SUPPLEMENTAL DATA

Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal

Kathy K. Niakan, Hongkai Ji, Rene Maehr, Steven A. Vokes, Kit T. Rodolfa, Richard I. Sherwood, Mariko Yamaki, John T. Dimos, Alice E. Chen, Douglas A. Melton, Andrew P. McMahon, Kevin Eggan

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Preimplantation embryo culture
Six to eight week-old B6D2F1 female mice were super-ovulated using injection of 5 IU of pregnant mare serum gonadotrophin (PMS; Sigma). 46 hours post PMS, 5 IU of human chorionic gonadotrophin (hCG, Sigma) was administered. Superovulated females were set up for mating with eight week-old or older B6D2F1 males. (Mice were obtained from Charles River Laboratories). Embryos were flushed from the uteri with M2 media (Chemicon) or 18-20 hours after hCG, zygotes were pulled from the oviducts of the female mice and placed into incubation drops of KSOM (Chemicon) at 37°C in an atmosphere of 5% CO₂. Embryos were harvested at 23, 43, 55, 66, 80 and 102 hours post-hCG respectively for 1-cell, 2-cell, 4-cell, 8-cell, morula and blastocyst stages. All embryos were examined carefully and only those with excellent morphology were collected and used for the study.

Immunohistochemistry
Embryos, cells, or embryoid bodies were fixed with 4% paraformaldehyde in PBS overnight at 4°C. Samples were washed with 0.1% Triton X100 in PBS and permeabilized in PBS-0.5% Triton for 15-30 minutes. Blocking was performed in blocking solution consisting of 10% fetal bovine serum in PBS-0.1% Triton for 1 hour or overnight at 4°C. Primary antibodies were diluted in blocking solution and samples were incubated at 4°C overnight. Primary antibodies used are goat anti-Sox17 (R&D); murine anti-Oct4 (Santa Cruz Biotech); rabbit anti-Nanog (Kamiya); rabbit anti-Laminin (Sigma Aldrich); rabbit anti-Collagen IV (Abcam); rabbit anti-Dab2 (Santa Cruz Biotech); murine anti-Cdx2 (Biogenex); rabbit anti-Gata4 (Santa Cruz Biotech); rabbit or goat anti Gata6 (Abcam; R&D; or Santa Cruz Biotech); rabbit anti-Hex (Chemicon); goat anti-Pem (Santa Cruz Biotech); rabbit anti-Sparc (Santa Cruz Biotech); goat anti-Sox2 (Santa Cruz Biotech); rabbit anti-dsRed (MBL).

Secondary antibody incubation was performed at 1:300 concentration using fluorophore-conjugated antibodies from Jackson Immuno Research and Molecular Probes. Embryos were placed on coverslip dishes (MatTek) for confocal imaging in Vectashield media with DAPI (Vector laboratories) diluted 1:3 in PBS-0.1% Triton.

Imaging
Confocal immunofluorescence pictures were taken with a Zeiss LSM 510 META confocal...
microscope. Epifluorescence images were performed on an Olympus IX70 using an Olympus DP70 camera.

Genotyping
For XEN and mES cells, genotyping was performed following MEF depletion. DNA was harvested from cells following a previous published protocol (Laird et al. 1991). Embryos were placed in Direct PCR solution (Viagen) with proteinase K and placed at 50°C for 12 hours, followed by a 5 minute incubation at 95°C. 2µl of this solution was used in a 25µl PCR reaction. All PCR reactions were performed using two sets of primers, Sox17 and either the PLFe (flappase) or the dTomato primer pairs. PCR reaction was performed using Takara Taq polymerase (Takara) with annealing set at 57°C and amplification at 72°C for 1.5 minutes. Primers used are as follows:

FLP forward primer (5' - CACTGATATTGTAAGTAGTTTGC - 3');
FLP reverse primer (5' - CTAGTGGAGTGGATGATCAGG - 3');
Sox17 forward primer (5' - CTCGGGGATGTAAGGGTGAA - 3');
Sox17 reverse primer (5' - GGAGACATCGCAGGAGGA - 3');
dTomato forward primer (5' - GCGAGGAGGTCATCAAAGAG - 3');
dTomato reverse primer (5' - CCCATGGTCCTCTTCTGCAT - 3').

Embryonic stem cell derivation and culture
Sox17+/− and Sox17−/− embryonic stem cells were derived using a modification of a protocol previously described (Nagy 2003) following the harvesting of blastocysts from intercrossed Sox17::Tomato heterozygous mice. Blastocysts were treated briefly with acidic tyrode’s solution (Chemicon) solution to dissolve the zona pellucida. ES media consists of: 85% KO-DMEM; 15% fetal bovine serum (Hyclone); 1% L-glutamine (Invitrogen); 1% Non-essential amino acids (Invitrogen); 1% β-mercaptoethanol (Invitrogen); 1% Penicillin/Streptomycin (Invitrogen) 10³IU of LIF (Chemicon). Derivation media consisted of ES media supplemented with another 10³IU of LIF and 50µM of MEK inhibitor (PD98059, Cell Signaling). Blastocysts were plated onto gamma-irradiated murine embryonic feeder layers in mES derivation media. The ICM of the outgrowth was picked after 7-14 days, washed briefly in 1X PBS and placed briefly into 0.25% trypsin to dissociate the outgrowth. The ICM outgrowth was then expanded in ES media. Concentrations for selection are as follows: 1µg/mL of puromycine and 250µg/mL of neomycin. Cells were grown for a minimum of one week in selection media.

Embryoid body formation
Embryoid bodies (EBs) were generated from embryonic stem cells after MEF depletion as previously described (Coucouvanis and Martin 1995). EBs were generated by dissociating ES cells using trypsin. 1 x 10⁶ total ES cells were placed in 5mL of EB media (EB media: ES media described above, without LIF) and grown for 3-6 days in one well of ultra low cluster 6-well plates (Corning). After which, EBs were collected and used for RNA isolation or for fixation and embedding in OCT for sectioning and staining.
Teratoma formation and histological analysis

Sox17-inducible cells, Sox17-mutant, XEN cells, and wildtype control cells were used in teratoma assays. Approximate 1 x 10⁶ cells were injected into the kidney capsules of NOD-SCID mice. Cells were dissociated and washed, then resuspended with a minimum volume of PBS with 1% FBS. Sox17-inducible cells were treated with doxycycline for 7 days prior to the injection and the NOD-SCID mice injected with these cells and were continuously administered doxycycline. Three weeks after the injection, the tumors were dissected from the mice, fixed in 4% formaldehyde, and embedded in paraffin. Teratomas were then sectioned and stained with hematoxylin and eosin.

RNA isolation and cDNA synthesis

RNA from cells was isolated as previously described in Trizol reagent (Invitrogen) (Chomczynski and Sacchi 1987). Prior to cDNA synthesis, RNA samples were treated with DNase (Ambion) and analyzed on an Agilent 2100 Bioanalyzer using a RNA Nano protocol (Agilent). Only good quality RNA was used in cDNA synthesis using an iScript Select cDNA synthesis kit (Biorad). Approximately 5-10µg of RNA was reverse transcribed using oligo(dT)₂₀ primers.

Quantitative Real-Time PCR

PCR primer pairs were either designed using Primer3 software online (www.frodo.mit.edu) from MIT or from those previously published (Molkentin et al. 1997; Fujikura et al. 2002; Niwa et al. 2005). Gene specific primers are listed below (Supplementary Table 1, 2). Primers were designed to have an optimal annealing temperature of 56°C and to amplify a 200-250 basepair product. Primers to detect transcripts in ES, XEN, TS, EBs (Supplementary Table 1) were designed to span an intronic region of the gene or at the 3’ untranslated region. Primers for ChIP validation (Supplementary Table 2) of Sox17 binding were designed to amplify the region where Sox17 was identified to bind based on the microarray analysis.

Reactions were set up in a 25µL volume containing 12.5µL of SYBR Green PCR Master Mix (Biorad), 200nM final concentration of the forward and reverse primers and 5ng of cDNA. Thermal cycling was initiated with 95°C 5min, followed by 45 cycles at 95°C 10 seconds, 56°C 30 seconds, 72°C 30 seconds; then, 72°C for 10 minutes; and lastly, 81 cycles at 55°C for 30 seconds.

Lentiviral production

The dTomato coding sequence was subcloned into the lentiviral vector FUGW (Addgene plasmid 14883), replacing the GFP sequence. Lentiviral gag-pol (pHR'CMV 8.2 deltaR; Addgene plasmid 8455) and VSV-g envelope (pCMV-VSV-g; Addgene plasmid 8454) were obtained from Addgene. These plasmids were transiently co-transfected into 293FT packaging cells (ATCC) at a 1:2:1 mass ratio (transgene:gag-pol:VSV-g) using SuperFect (Qiagen). Viral supernatant fractions were harvested after 60 hours, filtered through a 0.45 µm low protein binding cellulose acetate filter, and concentrated by centrifugation.
**Genome-wide transcriptional array in Sox17-induced ES cells**

Sox17-inducible cells were either treated with doxycycline for 5 days or left untreated. KH2 control cells treated with doxycycline (cells lacking a cDNA insert in the Col1A1 locus) were also prepared. RNA was isolated using the Trizol protocol (Invitrogen). This RNA was amplified and cDNA biotinylated cRNA was prepared and hybridized onto Illumina MurineRef-8 c2 microarrays (Illumina), as previously described (Sherwood et al. 2009). The microarrays were scanned using Beadstation and the data analyzed using BeadStudio (Illumina). The raw data for the array may be obtained from GEO (GSE19028).

**De novo motif analysis**

An upgraded version of TileMap (Ji and Wong 2005; Ji et al. 2008) was used to analyze the Sox17 ChIP-chip data. 2206 peaks with a moving average t-statistic $\geq 3$ and FDR $\leq 10\%$ were reported. The top 300 peaks were used for de novo motif discovery. Gibbs motif sampler incorporated in CisGenome (Ji et al. 2008) was run three times independently, and 10 motifs were searched simultaneously within each run. After de novo motif discovery, the occurrence rates of the reported motifs in all 2206 peaks were compared to their occurrence rates in negative control regions which are carefully chosen to match the physical distribution of ChIP-chip peaks (Ji et al. 2008). Three motifs with a relative enrichment ratio $> 2.5$ were reported. The final motif site lists were obtained by mapping the three motifs to 2206 peaks using CisGenome, using a likelihood ratio $= 500$ as the cutoff.

**Gene ontology and classification of Sox17-bound genes**

Gene ontology analysis was performed by applying the WEB-based GEne ScT AnaLysis Toolkit (Webgestalt) (http://bioinfo.vanderbilt.edu/webgestalt) (Zhang et al. 2005). Gene identifications were used to generate a figure of statistically significant enrichment of biological systems.

**Analysis to determine Sox17, Sox2, Nanog, and Oct4 binding overlaps**

The Sox17 peaks identified in the Sox17-induced murine ES and the XEN cell samples were compared to the peak data for Sox2, Nanog, and Oct4 binding previously published (Chen et al. 2008; Kim et al. 2008). For example, 61 out of 2206 Sox17 peaks have $\geq 1$bp overlap with peaks in the ChenOct4 data. To determine the significance for this enrichment, 2206 random regions (which did not contain Sox17 peaks) were picked in regions of the genome that were covered by the promoter array. The overlap between these 2206 random regions (not enriched for Sox17-binding), and the 3761 ChenOct4 peaks were compared. This procedure of picking random regions and comparing to the peaks for each transcription factor was repeated ten times for each factor, then averaged to give a value for the number of overlapping peaks that would be expected by chance. The fold enrichment is the number of peaks identified over the number of peaks expected by chance, which indicates that the observed number of overlapping peaks we observed were much more than the number of overlapping expected by chance (note: the fold change here has nothing to do with IP/Input probe intensity ratio). p-value was determined by perform a t-test with d.f. = 9.
SUPPLEMENTAL TABLES and FIGURES

**Table S1.** Gene-specific primers used for quantitative PCR analysis.

| Gene | 5' Primer | 3' Primer |
|------|-----------|-----------|
| Cdx2 | CCTGCGACAAGGGCTTTGTTTAG | TCCCCGACTCCCTCACCATAC |
| Dab2 | GGCAACAGGCTGACCATATTAG | TTGGTGTCGATTTACAGGTAGAT |
| Sox17 | AAAGAAACCACTCAACAAAGACG | TTTGTGGAAGTGGGATCAAGAC |
| Gapdh | TCCCACTCTTCACCTTCCATG | GGGTCTGGGATGAAATAGG |
| Coupft1 | ACCCAATCCTGCTCCCCTGACTC | CTCATGGTCCTTCCAAATAG |
| Nanog | AGGCTTGGAGACAGTGAGGTGC | TACCCTGCAATCTGGATTC |
| Oct4 | GCAGGACAGAAGCTTCAAGACAGA | CAAAGGCCTGGAAGGACAGATG |
| Gata6 | ATGGCGTAGAATGCTGAGG | TGAAGTGGTCGCTGTTGAG |
| Gata4 | AAACGGAAGCCAAGAAGATCTAAT | GAGCTGCGCTGAGTTGAGTG |
| Sox17 | GTAAAGGTGAAAGGCGAGGTG | GTCAACGCCTTCCAAGACTTG |
| Sox7 | GCCCATGATGCCTCCATTAGT | CGCTGCTGCAATCAAGTCAAT |
| Lama1 | AGGTTCCTCGGCTGATTGTTCTG | CAGTACTATGCCTCGAGGAT |
| Sox7 | CTTACAGGAGCAAGAGCTG | GCTGCTGCTGCTTCTG |
| Lamb1-1 | GGGAGGAGAATATATACTGAA | CAGGTTTGGTCTAGGTTG |
| Lmca1 | ATGAGGCTGAAATGCTGCTG | CGATCTCATTGAGCCTG |
| Col4a1 | GTCTCTGTCGAGAAGAGTTT | AAATACCATGGGAGGAGA |
| Col4a2 | AGGAAATCAAGCCCTACATC | ACGGAAATGCTGATACAG |
| Hnf3b | GGCCTAGTCAGAACAAGAG | CCGTCTAGCTCATCAAGCATC |
| Sparc | CCTGCGAAGACACTATTCAAT | ATGAGCTCCATGAGGAGCTG |
| Hnf4 | CATACCTGCTAAAGATGAGC | GCAACCTCCTTGGCCGAATG |
| Afp | CGAGAGTGGTGCTCCAAGGAAA | CAGAAGCTAGTTGAGCAT |
| Amn | CTACTGCTCAAGGAAATGCA | CTGCTCTTACAGAGGTCTAG |
| Pem | CAAGGAGACTCGGAAGAACA | GAACAAATGGCCTCTTGC |

**Table S2.** Site specific primers used for quantitative PCR analysis of Sox17-enriched binding sites identified by ChIP-chip.

| Gene | 5' primer | 3' primer | Fold Enrichment | ST Error |
|------|-----------|-----------|----------------|----------|
| Ap2m1 | CCAGGCTGCTGAGTAGT | CAGGGAGAGGCAAGGATT | 12.0 | 4.2 |
| Col4a1 | TGCTCGCTGTGATTTGCTG | CAGCAGCATCTCTGTTAGG | 64.4 | 44.5 |
| Ilfnk | CCAAGATCGCTCCTGAGTCA | TGGTGTAATGGCCATAAC | 10.9 | 3.5 |
| Lama1 | TGAGATGATGATAGAGCTG | ATTTCCAGCTATCGAGG | 7.9 | 3.3 |
| Gap4a | TCCACTTCTGCCTCTTTG | CGAAGAATCAGGACTTAT | 7.6 | 1.9 |
| Col4a2 | GTCTCTGCTGAGGCTGAGT | CAGGGCTCCAGATACAAA | 61.0 | 27.1 |
| Lamb1-1 | GTCTGCTGAGGCTGAGG | AGATGCTACCTAAGATG | 7.9 | 3.2 |
| Sall4 | AAAGAAGGCTTGCTGAGTCA | TCTGAGGAGGTCTAGAA | 7.6 | 1.9 |
| S1c38a4 | TCCGGCTGAGGCTGAGT | CAGCTTAAAGCAGAGATG | 11.7 | 3.6 |
| Vitrin | TCAGGGAAATACAGGCCATG | GATGACATGGGATAGGAAT | 15.0 | 6.2 |
Supplementary Tables 3-10 are located in accompanying Excel files.

**Table S3.** Sox17-enriched genomic regions and the associated genes identified in extraembryonic endoderm stem cells. These XEN cells were chromatin immunoprecipitated with Sox17 antibody followed by genome-wide promoter tiling arrays.

**Table S4.** List of genes significantly overexpressed following 5 days of Sox17 induction as compared to control murine ES cells treated with doxycycline. Sox17-inducible murine embryonic stem cells treated with doxycycline for 5 days and analyzed using genome-wide transcriptional arrays (probes designed against the 3’ UTR).

**Table S5.** Sox17-enriched genomic regions and the associated genes in Sox17-inducible murine embryonic stem cells treated with doxycycline for 48hrs. These mES cells were chromatin immunoprecipitated with Sox17 antibody followed by genome-wide promoter tiling arrays.

**Table S6.** Sox17-enriched genomic regions and the associated genes in Sox17-inducible murine embryonic stem cells treated with doxycycline for 48hrs. These mES cells were chromatin immunoprecipitated with FLAG antibody followed by genome-spanning promoter tiling arrays.

**Table S7.** List of genes significantly underexpressed following 5 days of Sox17 induction as compared to murine ES cells treated with doxycycline. Sox17-inducible murine embryonic stem cells treated with doxycycline for 5 days and analyzed using genome-wide transcriptional arrays (probes designed against the 3’ UTR).

**Table S8.** Sox17-enriched genes in murine embryonic stem cells were compared to the binding targets of Sox2, Nanog, and Oct4 identified by previous groups in murine embryonic stem cells (Chen et al. 2008; Kim et al. 2008). (1=overlapping binding identified; 0=no overlapping binding identified).

**Table S9.** Sox17-enriched genes in XEN cells were compared to the binding targets of Sox2, Nanog, and Oct4 identified by previous groups in murine embryonic stem cells (Chen et al. 2008; Kim et al. 2008). (1=overlapping binding identified; 0=no overlapping binding identified).
Table S10. Probability of overlapping enrichment at the same genomic region was calculated. The significance of the overlaps was determined by comparing the fold-change at a number of specific Sox17-enriched regions overlapping with Oct4, Sox2 or Nanog to the random expectation of any given Sox17-enriched region overlapping with Oct4, Sox2 or Nanog.

Table S11. De novo motif enrichment identified in Sox17-enriched regions.

| Motif | Logo | Enrichment to matched control |
|-------|------|-----------------------------|
| Sox17 | ![Logo](image) | 4.10                         |
| Motif 2 | ![Logo](image) | 3.89                         |
| Motif 3 | ![Logo](image) | 2.63                         |

Supplemental References
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. *Cell* **133**(6): 1106-1117.
Chomczynski, P. and Sacchi, N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem* **162**(1): 156-159.
Coucouvanis, E. and Martin, G.R. 1995. Signals for death and survival: a two-step mechanism for cavitation in the vertebrate embryo. *Cell* **83**(2): 279-287.
Fujikura, J., Yamato, E., Yonemura, S., Hosoda, K., Masui, S., Nakao, K., Miyazaki Ji, J., and Niwa, H. 2002. Differentiation of embryonic stem cells is induced by GATA factors. *Genes Dev* **16**(7): 784-789.
Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. 2008. An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat Biotechnol*.
Ji, H. and Wong, W.H. 2005. TileMap: create chromosomal map of tiling array hybridizations. *Bioinformatics* **21**(18): 3629-3636.
Kim, J., Chu, J., Shen, X., Wang, J., and Orkin, S.H. 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132**(6): 1049-1061.
Laird, P.W., Zijderveld, A., Linders, K., Rudnicki, M.A., Jaenisch, R., and Berns, A. 1991. Simplified mammalian DNA isolation procedure. *Nucleic Acids Res* **19**(15): 4293.
Molkentin, J.D., Lin, Q., Duncan, S.A., and Olson, E.N. 1997. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev* **11**(8): 1061-1072.
Nagy, A. 2003. *Manipulating the mouse embryo : a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. 2005. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell* **123**(5): 917-929.

Sherwood, R.I., Chen, T.Y., and Melton, D.A. 2009. Transcriptional dynamics of endodermal organ formation. *Dev Dyn* **238**(1): 29-42.

Zhang, B., Kirov, S., and Snoddy, J. 2005. WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* **33**(Web Server issue): W741-748.
Figure S1. Tomato expression is localized to the primitive endoderm of Sox17::Tomato blastocysts.

Tomato expression was detected within cells of the inner cell mass (ICM) of Sox17::Tomato embryos at the blastocyst stage (either Sox17+/− or Sox17−/−). Initially tomato expression was observed in cells that were adjacent to the blastocoel cavity and in cells deeper within the ICM. Tomato expression was restricted to the primitive endoderm, adjacent to the blastocoel cavity in late-blastocysts that were either hatching or hatched from the zona pelucida.
Figure S2. Sox17::Tomato mutant embryos were reduced in size and failed to properly gastrulate.

Day 8 embryos were immunofluorescently labeled with antibodies to detect either Tomato (red) or Sox17 (green) proteins. Sox17::Tomato deficient embryos failed to undergo proper gastrulation and were reduced in size compared to Sox17+/+ and Sox17+/- litter-mates. Sox17-antibody staining was not detected in the Sox17::Tomato deficient embryos.
Figure S3. Microarray analysis at defined stages of preimplantation development.

(A) *Oct4* maternal transcript was down-regulated from the zygote to the 4-cell stage. Following zygotic genome activation *Oct4* becomes up-regulated and expressed at the 8-cell, morula and blastocysts stages.

(B) *Gata6* transcript was initially detected at the 8-cell stage and remained expressed in the morula and blastocyst stages.
Figure S4. Sox17-mutant embryos maintain Nanog localization within epiblast progenitors and continue to express Laminin in the primitive endoderm.

A) Sox17 initial expression co-localizes with that of Nanog. By the >128-cell stage Nanog and Sox17 no longer overlap in their expression within the inner cell mass.

B) The expression of Nanog and Sox17 in wildtype, Sox17-heterozygous and Sox17-mutant embryos. Nanog expression and localization in the mutant embryos resembles that of wildtype embryos.

C) The total number of cells within the inner cell mass and the number of Nanog-expressing cells in Sox17-mutant embryos is not significantly different from that of wildtype and heterozygous embryos.

D) Sox17-mutant embryos continue to express Laminin in presumptive primitive endoderm cells, indicated by white arrows.
Figure S5. Quantitative RT-PCR in ES, XEN, and trophoblast stem (TS) cell lines was normalized to levels of β-Actin (average of three independent experiments).

(A) Primers specific to Nanog and Oct4 in ES, XEN, and TS cells.
(B) Primers specific to Sox17 and Gata4 in ES, XEN, and TS cells.
(C) Primers specific to Cdx2 in ES, XEN, and TS cells.
(D-F) Phase-contrast images of mouse stem cell lines.
(D) Embryonic stem cells grown on mouse embryonic fibroblasts.
(E) Extraembryonic endoderm (XEN) stem cells.
(F) Trophoblast stem (TS) cells.
Figure S6. Sox17::Tomato deficient embryoid bodies (EBs) fail to sort Tomato expression to the periphery.

(A) Epifluorescent imaging of Sox17+/− heterozygous EBs cultured for 6 days revealed Tomato expression was localized to the periphery. Sox17−/− deficient EBs reveal failures in cell sorting. Instead, Tomato expression in mutant Sox17−/− EBs, was disorganized and mis-localized.

(B) Quantitation of Tomato expression indicated that an equivalent number of Sox17+/− and Sox17−/− EBs have Tomato expression, but that the localization to the periphery was compromised in Sox17−/− deficient EBs.
Figure S7. Gene ontology analysis of Sox17-bound genes in extraembryonic endoderm (XEN) cells.

Ontology categories are listed on the y axis and the p values for enrichment are represented on the x axis. Red bars highlight the gene ontology categories involving adhesion functions, which are enriched for Sox17-binding.
Figure S8. Genes upregulated in Sox17-induced ES cells.

(A) The treatment of Sox17-inducible cells with doxycycline for 48 hours resulted in a 220-fold induction of Sox17 compared to controls. The controls used to calculate the fold-induction were Sox17-inducible cells not treated with doxycycline and KH2 cells treated with doxycycline.

(B) The induction of Sox17 resulted in an over 2-fold upregulation of Col4a1 and 2, Lama1, Gata4, and Gata6.

(C) The induction of Sox17 did not result in a significant change in transcript levels of Oct4, Nanog, or Sox2.

(D) Gene ontology analysis of the genes upregulated over 1.9-fold after 5-days of Sox17-induction. The ontology categories are listed on the y-axis and the p values for enrichment are represented on the x-axis. Red bars highlight that genes functioning in pathways known to be involved in adhesion functions are upregulated in Sox17-induced ES cells.
Figure S9. Sox17 is enriched at the regulatory region of extracellular matrix genes in Sox17-induced mouse ES cells.

Sox17-inducible mouse ES cells were treated with doxycycline and a Sox17 antibody was used to perform chromatin immunoprecipitated followed by quantitative PCR assay. Sox17 was enriched for binding at the regulatory regions upstream of the Col4a1, Col4a2, Lama1, and Lamb1-1 genes. Ppil4 and Rosa26 were used as negative control loci as a comparison.
Figure S10. Embryoid bodies immunofluorescently analysed for proteins associated with embryonic stem cells or extraembryonic endoderm. Day 6 EBs were generated from Sox17, Gata6, and Gata4 deficient ES cells, Gata6 and Gata4 compound mutant ES cells and ES cells constitutively over-expressing Nanog.

(A-D) All EBs had Nanog, Oct4, and Sox2 expression with a center that lacked any protein expression. The lack of protein expression in the center of EBs is expected as apoptotic events there have been previously described (Couchman and Martin, 1995). (A) Sox17 is localized to the periphery of wildtype EBs, surrounding Nanog expressing cells. In Gata4-/- EBs, Sox17 is expressed, but mislocalized in the center, whereas no Sox17 is detected in the Sox17-/-, Nanog OE, Gata6-/- and Gata4-/-, Gata6-/- compound null EBs. (B) Gata4 and Gata6 are localized to the periphery of wildtype EBs. While Gata4 and Gata6 are expressed in Sox17-/- EBs, these cells were not detected in the periphery. Gata4-/- EBs have disorganized Gata6 expression. Nanog OE, Gata6-/- and Gata4-/-, Gata6-/- EBs fail to express Gata4 and Gata6. (C) Laminin and (D) Dab2 were both expressed on the periphery of wildtype EBs, while no expression was detected in Sox17-/-, Nanog OE, Gata4-/-, Gata6-/- and Gata4-/-, Gata6-/- compound null EBs.
Figure S11. Col4a1 and Col4a2 expression is down-regulated in Sox17-/- deficient embryoid bodies.

The loss of both alleles of Sox17 in day 6 EBs results in a significant loss of both Col4a1 and Col4a2. Similarly, Col4a1 and Col4a2 are down-regulated in the Nanog overexpressing, Gata4-/-, Gata6-/-, and Gata4-/-, Gata6-/- compound mutant EBs. Col4a1 and Col4a2 were significantly upregulated in the Sox17+/- EBs (N=4).
Figure S12. Gene ontology analysis of Sox17-bound genes in Sox17-induced mouse embryonic (mES) stem cells.

Ontology categories are listed on the y axis and the p values for enrichment are represented on the x axis. Red bars highlight that genes functioning in pathways known to be involved in ES cell maintenance and establishment are enriched for Sox17-binding.
Figure S13. Gene ontology analysis of genes suppressed in Sox17-induced ES cells compared to wildtype mES controls.

Ontology categories are listed on the y axis and the p values for enrichment are represented on the x axis. Red bars highlight that genes functioning in pathways known to be involved in ES cell maintenance and establishment are downregulated in Sox17-induced ES cells.
Figure S14. Sox17-induced and Sox17-deficient embryonic stem cell cultures give rise to teratomas comprised of all three germ cells.

Hematoxylin and eosin staining of teratomas derived from Sox17-induced and Sox17-deficient ES cells. Different cell types and tissues are marked with abbreviations of the following: Cartilage (C) (mesoderm), Muscle (Mu) (mesoderm), Nerve (N) (ectoderm), Gland (Gl) (endoderm), Bone (B) (mesoderm), Melanocyte (M) (ectoderm) and Goblet (G) endoderm.
Figure S15. Sox2 and Nanog binding overlaps with Sox17-bound regulatory regions.

(A) Both Sox2 and Nanog were enriched for binding that overlaps with Sox17-bound regulatory regions of Col4a1 and 2. (B) Sox2 was enriched for binding overlapping with the Ap2m1 regulatory region bound by Sox17.
Figure S15 continued. Sox2 and Nanog binding overlaps with Sox17-bound regulatory regions.

Both Sox2 and Nanog were enriched for binding that overlaps with the Sox17-bound regulatory regions of (C) Sall4 and (D) Vitrin.
Figure S16. Sox2 binding enrichment at Sox17-bound regulatory regions in Sox17-induced cells.

(A) In contrast to a model for Sox2 and Sox17 reciprocal binding, Sox2 is more highly enriched at Lama1 and Col4a1 regulatory regions in the Sox17-induced cells than in the uninduced cells.