent cells were passaged by gentle trituration onto new laminin-coated wells at a split ratio of 1:3. Cultures could be maintained in monolayer format for at least 5 passages.

Flow Cytometry

Single ES cells were isolated for gene expression analysis based on double (GCTM2, TG30) or triple (GCTM2, TG30, EPCAM) surface marker staining. Cultures were treated with 10 μM blebbistatin (Sigma B0560, Saint Louis, MO) for 1 hour prior to dissociation to single cells using TrypLE™ (Life Tech, cat. no. 12605). Cells were stained in solution using a mixture of GCTM-2 (mouse IgM) and TG30 (anti-CD9, mouse IgG2a) (double stain) and anti-EPCAM-PerCp-Cy5.5 (BD347199) (triple stain). Primary antibodies against GCTM-2 and TG30 were detected using goat anti-mouse IgM-AF647 (A21238) and goat anti-mouse IgG2a-AF488 (A21131), respectively (Life Tech, Carlsbad, CA). Anti-Thy1.2 PE (BD553006) was used to gate out any mouse embryonic fibroblasts.

Control samples included unlabeled cells, cells labeled with secondary antibody only and single fluorochrome labeled cells. Cells were sorted using a FACSAria (BD Biosciences) with a 100 μM nozzle and low pressure conditions. Cells were first gated based on forward and side scatter properties and then Thy1.2 negative cells were analyzed for levels of GCTM2, TG30 and EPCAM labeling. Double stained cells (GCTM2 and TG30) were sorted into four populations; those negative for both surface markers, and cells that exhibited low level, mid-level or high level expression of both surface markers. Triple stained cells (GCTM2, TG30, EPCAM) were sorted into triple negative, triple high (top 6%) and TOP (top 1.5%) populations. Sorted single cells were processed for gene expression analysis as described below.

Assays of Self-Renewal

Single cell colony formation in adherent culture
Single ES cells grown in mTeSR were isolated by FACS, placed back in culture, and allowed to reform colonies to determine the self-renewal capacity of the top, high and mid populations. Briefly, H9 hESC cultured in mTESR1 were triple stained for GCTM2, TG30 and EPCAM (as above). Single cells from each of three FACS populations (top, high and mid) were placed at 500 cells per well of a 48-well plate containing 14,000 MEF per well. Cultures were maintained in mTESR1 with medium changed daily after 48 hr post-plating. After 10 days, wells were fixed and stained for colony counting as follows. The culture media was removed from cells and each well rinsed once with PBS prior to fixing with 100% ethanol for 5 minutes at room temperature. Ethanol was removed and wells allowed to air dry for 30 min prior to staining for 30 minutes with GCTM2 hybridoma supernatant. Wells were then washed twice with 100 mM Tris-HCl, pH 8.0 prior to incubation for 30 minutes with secondary antibody AP-goat anti-mouse IgM (#62-6822 Life Tech) (1:500 in 100 mM Tris-HCl). Wells were washed 2x with Tris-HCl and AP developed using Vector Lab AP kit II (sk-5100). Each well was imaged using a Leica dissecting microscope fitted with an IC80 HD camera.

Assay for self-renewal capacity of clumps of cells from discrete colony regions
To assess the self-renewing capacity of ES cells under conditions that allow high survival, small pieces (~100 μM^2) were excised from the edge, middle and adjacent to center colony regions of HES3 grown in FCS and replated onto MEFs (density 60,000 cells / cm^2). After 48 hr, medium was replaced daily. Following 1 week in culture, resulting colonies were passaged again, using pieces excised from the same region (edge, middle, adjacent to center) as the parent colony. Passaging was continued in this manner for 5 weeks.

Immunofluorescence Microscopy

HES3 colonies cultured in FCS condition in 8-well chamber slides (BD Falcon cat. no. 354108) and neural stem cells and EB outgrowths cultured 24-well culture dishes were washed twice with PBS prior to fixing with 2% paraformaldehyde (PFA) for 30 minutes at room temperature. Cells were permeabilized with 0.3% Triton X-100 in PBS and blocked with 1% IgG-free BSA. Primary antibodies (Table S3) were diluted in 0.1% IgG-free BSA. Cells were incubated with primary antibodies for 4 hr at room temperature or overnight at 4°C. Cells were washed with PBS prior to addition of appropriate isotype-matched AlexaFluor –labeled secondary antibodies (Life Tech) (anti-mouse, anti-goat, anti-rabbit) diluted 1:1000 in PBS. Samples were then washed with PBS prior to fixation and nuclear counterstaining with Prolong Gold plus DAPI (Life Tech, cat. no. P36931). Images were acquired using Zeiss Axioimager.Z1 and Axiovert 200 microscopes, both with epifluoresence capability and fitted with an Axiocam MRm camera and Axiovision 4.6 software (Carl Zeiss, Inc.).

