Supplemental Data

Supplemental Information Inventory

Supplemental Figures with Captions

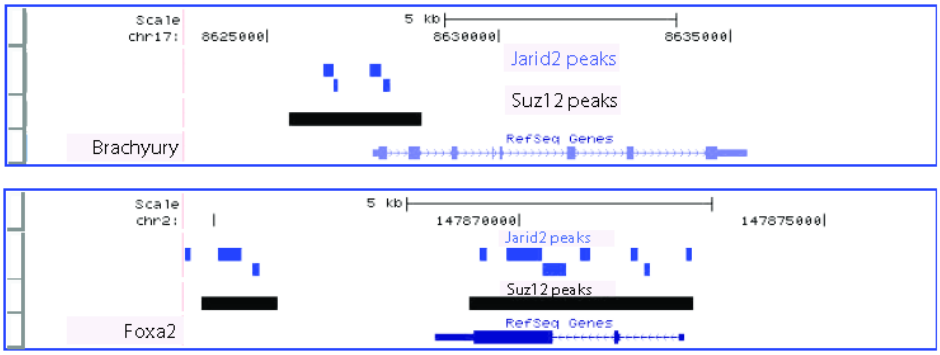
Supplemental Experimental Procedures

Supplemental References

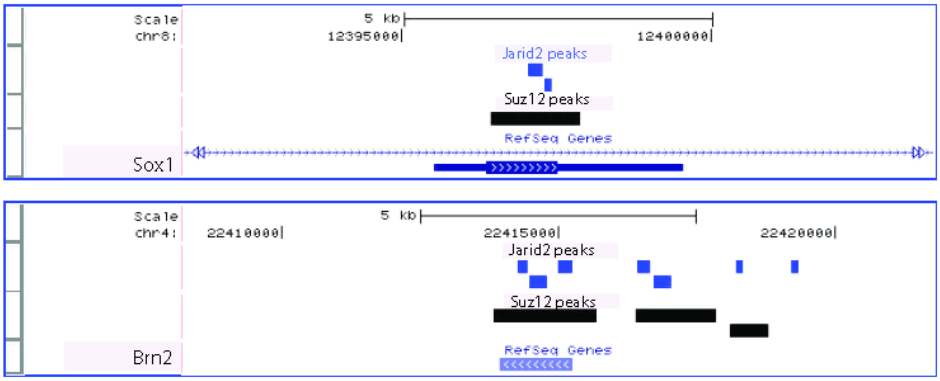
Supplemental Figures

Figure S1

A



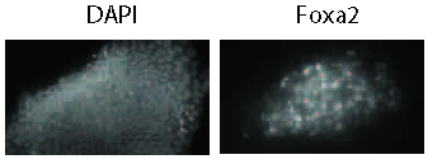
B



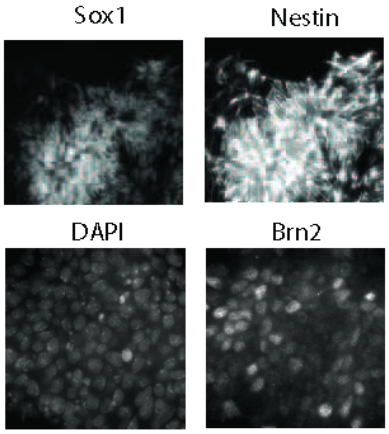
C

Gene	ES	ME	NE	ES/ME	ES/NE	DistDiag	p-value
Klf5	1894.3	1707.9	197.85	1.10914	9.574425	5.985696	8.53E-08
Tbx3	713.2	1100	104.65	0.648364	6.815098	4.36042	0.02047
Klf9	1036.9	838	205.95	1.237351	5.034717	2.685286	0.000313
Oct4	3853.3	1462.7	728.3	2.634375	5.290814	1.886674	6.62E-06
Nanog	1615.75	874.5	457.4	1.847627	3.532466	1.198371	9.32E-58
Sox2	3527.5	664.35	3496.4	5.309701	1.008895	1.152796	1.91E-78
Dnmt3a	1728.35	349.55	851.25	4.9445	2.030367	1.060754	1.45E-24
Zfp532	1308.2	309.35	956.1	4.228867	1.368267	0.93034	6.24E-37
Foxp1	859.65	207.35	799.75	4.145889	1.074898	0.910024	1.64E-10
Rbpj	1102.9	280.8	776.85	3.927707	1.419708	0.858864	4.94E-07

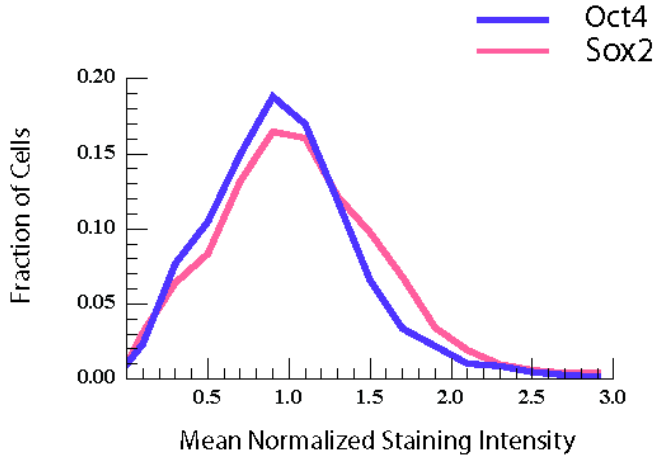
D



E



F



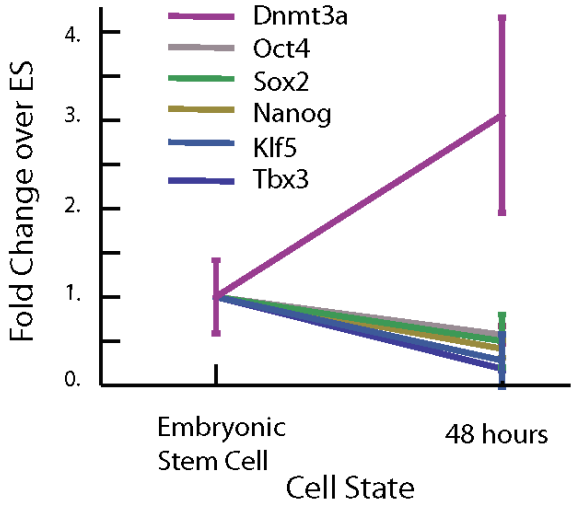
S1 related to Figure 1:

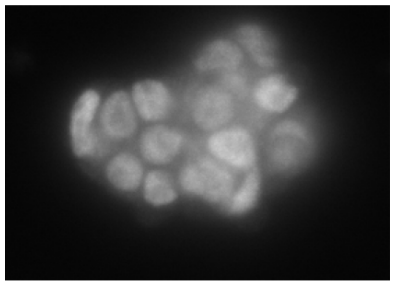
(A-B) Chromatin modifying factors Jarid2 and Suz12 bind near regulators of both the NE and ME lineage. UCSC Genome browser images showing ChIP-seq peak calls for Suz12 and Jarid2 in undifferentiated ES cells from published data (Chen et al., 2008; Pasini et al., 2010). Images show Jarid2 and Suz12 binding at the genomic loci of (A) *Foxa2* and *Brachyury*, core regulators of the ME lineage, as well as throughout the genomic loci of (B) *Sox1* and *Brn2*, core regulators of the NE lineage, suggesting *Brachyury*, *Foxa2*, *Sox1*, and *Brn2* are epigenetically repressed in ES cells. (C) **Analysis of microarray data;** First three columns show mean expression levels of ten differentially expressed transcription factors in ES, ME, and NE cell populations. All factors are expressed at levels significantly above the noise floor (~200). Columns 4 and 5 show fold down regulation for each factor in the ME and NE lineage as shown in Figure 1C. Column 6 shows distance of each point from diagonal of the plot. We used this number to measure differential expression of the factor between the ME and NE lineages. Column 7 gives the p-value for differential expression showing statistical significance of for all factors. (D-E) **Differentiation protocol induces multiple markers of ME or NE progenitor cells. (D)** Image of subpopulation of cells induced to the mesendodermal fate by Wnt3a (200ng/ml) immunostained for Foxa2. Cells stained 36 hours following Wnt addition. (E) Images of cells differentiated into neural ectodermal fate and immunostained for Sox1 and Nestin, 5 days after addition of retinoic acid (retinoic acid added after 48 hours in N2B27, see main text). Lower Panel, cells induced to the NE cell fate by retinoic acid immunostained for Brn2 36 hours after signal addition. (F) Histograms of mean normalized Oct4 and Sox2 staining intensity in ES

cells from data shown in Figure 1F. Histograms show that while Oct4 and Sox2 protein levels vary in the population of ES cells, the distribution is peaked at the mean. 72% of cells have Oct4 protein levels between .5-1.5 mean units.

Figure S2

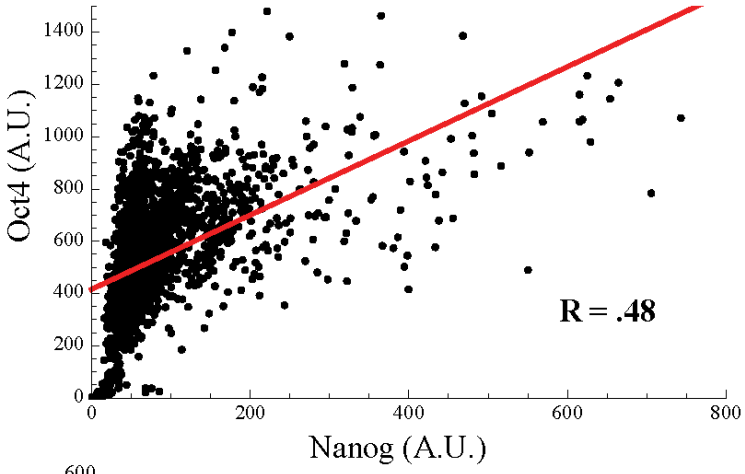
A



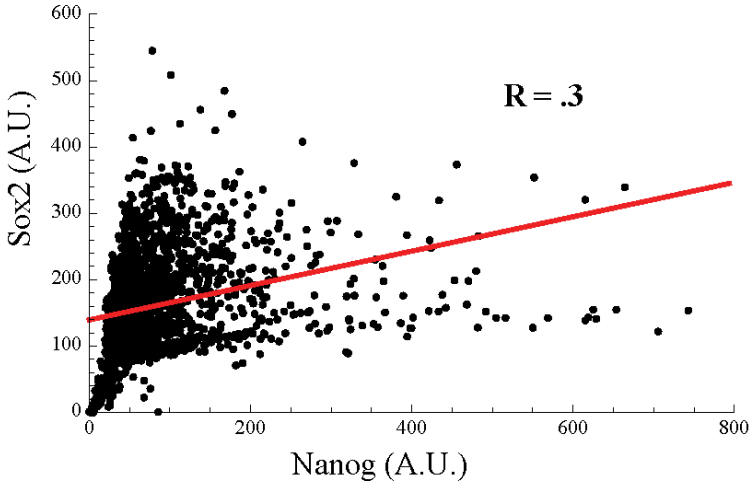


Klf4 ES Cells

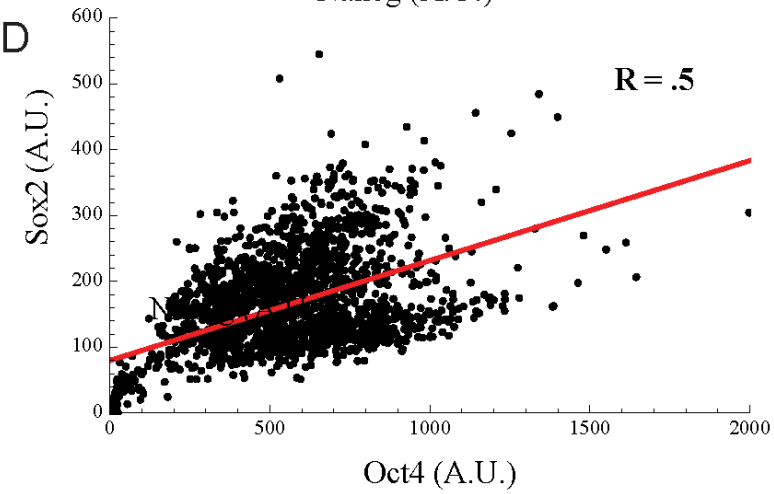
B



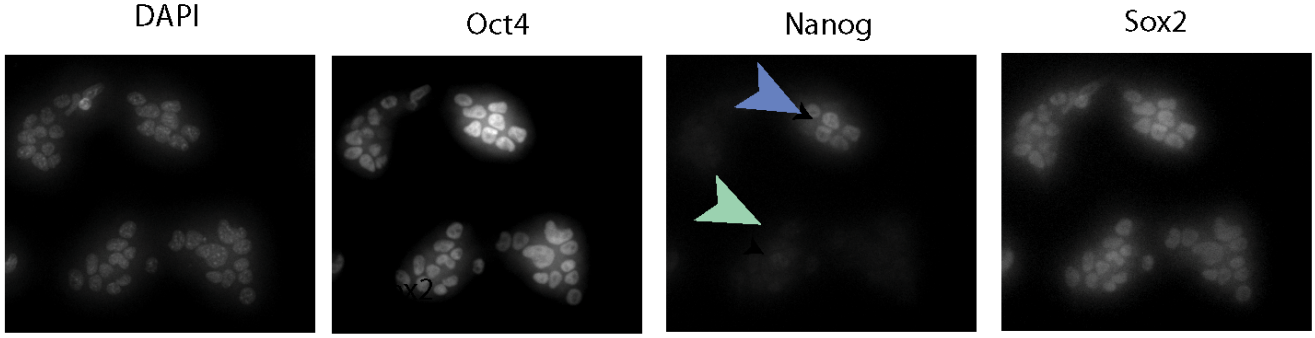
C



D



E

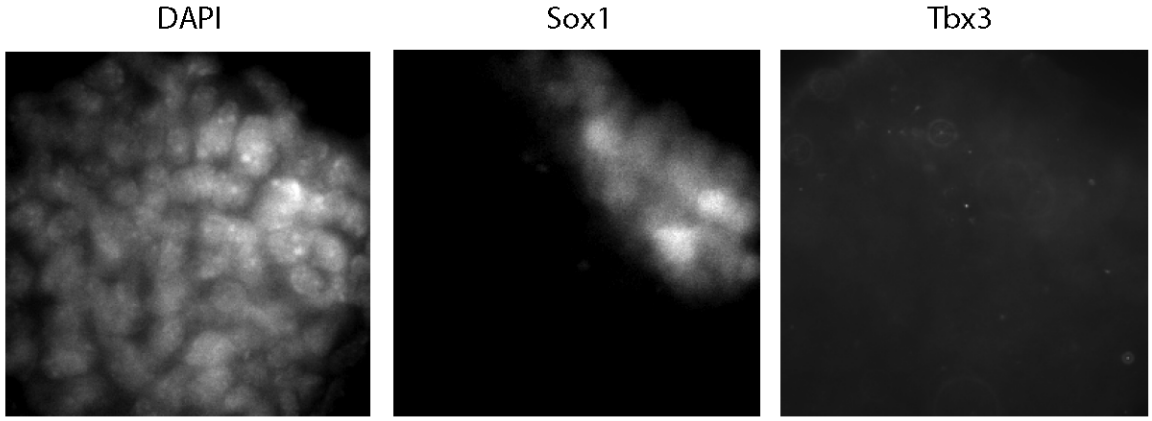


S2 related to Figure 2: Protein levels of Oct4, Sox2, Nanog, Tbx3, and Klf5 and Dnmt3a during loss of pluripotency.

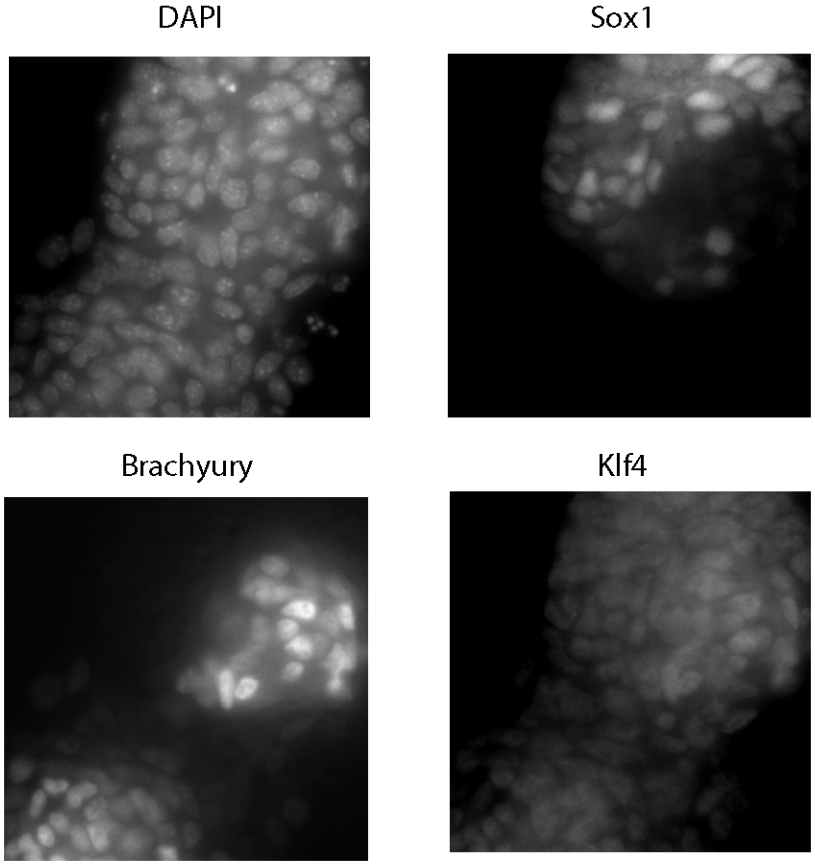
(A) Left Panel, mean and ±SD in nuclear protein levels, as measured by immunofluorescence and microscopy, for Oct4, Sox2, Nanog, Klf5, Tbx3, and Dnmt3a in pluripotency promoting conditions and after 48 hours in N2B27 for n>100 cells. Cells were stained for Oct4, Sox2, Nanog, Klf5, Tbx3, and Dnmt3a and nuclear protein levels were measured using microscopy. The level of each factor is shown relative to its mean level in pluripotency promoting conditions. Right panel, images of Klf4 immunofluorescence in ES cells growing under pluripotency promoting conditions for comparison with Figure 2C. (B-D) **Oct4, Sox2, and Nanog protein levels following Nanog siRNA.** Pair wise scatter plots and least squares fits to a linear model shown for pairs of Oct4, Sox2 and Nanog levels in cells immunostained for these proteins. Nanog protein levels were reduced using siRNA transfection, after which cells were grown for 48 hours in pluripotency promoting conditions. Oct4 and Sox2 protein levels fall but remain correlated in cells without Nanog. The mean Oct4 level is 60% and the mean Sox2 40% of the mean level in cells with nuclear Nanog protein. (E) Images of cells from (B-D), stained for DAPI, Oct4, Sox2, and Nanog. The field of view shows cells transfected cells (green arrow) and untransfected cells (blue arrow).

Figure S3

A

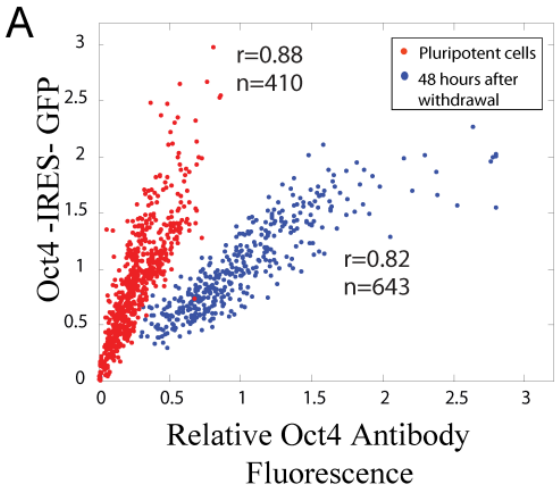


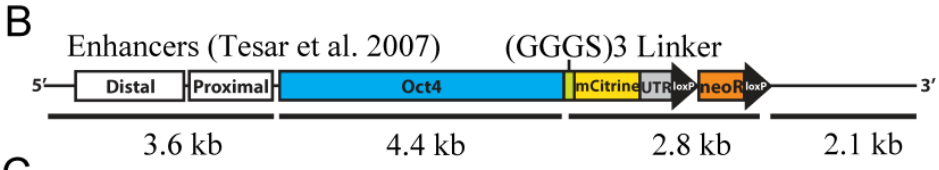
B

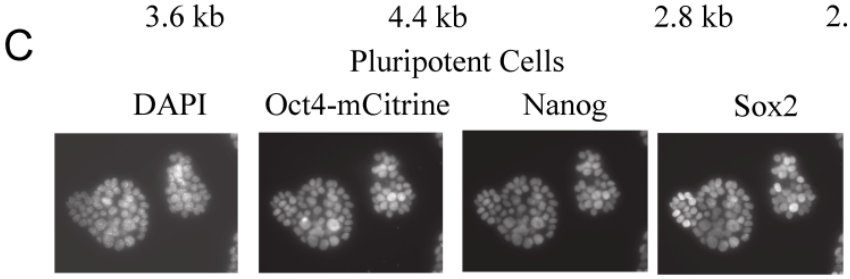


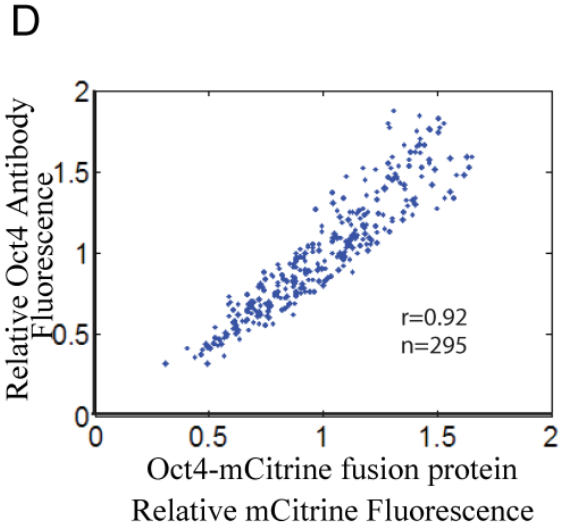
S3 related to Figure 3: Tbx3 is undetectable and Klf4 present at low and variable levels in both neural ectodermal (NE) and mesendodermal (ME) progenitor cells. (A) Images of cells exposed to 3μM CHIR (as in Figure 2 Main Text) showing Sox1-GFP expression, Tbx3 immunofluorescence, and DAPI staining. Some cells choose the ME fate while others choose NE fate (see main text). The Tbx3 and Brachyury antibodies were made in same host species and so could not be detected at the same time with secondary antibodies. However, Tbx3 was nearly undetectable in cells differentiated using CHIR, a condition in which 70% of cells activate Brachyury. Tbx3 is also absent in Sox1-GFP expressing cells. (B) Images of cells differentiated into either the ME or NE fate with 3μM CHIR. Images show Sox1-GFP expression and Brachyury and Klf4 immunofluorescence. Klf4 protein is present variably in both lineages and is present at ~25% of ES cell protein levels.

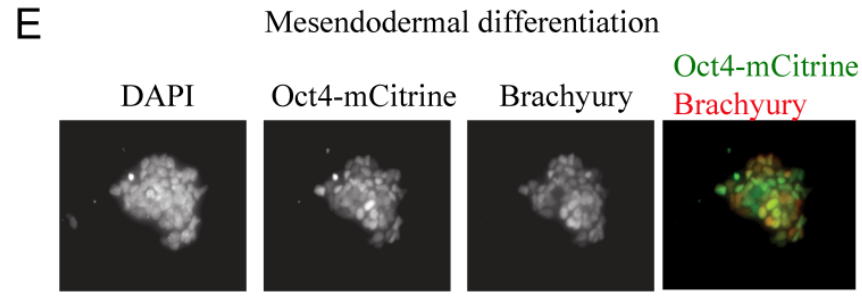
Figure S4

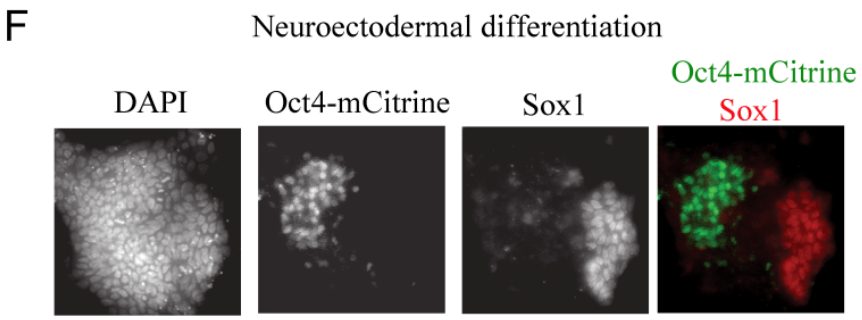

A
r=0.88
n=410
Pluripotent cells
48 hours after withdrawal
r=0.82
n=643
Oct4 -IRES- GFP
Relative Oct4 Antibody Fluorescence

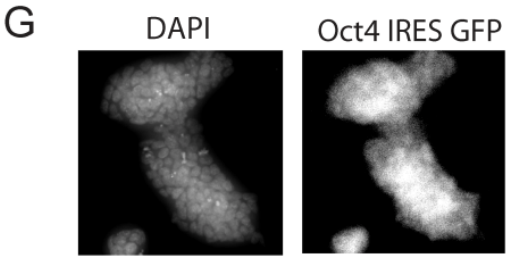

B
Enhancers (Tesar et al. 2007)
(GGGS)3 Linker
5'
Distal
Proximal
Oct4
mCitrine
UTR
loxP
neoR
loxP
3'
3.6 kb
4.4 kb
2.8 kb
2.1 kb

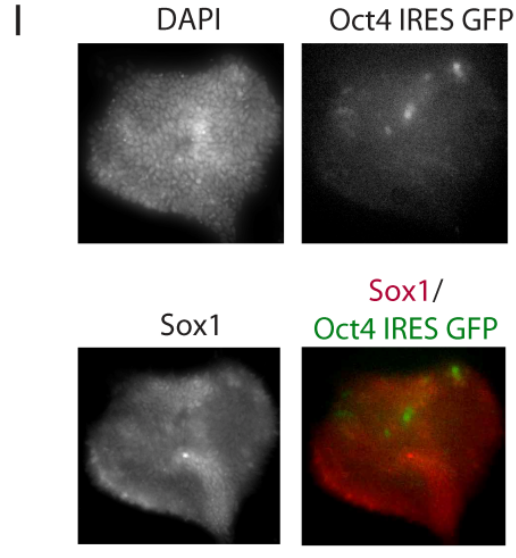

C
Pluripotent Cells
DAPI
Oct4-mCitrine
Nanog
Sox2

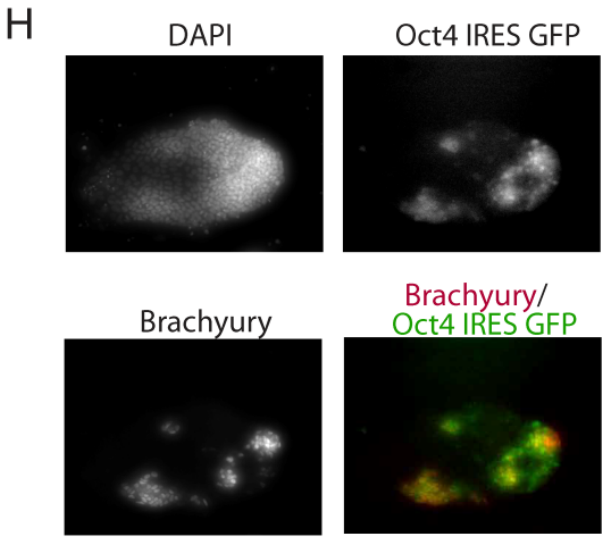

D
r=0.92
n=295
Relative Oct4 Antibody Fluorescence
Oct4-mCitrine fusion protein
Relative mCitrine Fluorescence


E
Mesendodermal differentiation
DAPI
Oct4-mCitrine
Brachyury
Oct4-mCitrine
Brachyury


F
Neuroectodermal differentiation
DAPI
Oct4-mCitrine
Sox1
Oct4-mCitrine
Sox1


G
DAPI
Oct4 IRES GFP

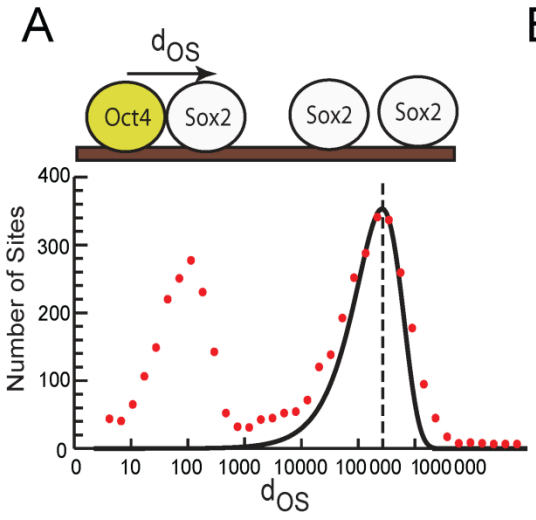

I
DAPI
Oct4 IRES GFP
Sox1
Sox1/
Oct4 IRES GFP

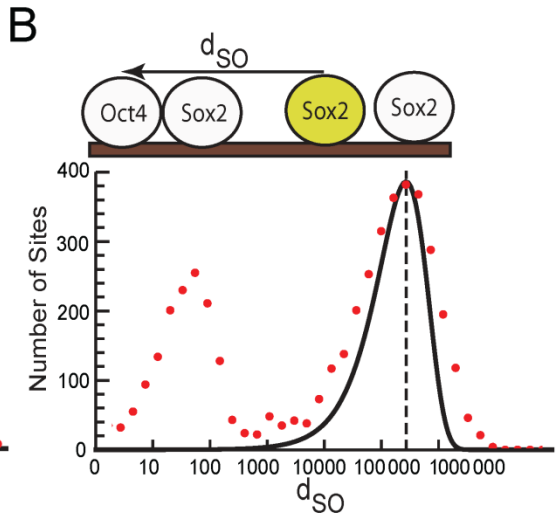

H
DAPI
Oct4 IRES GFP
Brachyury
Brachyury/
Oct4 IRES GFP

S4 related to Figure 4: Construction and validation of Oct4-mCitrine fusion transgenic cell line. (A) Scatter plot shows level of Oct4 protein in Oct4-IRES-GFP cell line from antibody staining and GFP level in pluripotency promoting conditions and 48 hours after pluripotency condition withdrawal. Withdrawal of pluripotency conditions produces a 73% drop in Oct4 protein level but only a 11% drop in GFP level. Since Oct4 and GFP are produced at the same rate, this result implies that Oct4 protein lifetime decreases during differentiation while the GFP lifetime remains fixed. (B) Diagram of the Oct4 transgene construct. Construct contains the Oct4 protein fused to mCitrine, a bright yellow fluorescent protein. The endogenous Oct4 3' UTR is also preserved. Construct was produced by engineering a BAC containing Oct4 and key enhancer regions (Tesar et al., 2007). (C) Images of Oct4-mCitrine cells growing under pluripotency promoting conditions (with LIF and BMP) immunostained for Nanog and Sox2. Images show that mCitrine signal is nuclear and that cells with Oct4-mCitrine also express ES cell markers Nanog and Sox2. (D) Oct4-mCitrine intensity correlates with Oct4 protein level in single cells. Scatter plot of nuclear mCitrine fluorescence signal versus Oct4 protein level measured by immunofluorescence for 295 cells growing under pluripotency promoting conditions (E) Images of Oct4-mCitrine cells differentiated into the mesendodermal (ME) lineage with CHIR show co-localization of Oct4-mCitrine protein and Brachyury, an ME marker, in the nuclei of individual cells. (F) Images of cells differentiated into the neural ectodermal (NE) lineage with retinoic acid show that Oct4-mCitrine and Sox1 (detected with immunofluorescence), an NE marker, are mutually exclusive. These data show that the Oct4-mCitrine cell line can be maintained in the pluripotent state and differentiated into both the ME and NE lineages. **(G-I) Cells**

of an Oct4-IRES-GFP line entering the mesendodermal lineage activate GFP while cells entering the neural lineage do not. We validated our results in a published Oct4-IRES-GFP line (Wernig et al., 2007). The cell line could not capture the fast dynamics of differentiation, due to long GFP lifetime. However, GFP expression patterns in terminal ME and NE cell populations are consistent with Oct4 activation in ME lineage and Oct4 absence in NE lineage. (G) Images of the Oct4-IRES-GFP cell line under pluripotency promoting conditions show expression of GFP. (H) Images of Oct4-IRES-GFP cells after mesendodermal differentiation with 200ng/ml Wnt3a for 24 hours stained for DAPI and Brachyury. Brachyury containing cells (red) have also activated Oct4-IRES-GFP (green), while many cells have undetectable levels of GFP. This result implies that cells entering the mesendodermal lineage have reactivated Oct4 transcription, since the IRES-GFP reports on transcription but not Oct4 protein level. (I) Images of Oct4-IRES-GFP cells differentiating into the neural ectodermal lineage stained for Sox1 and DAPI. Cells expressing Sox1 (red) have undetectable levels of GFP (green). The few cells that contain GFP do not express Sox1. GFP lifetime in mammalian cells is reported to be on the order of 30 hours. These results indicate that ES cells entering the neural lineage stop Oct4 transcription more than 30 hours before differentiation.

Figure S5



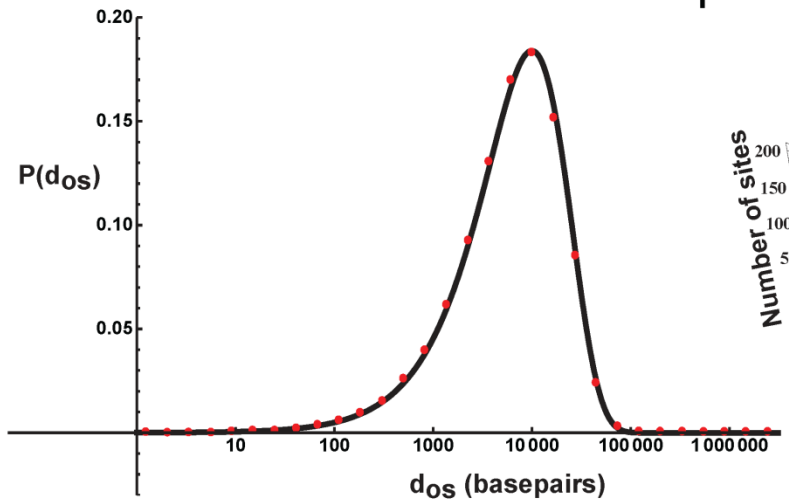


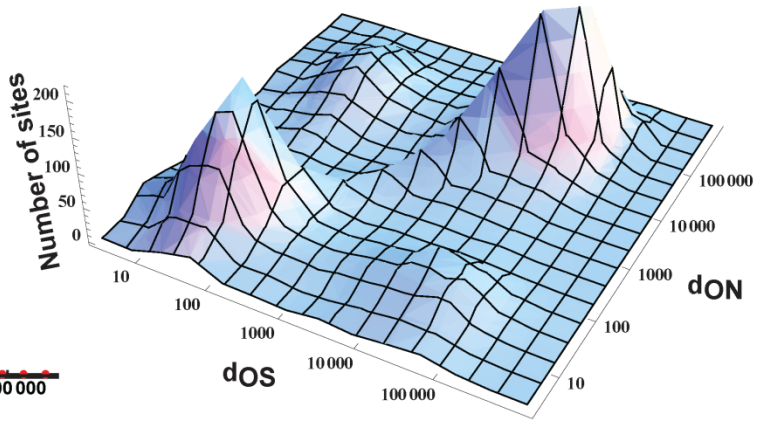
C

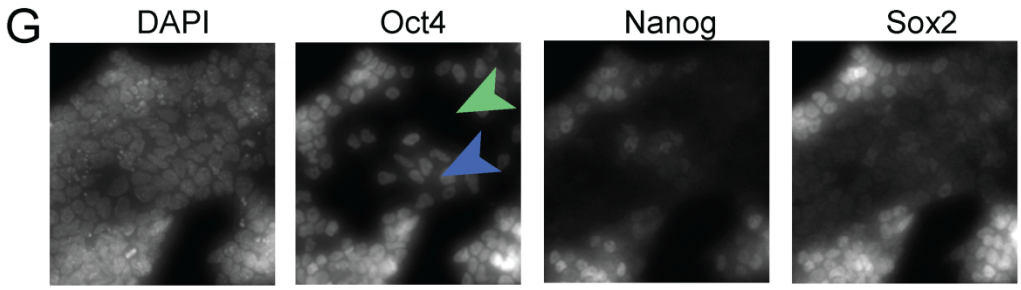
Gene	Monomeric Oct4 Site
Sox1	Coding Region(Exon)
Pax6	Coding Region(Intron)
Fgf8b	Coding Region(Exon)
Nkx6.2	Coding Region(Intron)

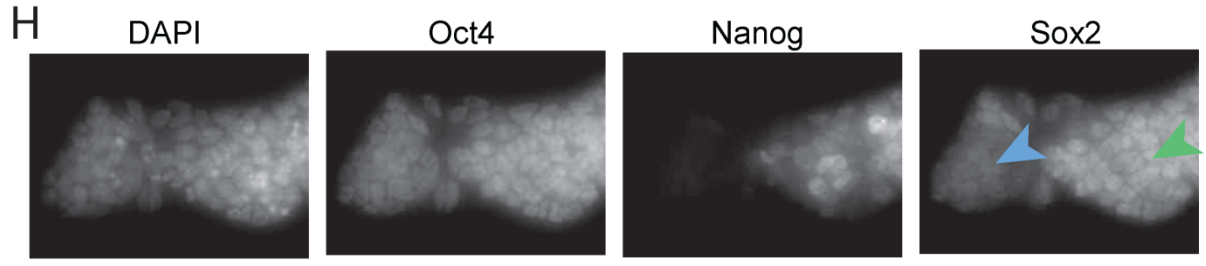
D

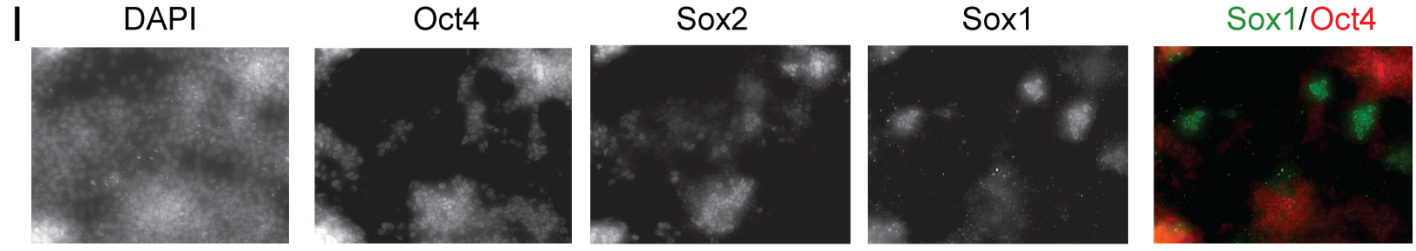
Gene	Monomeric Sox2 Site
Tcf2	Coding Region(Intron)
Wnt3a	Coding Region(Intron)
Foxa2	Nearest Gene
Brachyury	Nearest Gene









