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

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

Kathy K. Niakan, Hongkai Ji, René Maehr, et al.

Genes Dev. 2010 24: 312-326
Access the most recent version at doi:10.1101/gad.1833510

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

Kathy K. Niakan,[1,2,3,4,5,9] Hongkai Ji,[6,8] René Maehr,[2,3,4,5,8] Steven A. Vokes,[3,4,5,8,10] Kit T. Rodolfa,[1,2,3,5,7] Richard I. Sherwood,[3,4,5] Mariko Yamaki,[1,3,4,5] John T. Dimos,[1,3,4,5,11] Alice E. Chen,[2,3,4,5] Douglas A. Melton,[2,3,4,5] Andrew P. McMahon,[3,4,5] and Kevin Eggan[1,2,3,4,5,12]

[1]Stowers Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA; [2]Howard Hughes Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA; [3]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massacusetts 02138, USA; [4]Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massacusetts 02138, USA; [5]Harvard Stem Cell Institute, Harvard University, Cambridge, Massacusetts 02138, USA; [6]Department of Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland 21205, USA; [7]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massacusetts 02138, USA

In embryonic stem (ES) cells, a well-characterized transcriptional network promotes pluripotency and represses gene expression required for differentiation. In comparison, the transcriptional networks that promote differentiation of ES cells and the blastocyst inner cell mass are poorly understood. Here, we show that Sox17 is a transcriptional regulator of differentiation in these pluripotent cells. ES cells deficient in *Sox17* fail to differentiate into extraembryonic cell types and maintain expression of pluripotency-associated transcription factors, including *Oct4*, *Nanog*, and *Sox2*. In contrast, forced expression of Sox17 down-regulates ES cell-associated gene expression and directly activates genes functioning in differentiation toward an extraembryonic endoderm cell fate. We show these effects of Sox17 on ES cell gene expression are mediated at least in part through a competition between Sox17 and Nanog for common DNA-binding sites. By elaborating the function of Sox17, our results provide insight into how the transcriptional network promoting ES cell self-renewal is interrupted, allowing cellular differentiation.

[*Keywords*: Differentiation; self-renewal; transcriptional network; primitive endoderm; stem cell; Sox17]

Supplemental material is available at http://www.genesdev.org.

Received June 17, 2009; revised version accepted December 11, 2009.

A conserved feature of mammalian development is the formation of a blastocyst. Mouse and human blastocysts consist of a trophectodermal epithelium surrounding an inner cell mass (ICM), which contains both epiblast progenitor and primitive endoderm cells. Ultimately, the trophectoderm and primitive endoderm will give rise to the placenta and yolk sac. These extraembryonic tissues support the continued development of epiblast cells, which in turn form the fetus.

[8]These authors contributed equally to this work.
Present addresses: [9]Centre for Trophoblast Research, Anne McLaren Laboratory for Regenerative Medicine, University of Cambridge, Cambridge CB2 0SZ, United Kingdom; [10]Section of Molecular Cell and Developmental Biology, Institute for Cellular and Molecular Biology, The University of Texas at Austin, Austin, TX 78712, USA; [11]iPerian Bio, Inc., 951 Gateway Bldv., San Francisco, CA 94080, USA.
[12]Corresponding author.
E-MAIL Keggan@scrb.harvard.edu; FAX (617) 384-8234.
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.1833510.

The *Sox2*, *Nanog*, and *Oct4* transcription factors are expressed in the ICM and are required for proper formation of the epiblast (Nichols et al. 1998; Avilion et al. 2003; Chambers et al. 2003; Mitsui et al. 2003). It has been demonstrated recently that Sox2, Nanog, and Oct4 also form the core of a transcription factor network that promotes embryonic stem (ES) cell pluripotency and self-renewal while simultaneously repressing genes required for differentiation (Boyer et al. 2005). Although this network promoting pluripotency is well elaborated, the transcription factors that act to disrupt it and induce differentiation are poorly understood.

The earliest known event influencing differentiation of the ICM into its two derivatives (the epiblast and primitive endoderm) is Fgf signaling through the Grb2–Ras–MAP kinase pathway (Cheng et al. 1998). Activation of Grb2 leads to expression of the Gata6 and Gata4 transcription factors in ICM cells that will later form the primitive endoderm (Chazaud et al. 2006). It has been

proposed that Gata6 and Gata4 expression reinforce a transcriptional network that antagonizes *Nanog* expression (Yamanaka et al. 2006). These inhibitory interactions in turn seem to initiate a cell sorting process that compartmentalizes the ICM into distinct epiblast and primitive endoderm domains (Chazaud et al. 2006; Plusa et al. 2008). This cell sorting is likely facilitated by specific expression of cell surface proteins, including Dab2 and Laminin 1, within the primitive endoderm (Yamanaka et al. 2006).

It is also clear that *Gata6* and *Gata4* act in a genetic pathway controlling differentiation of ES cells into extra-embryonic endoderm (ExEn) (Fujikura et al. 2002). It has been suggested recently that the transcription factor *Sox17* also functions in the differentiation of mouse ES cells toward the ExEn (Shimoda et al. 2007). However, it is not known whether any of these factors directly regulate genes functioning in ExEn differentiation, including those encoding ExEn-specific cell surface proteins. Finally, it is not understood how the activation of this differentiation program leads to disengagement of the transcriptional network driving pluripotency and self-renewal.

While Sox17's functions in the formation and maintenance of definitive endoderm, vascular endothelium, and fetal hematopoietic stem cells are well established (Kanai-Azuma et al. 2002; Matsui et al. 2006; Kim et al. 2007), its role in ExEn differentiation is more poorly understood. We sought to clarify the function of *Sox17* in stem cell differentiation and understand how its expression affects genes involved in the maintenance of self-renewing ES cells.

We found that *Sox17* is expressed within the ICM of the mouse blastocyst and within ES cell cultures, where it is a central component of the transcriptional network governing differentiation. Our data elaborate three distinct mechanisms by which *Sox17* functions to promote differentiation. First, Sox17 binds to and activates a suite of genes that function in ExEn development and ICM cell sorting. Second, Sox17 consolidates the transcriptional network driving differentiation by directly stimulating the expression of *Gata6* and *Gata4*. Last, Sox17 inhibits the transcription of many genes normally expressed in ES cells and is required to properly antagonize the expression of *Sox2*, *Nanog*, and *Oct4*, thus aiding in the disengagement of the self-renewal program.

Results

Sox17 is expressed in the preimplantation embryo

Previous reports have suggested that *Sox17* mRNA initially appears at embryonic day 6 (E6) in the ExEn (Kanai-Azuma et al. 2002; Shimoda et al. 2007). However, its function in that context has not been studied extensively, nor has the possibility of earlier *Sox17* expression been explored. While examining published microarray data (Wang et al. 2005), we noted that *Sox17* transcription was induced between the morula and blastocyst stages of development (Fig. 1A). Clustering analysis of transcripts showed that the expression profile of *Sox17* was similar to those of several genes functioning within the primitive endoderm, including *Dab2*, *Lama1*, and *Sparc* (Fig. 1A).

To determine the expression domain of *Sox17* in the preimplantation embryo, we performed immunofluorescent staining with antibodies specific to both Sox17 and Oct4 (Fig. 1B). Sox17 was first observed at the 32-cell stage, where it was coexpressed with Oct4 (Fig. 1B). As the early blastocyst expanded, Sox17 expression was specific to a subset of ICM cells that increased in number over time. Sox17 was found in cells immediately adjacent to the blastocoel cavity and in cells deeper within the ICM (Fig. 1B). As development continued to the late blastocyst stage, cells expressing Sox17 became restricted to the primitive endoderm and were adjacent to the blastocoel cavity (Fig. 1B).

To confirm the specificity of our antibody recognizing Sox17, we used a mouse strain in which a cDNA encoding the *Tomato* red fluorescent protein was placed in-frame with the *Sox17* start codon via homologous recombination (Fig. 1C; Borowiak et al. 2009). *Sox17*::*Tomato* heterozygous mice ($Sox17^{+/-}$) were intercrossed and the resulting preimplantation embryos analyzed for Tomato expression (Fig. 1C; Supplemental Fig. S1). Control $Sox17^{+/+}$ embryos expressed Sox17, but not the Tomato protein (Fig. 1C). In heterozygous $Sox17^{+/-}$ embryos, the Tomato and Sox17 proteins were coexpressed, with Sox17 found in the nucleus and the Tomato protein localized throughout the cell (Fig. 1C). In $Sox17^{-/-}$ embryos, only the Tomato protein was present, with no detectable Sox17 protein, suggesting that this allele is a null mutation (Fig. 1C).

To confirm that the *Sox17* reporter allele was indeed a null, Tomato and Sox17 protein expression were examined by antibody staining in $Sox17^{+/+}$, $Sox17^{+/-}$, and $Sox17^{-/-}$ embryos at day 8 post-conception (Supplemental Fig. S2). As reported previously, $Sox17^{-/-}$ embryos were growth-retarded and failed to gastrulate properly (Supplemental Fig. S2; Kanai-Azuma et al. 2002; Kim et al. 2007).

Since Sox17 expression appeared to localize to the primitive endoderm lineage of the preimplantation embryo, we asked whether it was coexpressed with proteins known to control development of this cell type. We observed that Sox17 expression at the 32-cell stage preceded the appearance of Gata4 (Fig. 1D). Once Gata4 was expressed, around the 64-cell stage, it was always coincident with Sox17 (Fig. 1D). Gata6 expression was difficult to assess at the 32-cell stage due to high background staining; therefore, we could not determine the order of Gata6 and Sox17 expression. However, our microarray data (Supplemental Fig. S3) and recent reports (Plusa et al. 2008) suggested that *Gata6* is transcribed before *Sox17*, perhaps as early as the eight-cell stage. In the blastocyst, where we could accurately resolve Gata6 expression by immunofluorescence, it was always coincident with Sox17 (Fig. 1E).

We also found coexpression of Sox17 and the protein Laminin, which is thought to confer an adherence difference to cells of the primitive endoderm (Chazaud et al. 2006). At the 32-cell stage, only one cell expressing *Sox17* was ordinarily observed, and Laminin was present only on the surface of that cell (Fig. 1F). No Laminin expression was observed prior to the 32-cell stage. By the

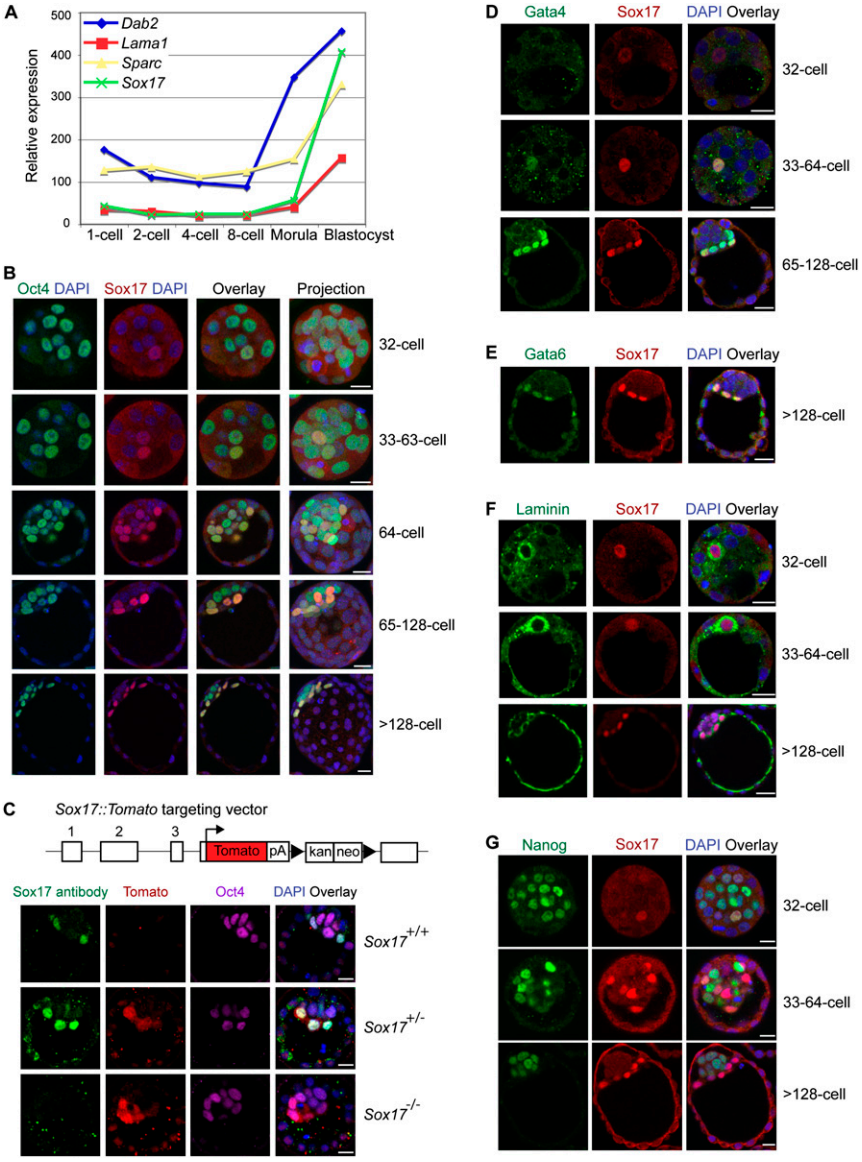

Figure 1. Sox17 is expressed in the preimplantation embryo. (*A*) Microarray analysis of *Sox17*, *Dab2*, *Laminin*, and *Sparc* expression at defined stages of mouse preimplantation development ($n = 3$). (*B–G*) Optical sections of immunostained mouse preimplantation embryos. (*B*) Optical sections and reconstructed projections of Oct4 and Sox17 expression. (*C*) Schematic representation of the *Sox17*::*Tomato* reporter gene. *Sox17*::*Tomato* heterozygous mice were intercrossed, and the resulting embryos were immunostained with Tomato, Oct4, and Sox17 antibodies. Optical sections of mouse embryo immunofluorescently stained with Sox17 and Gata4 (*D*), Gata6 (*E*), Laminin (*F*), and Nanog (*G*) antibodies.

late-blastocyst stage, Laminin was found in both the primitive endoderm, overlapping with Sox17, and in the trophectoderm, where Sox17 was not expressed (Fig. 1F).

During preimplantation development, Gata6 and Nanog are initially coexpressed in the ICM, and then gradually become partitioned into distinct expression domains within the primitive endoderm and epiblast (Chazaud et al. 2006; Plusa et al. 2008). Following this compartmentalization, cells that express *Nanog* give rise to the epiblast, while cells that express *Gata6* form the primitive endoderm. As Sox17 was also expressed early in development of the primitive endoderm, we asked whether it, too, showed a mosaic expression pattern with Nanog. Staining of embryos with antibodies specific to Nanog and Sox17 indicated that all cells within the 32-cell stage embryo that expressed Sox17 also expressed Nanog (Fig. 1G; Supplemental Fig. S4). Around the 64-cell stage, the ICM of the blastocyst was a mosaic of cells coexpressing both Nanog and Sox17, interspersed with cells that expressed either Nanog only or Sox17 only (Fig. 1G; Supplemental Fig. S4). By the 128-cell stage, this mosaic expression pattern had resolved, with the ICM comprised of cells that expressed only Sox17 or Nanog (Fig. 1G; Supplemental Fig. S4). These results demonstrate that, similar to Gata6, Sox17 is initially coexpressed with Nanog, but then gradually becomes partitioned to the primitive endoderm. This expression pattern for *Sox17* is consistent with the previously proposed cell sorting model for the segregation of epiblast progenitors and primitive endoderm cells within the ICM (Chazaud et al. 2006; Plusa et al. 2008).

Sox17 is required for blastocyst-derived ExEn (XEN) cell derivation

It has been demonstrated previously that the three cell lineages within the developing blastocyst can be modeled using in vitro cultured cell types. These include ES cells,

trophoblast stem cells, and XEN cells (Evans and Kaufman 1981; Martin 1981; Tanaka et al. 1998; Kunath et al. 2005). We took advantage of these in vitro cultured cell types to investigate the function of *Sox17* in the preimplantation embryo.

To confirm previous reports suggesting that Sox17 is expressed in XEN cells (Kunath et al. 2005), we performed immunofluorescent staining and quantitative RT–PCR. Sox17 was coexpressed in XEN cells along with Gata4, Laminin, Dab2, and Sparc (Fig. 2A; Kunath et al. 2005). As reported previously, XEN cells did not express Nanog or Oct4 (Fig. 2A; Supplemental Fig. S5; Kunath et al. 2005). Quantitative RT–PCR confirmed transcription of *Sox17* and *Gata4* in these cells (Supplemental Fig. S5).

To investigate whether *Sox17* has an important function in the ExEn, we asked whether XEN cell lines could be derived from $Sox17^{-/-}$ embryos. We intercrossed $Sox17^{+/-}$ animals and attempted to generate XEN cell lines from the resulting embryos. While $Sox17^{+/+}$ and $Sox17^{+/-}$ XEN cell lines could be generated, we were unable to derive $Sox17^{-/-}$ XEN cell lines (Fig. 2B), indicating that *Sox17* is required for the establishment of this cell type.

Our in vitro results suggested that *Sox17* deficiency might have consequences for the establishment of functional primitive endoderm cells in vivo. We next investigated whether the loss of *Sox17* in the blastocyst has consequences for expression of pluripotency genes such as Nanog or the expression of ExEn cell surface proteins such as Laminin. We collected *Sox17* mutant preimplantation embryos generated from heterozygous crosses and performed immunofluorescence analysis followed by PCR to confirm genotypes. *Sox17* mutant embryos appeared to properly segregate Nanog-expressing epiblast progenitors and maintained Laminin expression in the primitive endoderm (Supplemental Fig. S4). Our observations confirm previous reports suggesting a lack of an obvious defect in the establishment of the primitive endoderm in *Sox17*-deficient embryos (Kanai-Azuma et al. 2002; Shimoda et al. 2007). This result suggests there may be functional redundancy with other transcription factors, as has been suggested (Kanai-Azuma et al. 2002; Shimoda et al. 2007).

Sox17 regulates XEN gene expression

As $Sox17^{-/-}$ XEN cell lines could not be generated and the functional redundancy of the transcription factor network in the blastocyst seemed to preclude further functional studies in vivo, we proceeded to investigate the functional importance of *Sox17* in an alternative in vitro model of ICM differentiation. Since their first introduction as a model for studying cavitation in the early embryo, ES cell-derived embryoid bodies (EBs) have proven useful in understanding the formation of the XEN (Coucouvanis and Martin 1995). As in the ICM, cells within a newly formed EB express *Oct4* and *Nanog*. As

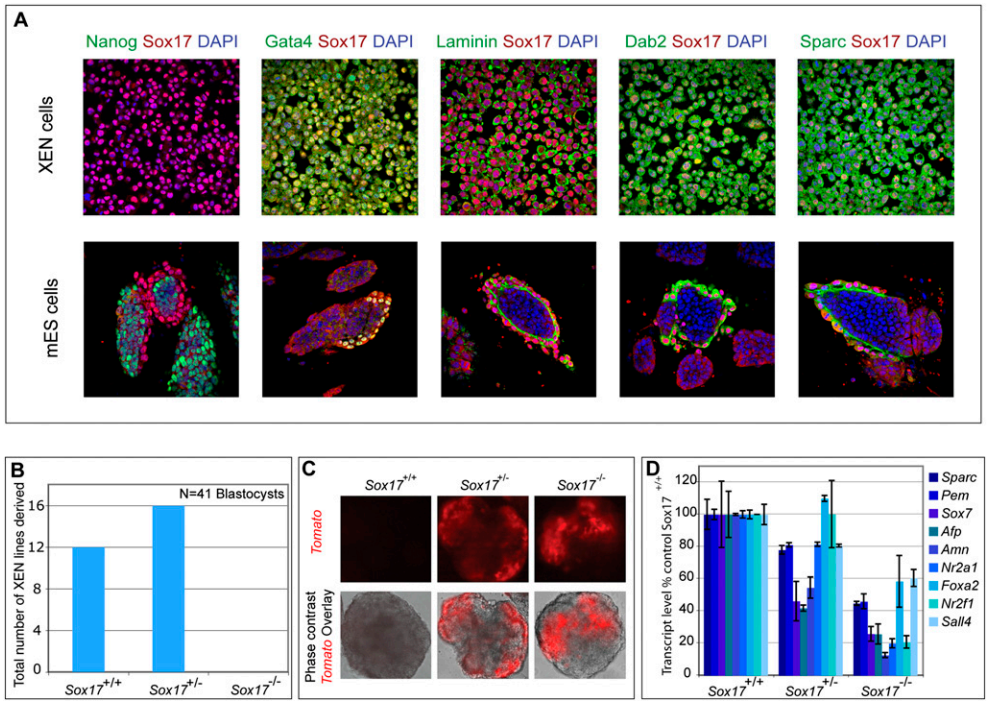

Figure 2. *Sox17* is required for XEN cell derivation and regulates ExEn gene expression. (*A*) Sox17, Gata4, Laminin, Dab2, Sparc, and Nanog immunostaining in XEN and ES cell cultures. (*B*) Forty-one embryos resulting from a $Sox17^{+/-}$ heterozygous intercross were used to derive XEN cell lines. (*C*,*D*) Analysis of day 6 EBs generated from $Sox17^{+/+}$, $Sox17^{+/-}$, and $Sox17^{-/-}$ ES cells. (*C*) Epifluorescence images of Tomato expression and localization in $Sox17^{+/-}$::*Tomato* and $Sox17^{-/-}$::*Tomato* EBs. (*D*) Quantitative RT–PCR analysis of *Afp*, *Amn*, *Nr2a1*, *Sox7*, *Sparc*, *Pem*, *Foxa2*, *Nr2f1*, and *Sall4* transcript levels, with values adjusted to *Gapdh* and relative expression reflected as a percent of the expression observed in wild-type $Sox17^{+/+}$ EBs ($n = 4$).

differentiation proceeds, a subset of cells activate genes found in the primitive endoderm, such as *Gata6* and *Gata4*. These cells are then thought to migrate to the margin of the EB, where they form an epithelial cell layer expressing additional genes functioning in the ExEn (Soudais et al. 1995; Morrisey et al. 1998). It has been shown previously that EBs generated from *Gata6* and *Gata4* mutant ES cells fail to form this epithelialized cell layer, suggesting this might be a more amenable system for studying the role of *Sox17* in differentiation (Capo-Chichi et al. 2005).

To test whether *Sox17* was required in this assay for differentiation of ES cells, we derived $Sox17^{+/-}$ and $Sox17^{-/-}$ ES cell lines and then used them to generate EBs (Fig. 2C; Supplemental Fig. S6). Epifluorescence imaging and quantification of the localization of *Sox17* expression revealed that *Sox17*::*Tomato* was always found at the periphery of $Sox17^{+/-}$ control EBs, adjacent to the epithelialized basement membrane (Fig. 2C; Supplemental Fig. S6). In contrast, *Sox17*::*Tomato* expression was highly disorganized and mislocalized to the center of mutant $Sox17^{-/-}$ EBs (Fig. 2C; Supplemental Fig. S6). These observations suggest that loss of *Sox17* leads to a failure in cell sorting to the periphery of the EB.

Since cells in $Sox17^{-/-}$ EBs attempting to express *Sox17* failed to sort to the exterior of EBs, we next tested whether loss of *Sox17* effected expression of genes functioning in ExEn differentiation (Shimoda et al. 2007). Loss of both alleles of *Sox17* resulted in a significant reduction in the expression of ExEn transcripts *Sox7*, *Pem*, *Sparc*, *Afp*, *Amn*, *Nr2a1*, *Foxa2*, and *Nr2f1* ($P < 0.01$) (Fig. 2D). Transcript levels for the majority of these genes depended on the dose of *Sox17*, such that many were reduced significantly in heterozygous $Sox17^{+/-}$ EBs ($P < 0.01$). In contrast, *Foxa2* and *Nr2f1* were up-regulated and unchanged, respectively, in heterozygous $Sox17^{+/-}$ EBs, while both were significantly down-regulated in mutant $Sox17^{-/-}$ EBs, suggesting that these genes are subject to more complex regulation by *Sox17* (Fig. 2D). Consistent with previous reports, *Sox17*-deficient ES cells also exhibited a dose-dependent down-regulation of the transcription factor *Sall4*, which has a role in both embryonic and extraembryonic stem cell maintenance (Lim et al. 2008).

Cells expressing Sox17 are committed to differentiate

During routine culture of mouse ES cells, we observed that Sox17 was transcribed at a low level (Supplemental Fig. S5). Immunofluorescence revealed that Sox17 was highly expressed only in a subset of cells on the outside of otherwise undifferentiated stem cell colonies (Fig. 2A). Quantitative RT–PCR confirmed that ExEn genes such as *Gata4* were also transcribed at a low level in these ES cell cultures (Supplementa Fig. S5). We found that cells either expressed Nanog or high levels of Sox17 but not both, while Sox17 expression colocalized with ExEn-associated proteins such as Gata4, Laminin, Dab2, and Sparc (Fig. 2A; Supplemental Fig. S4).

We next investigated whether cells that spontaneously expressed high levels of *Sox17* within an ES cell culture were committed to differentiation and were no longer pluripotent. To address this issue, we used a *Sox17*::*GFP* knock-in reporter ES cell line to perform in vitro clonal analysis and in vivo transplant assays into blastocysts (Fig. 3A; Kim et al. 2007). In order to follow the fate of the *Sox17*::*GFP* cells irrespective of *Sox17*::*GFP* expression, we further labeled these cells by transduction with a lentivirus carrying a coding sequence for the red fluorescent Tomato protein under control of the widely expressed *Ubiquitin* promoter (Fig. 3B). ES cells were dissociated into a single-cell suspension, and both *Sox17*::*GFP*-high and *Sox17*::*GFP*-low cells were identified under epifluorescence, picked by hand, and used for clonal analysis (Fig. 3C).

As might be expected, and similar to the parental cultures, clonally expanded *Sox17*::*GFP*-low cells formed typical ES cell colonies composed of cells that did not express *Sox17*::*GFP*, and were occasionally surrounded by cells that did (Fig. 3D). In contrast, when *Sox17*::*GFP*-high cells were clonally expanded under routine ES cell growth conditions, typical ES cell colonies were not observed. Instead, cultures of proliferating homogeneous *Sox17*::*GFP*-high cells were produced that morphologically resembled endoderm (Fig. 3E; Fujikura et al. 2002). These results suggest that *Sox17* expression marks ES cells committed to differentiation.

To confirm these observations in vivo, blastocysts were injected with either *Sox17*::*GFP*-high or *Sox17*::*GFP*-low cells. The embryos were then transferred into pseudopregnant recipient mice. The developmental potency of the injected cells was examined by isolating the resulting E6–E6.5 embryos, then determining the degree to which transplanted cells had incorporated into the embryonic and/or extraembryonic lineages (summarized in Table 1). The injection of *Sox17*::*GFP*-low-expressing ES cells into blastocysts (Fig. 3F) resulted in cells that remained *Sox17*::*GFP*-low and that only contributed to the epiblast (Fig. 3H). This suggests that *Sox17*::*GFP*-low cells, which are presumably undifferentiated ES cells, are more restricted in their potential in vivo than in vitro. One possible explanation for this finding is that, once injected into blastocysts, cells within the host embryo suppress the ability of the ES cells to differentiate into the ExEn.

In contrast, when single cells with high *Sox17*::*GFP* expression (Fig. 3G) were injected, they showed incorporation into the ExEn and failed to give rise to epiblast (Fig. 3I). Often, we observed several labeled cells in the ExEn, suggesting that the injected cells had divided (Fig. 3I). We also observed that these *Sox17*::*GFP*-high cells regularly contributed to the parietal and visceral endoderm, but we did not observe contribution to the extraembryonic visceral endoderm overlaying the extraembryonic ectoderm. Taken together, these data demonstrate that ES cells with high *Sox17* expression are committed to differentiate down the extraembryonic lineage.

Sox17 binds to target genes required for extraembryonic differentiation

The experiments described thus far show that *Sox17* is expressed in cells committed to differentiate into the

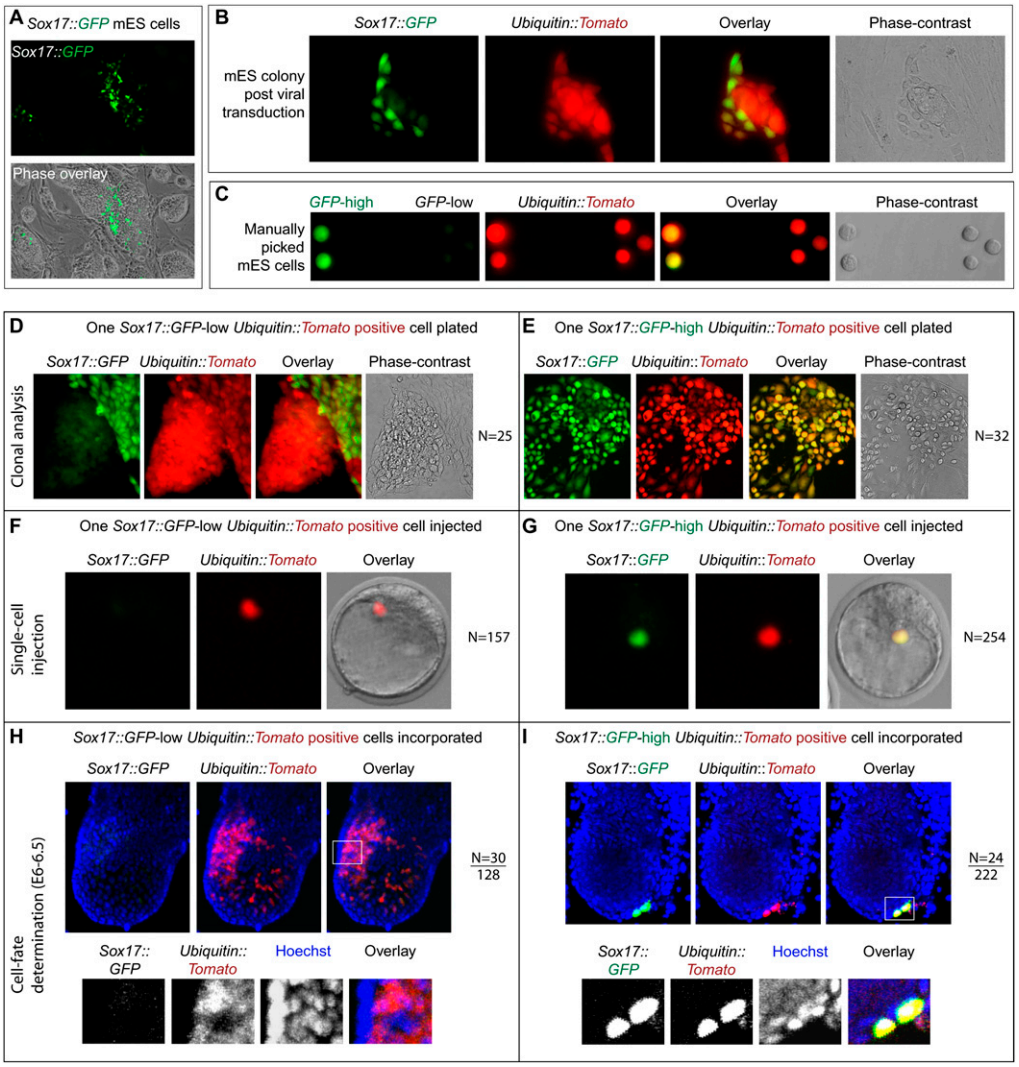

Figure 3. Cells expressing *Sox17* are committed to differentiate. (*A*) Epifluorescent image of *Sox17*::*GFP* expression under standard mouse ES cell culture conditions. (*B*) Epifluorescent image of *Sox17*::*GFP* ES cell following lentiviral transduction with *Ubiquitin*::*Tomato*. (*C*) Manually picked *Sox17*::*GFP*-high or *Sox17*::*GFP*-low cells expressing *Ubiquitin*::*Tomato* prior to clonal culture or blastocyst injection. (*D*,*E*) Single *Sox17*::*GFP*-low cells (*D*) and single *Sox17*::*GFP*-high subclones (*E*) plated under standard mouse ES culture conditions. (*F*, *G*) One hour post-injection of a single *Sox17*::*GFP*-low (*F*) or *Sox17*::*GFP*-high (*G*) cell into wild-type blastocysts. (*H*, *I*) Reconstructed projections of E6–E6.5 representative embryos that had incorporated a single or multiple *Sox17*::*GFP*-low (*H*) or *Sox17*::*GFP*-high (*I*) cell(s) prior to the establishment of the definitive endoderm (Hoechst nuclear overlay). (*H*) Representative dissected embryo from the transfer of a blastocyst injected with eight to 10 *Ubiquitin*::*Tomato*-positive/*Sox17*::*GFP*-low cells. (*I*) Representative dissected embryo from the transfer of a blastocyst injected with a single *Ubiquitin*::*Tomato*-positive/*Sox17*::*GFP*-high cell.

extraembryonic lineage, where it is functionally required for the regulation of gene expression and the establishment of XEN cell lines. We next investigated whether *Sox17* directly or indirectly regulates ExEn gene expression by identifying Sox17 DNA-binding sites using chromatin immunoprecipitation (ChIP) coupled with whole-genome promoter tiling array analysis (ChIP–chip).

We used the Sox17 antibody to ask whether Sox17 was binding directly to the regulatory regions of genes in XEN cell lines (Fig. 4; Supplemental Table S3; Kunath et al. 2005). In XEN cells, Sox17-binding sites were located

Table 1. *Cells expressing Sox17 are committed to differentiate*

	Blastocysts injected	Embryos recovered	Incorporated cells	
Sox17-GFP-high *Tomato*-positive	254	222	24	All ExEn
Sox17-GFP-low *Tomato*-positive	157	128	30	All epiblast

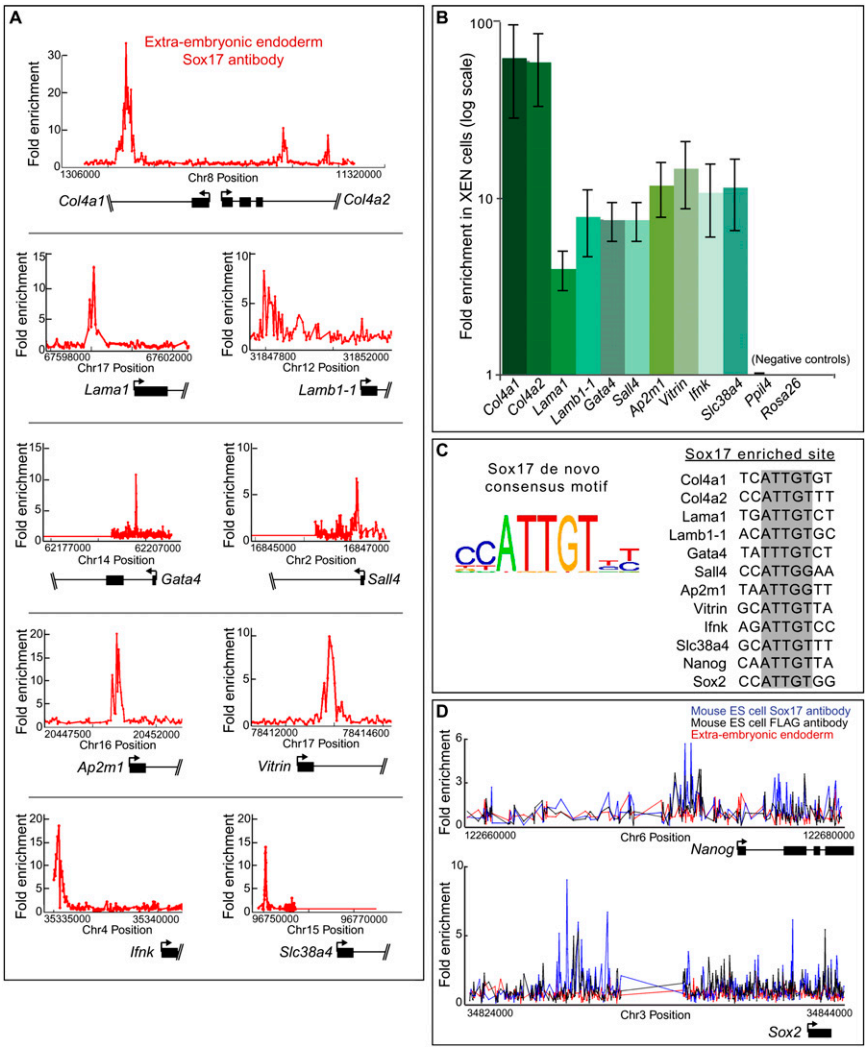

Figure 4. Sox17 binds to target genes required for extraembryonic differentiation. (*A*) Examples of DNA sequences bound by Sox17. The unprocessed ChIP enrichment ratio was plotted with the associated chromosome and gene location, exon (box), intron (horizontal line), transcription direction, and start site (arrow) for each Sox17-bound region. (*B*) ChIP followed by quantitative PCR using primers spanning the Sox17-bound regions. Fold enrichment was normalized to Ct value of *Ppil4* and *Rosa26* loci (n = 3). (*C*) The Sox17-binding consensus motif of ATTGT was identified using de novo motif analysis. The gray box indicates similarities in the genomic consensus motif at Sox17-enriched sites. (*D*) Sox17 occupies the promoter sequence of *Nanog* and *Sox2* in *Sox17*-induced mouse ES cells (black and blue) but not in XEN cells (red).

within the promoters or the introns of 2206 (3%) genes (Supplemental Table S3). We performed an ontology analysis for the genes with Sox17-binding sites, and found that a significant number had functions in basement membrane establishment and maintenance (Supplemental Fig. S7). Genes with significant binding sites included *Col4a1*, *Col4a2*, *Lama1*, *Lamb1-1*, and *Lamc1*, each of which are necessary to establish the extracellular matrix (ECM) during basement membrane formation between the epiblast and ExEn (Fig. 4A; Supplemental Table S3; Miner and Yurchenco 2004). We also found Sox17 bound the promoter regions of genes encoding additional basement membrane components not identified previously as being expressed in the XEN, including *Vitrin* (Fig. 4A; Supplemental Table S3; Manabe et al. 2008). ChIP followed by quantitative PCR using primers spanning the enriched Sox17-binding sites confirmed our microarray results (Fig. 4B). These studies suggest that one of Sox17's functions in the differentiation of ICM and ES cells is to bind the regulatory regions of many genes that encode basement membrane components, thus leading to their activation.

In addition to these ECM genes, Sox17 was also bound to promoter regions of other genes implicated in ExEn development, including *Pdgfrα*, *Fgfr2*, *Sall4*, *Slc38a4*, *Ifnk*, and *Ap2m1* (Fig. 4A,B; Supplemental Table S3; Arman et al. 1999; Mitsunari et al. 2005; Lim et al. 2008; Plusa et al. 2008). These results point to the significance of Sox17 as a critical and direct effecter of extraembryonic differentiation. Consistent with this conclusion, we noted that Sox17 was also bound to the regulatory regions of *Gata4* and *Gata6* (Fig. 4A,B; Supplemental Table S3). Thus, in addition to directly activating genes required for primitive endoderm differentiation, it seemed that *Sox17* might also function to activate and reinforce the transcriptional network governing differentiation.

Sox17 expression is sufficient to activate its target genes

The direct binding of Sox17 to the regulatory regions of genes expressed within the ExEn suggests that forced expression of this transcription factor might be sufficient

to directly activate its target genes. To address this question, we performed a gain-of-function analysis by generating ES cell lines carrying a doxycycline-inducible Flag-tagged *Sox17* transgene (Fig. 5A). After treatment of transgenic ES cells with doxycycline, there was robust induction of the *Sox17* transcript (Fig. 5C; Supplemental Fig. S8). Sox17 protein expression, which colocalized with the Flag antigen, was observed in a subset of cells within the induced cultures (Fig. 5B). However, we did note that some cells did not seem to respond to the doxycycline treatment and expressed neither Sox17 nor the Flag antigen (Fig. 5B).

To determine whether Sox17 overexpression could directly activate its target genes, we treated *Sox17*-inducible ES cells with doxycycline, collected RNA, and performed genome-wide transcriptional analysis and quantitative RT–PCR. Following Sox17 induction, we found increased expression of both the *Lama1* gene and genes encoding basement membrane components *Col4a1* and *Col4a2* (Fig. 5D; Supplemental Fig. S8; Supplemental Table S4). Immunofluorescence analysis confirmed that ectopic expression of *Sox17* in ES cells was sufficient to induce the expression of Laminin and Collagen proteins (Fig. 5B). ChIP–chip analysis using these doxycycline-treated ES cells further demonstrated that Sox17 occupied the regulatory regions upstream of *Col4a1*, *Col4a2*, *Lama1*, and *Lamb1-1* (Supplemental Fig. S9; Supplemental Tables S5,S6).

We next wanted to determine whether Sox17 can positively regulate the key endoderm transcription factors *Gata6* and *Gata4*, whose induction alone has been demonstrated to be sufficient to differentiate ES cells into XEN stem cells (Fujikura et al. 2002; Shimosato et al. 2007). Following doxycycline treatment of ES cells, we found that *Sox17* expression was sufficient to activate transcription of *Gata6* and *Gata4* (Fig. 5D; Supplemental Fig. S8; Supplemental Table S4). *Sox17* induction also resulted in the down-regulation of *Sall4*.

These results suggest that Sox17 not only regulates expression of cell surface proteins required for differentiation, but it is also a critical component of the transcriptional network that governs the decision that ES cells make to differentiate into ExEn. If this is the case, then the loss of *Sox17* might have ramifications for the ability of ES cells to exit their pluripotent transcriptional program and to differentiate into the ExEn. We began to investigate this possibility by analyzing the epistatic relationship between Sox17 and the other transcription

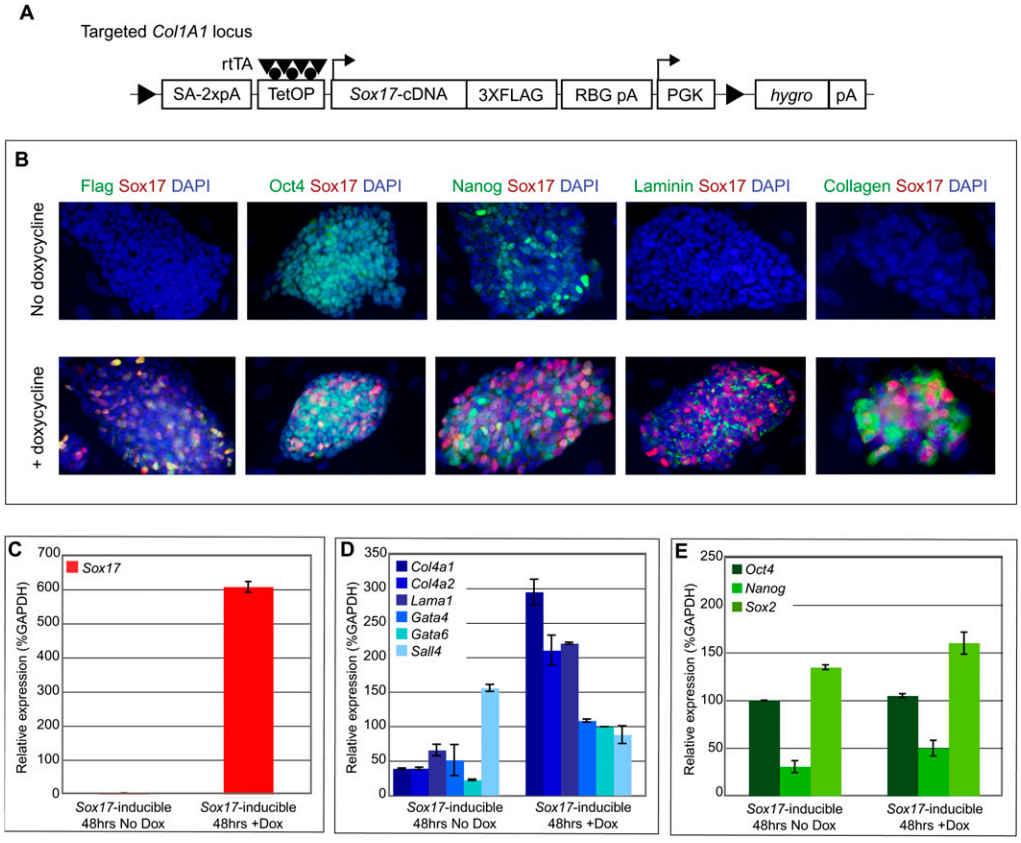

Figure 5. *Sox17* expression is sufficient to activate ExEn target genes. (*A*) Schematic of the inducible *Sox17* construct. (SA) Splice acceptor; (pA) polyadenylation signal; (TetOP) tetracycline/doxycycline-responsive element; (3XFLAG) Flag epitope. (*B*) Immunostaining of *Sox17*-inducible ES cells in the absence or presence of doxycycline. Quantitative RT–PCR of *Sox17* (*C*); *Col4a1*, *Col4a2*, *Lama1*, *Gata4*, *Gata6*, and *Sall4* (*D*); and *Oct4*, *Nanog*, and *Sox2* (*E*) in *Sox17*-inducible ES cells without or with doxycycline treatment (48 h). The relative expression was reflected as a percent of *Gapdh* expression ($n = 4$).

factors known to influence cell fate decisions within ES cells.

Epistatic relationship between Sox17 and other transcriptional regulators

We examined the epistatic relationship between *Sox17*, *Gata6*, *Gata4*, and *Nanog* by examining gene expression and extraembryonic differentiation in mutant EBs. EBs were generated from *Sox17*-, *Gata6*-, and *Gata4*-deficient ES cells (Morrisey et al. 1998; Watt et al. 2004); *Gata6* and *Gata4* compound mutant ES cells (Zhao et al. 2008); and ES cells constitutively overexpressing *Nanog* (Silva et al. 2006).

Notably, we observed that loss of *Sox17* resulted in a significant increase in expression of the pluripotency-associated genes *Sox2*, *Nanog*, and *Oct4* ($P < 0.05$), and that *Sox17*$^{-/-}$ EBs possessed a morphology very similar to EBs generated from ES cells with forced *Nanog* overexpression (Fig. 6A,B; Supplemental Fig. S10). These phenotypes were also shared with *Gata6*$^{-/-}$, *Gata4*$^{-/-}$, and *Gata4*$^{-/-}$ *Gata6*$^{-/-}$ compound-null EBs (Fig. 6A,B; Supplemental Fig. S10), suggesting that, like *Gata6* and *Gata4*, *Sox17* expression is also required for normal down-regulation of *Nanog*. However, we observed that a dramatic increase in *Oct4* expression was specific to the removal of *Sox17*, as this effect was not observed in *Gata4* mutants and was more variable following *Gata6* loss of function (Fig. 6B). These results indicate that *Sox17* is functionally necessary to antagonize the expression of three key pluripotency genes.

We observed that *Gata6* and *Gata4* expression persisted in *Sox17*$^{-/-}$ EBs, but that cells expressing these genes did not properly localize to the periphery of EBs to form an epithelialized margin (Fig. 6A,C; Supplemental Fig. S10). Despite greater variance in *Gata6* expression in the *Sox17*$^{+/-}$ EBs, the absolute level of *Gata6* in this background was not significantly different from the level in wild-type EBs ($P = 0.2$) (Fig. 6C). While the levels of *Gata6* expression remained unchanged in *Sox17*$^{-/-}$ EBs, *Gata4* expression levels were significantly decreased ($P < 0.01$) (Fig. 6C). These results suggest that the relative abundance and correct localization of *Gata4* during extraembryonic development were *Sox17*-dependent, while the initiation of *Gata4* expression was *Sox17*-independent (Fig. 6A,C).

As expected from previous findings, *Gata6* mutant EBs lacked *Gata4* expression and resembled the *Gata6* and *Gata4* compound-null EBs, while *Gata4* mutant EBs maintained *Gata6* expression. Thus, *Gata6* lies upstream

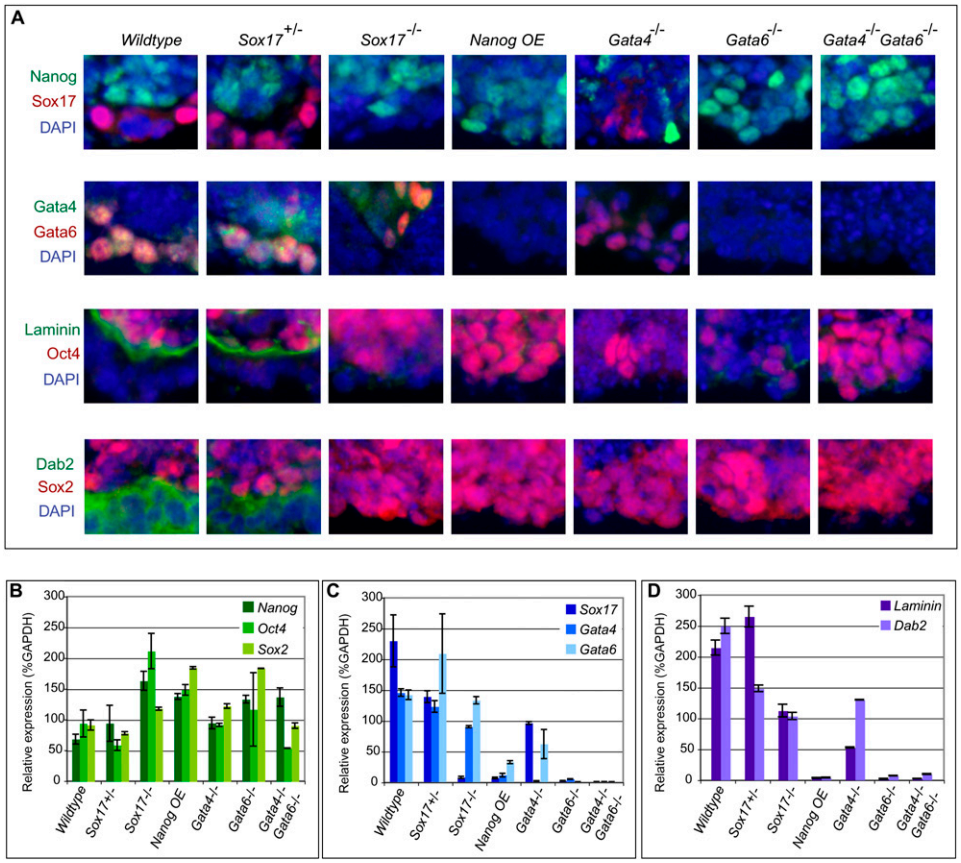

Figure 6. Epistatic relationship between *Sox17* and other transcriptional regulators. (*A*) Wild-type, *Sox17*$^{+/-}$, *Sox17*$^{-/-}$, *Nanog*-overexpressing (OE), *Gata4*$^{-/-}$, *Gata6*$^{-/-}$, and *Gata4*$^{-/-}$ and *Gata6*$^{-/-}$ compound mutant day 6 EBs were examined by immunostaining for Oct4, Nanog, Sox2, Sox17, Laminin, Dab2, Gata4, and Gata6. Quantitative RT–PCR analysis of *Nanog*, *Oct4*, and *Sox2* (*B*); *Sox17*, *Gata6*, and *Gata4* (*C*); and *Laminin* and *Dab2* (*D*), with relative expression reflected as a percent of *Gapdh* expression ($n = 4$).

of *Gata4* in this genetic pathway (Fig. 6A,C; Supplemental Fig. S10). *Gata6*$^{-/-}$ EBs lacked detectable *Sox17* expression, suggesting that *Gata6* also lies upstream of *Sox17* (Figs. 6A,C, 7D). In contrast, *Gata4*$^{-/-}$ EBs maintained *Sox17* expression, although *Sox17* expression in these EBs was largely disorganized and quantitatively lower than in wild-type controls ($P < 0.001$) (Fig. 6A,C; Supplemental Fig. S10). These results suggest that, while migration of cells expressing *Sox17* to the periphery of EBs required *Gata4*, the initiation of *Sox17* expression was *Gata4*-independent. This finding is consistent with the notion that *Gata4* and *Sox17* are not dependent on one another for initiation of their expression, but that, once activated, they reciprocally stimulate each other's transcription.

We next asked whether failures in cell migration might be due to reduced expression of ExEn-associated cell surface proteins. Loss of Sox17 resulted in a significant reduction in *Lama1*, *Col4a1*, *Col4a2*, and *Dab2* transcripts ($P < 0.04$), and we failed to detect Laminin and Dab2 staining in *Sox17*$^{-/-}$ EBs (Fig. 6A,D; Supplemental Figs. S10, S11). This was very similar to the phenotype that we observed in EBs upon loss of the Gata transcription factors or overexpression of *Nanog*. Interestingly, we consistently observed that the expression of *Lama1*, *Col4a1*, and *Col4a2* were up-regulated significantly ($P < 0.02$) in *Sox17*$^{+/-}$ EBs compared with wild-type controls, suggesting compensatory induction of these transcripts, resulting from the loss of one *Sox17* allele (Fig. 6D; Supplemental Fig. S11).

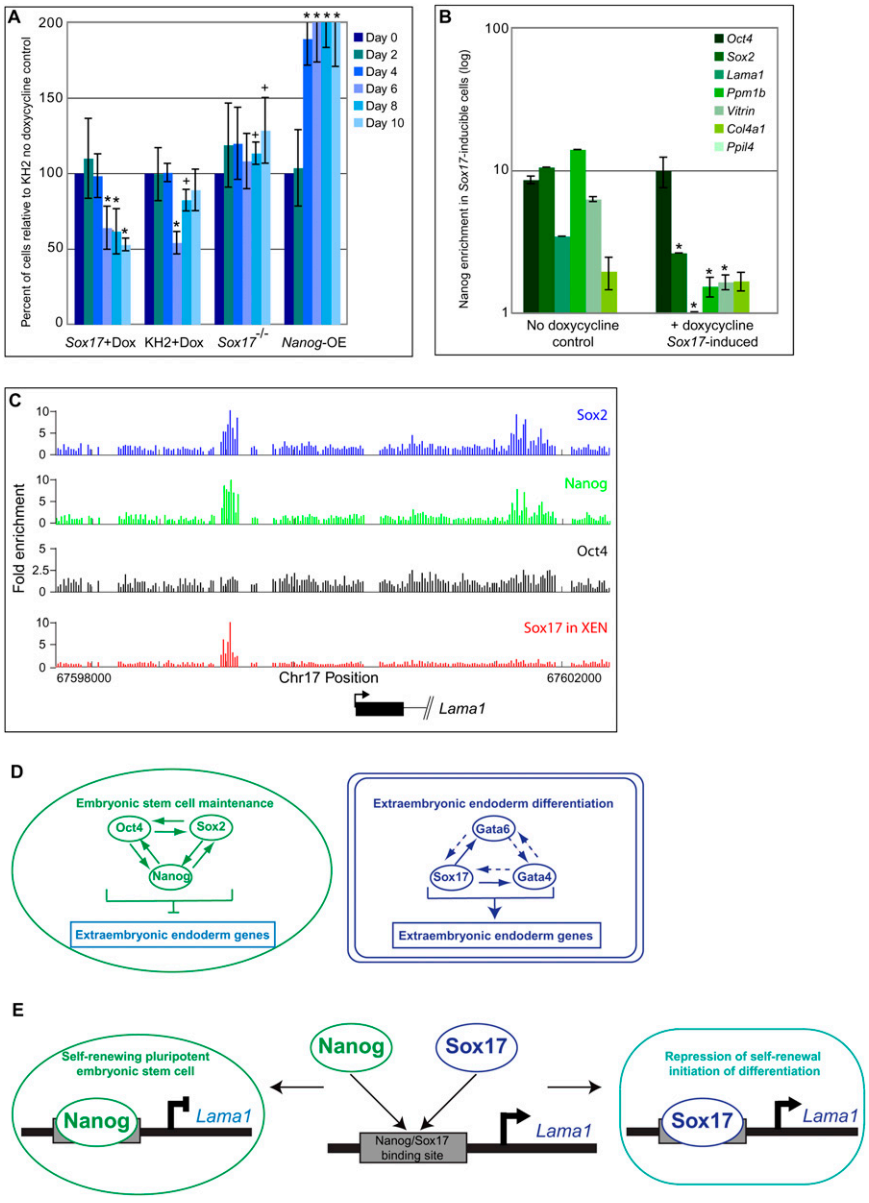

Figure 7. *Sox17* inhibits ES cell self-renewal by displacing Nanog. (*A*) Control ES (KH2), *Sox17*-inducible, *Sox17* mutant, and *Nanog*-overexpressing cells were plated at the same initial cell density in the presence or absence of doxycycline ($n = 3$). (*) $P < 0.01$; (+) $P < 0.05$. (*B*) ChIP followed by quantitative PCR analysis of Nanog enrichment at the *Oct4*, *Sox2*, *Lama1*, *Ppm1b*, *Vitrin*, and *Col4a1* regulatory regions was compared in *Sox17*-uninduced and *Sox17*-induced cells following 48 h of doxycycline induction ($n = 3$). (*) $P < 0.01$. (*C*) Sox17 ChIP–chip targets were compared with the regulatory regions bound by Sox2, Nanog, and Oct4 (as published previously in Chen et al. 2008; Kim et al. 2008). Representative example of the target gene, *Lama1*. Unprocessed ChIP enrichment ratios were plotted for each probe, together with the associated chromosome and genomic location, exon (box), intron (horizontal line), transcription direction, and start site (arrow). (*D*) A model for Sox17 gene regulation in ExEn differentiation. The Oct4, Sox2, and Nanog transcription factor network is a feed-forward loop maintaining mouse ES cell pluripotency (Boyer et al. 2006), while inhibiting genes involved in differentiation. *Gata6* lies upstream of *Gata4* in the differentiation cascade (depicted as dashed lines, indicating no evidence to date that these interactions are direct) (Chazaud et al. 2006). *Sox17* lies downstream from *Gata6* and directly regulates the expression of *Gata6* and *Gata4* (solid lines). Sox17 binds directly to and activates the transcription of genes known to function in ExEn differentiation. (*E*) Nanog and Sox17 were bound reciprocally to shared Nanog/Sox17 genomic sites, represented here by the region upstream of the *Lama1* start site (arrow). This suggests a mechanism for initiation of differentiation in which Sox17 displaces repressive *Nanog* complexes from shared binding sites and, in turn, activates gene expression.

Sox17 expression is sufficient to down-regulate aspects of the ES cell transcriptional program

Sox17 is required for the proper reduction of *Sox2*, *Nanog*, and *Oct4* expression in EBs, suggesting that it might functionally repress the ES cell transcription program during differentiation. To test this, we asked whether Sox17 interacted with *Sox2*, *Nanog*, *Oct4*, or other genes known to function in the maintenance of ES cell pluripotency, and whether Sox17 expression was sufficient to inhibit their expression (Boiani and Scholer 2005).

We induced the expression of *Sox17* in ES cells and identified Sox17-binding sites by ChIP–chip. Indeed, we found that Sox17 was bound upstream of *Sox2* and *Nanog*, suggesting that it regulates the expression of these genes directly (Fig. 4D; Supplemental Table S5). However, we did not find convincing binding for Sox17 near *Oct4* (Supplemental Tables S5, S6), suggesting either indirect regulation or that the binding site was not represented on our array.

Sox17 was also observed to bind directly to the regulatory regions of many genes acting in pathways known to function in ES cell pluripotency and self-renewal. Ontology analysis of all of the Sox17 ChIP–chip binding targets identified in ES cells demonstrated a significant enrichment near genes involved in the cell cycle, as well as the *Wnt*, *Tgfβ*, and *Jak–STAT* signaling pathways (Supplemental Fig. S12; Supplemental Table S5; Boiani and Scholer 2005). We also performed a genome-wide analysis of transcription after *Sox17* induction in ES cells and observed that transcription of pluripotency-associated genes was suppressed (Supplemental Fig. S13; Supplemental Table S7; Boiani and Scholer 2005). However, *Sox17* expression was not sufficient to rapidly down-regulate *Sox2*, *Nanog*, and *Oct4* (Fig. 5B,E; Supplemental Fig. S8), a finding consistent with the overexpression of *SOX17* in human ES cells (Seguin et al. 2008).

We next performed a teratoma analysis to evaluate the pluripotency of cells induced to express *Sox17*, and to determine whether they had a propensity to differentiate into ExEn cell types. We injected *Sox17*-inducible cells that had been treated with doxycycline for 5 d into NOD-SCID mice and continued to administer doxycyline during the course of the teratoma assay. We found that, following *Sox17* induction, ES cultures maintained the potential to give rise to all three embryonic germ layers in teratomas (Supplemental Fig. S14), and we did not observe obvious skewing in the various cell types produced. While these results suggest that *Sox17* expression might not be sufficient to prevent teratoma formation, we did observe that doxycyline treatment of our Sox17-inducible cells did not result in homogeneous induction of Sox17 expression (Fig. 5B). To more accurately resolve whether *Sox17* induction has consequences for ES cell pluripotency in a teratoma assay, it will be important to purify cells that have successfully induced *Sox17*.

We also performed a teratoma analysis to investigate whether *Sox17*-deficient ES cells maintained the capacity to give rise to XEN cells. *Sox17* mutant ES cells maintained the capacity to give rise to all three germ layers (Supplemental Fig. S14). The presence of large epithelial structures reminiscent of the ExEn have been observed previously in teratomas derived from human ES cells overexpressing *SOX7* (Seguin et al. 2008). It was difficult to discern these structures in either the wild-type control or *Sox17*-deficient teratomas, and we therefore cannot conclude from these results whether *Sox17* deficiency prevents formation of ExEn within a teratoma.

Sox17 expression has significant consequences for ES cell self-renewal

We observed that *Sox17* was both necessary and sufficient for the induction of basement membrane components in ES cell cultures, and antagonized pathways promoting ES cell self-renewal and pluripotency. We next asked whether *Sox17* expression had functional consequences for ES cell self-renewal. Indeed, induction of *Sox17* with doxycycline resulted in a significant and sustained inhibition of ES cell self-renewal, while *Sox17* deficiency led to a significant increase in self-renewal.

Sox17-inducible, *Sox17* mutant, Nanog-overexpressing, and wild-type ES cell controls were plated at the same density, and the total cell number was compared at 2, 4, 6, 8, and 10 d following treatment with or without doxycycline. At 6 d, doxycycline-treated control and *Sox17*-induced cells both resulted in a significant reduction in cell number compared with the untreated control (Fig. 7A). At 8 and 10 d, the *Sox17*-induced cells maintained a significant reduction in cell number compared with both the doxycycline-treated and untreated controls (Fig. 7A). Reciprocally, *Sox17*-deficient ES cells exhibited a significant increase in self-renewal compared with control ES cells (Fig. 7A).

Sox17, Nanog, Oct4, and Sox2 share a number of significant overlapping targets

The ability of *Sox17* to inhibit genes involved in ES cell self-renewal and induce genes required for ExEn differentiation suggests that Sox17 might compete with Sox2, Nanog, and Oct4 for regulation of target genes within either or both of these pathways. We sought to investigate the potential for the competition of Sox17 and any of the core pluripotency factors by comparing the Sox17-bound site we identified by ChIP–chip in XEN and ES cells to those published previously for Sox2, Nanog, and Oct4 (Chen et al. 2008; Kim et al. 2008). Indeed, we found a number of Sox17 targets were also bound by Sox2, Nanog, and Oct4 (Fig. 7C; Supplemental Fig. S15; Supplemental Tables S8, S9), and that the overlap in targets was significantly greater than expected by chance (Supplemental Table S10). Of particular note, we found that the Sox17-bound regulatory regions we identified upstream of *Lama1*, *Col4a1*, *Ap2m1*, and *Vitrin* were also bound by Sox2 and Nanog, suggesting that Sox17 might function by disengaging the repression of Sox2 or Nanog at these sites (Fig. 7B,C; Supplemental Fig. S15; Supplemental Tables S8, S9).

If Sox17 competes with either Nanog or Sox2 for the same DNA-binding sites, then we would predict that these transcription factors might share a consensus

DNA-binding sequence. Indeed, it had been predicted previously that Sry homeobox-containing genes like Sox2 and Sox17 should share an identical consensus-binding motif (Kamachi et al. 2000). We investigated whether Sox17 was interacting with a conserved sequence in stem cells by performing de novo motif analysis using the top 300 Sox17-binding sites. We found that Sox17 was generally binding to sites containing the consensus motif of 5′-ATTGT-3′ (Fig. 4C; Supplemental Table S11). Interestingly, this motif was similar to the Sox2 portion of the Oct4–Sox2 composite motif that has been reported previously (Ji et al. 2006). The Sox17-binding sites that we experimentally identified in *Col4a1*, *Col4a2*, *Lama1*, *Lamb1-1*, *Gata4*, *Sall4*, *Ap2m1*, *Vitrin*, *Ifnk*, *Slc38a4*, *Nanog*, and *Sox2* all contained this motif (Fig. 4C). This further suggests that Sox17 and Sox2 could be acting antagonistically at these shared targets by either directly competing for these target sequences or reciprocally influencing the binding of Nanog or Oct4 at these locations.

To directly test this model, we performed ChIP analysis using Sox2 and Nanog antibodies on mouse ES cells both before and after doxycycline induction of *Sox17*. Following ChIP, we performed quantitative PCR using primers that flank binding sites shared by Sox17, Nanog, and Sox2. If Sox17 was interfering with either Sox2 or Nanog binding at these shared sites, then we would predict a reduction in enrichment of one of these pluripotency factors following *Sox17* induction. Although we did not observe a reduction in Sox2 enrichment at these sites (Supplemental Fig. S16), we found that *Sox17* induction did interfere with binding of Nanog at its Sox17 shared binding sites (Fig. 7B; Supplemental Tables S8, S9). Nanog binding at *Lama1*, *Ppm1b*, and *Vitrin* was reduced fivefold in ES cells following 48 h of *Sox17* induction ($P < 0.01$) (Fig. 7B). Nanog enrichment was also reduced significantly at the shared Sox17-binding site in the regulatory region of the *Sox2* gene ($P < 0.01$) (Fig. 7B). Notably, and consistent with continued expression of *Oct4* in the primitive endoderm, Nanog enrichment at the Oct4 regulatory region was not affected significantly by *Sox17* induction (Fig. 7B). These results suggest that Sox17 acts to promote differentiation and inhibit self-renewal at least in part by disengaging Nanog from a subset of its transcriptional targets.

Discussion

Here we present evidence demonstrating that *Sox17* is an integral component of a transcriptional network that drives differentiation of pluripotent cells toward the ExEn. Our findings raise a significant note of caution for the use of *Sox17* as a marker of definitive endoderm, since it is strongly expressed in extraembryonic cells, which are readily generated by mouse ES cells. Our findings in mouse stem cells contrast with studies that suggest that overexpression of *SOX17* induced differentiation of human ES cells toward a definitive endoderm phenotype (Seguin et al. 2008). This difference in the response of mouse and human ES cells to *Sox17*/*SOX17* could be explained by the recent observation that human ES cells seem more reminiscent of cells within the epiblast than cells within the ICM (Brons et al. 2007; Tesar et al. 2007).

It is thought that symmetry of self-renewal within the ICM and within ES cell cultures is first broken by Fgf signaling through the Grb2–kinase pathway, which in turn activates *Gata6* expression (Cheng et al. 1998; Chazaud et al. 2006). Once Gata6 protein accumulates, it either directly or indirectly triggers expression of *Gata4* and *Sox17* (Fig. 7D). Our data suggest that the Gata4 and Sox17 proteins then act to consolidate this transcriptional network for differentiation by creating a positive feed-forward circuit that can amplify the levels of the Gata6, Gata4, and Sox17 transcription factors (Fig. 7D). We provided evidence that genetic interactions with the remainder of the transcriptional network for differentiation can be mediated by direct binding of the Sox17 protein, accompanied by a positive effect on transcription. This transcriptional feed-forward loop may also explain the cell fate commitment of differentiating ES cells that express Sox17. Once ES cells accumulate sufficient quantities of Sox17 protein, it is no longer possible for them to disengage this differentiation program.

The preimplantation expression of Sox17 and its function in stem cells implies that its loss may have consequences for XEN development in vivo. Notably, we failed to observe any significant effects on ExEn development in *Sox17*-deficient embryos. We speculate that the lack of a phenotype in vivo relates to the highly regulative environment of the mouse embryo, and therefore reflects compensatory changes in gene expression networks controlling development. This model is supported by the observation that both *Gata4* and *Sox7* continue to be expressed in *Sox17*-deficient embryos, which also maintain expression of important extraembryonic factors such as *Afp* and *Hnf4* (Kanai-Azuma et al. 2002; Shimoda et al. 2007). This hypothesis would be best addressed by generating compound mutant mouse embryos.

Our genetic analysis in stem cells indicates that *Sox17* is both necessary and sufficient to directly activate the transcription of a wide variety of genes known to function in basement membrane establishment. It may be that Gata6 and Gata4 collaborate to activate these targets, as loss of either *Gata6* or *Gata4* has been demonstrated to influence their expression (Capo-Chichi et al. 2005). However, it has not been investigated whether there are direct interactions between the Gata transcription factors and regulatory regions of these target genes. Therefore, it is a formal possibility that Gata6 and Gata4 act primarily to amplify the expression of *Sox17*, which is then the primary player in activating the effecter genes for differentiation. Genetic analysis demonstrating that Laminin and Collagen IV can become expressed even in the absence of *Gata4* support this model.

The observation that Sox17 interacts with many of the same promoter sequences to which Sox2, Nanog, and Oct4 also bind suggested that the transcription factors compete for the same DNA targets, providing a potential molecular mechanism by which the expression of Sox17 promotes differentiation at the expense of self-renewal

(Fig. 7E). Our demonstration that forced overexpression of Sox17 in mouse ES cells is sufficient to displace Nanog from its binding sites supports this model. Thus, Sox17 seems to induce differentiation of ES cells and inhibit self-renewal by antagonizing the action of Nanog (Fig. 7E). As we observed that there is compensatory redundancy in the function of ExEn-specific transcription factors in vivo, it will be important to determine the extent to which Sox17, Gata4, Gata6, and Nanog compete for DNA-binding sites within the blastocyst.

In conclusion, these findings suggest that Sox17 regulates lineage commitment in several distinct cell types at different stages of development. These observations, in turn, raise an important question to be addressed in the future. When Sox17 is subsequently expressed in other cell types, such as the definitive endoderm and fetal blood (Kanai-Azuma et al. 2002; Kim et al. 2007), does it bind to and activate distinct sets of genes, or does it recycle the same transcriptional targets to specify these additional cell types? If Sox17 regulates similar sets of genes in these dissimilar cell types, it may shed light on how an entire transcriptional network can be recycled during development to accomplish distinct tasks in the construction of an embryo.

Materials and methods

All experiments were performed in accordance with the Harvard University International Animal Care and Use Committee guidelines for humane care and use of animals.

XEN cell derivation

XEN cells were generated as described previously (Kunath et al. 2005). Blastocysts were treated with acidic tyrode's solution (Chemicon) to remove the zona pellucida, and were then plated onto γ-irradiated mouse embryonic feeder layers in TS cell derivation conditions. After 7–14 d, XEN cells were passaged in XEN media. XEN media consists of 85% KO-DMEM (Invitrogen), 15% fetal bovine serum (Hyclone), 1% L-glutamine (Invitrogen), 1% β-mercaptoethanol (Invitrogen), and 1% penicillin/streptomycin.

Sox17-inducible mouse ES cell lines

Sox17-inducible ES cell lines were generated using a modification of a previously described protocol (Hochedlinger et al. 2005; Beard et al. 2006). *Sox17* cDNA (Harvard Institute of Proteomics) was PCR-amplified and ligated to a 3XFlag sequence (Sigma), then cloned into a flip-in vector consisting of a frt-flanked neomycin-selectable marker and a promoterless, ATG-less hygromycin resistance gene, downstream from a tet-operon. The *Sox17*-flip-in vector was electroporated with 50 μg of *Sox17*-flip-in vector and 25μg of *pCAAGS*-FLPe-puro into 1×10^7 KH2 ES cells at 500 V and 25 μF using Gene PulserII (Bio-Rad). KH2 ES cells allow for the expression of the M2-rtTA tetracycline-responsive transactivator under the control of the *ROSA26* promoter. To induce *Sox17* expression, doxycycline was diluted in ES cell culture media to a final concentration of 1 mg/mL.

Embryo injections

E3.5 B6D2F2 blastocysts were flushed using M2 media (Specialty Media). Manipulations were done on a heated stage (37°C) in HEPES-buffered CZB (HZCB) (81 mM NaCl, 5 mM KCl, 1.18 mM $MgSO_4$, 1.17 mM KH_2PO_4, 0.1 mM EDTA, 5.5 mM glucose, 0.1 mg/mL polyvinyl alcohol, 0.53% lactic acid) under mineral oil. A hydraulic micromanipulator (Narishige) was used on an inverted microscope (Nikon TE200). A piezo micromanipulator (PrimeTech) was used for drilling the zona pellucida using a 17-μm flat-tipped pipette (Humagen).

Embryo transfer

Blastocysts were injected into day 2.5 pseudopregnant ICR females mated with vasectomized males as described previously (Nagy 2003). Embryos were dissected in PBS with 0.5% FBS (Hyclone) at E6–E6.5. Embryos were incubated in 1% Hoechst for 10 min and placed on a coverslip bottom dish (MatTec) for imaging using a confocal microscope (Zeiss LSM 510 META).

ChIP–chip

Sox17-inducible ES cells were treated with 1 μg/mL doxycycline for 48 h prior to harvesting. XEN cells were grown for 48 h prior to harvesting. Immunoprecipitation was performed on $\sim 1 \times 10^7$ to 2×10^7 cells as described previously (Vokes et al. 2007). Nine micrograms of goat anti-Sox17 (R&D), rabbit anti-Nanog (Cosmo-Bio), goat anti-Sox2 (Santa Cruz Biotechnologies), and mouse anti-M2 Flag (Sigma) antibodies were used for ChIP. ChIP samples were amplified by ligation-mediated PCR as described previously (Vokes et al. 2007). A total of three biological replicates versus input samples were performed for the XEN sample. Of each sample, 7.5 μg were hybridized to Affymetrix promoter 1.0R arrays, MOE430 A and B (Affymetrix). Detailed micorarray analysis can be found in the Supplemental Material. The raw data for the array may be obtained from Gene Expression Omnibus (GSE19026).

Acknowledgments

We thank Jennifer Couget, Jiangwen Zhang, and Kathryn Koszka for technical assistance. We are grateful to Stephen Duncan for *Gata4*-, *Gata6*-, and *Gata4*/*Gata6*-null ES cells; Sean Morrison for *Sox17*-GFP reporter ES cells; Austin Smith for *Nanog*-overexpressing ES cells; Janet Rossant for XEN and trophectoderm cell lines; and Rudolf Jaenisch for KH2 ES cells and constructs. We thank the Harvard Institute of Proteomics for the *Sox17*-cDNA clone. We thank Ian Henderson and William Anderson for helpful discussions and critical review of the manuscript, and members of K.E.'s and D.A.M.'s laboratories for helpful discussion. We acknowledge support from Harvard Stem Cell Institute and Stowers Medical Institute (to K.E.); Howard Hughes Medical Institute (to K.E. and D.A.M.); and National Institutes of Health grants HD045732-03 (to K.E.) and HG003903 (to A.P.M.). K.K.N. is supported by a Centre for Trophoblast Research Next Generation Fellowship. H.J. is supported by the Faculty Innovation Fund of Johns Hopkins Bloomberg School of Public Health. S.A.V. is supported by the Helen Hay Whitney Foundation and Charles A. King Trust, Bank of America, Co-Trustee. K.T.R. is supported by the National Science Foundation. K.T.R. and R.I.S. are Sternlicht Fellows of the Harvard Stem Cell Institute. A.E.C. is a fellow of the Jane Coffin Childs Memorial Fund for Medical Research. K.E. is a fellow of the John D. and Katherine T. MacArthur Foundation.

References

Arman E, Haffner-Krausz R, Gorivodsky M, Lonai P. 1999. Fgfr2 is required for limb outgrowth and lung-branching morphogenesis. *Proc Natl Acad Sci* **96:** 11895–11899.

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. 2003. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes & Dev* **17:** 126–140.

Beard C, Hochedlinger K, Plath K, Wutz A, Jaenisch R. 2006. Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis* **44:** 23–28.

Boiani M, Scholer HR. 2005. Regulatory networks in embryo-derived pluripotent stem cells. *Nat Rev Mol Cell Biol* **6:** 872–884.

Borowiak M, Maehr R, Chen S, Chen AE, Tang W, Fox JL, Schreiber SL, Melton DA. 2009. Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell Stem Cell* **4:** 348–358.

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al. 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122:** 947–956.

Boyer LA, Mathur D, Jaenisch R. 2006. Molecular control of pluripotency. *Curr Opin Genet Dev* **16:** 455–462.

Brons IG, Smithers LE, Trotter MW, Rugg-Gunn P, Sun B, Chuva de Sousa Lopes SM, Howlett SK, Clarkson A, Ahrlund-Richter L, Pedersen RA, et al. 2007. Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* **448:** 191–195.

Capo-Chichi CD, Rula ME, Smedberg JL, Vanderveer L, Parmacek MS, Morrisey EE, Godwin AK, Xu XX. 2005. Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells. *Dev Biol* **286:** 574–586.

Chambers I, Colby D, Robertson M, Nichols J, Lee S, Tweedie S, Smith A. 2003. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* **113:** 643–655.

Chazaud C, Yamanaka Y, Pawson T, Rossant J. 2006. Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. *Dev Cell* **10:** 615–624.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, et al. 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* **133:** 1106–1117.

Cheng AM, Saxton TM, Sakai R, Kulkarni S, Mbamalu G, Vogel W, Tortorice CG, Cardiff RD, Cross JC, Muller WJ, et al. 1998. Mammalian Grb2 regulates multiple steps in embryonic development and malignant transformation. *Cell* **95:** 793–803.

Coucouvanis E, Martin GR. 1995. Signals for death and survival: A two-step mechanism for cavitation in the vertebrate embryo. *Cell* **83:** 279–287.

Evans MJ, Kaufman MH. 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* **292:** 154–156.

Fujikura J, Yamato E, Yonemura S, Hosoda K, Masui S, Nakao K, Miyazaki Ji J, Niwa H. 2002. Differentiation of embryonic stem cells is induced by GATA factors. *Genes & Dev* **16:** 784–789.

Hochedlinger K, Yamada Y, Beard C, Jaenisch R. 2005. Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell* **121:** 465–477.

Ji H, Vokes SA, Wong WH. 2006. A comparative analysis of genome-wide chromatin immunoprecipitation data for mammalian transcription factors. *Nucleic Acids Res* **34:** e146. doi: 10.1093/nar/gkl803.

Kamachi Y, Uchikawa M, Kondoh H. 2000. Pairing SOX off: With partners in the regulation of embryonic development. *Trends Genet* **16:** 182–187.

Kanai-Azuma M, Kanai Y, Gad JM, Tajima Y, Taya C, Kurohmaru M, Sanai Y, Yonekawa H, Yazaki K, Tam PP, et al. 2002. Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* **129:** 2367–2379.

Kim I, Saunders TL, Morrison SJ. 2007. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. *Cell* **130:** 470–483.

Kim J, Chu J, Shen X, Wang J, Orkin SH. 2008. An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* **132:** 1049–1061.

Kunath T, Arnaud D, Uy GD, Okamoto I, Chureau C, Yamanaka Y, Heard E, Gardner RL, Avner P, Rossant J. 2005. Imprinted X-inactivation in extra-embryonic endoderm cell lines from mouse blastocysts. *Development* **132:** 1649–1661.

Lim CY, Tam WL, Zhang J, Ang HS, Jia H, Lipovich L, Ng HH, Wei CL, Sung WK, Robson P, et al. 2008. Sall4 regulates distinct transcription circuitries in different blastocyst-derived stem cell lineages. *Cell Stem Cell* **3:** 543–554.

Manabe R, Tsutsui K, Yamada T, Kimura M, Nakano I, Shimono C, Sanzen N, Furutani Y, Fukuda T, Oguri Y, et al. 2008. Transcriptome-based systematic identification of extracellular matrix proteins. *Proc Natl Acad Sci* **105:** 12849–12854.

Martin GR. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci* **78:** 7634–7638.

Matsui T, Kanai-Azuma M, Hara K, Matoba S, Hiramatsu R, Kawakami H, Kurohmaru M, Koopman P, Kanai Y. 2006. Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. *J Cell Sci* **119:** 3513–3526.

Miner JH, Yurchenco PD. 2004. Laminin functions in tissue morphogenesis. *Annu Rev Cell Dev Biol* **20:** 255–284.

Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. 2003. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **113:** 631–642.

Mitsunari T, Nakatsu F, Shioda N, Love PE, Grinberg A, Bonifacino JS, Ohno H. 2005. Clathrin adaptor AP-2 is essential for early embryonal development. *Mol Cell Biol* **25:** 9318–9323.

Morrisey EE, Tang Z, Sigrist K, Lu MM, Jiang F, Ip HS, Parmacek MS. 1998. GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes & Dev* **12:** 3579–3590.

Nagy A. 2003. *Manipulating the mouse embryo: A laboratory manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Nichols J, Zevnik B, Anastassiadis K, Niwa H, Klewe-Nebenius D, Chambers I, Scholer H, Smith A. 1998. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **95:** 379–391.

Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis AK. 2008. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development* **135:** 3081–3091.

Seguin CA, Draper JS, Nagy A, Rossant J. 2008. Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* **3:** 182–195.

Shimoda M, Kanai-Azuma M, Hara K, Miyazaki S, Kanai Y, Monden M, Miyazaki J. 2007. Sox17 plays a substantial role in late-stage differentiation of the extraembryonic endoderm in vitro. *J Cell Sci* **120:** 3859–3869.

Shimosato D, Shiki M, Niwa H. 2007. Extra-embryonic endoderm cells derived from ES cells induced by GATA factors acquire the character of XEN cells. *BMC Dev Biol* **7:** 80.

Silva J, Chambers I, Pollard S, Smith A. 2006. Nanog promotes transfer of pluripotency after cell fusion. *Nature* **441:** 997–1001.

Soudais C, Bielinska M, Heikinheimo M, MacArthur CA, Narita N, Saffitz JE, Simon MC, Leiden JM, Wilson DB. 1995. Targeted mutagenesis of the transcription factor GATA-4 gene in mouse embryonic stem cells disrupts visceral endoderm differentiation in vitro. *Development* **121:** 3877–3888.

Tanaka S, Kunath T, Hadjantonakis AK, Nagy A, Rossant J. 1998. Promotion of trophoblast stem cell proliferation by FGF4. *Science* **282:** 2072–2075.

Tesar PJ, Chenoweth JG, Brook FA, Davies TJ, Evans EP, Mack DL, Gardner RL, McKay RD. 2007. New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448:** 196–199.

Vokes SA, Ji H, McCuine S, Tenzen T, Giles S, Zhong S, Longabaugh WJ, Davidson EH, Wong WH, McMahon AP. 2007. Genomic characterization of Gli-activator targets in sonic hedgehog-mediated neural patterning. *Development* **134:** 1977–1989.

Wang S, Cowan CA, Chipperfield H, Powers RD. 2005. Gene expression in the preimplantation embryo: In-vitro developmental changes. *Reprod Biomed Online* **10:** 607–616.

Watt AJ, Battle MA, Li J, Duncan SA. 2004. GATA4 is essential for formation of the proepicardium and regulates cardiogenesis. *Proc Natl Acad Sci* **101:** 12573–12578.

Yamanaka Y, Ralston A, Stephenson RO, Rossant J. 2006. Cell and molecular regulation of the mouse blastocyst. *Dev Dyn* **235:** 2301–2314.

Zhao R, Watt AJ, Battle MA, Li J, Bondow BJ, Duncan SA. 2008. Loss of both GATA4 and GATA6 blocks cardiac myocyte differentiation and results in acardia in mice. *Dev Biol* **317:** 614–619.