Reaggregate Assay imaging and data processing

Colonies in each well of the 24 well plates were imaged using a Zeiss Axio Observer A1 equipped with a 10x EC Plan-Neofluor objective (Zeiss 0.3 NA, 420341-9911) and Antivibration Pad L (920007970, Eppendorf, http://www.eppendorf.ca). Images were collected with a Zeiss 12 Bit CCD camera (MRm ver 3.0 Zeiss) at a size of 1024 x 1024 pixels (1.048 megapixel) using the multidimensional acquisition procedure in Axiovision software release 4.8. The size of the pixels was measured by imaging a 10 line-pair per millimeter slide (N38-258, Edmund Optics).

Statistical analysis of single cell expression data

The cluster package for R was used to perform the fuzzy cluster analysis for samples in each media condition (FCS, KSR/FGF2, mTeSR). The dist function computed a distance matrix using euclidean (root sum-of-squares) distance to calculate the dissimilarities between samples in each media condition. The FANNY function was used to prioritize clustering of the samples and the output of this function was displayed in a principle component plot, where each data point can belong to up to 2 groups. The mean and covariance matrix of the observations in each cluster were used to compute the boundary of each ellipse. Agglomerative hierarchical cluster analysis was used to

identify gene groups sharing similar expression profiles across different populations. A $\log_2$-transformation was applied to the data. We used the statistical package for R (version 2.15.0) to generate heatmaps. Genes were grouped according to a $(1-R^2)$ distance metric where R^2 represents the Spearman rank correlation coefficient (Alarcon and van Santen, 2010), and clusters were joined using average linkage. Violin plots were generated using the CRAN package vioplot.

Reaggregation Limiting Dilution Assay imaging and data processing

Colonies in each well of the 24 well plates were imaged using a Zeiss Axio Observer.A1 equipped with a 10x EC Plan-Neofluar objective (Zeiss 0.3 NA, 420341-9911) and Antivibration Pad L (920007970, Eppendorf, http://www.eppendorf.ca). Images were collected with a Zeiss 12 Bit CCD camera (MRm ver 3.0 Zeiss) at a size of 1024 x 1024 pixels (1.048 megapixel) using the multidimensional acquisition procedure in Axiovision software release 4.8. The size of the pixels was measured by imaging a 10 line-pair per millimeter slide (N38-258, Edmund Optics).
It was necessary to use long exposure times (~2 s) to maximize the fluorescence intensities in the collected images. These images were then corrected for blur by the method of Shan, et. al., which optimizes an estimated blur kernel (PSF) and image debluring by a unified probabilistic model.(Shan et al., 2008). The deblurred images were then analyzed using CellProfiler, which is open source image analysis software (Carpenter et al., 2006; Kamentsky et al., 2011). After debluring and prior to thresholding additional low-cut filtering in the Fourier domain was implemented via an ImageJ plugin utilizing the procedure of Bright (Bright, 2004; Rasband, 1997 - 2011; Russ, 2010). Multidimensional images were separated by channel and associated with metadata in the CellProfiler pipeline. Identification of objects in processed images was accomplished through use of the ilastik open source software (www.ilastik.org). Ilastik uses interactive machine learning algorithms to classify pixels and perform segmentation(Sommer et al., 2011). The software generates a classifier file which can be utilized by CellProfiler to identify objects. Outlines of objects identified for each stain were overlaid with the enhanced image for visual inspection and confirmation. Objects indentified in each channel associated by metadata were then measured, grouped, and categorized. (Supplemental – CellProfiler pipeline).
Statistics were generated with the SAS statistical analysis software (SAS institute, Cary NC, http://www.sas.com) and with the R statistical software package (R project, http://www.r-project.org).