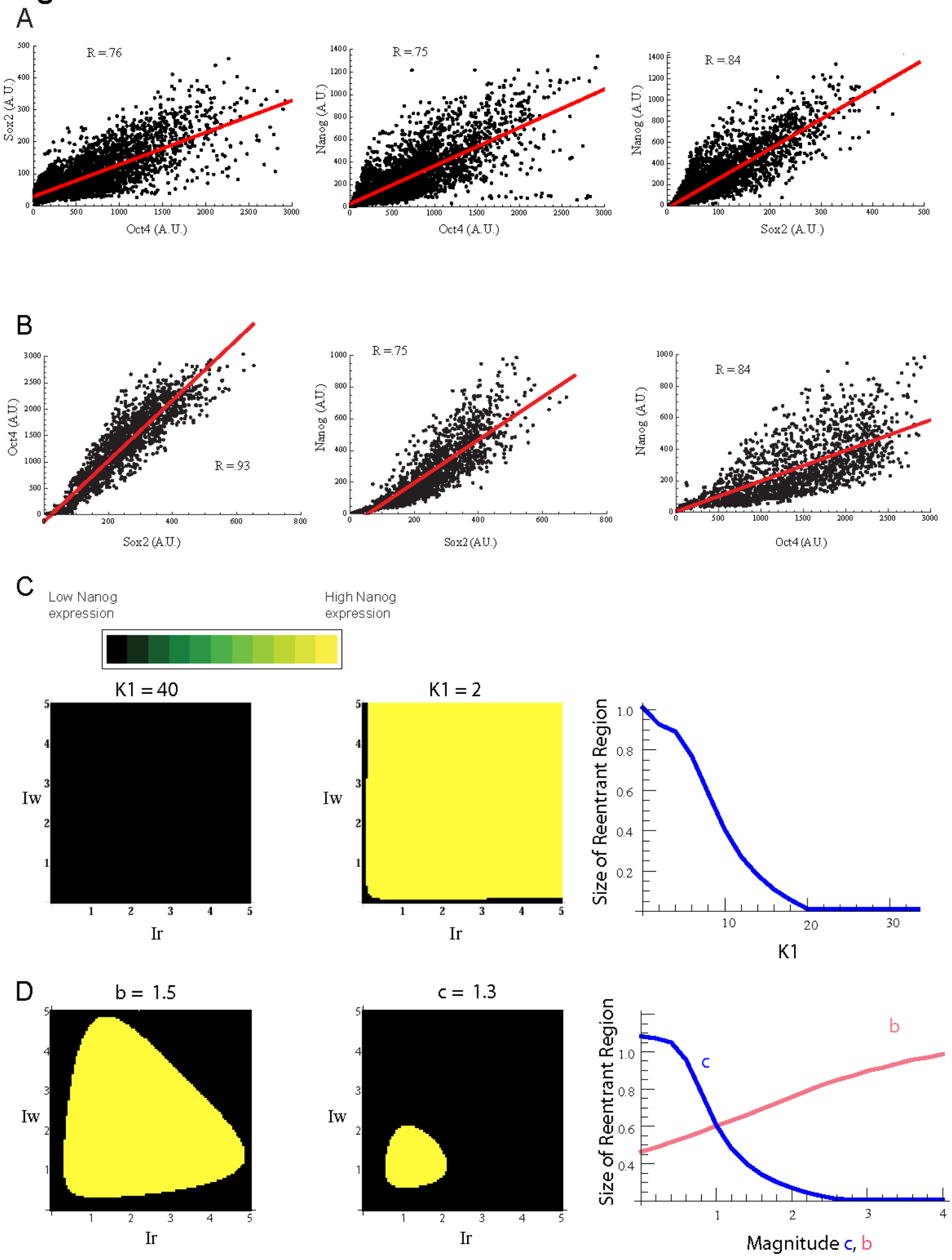


S5 related to Figure 5 and Figure 6: Bioinformatic analysis of ChIP-seq data reveals two classes of Oct4 and Sox2 binding sites and suggests that Oct4 and Sox2 bind both together and asymmetrically in ES cells. We analyzed published ChIP-seq data that measured Oct4 and Sox2 binding events in the ES cell genome to develop a hypothesis regarding the functional roles for independent Oct4 or Sox2 protein expression during ME and NE differentiation. The analysis is described in the Supplemental Experimental Procedures. (A,B) Histograms of the distance between each Oct4 binding site and the nearest Sox2 site, d_{OS} (A), or the distance between each Sox2 site and the nearest Oct4 site**,** d_{SO}, obtained from Chip-seq data (Chen et al., 2008). (B) shows two peaks, first one at 80bp, corresponding to dimeric Oct4-Sox2 binding sites, and a second larger one near 200000 base pairs, which fits a model with independent Oct4 and Sox2 sites (solid line shows fit to a single exponential, Supplemental Experimental Procedures). (C) Table of genes adjacent to isolated Oct4 binding sites illustrates binding near regulators of the neural ectodermal lineage (Sox1, Pax6). (D) Table of genes adjacent to isolated Sox2 binding events illustrates binding near key regulators of the mesendodermal lineage (Wnt3a, Brachyury, Foxa2). We performed the same analysis between Oct4 and Suz12, a member of the polycomb protein family that does not physically interact with Oct4 or Sox2, as a control. Oct4 and Suz12 distributions (as well as Sox2 and Suz12) had a single peak (data not shown). (E) Numerical validation of the random independent binding model (Supplemental Experimental Procedures). Red dots represent data from a numerical experiment, and the black line is $P_r(z) = \lambda \exp(-\lambda \exp(z)) \exp(z)$, for $\lambda = \frac{2x5000}{10^8} = \frac{1}{10000}$. The maximum of this distribution occurs at 10000, as we would expect from our analytical

results. (F) Two dimensional histogram of the distance between each Oct4 binding event and the nearest Sox2 binding event (d_{os}), and the distance between each Oct4 binding event and the nearest Nanog binding event (d_{ON}). This figure shows that Oct4, Sox2, and Nanog bind together (within 100 bp) at a large number of locations as seen by the peak near the origin of this plot. The plot also has a large peak corresponding to monomeric Oct4 sites, which are far away from the nearest Sox2 or Nanog Sites. Additionally, the plot has two smaller local maxima away from the diagonal, showing that Oct4 binds with Sox2 without Nanog and also with Nanog without Sox2. This would suggest at least 4 different kinds of Oct4 species: momomeric Oct4, Oct4-Sox2-Nanog trimer, Oct4-Sox2 dimer, and Oct4-Nanog dimer. **(G-H) Validation of Oct4 and Nanog siRNA constructs using immunofluorescence and microscopy.** We selected siRNA constructs that had been previously validated in the literature (Hu et al., 2009). We directly validated the constructs in our experimental system by probing for Oct4, Sox2, and Nanog expression in ES cells that had been transfected with siRNA against Oct4 or Nanog. Cells were grown in the presence of LIF and BMP and transfected with siRNA for 48 hours. Cells were fixed and stained for Oct4, Sox2, Nanog, and DAPI. (G) Images of cells transfected with Oct4 siRNA show that many cells in the transfected cell population contain undetectable Oct4 levels. Oct4 knock-down also induced down regulation of Sox2 and Nanog protein: cells without Oct4 have uniformly low levels of Sox2 and Nanog. Green arrowhead, transfected. Blue arrowhead, untransfected. (H) Images of cells transfected with the siRNA against Sox2 validate that the Sox2 siRNA construct induces knock down of Sox2 protein. Sox2 knock-down resulted in a decrease in Oct4 and Nanog protein levels. Blue arrowhead, transfected. Green arrowhead,

untransfected. (I) **Over expression of Sox2 and elimination of Oct4 leads to Sox1 expression even in pluripotency supporting conditions.** Images of cells grown in the presence of LIF and BMP. Sox2 expression is up-regulated by a constitutive Sox2 expressing plasmid in some cells, while Oct4 expression is abolished through siRNA is some cells. Cells were stained for DAPI, Oct4 (red), Sox2, and neural ectodermal marker Sox1 (green). Cells without Oct4 express the neural ectodermal marker Sox1, suggesting that appropriate modulation of Oct4 and Sox2 can induce neural ectodermal differentiation. Normally, no cell expresses Sox1 in the presence of LIF and BMP, as these factors strongly inhibit NE differentiation.

Figure S6


A
R = .76
Sox2 (A.U.)
Oct4 (A.U.)
R = .75
Nanog (A.U.)
Oct4 (A.U.)
R = .84
Nanog (A.U.)
Sox2 (A.U.)
B
Oct4 (A.U.)
R = .93
Sox2 (A.U.)
R = .75
Nanog (A.U.)
Sox2 (A.U.)
R = .84
Nanog (A.U.)
Oct4 (A.U.)
C
Low Nanog expression
High Nanog expression
K1 = 40
Iw
Ir
K1 = 2
Iw
Ir
Size of Reentrant Region
K1
D
b = 1.5
Iw
Ir
c = 1.3
Iw
Ir
Size of Reentrant Region
c
b
Magnitude c, b

S6 related to Figure 7:

siRNA knock down of Oct4 and Sox2 in ES cells reveals positive regulatory interactions between Oct4, Sox2, and Nanog. (A-B) Pair wise scatter plots of cells immunostained for Oct4, Sox2, and Nanog following siRNA knock down, including linear least square fits. Cell images are shown in S5G (Oct4 siRNA) and S5H (Sox2 siRNA). (A) siRNA mediated down regulation of Oct4 in pluripotency conditions induces down regulation of Sox2 and Nanog. After siRNA transfection, cells were grown for 48 hours in pluripotency promoting conditions. Since siRNA transfection is variable, some cells in the population are transfected with the siRNA while others contain unperturbed Oct4 protein. Thus, in a single scatter plot, we can compare Oct4 expressing cells to cells that lack Oct4. Plots demonstrate that Oct4, Sox2, and Nanog remain strongly correlated after Oct4 siRNA, and the Pearson correlation coefficient is shown in each case. On average, cells with low Oct4 contain lower Sox2 and lower Nanog than in untransfected cells. This implies that Oct4 drives Sox2 and Nanog, so that interference of Oct4 leads to down regulation of these factors. (B) siRNA mediated down regulation of Sox2 in pluripotency promoting conditions induces down regulation of Oct4 and Nanog. After siRNA transfection, cells were grown for 48 hours in pluripotency promoting conditions. Since transfection is variable, some cells in the population are affected by siRNA while others remain pluripotent. Thus, in a single scatter plot, we can compare Sox2 expressing cells to cells that lack Sox2. Plots demonstrate that Oct4, Sox2, and Nanog remain strongly correlated after Sox2 siRNA. Pearson correlation coefficients are shown. Thus, cells with low Sox2 have lower Oct4 and Nanog than in untransfected cells. This implies that Sox2 drives Oct4 and Nanog so that interference

with Sox2 leads to down regulation of these factors. Results are as predicted from a model of positive regulation between all factors. **(C-D) Analysis of mathematical model explores modulation of Nanog activation in presence of differentiation signals.** Plots of Nanog concentration as a function of signaling inputs generated from simulation of mathematical model. The x-axis is the level of signal I_r (RA, NE signal) and the y-axis is the level of signal I_w (Wnt, ME signal). Steady state Nanog levels were found for 10,000 points in (I_r , I_w) space. For each (I_r , I_w), Nanog level is illustrated with color (high Nanog in yellow and low Nanog in black; color bar). Yellow regions of Nanog expression correspond to differentiation inhibition, while black regions correspond to differentiation. (C) In the mathematical model described in the Supplementary Experimental Procedures and in Figure 7, Nanog production is regulated by Oct4, Sox2, and Nanog concentration through a Hill function that contains a threshold, K_1, that governs, in part, the size of the reentrant region: the region in which the pluripotency circuit is re-activated and the cells do not adopt either the ME or NE lineage when exposed simultaneously to the two differentiation signals. Left Panel, when the threshold is high (K_1= 40), the reentrant region disappears. In the presence of Wnt and retinoic acid, Oct4 and Sox2 do not reach high enough levels to activate Nanog expression. Middle panel, at lower threshold values (K_1 decreases to K_1= 2) the reentrant region expands to fill the plot. At this level of threshold, signals drive Oct4 and Sox2 to reactivate Nanog over a large region of the (I_r, I_w) space. Right Panel, modulation of K_1 continuously adjusts the size of the reentrant region. Here, we plot the fraction of (I_r, I_w) space in which Nanog becomes activated as a function of K_1. Increasing K_1 shrinks the size of this region. This is consistent with the plots shown in

the left and middle panels. (D) Modulation of relative strength of signal inhibition and Nanog repression changes the size of the reentrant region. Left Panel, increasing the strength of Nanog inhibition of signal by 50% ($b= b_w= b_r = 1.5$, Supplemental Experimental Procedures) increases the size of the re-entrant region relative to that shown in yellow in Figure 7. Increasing b represses the inhibitory effects of differentiation signals. This increases the region of (I_r, I_w) space in which the pluripotency circuit is re-activated upon addition of both signals. Middle panel, increasing strength of repressive effects of signal ($c = c_w= c_r = 1.3$) shrinks the reentrant region. Here, the inhibitory effects of signal become stronger. Signals, then, drive Oct4 and Sox2 to levels that cannot reactivate the pluripotency circuit. Right panel, by continuously adjusting the magnitude of $c = c_r = c_w$ (blue line), the size of the reentrant region is modulated. By increasing the strength of differentiation signals (increasing c), we can shrink the size of the reentrant region to 0. On the other hand, by modulating $b = b_r = b_w$ (pink line), the region of reentrance is continuously enlarged. Together, these figures show that the size (including presence or absence) of the reentrant region can be modulated continuously by adjusting the strengths of differentiation signals, Nanog repression, and the threshold for pluripotency circuit activation.

Supplemental Experimental Procedures

ES Cell Culture: Mouse embryonic stem cells were maintained and passaged on gelatin coated plates without feeder cells in N2B27 media containing LIF and BMP

without serum as described previously (Ying and Smith, 2003; Ying et al., 2008). ES cell media was supplemented with the FGF receptor inhibitor PD173074 (Sigma, StemGent) at 100 nM to suppress background differentiation. Similar results were obtained using the MEK inhibitor PD0325901 (Sigma, StemGent) at 1 μM. Cells were routinely propagated by trypsinization and re-plating every two to three days, with a split ratio of 1 in 10. N2B27 was prepared with B27 without vitamin A. Cell lines used in this study were Sox1-GFP,Oct4-IRES-GFP (Lengner et al., 2007; Wernig et al., 2007; Ying et al., 2003), and Oct4-mCitrine. In the Sox1-GFP reporter line, the open reading frame of one copy of Sox1 has been replaced with GFP. All cell lines were depleted of feeder cells and transitioned to serum free medium over several passages prior to experiments (Ying and Smith, 2003). N2B27 was prepared as described in (Gaspard et al., 2008; Ying and Smith, 2003; Ying et al., 2003).

ES Cell Differentiation: Differentiation protocols were adapted from established neural ectoderm and mesendoderm differentiation procedures to allow differentiation into either progenitor cell population under controlled monolayer conditions (ten Berge et al., 2008; Ying et al., 2003). The method described here enables NE or ME differentiation with cell fate determined by the addition of a single signal to the cell culture. ES cells were trypsinized, washed, and replated into N2B27 media at a density of ~ 1.5×10^4 $cells / cm^2$. After 48 hours, the old media was replaced with fresh N2B27. A single signal was added to this media to drive lineage specific differentiation. For mesendoderm differentiation, Wnt3a (200 ng/ml, R&D) or CHIR99021 (3 μM) was added to the N2B27

media after 48 hours in N2B27 (Jackson et al., 2010; ten Berge et al., 2008). CHIR is a specific Gsk3-beta inhibitor that is often used as a Wnt agonist due to its stability. Wnt3a and CHIR produced identical results in our experiments, but we favored CHIR because of its stability. For neural ectoderm differentiation, retinoic acid was added at 500 nM (Sigma) after 48 hours in N2B27 (Greber et al., 2010; Sterneckert et al., 2010; Ying et al., 2003). In this way, cells are treated identically for the first 48 hours, and the addition of a single signal drives cell fate choice. Cells were imaged in glass bottom dishes from MatTek or tissue culture treated imaging plates from Midsci. For differentiation experiments, cells were fixed and stained 36 hours after addition of the differentiation signal.

Immunofluorescence: Cells were grown on tissue culture treated plastic or glass bottom plates and fixed with 4% paraformaldehyde. Cells were permeabilized in PBS/0.3% Triton, blocked with 10% donkey serum, incubated with primary antibody, washed, and incubated with secondary antibody conjugated to either Alexa488, Alexa568, or Alexa647 (Invitrogen). After washing, cells were stained with DAPI and imaged. Images were acquired with a Zeiss Axiovision inverted microscope with a 20x Plan-Apochromat objective (NA .8) or 40× Plan-Apochromat objective (NA 1.3) with an Orca ER camera, CFP/YFP/DsRed triple band filter set (Chroma 61008) and DAPI filter set (Chroma 31000v2, 41008 647). Automated segmentation was performed in MetaMorph. Data was analyzed using custom written code in Mathematica. Antibodies in this study have all been validated in the literature: Oct4 (Santa Cruz sc-8628, 1:300; Stemgent 09-0023, 1:500; Cell Signaling 2840, 1:400 (Figure 6G)); Sox2 (Santa Cruz

,1:300; Cell Signaling 4900S, 1:500); Klf4 (GKLF (H-180): sc-20691, 1:300); Klf5 (Abcam ab24331, 1:600); Tbx3 (Santa Cruz sc-17871, 1:300); Nanog (COSMO Bio, 1:600); Sox1 (Santa Cruz sc-17317, 1:300); Nestin (Covance, 1:400); Foxa2 (Santa Cruz sc-6554, 1:300); Dnmt3a (Imgenex Corp, monoclonal, 1:800); Brachyury (Santa Cruz sc-17745, 1:300); FoxP1 (Cell Signaling #2005, 1:400); Klf9 (Santa Cruz sc-28195, 1:400); Jarid2 (ABCAM ab48137, 1:600); Brn2 (Santa Cruz sc-6029, 1:300); Gbx2 (Santa Cruz sc-22230, 1:600). Oct4, Sox2, and Nanog antibodies were validated by immunofluorescence staining of HEK cells that had been transfected with cDNA encoding these three genes.

Quantification and Statistical Inference for Immunofluorescence: Semi-automated image analysis was performed through custom written code in MATLAB (MathWorks) and Mathematica (Wolfram). Points on the conditional probability plot (Figure 3H) were quantified by measuring the probability of observing a given protein, conditioned on the presence of each lineage marker. This plot was produced by making automated presence or absence calls of protein expression after image analysis. On this plot, variably expressed proteins like Dnmt3a lie on the diagonal and are present with high probability in both Sox1 and Brachyury expressing cells. Lineage specific proteins like Oct4 and Sox2 fall on the extreme off-diagonal of the plot and are present with high probability in either ME or NE progenitor cells but not both, consistent with the images in Figures 3A,C. The results were consistent with those obtained by the analysis of the Kullback-Leibler distance between the full distribution function of the expression level of each protein in Brachyury positive and Sox1 positive cells.

Mutual information analysis was used to quantify the amount of information that each transcription factor provided about the cell's underlying lineage choice. Mutual information measures the ability of one variable (transcription factor expression) to explain variation in another (ME or NE lineage selection). We found that Oct4 and Sox2 together explain up to 89% of the cell's lineage decision while factors like Dnmt3a, Klf9, and Tbx3 explain less than 1%. Further, Oct4 and Sox2 expression patterns are highly unlikely to be observed by chance. For example, if Oct4 was expressed with equal probability in ME and NE cells, then the probability of observing lineage choice dependent differential expression greater than or equal to that shown in Figure 3 is less than $p=10^{-19}$. Similarly, the p-values for obtaining the observed lineage specific pattern by chance are $p< 10^{-19}$ for Sox2 and $p < 10^{-13}$ for Klf5. This analysis enabled the placement of proteins into one of the three groups of transcription factors based on protein expression patterns.

Live Cell Fluorescence Microscopy: For live cell fluorescence microscopy, cells were maintained in differentiation media as described above without phenol red and were seeded on glass bottom dishes (MatTek) that were deposited with a nanofilm composed of alternating layers of polystyrene sulfonate (PSS) and polyethyleneimine (PEI) (Ai et al., 2003). Cells were imaged on a Zeiss Axiovision inverted microscope with a 20x Plan-Apochromat objective (NA .8) or 40× Plan-Apochromat objective (NA 1.3) with an Orca ER camera, CFP/YFP/DsRed triple band filter set (Chroma 61008) and DAPI filter set (Chroma 31000v2, 41008 647). The microscope stage was enclosed with an environmental chamber incubator that maintained a temperature of 37°C (Pecon). Images were acquired every 10 or 30 min for 6–48 hr. Image acquisition was controlled

by MetaMorph Software (Molecular Devices); image analysis was done with ImageJ (NIH), MetaMorph, Mathematica (Wolfram Research), and MATLAB (MathWorks).

Analysis of Published Microarray Data:

Microarray data in Figure 1 is from (Shen et al., 2008) and was obtained from GEO (GSE12982). All arrays probes were considered independently. We used GO categories to identify all transcription factors and DNA binding proteins in the mouse genome. From this set, we identified a gene as being expressed in the ES cell if its mean expression level across ES cell microarray replicates was above 600. To select the threshold, we studied the mean expression level of a set of transcription factors known to be expressed in the ES cell. We selected 600 as a number that was below the mean level of all factors but also several fold above the noise floor (~200) on the Affymetrix array. A higher threshold of 800 would result in an identical analysis except that Tbx3 would be removed from our list of candidate genes. We structured the analysis to remove spurious signatures of expression and fold changes due to low expression levels or microarray noise.

Using the threshold, we identified genes that are expressed in ES cells, and calculated their fold down regulation in the ME and NE lineages. We calculated the distance of each transcription factor from the diagonal of plot Figure 1C and selected the top ten genes for further analysis. Each selected gene was expressed above 600 in ES cells and in at least one of the ME or NE progenitor cell samples. Again, our threshold ensured that we were studying genes that were likely to be present in the ME or NE cell

population and not measurement artifacts. We calculated the p-value for the distance of these genes from the diagonal seen in Figure 1C by estimating the probability of this distance occurring solely due to statistical fluctuation. Thus, our null model was that the gene was not asymmetrically expressed in the two lineages but was positioned away from the diagonal solely due to noise. We used the replicates of microarray data in the ME lineage to estimate mean and variance of the distribution function for the level of expression of a given gene. We assumed that the statistics of the expression level of was normally distributed with this mean and variance. Using this distribution, we calculated the probability of obtaining the observed NE expression value for that gene. Similarly, we calculated the probability of obtaining the ME data based on the distribution of the expression levels of the gene in the NE microarray replicates. We used the largest of these probabilities as the p-value for finding the observed differential expression level of the genes by chance. These p-values are tabulated in Figure S1C.

Gene Expression Microarrays for ES vs. 48 hr state:

ES and differentiated cells were cultured and harvested in triplicates. RNA was extracted using Trizol, phenol:chloroform extraction, and isopropanol precipitation. Biotinylated and amplified cRNA was prepared by the Illumina TotalPrep RNA Amplification Kit (Applied Biosystems/Ambion). Biotinylated cRNA was hybridized onto Illumina Mouse-Ref8 BeadChip microarrays, according to the manufacturer's protocol. Microarrays were analyzed by an Illumina BeadArray Reader, and raw data was normalized in BeadStudio. We ensured that our statistical analysis (Figure 2B) did not

depend on normalization scheme (quantile versus average normalization). Quantile normalization and average normalization produced similar results. In Figure 2B, we plot statistics of the quantile normalized data from BeadStudio. No additional thresholding or post-processing was performed on this data. Microarray data were deposited at GEO GSE29005.

Bioinformatic Analysis of Distance between Oct4 and Sox2 Binding Events:

To gain insight into a possible functional role for Oct4 and Sox2 protein expression patterns, we analyzed the pattern of Oct4 and Sox2 binding in the ES cell genome using published data to ascertain whether Oct4 and Sox2 always regulate gene expression as a dimer or whether other monomeric modes of regulation might occur even prior to differentiation.

If Oct4 and Sox2 act as a dimer, Oct4 binding sites, as determined by ChIP-Seq measurements, should occur very close to Sox2 binding sites. Therefore, we studied the distance, d_{OS}, between each measured Oct4 binding event and the Sox2 binding event nearest to this site. From the ChIP-Seq data (Chen et al., 2008; Marson et al., 2008), we can measure the distribution of d_{OS} and ask if this distribution, $P_{ChIP}(d_{OS})$, is consistent with dimeric binding. If Oct4 and Sox2 act primarily as a dimer, the distribution should contain a length scale that is characteristic of this dimeric interaction. On the other hand, if Oct4 and Sox2 bind independently in the genome, the distribution of d_{OS} will have a form characteristic of a model in which Oct4 and Sox2 bind independently and randomly in the genome.

When we extracted the distribution, $P_{ChIP}(d_{OS})$, from published data, we found that the distribution has two peaks, so that d_{OS} has two characteristic length scales. The first peak occurs at a length of d_{OS} = 47 base pairs and is consistent with short range, dimeric interactions between Oct4 and Sox2. The second peak, however, occurred at d_{OS} = 200000 base pairs. This length scale suggested independent binding between Oct4 and Sox2. To probe this hypothesis, we developed an analytical model of independent binding between Oct4 and Sox2 and studied the structure and length scales associated with such a model distribution.

Distribution of d_{OS} for Independent Binding Model

In the random and independent model, Oct4 and Sox2 occupy randomly placed sites on the genome, and the factors bind these sites independently of one another. For this model, we can analytically calculate, $P_r(d_{OS})$, the distribution function for d_{OS} where the subscript r denotes the random model. As before, d_{OS} is the distance between an Oct4 binding event and the nearest Sox2 binding event.

To calculate $P_r(d_{OS})$, we consider a model with two parameters: G, the genome size, and, N, the number of Sox2 binding events. For each Oct4 binding event in the genome, we calculate the probability of binding the nearest Sox2 binding event to be $d_{OS}+1$ units away. This probability includes two conditions: **Condition A**: no Sox2 binding event will occur within d_{OS} base pairs (on either side) of the Oct4 event, and **Condition B**: that one Sox2 binding event will occur at exactly $d_{OS}+1$ base pairs away.

These two conditions lead to two terms in the probability distribution function. Condition A requires that $(N-1)$ Sox2 binding events occur outside of d_{OS}. For a single Sox2 binding event, the probability that the event falls within a region of size d_{OS} is $\frac{d_{OS}}{G}$. $\frac{2d_{OS}}{G}$ is the probability that a Sox2 binding event lies within a region of size d_{OS} on either side of Oct4. Thus, $1-\frac{2d_{OS}}{G}$ is the probability that a given Sox2 binding event falls outside this region. Thus, for a single Sox2 binding event, the probability that this binding event falls beyond the d_{OS} region depends only on the size of this region relative to the size of the genome, G.

The probability that $(N-1)$ Sox2 sites fall outside a region of size $2d_{OS}$ is thus:

$$(1-\frac{2d_{OS}}{G})^{N-1} \qquad (1)$$

Since G, the genome size, is much larger than the typical distance between Oct4 and Sox2 binding events, we can approximate this term as:

$$(1-\frac{2d_{OS}}{G})^{N-1} \approx (1-\frac{2d_{OS}N}{GN})^{N} \approx \exp\left(-2d_{OS}\frac{N}{G}\right), \qquad (2)$$

where we have used the fact that $(1-\frac{x}{m})^{m} \to \exp(-x)$ as m becomes large. In our model, we are interested in $d_{OS} \ll G$ and $\frac{Nd_{OS}}{G}$ of order 1.

Finally, from condition B above, we must account for a Sox2 binding event occurring at exactly $d_{OS}+1$ base pairs from the Oct4 binding event. The probability that a given Sox2 binding event occurs $d_{OS}+1$ from the Oct4 binding event is $\frac{2}{G}$ as there are 2 sites $d_{OS}+1$ base pairs from a given Oct4 binding event. Such a configuration can occur in

N different ways since any of the N Sox2 events can occupy this position. In this way, the total probability distribution meeting both conditions A and B is given in the limit of large N by

$$P_r(d_{OS}) = \frac{2N}{G}\exp\left(-2d_{OS}\,\frac{N}{G}\right) = \lambda\exp(-\lambda d_{OS}), \quad (3)$$

and $\lambda = \frac{2N}{G}$. The mean density of binding events in the genome is $\frac{N}{G}$ and 2 is a geometric factor. Thus, the null model for $P_r(d_{OS})$ is an exponential distribution where the typical distance between an Oct4 binding event and the nearest Sox2 binding event is $\frac{G}{2N}$.

Given the small number of Oct4 and Sox2 binding events relative to the genome size, we binned d_{OS} on a log scale. On such a scale, the exponential distribution has the form:

$$P_r(z) = \lambda\exp(-\lambda\exp(z))\exp(z), \quad (4)$$

where $z = \log(d_{OS})$. The distribution, $P_r(z)$, will have a maximum at $z = -\log(\lambda)$.

Numerical Test of Analytical Model

Before using our random and independent models to analyze the ChIP-Seq data, we evaluated the model on simulated binding data through numerical experiments. We performed numerical tests with $N = 5000$ and $G = 10^8$ so that these parameters were similar to the number of measured Sox2 binding sites and the size of the mouse genome.

In these numerical tests, we placed 5000 points at random on a line of length $G = 10^8$. These points represent randomly placed Sox2 binding sites. We then selected an additional point at random, simulating the placement of an independent Oct4 binding event, and found the Sox2 point nearest to this randomly placed Oct4 binding event. We calculated the distance between the Oct4 site and this closest Sox2 site, d_{OS}.

We performed $50{,}000$ iterations of this numerical experiment, plotted the distribution of d_{OS} on a log scale, and compared this to the analytically calculated density function, $P_r(z)$ for $G = 10^8$ and $N = 5000$. We selected a genome size with $G = 10^8$ in the numerical example because our model becomes more accurate as the genome size grows. Therefore, as a stringent test of the model, we selected a genome size ten times smaller than $G = 10^9$ for the mouse genome. The agreement between the numerical simulation and our model is shown in Figure S5E.

Bioinformatic Analysis of Published ChIP-seq data:

Using the published data (Chen et al., 2008; Marson et al., 2008), we extracted the distribution of d_{OS}, $P_{ChIP}(d_{OS})$. We found that $P_{ChIP}(d_{OS})$ has two peaks rather than a single peak. The first peak accounts for dimeric sites where Oct4 and Sox2 bind close to one another in the genome. This peak occurred at 47 base pairs, consistent with short range dimeric binding. However, the second peak fit the independent binding model described above. The distribution derived from the data thus has an exponential shape, characteristic of the random binding model. Further, the peak of this portion of the distribution occurred at 200000. In the independent binding model, the peak occurs

at $\frac{1}{\lambda} = \frac{G}{2N}$. Our measured decay length of 200000 is consistent with such a prediction. Given that $N \approx 5000$ and $G \approx 2.5 * 10^9$for the mouse genome, the peak should occur precisely around 200000 base pairs.

In this way, analysis of previous ChIP-seq data suggests that Oct4 and Sox2 bind both as a dimer and independently in the mouse genome.

Plasmid Transfection: Overexpression studies used plasmids from Yamanaka with Oct4 or Sox2 driven from the CAG promoter. (Oct4, Addgene 13461; Sox2, Addgene 13462) (Mitsui et al., 2003). Overexpression constructs were added to cells 24 hours prior to addition of differentiation signals. Transfections were carried out using Lipofectamine 2000 (Invitrogen) as described (Hu et al., 2009).

siRNA: We selected siRNA constructs that had been described and validated by Elledge (Hu et al., 2009). Constructs were validated by transfection in ES cells followed by immunofluorescence and microscopy. siRNA (Invitrogen) transfection was performed using Lipofectamine 2000 (Invitrogen) as described by Elledge. Cells were transfected in 24 well plates with 10 pMol of siRNA. Constructs used: Nanog: MSS231180 (UACGUAAGGCUGCAGAAAGUCCUCC); Oct4: MSS237605 (CCAAUGCCGUGAAGUUGGAGAAGGU). For siRNA transfection during mesendodermal differentiation, Oct4 siRNA was added with Sox2 plasmid 24 hours

prior to differentiation signal addition. For Nanog siRNA in Figure 1, Nanog siRNA was added to cells at 24 hours prior to CHIR addition.

Generation of Oct4-mCitrine Transgenic ES Cell Line: The Oct4-mCitrine fusion construct was generated using Red/ET Recombination (Gene Bridges, Dresden Germany). An insert containing homology to 50bp upstream of the Oct4 stop codon, the glycine/serine-rich linker, mCitrine cDNA, the native 3' UTR and pA site of Oct4, the loxP-PGK-gb2-neo-loxP selection cassette (Gene Bridges), and homology to 50bp downstream of the Oct4 3' UTR was generated by fusion PCR and validated by sequencing the final amplicon. This fusion insert was electroporated into BAC strain RP24-248K18 (CHORI) harboring Oct4, and Red/ET Recombination was performed. Properly modified BAC clones were screened by PCR. A BAC fragment containing the distal and proximal enhancers of Oct4 (Minucci et al., 1996; Tesar et al., 2007), the Oct4 ORF, the fusion insert, and 2.1kb of 3' homology was subcloned into pColE1 using Red/ET Recombination. DNA linearized by I-PpoI was electroporated into V6.5 ES cells, and colonies were selected on G418 (200 µg/mL). Clones that stably retained mCitrine fluorescence in pluripotency maintaining conditions for many passages were selected and further characterized by immunofluorescence, PCR, Southern blotting, and flow cytometry. The mCitrine signal was confirmed to be nuclear and also correlated with Oct4 immunofluorescence staining both under pluripotency promoting conditions and during differentiation with a correlation coefficient of R= .92 (Figure S4).

FACS: ES cells were trypsinized and fixed in suspension with formaldehyde (1.5% final concentration), permeabilized with ice cold methanol (90%), and washed twice with PBS containing 5% normal donkey serum. Finally, cells were stained with primary antibodies (listed above) diluted in PBS containing 1% BSA, and detected using fluorescent-conjugated secondary antibodies raised in donkeys. Flow cytometry was performed on a BD LSRII flow cytometer equipped with 405 nm, 488 nm, 561 nm, and 633 nm lasers. The data acquired were analyzed using custom programs written in MATLAB (MathWorks) and Pascal.

Chromatin Immunoprecipitation and qPCR:

Chromatin Immunoprecipitation (ChIP) was adapted from (Mikkelsen et al., 2007). Cells were trypsinized, crosslinked in 1% formaldehyde for 12 min at 37°C, followed by quenching in 125 mM glycine for 5 min. Cells were pelleted and washed 3x in PBS supplemented with Complete protease inhibitors (Roche) and then flash frozen in liquid nitrogen. Individual pellets were resuspended in 1mL SDS lysis buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl, pH 8.1) and incubated for 10 min on ice. Sonication of crosslinked DNA was performed using a Branson 250D Sonifier at 40% power (0.7 s pulse, 1.3 s pause) for 3 min. Sheared DNA samples were diluted into 10 mL ChIP dilution buffer supplemented with Complete protease inhibitors (Roche). Samples were incubated overnight with rotation along with 2 µg/mL polyclonal goat anti-Oct4 (Santa Cruz Biotechnology, sc8628-x) or 2 µg/mL polyclonal goat anti-Sox2 (Santa Cruz Biotechnology, sc17320-x) antibodies as described in (Chen et al., 2008). Chromatin targets were immunoprecipitated using 100 µL aliquots of Protein-G conjugated agarose bead slurry (Sigma) for 1 hr, followed by sequential washing using

the following buffers: twice with low salt immune complex buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl, pH 8.1, 150 mM NaCl), LiCl wash buffer (0.25 M LiCl, 1% NP40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl, pH 8.1), and TE buffer for 5 min at 4ºC per wash. Precipitated DNA was eluted in two iterations of 125 µL 0.2% SDS, 0.1 M NaHCO3, 5 mM DTT at 65ºC for 10 min followed by 5 hr reverse crosslinking at 65ºC, Proteinase K (Invitrogen) treatment, phenol:chloroform extraction, and ethanol precipitation. Resuspended DNA was normalized by concentration using a Qubit Fluorometer (Invitrogen) such that the concentrations of the ChIP samples matched the concentrations of the whole cell extract inputs, which were used to quantitatively normalize enrichment scores for downstream qPCR. To compute p-values for key peaks of enrichment, we used as our null model, the distribution of qPCR fold enrichments over all the loci probed in a given ChIP sample. Then, we computed the probability of obtaining the enrichment values at a region of interest using this null model.

QPCR was performed in triplicates on an ABI 7900HT using primers from ChIP-qPCR tiling arrays (SABiosciences) and custom primer sets (Operon) as listed below. Thermal cycling was performed as follows: 95ºC 10 min; 50 cycles of 95ºC 20 s, 55ºC 30 s, 72ºC 45 s. Fold enrichment was calculated using the $2^{\Delta Ct}$ method, where ΔCt = Ct(normalized Whole Cell Extract) – Ct(normalized ChIP DNA). Dissociation curve analysis was performed for each reaction, and replicates that produced multiple or incorrect products were omitted from enrichment calculations. Error bars represent mean ± SEM of technical or biological replicates, as stated in figure captions.

We observed variable qPCR results for the Sox1 locus in the whole cell lysate as well

as in the pull downs in the ES, ME and NE cell types. This could be due to the high CpG density across regions of the Sox1 promoter probed by our tiling primers.

Primers for ChIP-qPCR

Gene	Location	F	R	SABiosciences ID
Oct4	-6660	N/A	N/A	GPM1033119(-)07A
Oct4	-5388	N/A	N/A	GPM1033119(-)06A
Oct4	-4387	N/A	N/A	GPM1033119(-)05A
Oct4	-3384	N/A	N/A	GPM1033119(-)04A
Oct4	-2300	agggcacatctgtttcaagc	Ctggccaggacaagagacat	N/A
Oct4	-1800	Ctctcgtcctagcccttcct	Atctctctggccctctccat	N/A
Oct4	-380	N/A	N/A	GPM1033119(-)01A
Sox2	-5606	N/A	N/A	GPM1035441(-)06A
Sox2	-4578	N/A	N/A	GPM1035441(-)05A
Sox2	-3526	N/A	N/A	GPM1035441(-)04A
Sox2	-2264	N/A	N/A	GPM1035441(-)03A
Sox2	-1010	N/A	N/A	GPM1035441(-)02A
Sox2	-500	N/A	N/A	GPM1035441(-)01A
Sox2	3500	Ctcagcctctaggcctgtgt	Cccttcccagtaccttaccc	N/A
Sox2	3750	Gcacagtcgacagttcttgc	Aggctgagtcgggtcaatta	N/A
Sox2	3905	N/A	N/A	GPM1035441(+)04A
Sox2	4250	Gataaactgcagcgctaccc	cctcggaaagaagtcacagg	N/A
Nanog	-5280	N/A	N/A	GPM1038247(-)06A
Nanog	-4582	N/A	N/A	GPM1038247(-)05A
Nanog	-3514	N/A	N/A	GPM1038247(-)04A
Nanog	-2228	N/A	N/A	GPM1038247(-)03A
Nanog	-1012	N/A	N/A	GPM1038247(-)02A
Nanog	-398	N/A	N/A	GPM1038247(-)01A
Nanog	1460	N/A	N/A	GPM1038247(+)02A
Brachyury	-4500	ctttggactcatccctctgc	Ggaggagaatggcaaactga	N/A
Brachyury	-4250	ttccagtggtcggtacacct	Atccaaaccacacgtcacag	N/A
Brachyury	-4000	tgagcgcctgtaatgatttct	taaagtcaggtgccacagca	N/A
Brachyury	-3288	N/A	N/A	GPM1032778(-)04A
Brachyury	-2496	N/A	N/A	GPM1032778(-)03A
Brachyury	-1282	N/A	N/A	GPM1032778(-)02A
Brachyury	-521	N/A	N/A	GPM1032778(-)01A
Sox1	-3508	N/A	N/A	GPM1039702(-)04A
Sox1	-2544	N/A	N/A	GPM1039702(-)03A

Sox1	-1382	N/A	N/A	GPM1039702(-)02A
Sox1	-188	N/A	N/A	GPM1039702(-)01A
Sox1	1324	N/A	N/A	GPM1039702(+)02A

Mathematical Modeling of Pluripotency Circuit:

A key feature of the circuit diagram shown in Figure 7a is the existence of positive regulatory interactions between Oct4, Sox2, and Nanog that compete with the asymmetric influence of differentiation signals on Oct4 and Sox2. We hypothesized that a transcriptional regulatory circuit containing such competing regulatory interactions could produce non-intuitive dynamic behaviors in response to multiple simultaneous signals. Therefore, we developed a mathematical model to probe the response of this circuit to combinations of differentiation signals.

As in many biological systems, we lack detailed biochemical information about the structure of interactions between circuit components and promoter regulatory elements, let alone the rate constants for these interactions. For example, we do not know exactly how Oct4, Sox2, and Nanog bind together to regulate transcription and whether this binding occurs through a multimeric complex. Based on ChIP-seq data (Chen et al., 2008), there are multiple binding elements for each of these factors in the promoters and enhancers of the others. While we lack detailed biochemical information, our experiments provide information regarding the topology of the transcriptional circuit. We built a mathematical model of the system that makes as few assumptions as possible about circuit details, while probing the consequences of circuit topology. We used the

model to probe the potential signal integration properties of a transcriptional circuit in which positive feedback between components competes with the asymmetric influence of external differential signals.

Model Captures Key Features of Circuit Topology

Our experiments show that Oct4 and Sox2 individually inhibit differentiation. Therefore, we use our model to study the concentration of Oct4 and Sox2 in the presence of conflicting signals. In ES cells, biochemical evidence suggests that Oct4 and Sox2 might control transcription through a heterodimeric interaction. In our model, we assumed that such a dimer forms rapidly on the time-scale of transcription, so that the concentration of this heterodimer is proportional to the product of monomer concentrations. This assumption enables us to construct a model that is compatible with potentially complex biochemical interactions.

Our model has three key features based on the experimental results in the main text. (1) Differentiation signals, Wnt and RA, asymmetrically regulate Oct4 and Sox2. (2) Oct4, Sox2, and Nanog drive one another through a positive regulatory loop (Figure S6). (3) Nanog blocks the influence of differentiation signals on Oct4 and Sox2 (Figure 2). In our experiments, Nanog falls during the early phases of differentiation along with many other factors, like Klf4 and Tbx3, that are important for pluripotency but do not play a role during lineage choice. Therefore, we include Nanog in our mathematical model as a proxy for its role and the roles of factors like Tbx3 and Klf4 in positive regulation of the pluripotent state.

The model explores the dynamic response of the pluripotency circuit to differentiation signals when signals are added to cells 48 hours after the withdraw of pluripotency promoting factors LIF and BMP. In the model, we focus on the response of Oct4, Sox2 and Nanog to signal and do not model the activation of Brachyury and Sox1. Our experiments suggest that Brachyury and Sox1 can become activated if Oct4 or Sox2 is down-regulated. Therefore, we sought to explore the relative levels of Oct4 and Sox2 when the circuit is driven by multiple input signals.

The three key structural features described above can be modeled using the following equations:

$$\frac{d[Nanog]}{dt} = f([Oct4],[Sox2],[Nanog]) - \frac{[Nanog]}{\tau_n} \quad (5)$$

$$\frac{d[Oct4]}{dt} = \frac{I_w}{K_w + I_w} + f([Oct4],[Sox2],[Nanog]) - \frac{[Oct4]}{\tau_o(I_r,[Nanog])} \quad (6)$$

$$\frac{d[Sox2]}{dt} = \frac{I_r}{K_r + I_r} + f([Oct4],[Sox2],[Nanog]) - \frac{[Sox2]}{\tau_s(I_w,[Nanog])} \quad (7)$$

where [Oct4], [Sox2] and [Nanog] are the concentrations of Oct4, Sox2, and Nanog, respectively. I_w is the concentration of Wnt signal, and I_r is the concentration of retinoic acid.

The model has three key terms. First, Wnt signal, I_w, drives Oct4 production, and retinoic acid, I_r, drives Sox2. The terms $\frac{I_w}{K_w + I_w}$ and $\frac{I_r}{K_r + I_r}$ represent this drive, and these terms saturate at a level of signal, K_w or K_r. Second, the term $f([Oct4],[Sox2],[Nanog])$ accounts for the transcriptional influence of [Oct4],[Sox2],

and [Nanog] on one another. Published biochemical data suggest that Oct4, Sox2, and Nanog interact with one another both in solution and on DNA. For example, Oct4 and Sox2 form a heterodimer that regulates expression of genes including Oct4, Sox2, and Nanog (Chew et al., 2005; Rodda et al., 2005). However, little is known regarding the precise role of each biochemical complex or the kinetics of complex formation. Given this, we assumed that all complexes form rapidly compared to the time scale of transcription. This separation of time scales allows us to treat the complexes as being at equilibrium, so that the concentration of a complex (eg [Oct4-Sox2]) is proportional to the concentration of individual monomers, $[Oct4 - Sox2] = K_{OS}\ [Oct4][Sox2]$, so that the equilibrium binding constant K_{OS} relates the monomer concentration to the heterodimer concentration. Therefore, our model needs to account only for the concentrations of monomers.

Our functional experiments suggest that Oct4, Sox2, and Nanog regulate one another's protein expression level. For example, siRNA experiments demonstrate that down-regulation of Oct4, Sox2, or Nanog individually leads to the down-regulation of the other two factors. Such functional evidence is consistent with published Chip-seq data that show Oct4, Sox2, and Nanog binding at one another's promoters. Therefore, we chose functional forms that do not make assumptions about detailed biochemistry, but that are compatible with models that account for biochemical complex formation. Our experiments suggest that f is a monotonically increasing function of its arguments, so that Oct4, Sox2, and Nanog drive the transcription of one another. We can think of this drive as occurring through a biochemical complex whose concentrations is proportional to the product $([Oct4] * [Sox2] * [Nanog])$.

Third, each protein has a characteristic lifetime. Nanog lifetime, τ_n is constant. However, the lifetimes of Oct4 and Sox2, $\tau_o$and τ_s, are functions of differentiation signals and Nanog concentration. The terms $\frac{[Oct4]}{\tau_o(I_r,[Nanog])}$ and $\frac{[Sox2]}{\tau_s(I_w,[Nanog])}$ model the asymmetric influence of differentiation signals on Oct4 and Sox2. The lifetime of Oct4, τ_o is a function of retinoic acid concentration since we observed experimentally that retinoic acid appears to decrease the Oct4 protein lifetime. τ_o is also a function of Nanog concentration as, in our experiments, Nanog inhibits the influence of retinoic acid on Oct4 lifetime. Experimentally, we know that cells expressing Nanog do not respond to differentiation signals. This term accounts for the opposing effects of retinoic acid and Nanog on Oct4 lifetime. τ_s, the Sox2 lifetime, is a function of Wnt signal as well as Nanog. Wnt acts to decrease Sox2 lifetime while Nanog inhibits this effect. The equations above capture the essential topological features of the transcriptional regulatory circuit that we found in our experimental data. The key feature of the model is the interaction between $f([Oct4],[Sox2],[Nanog])$, $\tau_s(I_w,[Nanog])$, and $\tau_o(I_r,[Nanog])$.

Implementation of Pluripotency Circuit Model and Numerical Simulation

To simulate the circuit, we selected specific forms for f, $\tau_o$and τ_s. We selected functional forms that do not commit to a specific biochemical implementation. Specifically, the action of transcription factors at a promoter is often modeled as a Hill function. The Hill function accounts for cooperativity, thresholding and saturation, three common features of transcriptional regulation (Bintu et al., 2005; Carey et al., 1990; Elowitz and Leibler, 2000; Francois et al., 2007). Hill functions can arise in

transcriptional responses due to multiple transcription factor binding sites in a promoter region. ChIP-seq data (Chen et al., 2008) indicate that Oct4, Sox2, and Nanog all exhibit multiple pluripotency factor binding events. Such a binding pattern is consistent with cooperative regulation.

Using Chip-seq data analysis, we studied higher order interactions of Oct4, Sox2, and Nanog in the genome by finding the closest Nanog and Sox2 site to each Oct4 binding event (Figure S5F).

Further, published experiments have suggested that Oct4 and Sox2 form a heterodimer and that Nanog might interact with this complex (Chew et al., 2005; Rodda et al., 2005; Wang et al., 2006). The detailed biochemical structure of these complexes and their interactions with promoter elements have not been characterized.

For simplicity, we assumed that biochemical complex formation is often fast compared to the timescales of transcription and that the different complexes exist in equilibrium. Thus, complex equilibration can be treated as an equilibrium process, and the concentration of any single complex (eg an Oct4-Sox2 heterodimer) is proportional to the products of concentration of its constituents. Given our limited knowledge, we chose to use the simplest model which might be consistent with higher levels of complexity. While complexes might be important for the mechanistic details of regulation, our experiments show that Oct4 and Sox2 alone can repress differentiation. The model, however, is consistent with a picture in which Oct4- Sox2 heterodimers form on a rapid timescale, so that their concentration is proportional to the concentration of Oct4, Sox2 monomers. Therefore, for example as mesendodermal induction signals destroy Sox2,

they also destroy any complex containing Sox2 by mass action. The Oct4 monomers and homodimers still survive, and this is captured in our model (Figure 7C) (except that we do not distinguish between Oct4 monomers and homodimers). Given the lack of detailed information, we selected a Hill function that is symmetric for the three factors.

We took:

$$f([Oct4],[Sox2],[Nanog]) = \frac{[Oct4]^m[Sox2]^m[Nanog]^m}{K_1 + [Oct4]^m[Sox2]^m[Nanog]^m}, \quad (8)$$

where $K_1$is the value of the threshold, and m is the value of the Hill coefficient. We explore modulation of K_1 in detail below. In simulations, we found that a Hill coefficient (m) of ≥ 2 was required to achieve stable fixed points of pluripotency circuit activation. For $m > 2$, we found similar circuit dynamics, so we selected $m = 2$.

We selected

$$\tau_o(I_r,[Nanog]) = \frac{\alpha_0(1 + b_r[Nanog])}{(1 + c_r I_r)}, \quad (9)$$

which captures the effect that retinoic acid (I_r) reduces the lifetime of Oct4, τ_o, and that this effect is blocked by Nanog. The parameter c_r controls the strength with which retinoic acid decreases Oct4 lifetime, and $b_r$controls the strength of Nanog inhibition of this effect. Here, α_0 is the Oct4 lifetime in the absence of both external signal and Nanog.

Similarly,

$$\tau_s(I_w, [Nanog]) = \frac{\alpha_s(1+ b_w[Nanog])}{(1+ c_w I_w)}, \qquad (10)$$

Here, $\alpha_s$sets the basal Sox2 lifetime, and this lifetime decreases in the presence of Wnt, and the presence of Nanog inhibits this effect. The parameters b_w and c_w modulate the relative strengths of these effects, respectively. Our experiments point to a change in protein lifetime as the mechanism of signal induced down-regulation. Live cell imaging and FACS experiments showed that down-regulation of Oct4 by RA and Sox2 by Wnt occur on a rapid ~6 hour time scale.

We used numerical integration to study the dynamics of this system in response to signal. To perform numerical integration, we used the function NDSolve in Mathematica and implemented a Runge-Kutta algorithm. Each simulation allowed the simulation to reach a steady state before retrieving the output. The values of the parameters defined above that were used to produce the plots in Figure 7b,c,d,e,f are:

K_1=10 , K_r = K_w= 1, τ_n = α_o= α_s = 10 , c_r = c_w = 1, b_r = b_w = 1, and $m = 2$. In the absence of any differentiation signal (Figure 7c), I_w = I_r= 0. For ME differentiation (Figure 7d), I_w = 2 and I_r = 0. For NE differentiation (Figure 7e), I_w = 0 and I_r = 2. For simultaneous signal addition (Figure 7f), I_w = I_r= 2. For the 2D plot of Nanog activation over a space of inputs shown in Figure 7, parameters are as in c,d,e, but 10,000 points in (I_w, I_r) space are sampled on a lattice.

Pluripotency Circuit Simulation Results

As in our experiments, we found that in the absence of differentiation signals (I_w = I_r = 0), all three factors are driven to low levels as shown in Figure 7. Under these conditions, there is a stable fixed point for Oct4, Sox2, and Nanog expression at the origin. In fact, the system of differential equations (equations 5-7) has a fixed point at [Nanog] = [Oct4] = [Sox2] = 0 when I_w = I_r = 0 for our choice of f, τ_s, τ_o.

Further, simulations show that each differentiation signal drives differentiation. For example, in the presence of Wnt, when I_w = 2 and I_r = 0, the system is driven toward a stable fixed point in which Oct4 is present while Sox2 remains near 0. The vector fields in Figure 7 demonstrate the stability of this fixed point, and the structure of the flow shows that all initial conditions flow towards this point. At this fixed point, Sox2 concentration is near 0, so that the concentration of any complexes containing Sox2 will also be 0. In the presence of retinoic acid alone, I_r = 2 and I_w = 0, Sox2 is driven up and Oct4 remains near 0. This point is similarly a stable and attractive fixed point as shown in Figure 7. At this fixed point, the concentration of Oct4 is near 0, so that the concentration of any Oct4 containing complex (eg an Oct4-Sox2 heterodimer) will also go to 0 at this point.

In the presence of both signals, I_w = I_r= 2, we found that the positive feedback between the pluripotency factors and negative influence of signals compete. These interactions lead to a fixed point at relatively high Oct4, Sox2, and Nanog levels even in the presence of differentiation signals. Nanog reactivation is the result of a competition between circuit elements. Under this condition, the terms $\frac{I_w}{K_w + I_w}$ and $\frac{I_r}{K_r + I_r}$ drive production of Oct4 and Sox2, respectively. At the same time, signals decrease the

lifetime of both Oct4 and Sox2 through $\tau_s(I_w, [Nanog])$, $\tau_o(I_r, [Nanog])$. If Oct4 and Sox2 remain high enough to activate Nanog expression, Nanog then inhibits the influence of differentiation through the term (1+ [Nanog]) in the lifetime of both Oct4 and Sox2. In this way, conflicting signals can lead to a jamming of the circuit, so that differentiation is prevented by a reactivation of Nanog expression.

To characterize the reentrance or jamming in more detail, we explored the steady state levels of Oct4, Sox2, and Nanog over a wide range of I_w and I_r inputs. We found that the region of reentrance occupies a finite area of the (I_w, I_r) parameter space, and is a stable and reproducible behavior of this transcriptional circuit. At high absolute levels of signal, the repressive action of signal overwhelms Nanog, and the system returns to differentiation (Figure 7). Thus, the size of the reentrant region is finite. Figure 7 explores the reentrant regime experimentally.

Size of Reentrant Region Determined by Key Parameters

Intuitively, reentrance arises due to a competition between the inherent positive interactions between pluripotency circuit factors and the cross repressive effects of signals. Differentiation signals drive Oct4 and Sox2 independently, and this drive can resurrect Nanog and thus block the effects of signals. The existence and size of the reentrant region depend on three classes of parameters: K_1, the threshold at which positive feedback between factors becomes active; b_r, b_w, the strength of Nanog repression of differentiation signals; c_r, c_w, the strength at which signals repress Oct4 or Sox2. In the experimental system, we do not know the relative values of these parameters. Therefore, we explored the effects of modulating each parameter

individually. By tuning the magnitudes of these parameters, we adjusted the size of the reentrant region.

First, K_1 determines the threshold at which Nanog becomes activated by Oct4, Sox2, and Nanog. In this way, this parameter determines the threshold at which the positive interactions between Oct4, Sox2, and Nanog become self-sustaining. When this threshold increases, more Oct4 and Sox2 are required to re-activate Nanog. Thus, by changing this threshold, we can alter the size of the reentrant regime. At high threshold, K_1 = 40, (Figure S6C,Left Panel), the region does not exist because signals cannot drive Oct4 and Sox2 to high enough levels to activate Nanog. At a lower threshold, K_1 = 10, (Figure 7F), the region appears. As the threshold decreases further, K_1 = 2, the region occupies a larger fraction of signal space (Figure S6C, Right Panel). We quantify this effect by modulating K_1 and measuring the fraction of points in (I_1, I_2) space at which Nanog is active at steady state (Figure S6C, right panel). This plot shows that the size of the reentrant region decreases continuously as K_1 increases.

Second, Nanog represses differentiation by inhibiting signal-induced repression of Oct4 and Sox2. In our model, the strength of this effect is modulated by the parameters b_r, b_w. Increasing the strength of Nanog repression increases the size of the reentrant region by enabling Nanog to counteract higher doses of signal. For example, by increasing the strength of Nanog inhibition from b_r, b_w = 1 to b_r, b_w = 1.5, we expand the area of reentrance (Figure S6D, Left Panel) compared to that shown in Figure 7. Further, we can tune this effect continuously, so that increasing b_r and b_w continuously expands the size of the reentrant region. The plot in Figure S6D (Right Panel) shows that the size of the region changes continuously with b_r, b_w. As the strength of Nanog

repression goes to 0, we find that reentrance occurs in a small region of parameter space due to the effects of signal alone.

Finally, Wnt and retinoic acid individually repress Sox2 and Oct4, respectively. As the strength of this repressive effect increases, signals, both individually and in combination, drive Oct4 and Sox2 to lower levels. Thus, as we increase the strength of signal-induced down-regulation by adjusting c_r and c_w, we decrease the size of the reentrant region. If we set c_r, c_w = 1.3, we shrink the size of the reentrant region (Figure S6D, Middle Panel) as compared to that in Figure 7 where c_r = c_w= 1. In Figure S6D (Right Panel), we show that continuous modulation of c_r, c_w shrinks the size of the reentrant region to 0. In this way, the balance between the biochemical parameters controls the size and presence of the reentrant regime.

In this way, the size and existence of the region of reentrant Nanog activation is determined by the competition between the positive feedback between pluripotency circuit factors and the asymmetric influence of differentiation signals. In our model, three key parameter classes, K_1, b_i, c_i, determine the relative magnitude of these effects. The existence and size of the reentrant region depend on the magnitudes of these parameters. In the actual biological system, we do not know the value of these parameters. Therefore, the reentrant region could be present or absent. Thus, we performed experiments in which we titrated the levels of CHIR and retinoic acid and studied the fate choice of cells. These experiments identified concentrations of CHIR and retinoic acid that together did not activate Brachyury or Sox1 on the time scale of our experiments.

Further, we found that cells in this cell population contained Oct4, Sox2, and Nanog. These experiments suggest that signals can drive Oct4 and Sox2 to high enough levels to reactivate the pluripotency circuit, so that the biological system exists in a region of parameter space in which the Nanog reactivation is possible.

Alternate Pluripotency Circuit Models

An equivalent model based on transcriptional repression predicts similar jamming effects as those explored above and in Figure 7. Jamming results from the circuit topology and does not depend on the detailed form of the differential equations.

Supplemental References

Ai, H., Lvov, Y.M., Mills, D.K., Jennings, M., Alexander, J.S., and Jones, S.A. (2003). Coating and selective deposition of nanofilm on silicone rubber for cell adhesion and growth. Cell Biochem Biophys *38*, 103-114.

Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005). Transcriptional regulation by the numbers: models. Curr Opin Genet Dev *15*, 116-124.

Carey, M., Lin, Y.S., Green, M.R., and Ptashne, M. (1990). A mechanism for synergistic activation of a mammalian gene by GAL4 derivatives. Nature *345*, 361-364.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., *et al.* (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. In Cell, pp. 1106-1117.

Chew, J.L., Loh, Y.H., Zhang, W., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P., *et al.* (2005). Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol Cell Biol *25*, 6031-6046.

Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. Nature *403*, 335-338.

Francois, P., Hakim, V., and Siggia, E.D. (2007). Deriving structure from evolution: metazoan segmentation. Mol Syst Biol *3*, 154.

Gaspard, N., Bouschet, T., Hourez, R., Dimidschstein, J., Naeije, G., van den Ameele, J., Espuny-Camacho, I., Herpoel, A., Passante, L., Schiffmann, S.N., *et al.* (2008). An intrinsic mechanism of corticogenesis from embryonic stem cells. Nature *455*, 351-357.
Greber, B., Wu, G., Bernemann, C., Joo, J.Y., Han, D.W., Ko, K., Tapia, N., Sabour, D., Sterneckert, J., Tesar, P., *et al.* (2010). Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. Cell Stem Cell *6*, 215-226.
Hu, G., Kim, J., Xu, Q., Leng, Y., Orkin, S.H., and Elledge, S.J. (2009). A genome-wide RNAi screen identifies a new transcriptional module required for self-renewal. Genes Dev *23*, 837-848.
Jackson, S.A., Schiesser, J., Stanley, E.G., and Elefanty, A.G. (2010). Differentiating embryonic stem cells pass through 'temporal windows' that mark responsiveness to exogenous and paracrine mesendoderm inducing signals. PLoS One *5*, e10706.
Lengner, C.J., Camargo, F.D., Hochedlinger, K., Welstead, G.G., Zaidi, S., Gokhale, S., Scholer, H.R., Tomilin, A., and Jaenisch, R. (2007). Oct4 expression is not required for mouse somatic stem cell self-renewal. Cell Stem Cell *1*, 403-415.
Marson, A., Levine, S.S., Cole, M.F., Frampton, G.M., Brambrink, T., Johnstone, S., Guenther, M.G., Johnston, W.K., Wernig, M., Newman, J., *et al.* (2008). Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell *134*, 521-533.
Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., *et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.
Minucci, S., Botquin, V., Yeom, Y.I., Dey, A., Sylvester, I., Zand, D.J., Ohbo, K., Ozato, K., and Scholer, H.R. (1996). Retinoic acid-mediated down-regulation of Oct3/4 coincides with the loss of promoter occupancy in vivo. EMBO J *15*, 888-899.
Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.
Pasini, D., Cloos, P.A., Walfridsson, J., Olsson, L., Bukowski, J.P., Johansen, J.V., Bak, M., Tommerup, N., Rappsilber, J., and Helin, K. (2010). JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature *464*, 306-310.
Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H., and Robson, P. (2005). Transcriptional regulation of nanog by OCT4 and SOX2. J Biol Chem *280*, 24731-24737.
Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell *32*, 491-502.
Sterneckert, J., Stehling, M., Bernemann, C., Arauzo-Bravo, M.J., Greber, B., Gentile, L., Ortmeier, C., Sinn, M., Wu, G., Ruau, D., *et al.* (2010). Neural Induction Intermediates Exhibit Distinct Roles of Fgf Signaling. Stem Cells.
ten Berge, D., Koole, W., Fuerer, C., Fish, M., Eroglu, E., and Nusse, R. (2008). Wnt signaling mediates self-organization and axis formation in embryoid bodies. Cell Stem Cell *3*, 508-518.
Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. Nature *448*, 196-199.
Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. Nature *444*, 364-368.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., Bernstein, B.E., and Jaenisch, R. (2007). In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. Nature *448*, 318-324.
Ying, Q.L., and Smith, A.G. (2003). Defined conditions for neural commitment and differentiation. Methods Enzymol *365*, 327-341.
Ying, Q.L., Stavridis, M., Griffiths, D., Li, M., and Smith, A. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. In Nat Biotechnol, pp. 183-186.
Ying, Q.L., Wray, J., Nichols, J., Batlle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. Nature *453*, 519-523.