Differentiation Assays

Embryoid Body (EB) Assays
HES3 cells maintained in KSR/FGF culture were dissociated, stained, and sorted into four populations based on staining with GCTM2 and TG30 as above. Following sorting, cells were washed once in KSR culture media containing blebbistatin and FGF2. Alternatively, for cultures grown in FCS, clumps of cells from different regions of the colony were isolated as described above under self-renewal assays. To induce EB

formation, cells (4000 per tube) were aggregated by centrifugation at 480*g* for 2 min in low-binding PCR tubes (Axygen 321-02-501) containing 100 uL KSR culture media containing blebbistatin and FGF2. After 48 hr, EBs were transferred to 96-well ultra-low attachment plates (Corning Costar 3474) in STEMdiff ™APEL™ medium (#05210 STEMCELL) without FGF2. Culture medium was changed every other day. After 6 days, EBs were transferred to gelatin-coated 48-well plates in hESC culture media containing 20% FBS to generate outgrowths. Medium was changed daily for 2 weeks. The embryoid bodie outgrowths were analysed by indirect immunofluorescence staining or by QRT-PCR.

NSC Differentiation
For differentiation to β-III tubulin positive neurons, NSC were cultured on laminin-coated 48-well plates in NSC medium without growth factors for 1 week.

Analysis of gene expression

Single cell qPCR
For analysis of gene expression in single ES cells, individual cells were immediately picked or sorted directly into low binding PCR tubes (Axygen 321-02-501) following microdissection or isolation by FACS. Tubes contained preamplification mix consisting of 5 uL Cells Direct 2X reaction mix (P/N 55175, Life Tech cat no 46-7201), 2.5 uL of a 0.2X concentration of pooled TaqMan primer sets (Tables 1 & 2, supplemental) in TE buffer, 0.5 uL of RT / Taq enzyme mix (Cells Direct kit, P/N 55548, Life Tech) and 2 uL TE buffer. Samples were snap frozen on dry ice to facilitate cell lysis followed by RT and multiplex preamplification. Cycling conditions were: RT at 50°C for 20 minutes, activation at 95°C for 2 minutes, followed by 40 cycles of preamplification each at 95°C for 15 seconds followed by 60°C for 4 minutes. Reactions were stopped by heating at 99.9°C for 10 minutes prior to storage at -20°C. Preamplified samples were diluted 1:5 with TE buffer prior to qPCR analysis on microfluidic chips. Single gene-specific qPCR reactions using individual Taqman primer sets (Supplemental Tables 1 & 2) were carried out on a Biomark™ HD instrument using 48x48 Dynamic Array Gene Expression microfluidic chips (Fluidigm Corp). Prior to loading onto the chip, individual TaqMan gene expression assays (20X) were diluted 1:1 with 2X assay loading reagent (Fluidigm PN 85000736). Pre-amplified samples (2.75 uL) were combined with 2.5 uL TaqMan Universal Master Mix (ABI 4304437) and 0.25 uL 20X GE sample loading reagent (Fluidigm PN 85000735). qPCR cycling conditions were 50°C for 2 minutes, 95°C for 10 minutes, followed by 40 cycles each at 95°C for 15 seconds followed by 60°C for 1 minute.

Gene expression in EB
Samples from EB assays were analysed following RNA extraction using individual Taqman probesets in accordance with the manufacturer's instructions.

Supplemental References

Alarcon, B., and van Santen, H.M. (2010). Two receptors, two kinases, and T cell lineage determination. Sci Signal *3*, pe11.

Bright, D.S. (2004). Digital Image Processing with NIH Image (Mac) / Scion Image (PC) / ImageJ.

Carpenter, A.E., Jones, T.R., Lamprecht, M.R., Clarke, C., Kang, I.H., Friman, O., Guertin, D.A., Chang, J.H., Lindquist, R.A., Moffat, J.*, et al.* (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome biology *7*, R100.

Kamentsky, L., Jones, T.R., Fraser, A., Bray, M.A., Logan, D.J., Madden, K.L., Ljosa, V., Rueden, C., Eliceiri, K.W., and Carpenter, A.E. (2011). Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. Bioinformatics *27*, 1179-1180.

Rasband, W.S. (1997 - 2011). ImageJ.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. Nat Biotechnol *18*, 399-404.

Russ, J.C. (2010). The image processing handbook, 6th edn (Boca Raton, CRC Press).

Shan, Q., Jia, J., and Agarwala, A. (2008). High-quality motion deblurring from a single image. Acm T Graphic *27*.

Sommer, C., Straehle, C., Kothe, U., and Hamprecht, F.A. (2011). Ilastik: Interactive learning and segmentation toolkit. Paper presented at: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium.