Please cite this article in press as: Lim et al., Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages, Cell Stem Cell (2008), doi:10.1016/j.stem.2008.08.004

Cell Stem Cell
Article

# Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages

Chin Yan Lim,[1,7] Wai-Leong Tam,[1,7,*] Jinqiu Zhang,[1,7] Haw Siang Ang,[4] Hui Jia,[2,8] Leonard Lipovich,[2,8] Huck-Hui Ng,[1,5] Chia-Lin Wei,[3] Wing Kin Sung,[2] Paul Robson,[1,5] Henry Yang,[4] and Bing Lim[1,6,*]
[1]Stem Cell and Developmental Biology
[2]Information and Mathematical Sciences Group
[3]Genome Technology and Biology Group
Genome Institute of Singapore, 138672 Singapore
[4]Bioinformatics Institute, 138671 Singapore
[5]Department of Biological Sciences, National University of Singapore, 117543 Singapore
[6]Harvard Institutes of Medicine, Harvard Medical School, Boston, MA 02115, USA
[7]These authors contributed equally to this work
[8]Present address: Center for Molecular Medicine and Genetics, School of Medicine, Wayne State University, Detroit, MI 48201, USA
*Correspondence: tamwl@gis.a-star.edu.sg (W.-L.T.), limb1@gis.a-star.edu.sg (B.L.)
DOI 10.1016/j.stem.2008.08.004

## SUMMARY

**Stem cells self-renew or differentiate under the governance of a stem-cell-specific transcriptional program, with each transcription factor orchestrating the activities of a particular set of genes. Here we demonstrate that a single transcription factor is able to regulate distinct core circuitries in two different blastocyst-derived stem cell lines, embryonic stem cells (ESCs) and extraembryonic endoderm (XEN) cells. The transcription factor Sall4 is required for early embryonic development and for ESC pluripotency. Sall4 is also expressed in XEN cells, and depletion of Sall4 disrupts self-renewal and induces differentiation. Genome-wide analysis reveals that Sall4 is regulating different gene sets in ESCs and XEN cells, and depletion of Sall4 targets in the respective cell types induces differentiation. With Oct4, Sox2, and Nanog, Sall4 forms a crucial interconnected autoregulatory network in ESCs. In XEN cells, Sall4 regulates the key XEN lineage-associated genes Gata4, Gata6, Sox7, and Sox17. Our findings demonstrate how Sall4 functions as an essential stemness factor for two different stem cell lines.**

## INTRODUCTION

The development of the preimplantation embryo and lineage commitment of early blastocyst-derived stem cells is highly regulated. From the mouse blastocyst, three stem cell types can be isolated and cultured in vitro. These include embryonic stem cells (ESCs) from the inner cell mass (ICM), trophoblast stem (TS) cells from the trophectoderm, and stem-cell-like extraembryonic endoderm (XEN; hereafter referred to as stem cells) cells from the primitive endoderm (Evans and Kaufman, 1981; Martin, 1981; Rossant, 2007; Yamanaka et al., 2006). When reintroduced into the embryo, these cells contribute specifically to the appropriate embryonic cell lineages (Beddington and Robertson, 1989; Kunath et al., 2005; Tanaka et al., 1998). The maintenance of these undifferentiated stem cell lines is dependent upon certain key factors that specify lineage identity, as well as support self-renewal.

During the course of development, these key transcription factors are present to establish specific networks in a spatial and temporal manner to facilitate cell-fate decisions and the establishment of different lineages in the early blastocyst. The first cell-fate decision is undertaken by cells in the morula to give rise to the trophectoderm and ICM. Cells of the ICM further undertake a second cell-fate decision to form either the epiblast, which produces the embryonic germ layers or the primitive endoderm, which forms the extraembryonic yolk sac (Rossant, 2007; Yamanaka et al., 2006). During this period, the ICM cells have started to express the epiblast and primitive endoderm determinants in a mutually exclusive manner, even though *Oct4* expression is maintained in both populations. *Nanog*, a key ESC-associated factor expressed from the morula stage, becomes localized to the epiblast layer of the blastocyst (Chambers et al., 2003; Mitsui et al., 2003). *Nanog* has been suggested to act as a selector of the epiblast fate over the primitive endoderm fate, as ESCs depleted of *Nanog* acquire endoderm-like characteristics coupled with the expression of endoderm-associated genes (Mitsui et al., 2003). *Gata6*, on the other hand, is a determinant of the primitive endoderm lineage. Mutations in *Gata6* lead to defective development of the visceral endoderm in the mouse embryo. However, the *Gata6* mutants do not exhibit extraembryonic endoderm defects until a few days after blastocyst formation (Koutsourakis et al., 1999; Morrisey et al., 1998). This strongly suggests that other factors act in the initial stage of primitive endoderm specification.

Sall4 belongs to the *Spalt-like* (*sall*) family of C2H2 zinc-finger transcription factors. The first member of this family, *Spalt*, was identified as an essential homeotic gene required for early development in *Drosophila* (Frei et al., 1988; Kuhnlein et al., 1994). *Sall4* has been shown to be involved in the proper development of the ICM, where its depletion in the early zygote causes an expansion of trophectoderm cells into the ICM of the blastocyst, and *Sall4* null mutants perish by E6.5 (Elling et al., 2006;

Please cite this article in press as: Lim et al., Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages, Cell Stem Cell (2008), doi:10.1016/j.stem.2008.08.004



Sakaki-Yumoto et al., 2006; Zhang et al., 2006). Presumably, Sall4 is required for a transcription program that maintains pluripotency of the epiblast within the ICM. In ESCs, a loss of *Sall4* leads to rapid loss of pluripotency and respecification into the trophoblast lineage (Zhang et al., 2006). These effects of Sall4 depletion could be attributed partly to the loss of Sall4-mediated transcriptional regulation of the developmentally important *Oct4* gene (Zhang et al., 2006).

Interestingly, neither ESCs nor XEN cells can be derived from *Sall4* null mutant blastocysts, suggesting that Sall4 could also be important for maintaining the primitive endoderm lineage (Elling et al., 2006; Sakaki-Yumoto et al., 2006). This led us to hypothesize that Sall4 might be central to the core transcription circuit in the primitive endoderm and its associated XEN cells. In this study, we delineated the core circuitries that are regulated by Sall4 in governing the distinctive blastocyst-derived stem cell lines. We dissected crucial components of the Sall4-driven circuitries through loss-of-function analyses and showed that these are required for maintaining the stem cell states. In doing so, we addressed the question of how a single transcription factor, Sall4, regulates distinctive networks in the maintenance of two developmentally divergent stem cell lineages.

## RESULTS

### Sall4 Is Required for the Maintenance of Blastocyst-Derived XEN Cells

In the absence of Sall4, normal development of the epiblast and primitive endoderm is disrupted in the early embryo, suggesting an essential role for Sall4 in these lineages. In ESCs, Sall4 is required for maintaining pluripotency and self-renewal, in part through its role in regulating *Oct4* expression (Zhang et al., 2006). However, its role in the primitive endoderm is unclear.

To facilitate studies of Sall4 function in the primitive endoderm, we derived a XEN cell line, designated as XEN-GIS, from E4.5 blastocysts using methods previously described (Kunath et al., 2005). The cultures of XEN-GIS cells contained mixed populations of cells with two distinctive rounded or epithelioid morphologies that were also observed in previously reported XEN cell cultures (Kunath et al., 2005). Examination of several known XEN marker genes, including Gata4, Gata6, and Sox17, by immunofluorescence analysis showed that these transcription factors are expressed and localized to the nuclei of all cells, regardless of cell morphology (see Figure S1 available online). We then examined the global gene expression profile of the XEN-GIS cell line by microarray analysis and compared that with three previously characterized XEN lines, XEN1-3, IM8A1-I, and IM8A1-II (Kunath et al., 2005). The expression profiles of four ESC lines and four MEF cell lines were also included in the comparison. Hierarchical clustering of the transcriptome profiles of these cell lines showed that the XEN-GIS cells clearly clustered with the established XEN cell lines, and the global transcriptome profile of the XEN-GIS cell line was very similar to that of previously characterized XEN cell lines (Figures S2A and S2B). Closer examination of specific parietal (PE) and visceral (VE) endoderm marker genes in these newly derived XEN-GIS cells revealed high levels of expression of known PE markers *Plat*, *Sparc*, and *Fst*; VE markers *Ihh* and *Cited1*/*Msg1*; and other extraembryonic endoderm genes such as *Tcf2*/*vHnf1*, *Vegfa*, and *Stra6* (Table S1). Consistent with cells of the extraembryonic endoderm lineage, genes specific to ESCs (*Oct4*, *Nanog*, *Zfp42*), trophectoderm (*Cdx2*, *Eomes*, and *Hand1*), mesoderm (*T-Brachyury* and *MyoD*) and ectoderm (*Sox1*, *Otx2*, and *Pax6*) were absent in the XEN-GIS cells (Table S1; data not shown). Thus, these results indicate that the transcriptome of XEN-GIS cells is similar to and corresponds well with that of previously characterized XEN cells (Kunath et al., 2005).

Next, we sought to examine the differentiation potential of the XEN-GIS cells. It has been reported that in vitro differentiation of XEN cells led to decreased levels of *Gata4*, *Gata6*, and *Sox7*, while the VE marker gene *Afp* was induced (Kunath et al., 2005). To investigate if our XEN-GIS cells exhibited similar differentiation potential as the established XEN lines, we cultured the cells in differentiation conditions for 4 days, followed by quantitative RT-PCR (qRT-PCR) analysis of marker genes to detect gene expression changes. Similar to previous reports, differentiation of the XEN-GIS cells led to strong induction of *Afp*, along with decreases in the levels of markers of PE (*Gata6*, *Gata4*, *Sox7*, *Sox17*, and *Pthr1*) and VE (*Ihh* and *Foxa2*) lineages (Figure S2C). Taken together, these results indicate that the XEN-GIS cells are highly similar to previously characterized XEN lines and are representative of the extraembryonic endoderm lineage. We will henceforth refer to these XEN-GIS cells as XEN cells.

We next examined the expression of Sall4 in these XEN cells and observed that Sall4 was expressed and localized to the nuclei (Figure 1A). Analysis of mRNA levels showed that *Sall4* was expressed at lower levels in XEN cells relative to ESCs, whereas no transcripts were detected in embryonic or adult fibroblasts (Figure 1B). As Sall4 levels appear abundant in XEN cells, we sought to further investigate if Sall4 has critical functions in these cells. To determine the requirement of Sall4 in XEN cells, we depleted the levels of Sall4 by RNAi. Upon *Sall4* depletion, the cells underwent morphological changes where the normally rounded, highly refractile cells adopted a large vacuolated morphology that was phenotypically similar to previously described in vitro differentiated XEN cells (Kunath et al., 2005). We also observed a concomitant reduction in the number of viable cells (Figures 1C and 1D). Analysis of annexin V-positive cells by flow cytometry indicated a marked increase in apoptosis within XEN cells treated with Sall4 shRNA (~17%) compared to control (~9%) (Figure 1E). Furthermore, qRT-PCR analysis of genes associated with the extraembryonic endoderm lineage in Sall4-shRNA-treated XEN cells showed that *Gata6*, *Gata4*, *Sox7*, and *Sox17* were downregulated by 30%–60% compared to control shRNA-treated cells (Figure 1F).

Further characterization of the differentiation induced in the Sall4-shRNA-treated XEN cells using microarray transcriptome profiling provided results consistent with the qRT-PCR analysis, in which expression of the extraembryonic endoderm markers *Gata4*, *Gata6*, *Sox7*, and *Sox17* were significantly downregulated (Table S2). In addition, markers of the PE lineage *Pdgfra*, *Sparc*, and *Pthr1* and VE lineage *Cited1*, *Ihh*, and *Acvr1* were also significantly reduced. The similarities between these Sall4 RNAi-induced gene expression changes and those observed in the in vitro differentiated XEN cells (Figure S2C; Kunath et al., 2005) led us to conclude that Sall4 is required for maintaining the undifferentiated state of XEN cells. This finding is also

Please cite this article in press as: Lim et al., Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages, Cell Stem Cell (2008), doi:10.1016/j.stem.2008.08.004



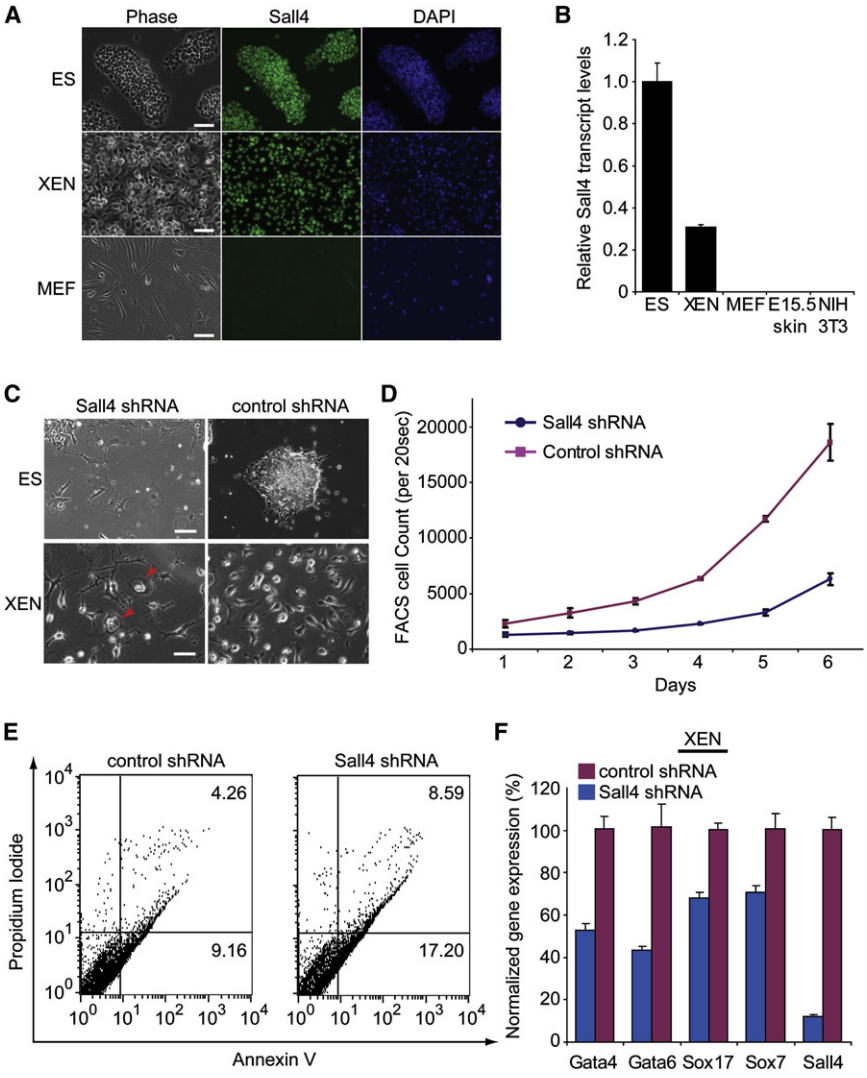


**Figure 1. Sall4 Is Required for Maintaining Undifferentiated XEN Cells**

(A) Expression of Sall4 in ESCs and XEN and MEF cells is analyzed by immunostaining. Scale bars, 200 μm.

(B) *Sall4* is expressed in ESCs and XEN cells but absent in other cell types. *Sall4* transcript levels in ESCs and XEN, MEF, embryonic skin (E15.5), and NIH 3T3 cells were measured by qRT-PCR. All values were normalized to β-actin and plotted relative to the expression levels in ESCs. Error bars indicate standard error of three replicates.

(C) Phase images of ESCs and XEN cells 4 days after transfection with Sall4 or control shRNA vectors. Sall4 knockdown ESCs differentiated into a flattened giant trophoblast cell-like morphology, while XEN cells depleted of Sall4 adopted large vacuolated cell morphology (indicated with arrows). Scale bars, 100 μm.

(D) FACS analysis to determine the growth of XEN cells after transfection with Sall4 or control shRNA vectors. The transfected cells were cultured under puromycin selection for 3 days prior to reseeding at the density of 10 000 cells per well. Proliferation of the treated XEN cells was determined by FACS counting of cell numbers from 1 to 6 days. Error bars indicate standard error of two technical replicates at each time point.

(E) Apoptotic cell death (lower right quadrant) in XEN cells transfected with Sall4 or control shRNA vectors was measured by Annexin V staining 4 days after transfection.

(F) Quantitative RT-PCR analysis of marker gene expression in XEN cells treated with Sall4 or control shRNA for 4 days. All values were normalized to β-actin and plotted relative to the expression levels in control shRNA-treated cells. Error bars indicate standard error of three replicates.

consistent with the observation that extraembryonic endoderm development is disrupted in *Sall4*$^{-/-}$ mice.

## Lineage-Specific Sall4 Occupancy of Target Genes

We then sought to understand how a single transcription factor, Sall4, which is essential in maintaining the two lineage-distinct but developmentally connected blastocyst-derived ESC and XEN cell populations, might exert its function in the two stem cell lines. To do so, we sought to identify the target genes regulated by Sall4 in both ESCs and XEN cells by genome-wide promoter occupancy analysis. We performed chromatin immunoprecipitation (ChIP) followed by hybridization to DNA microarrays containing probes that interrogate the region from 5.5 kb upstream to 2.5 kb downstream of the transcription start sites of ~17,000 mouse genes. Biological triplicates were examined for each cell line, and promoters bound by Sall4 were determined by statistical analyses that assessed factor enrichment at multiple neighboring probes for each gene across replicate microarrays. In both ESCs and XEN cells, Sall4 was found to bind ~1000 putative target genes each, with the false-positive rate estimated to be ~5%–6% based on ChIP-qPCR analyses (Figure 2A; Tables S3 and S4; Figures S3A–S3C). Comparison of Sall4-bound regions in the two cell lines delineated 814 ESC-specific and 751 XEN cell-specific binding sites, while only ~25% of Sall4-bound regions were common to both cell types (Figure 2A). For example, the promoters of *Sox2*, *Phc1*, and *Dppa3* exhibited ESC-specific Sall4-binding profiles, while *Gata6*, *Gata4*, and *Pdgfra* showed XEN-specific Sall4 occupancy (Figure 2B). In addition, of the promoters bound by Sall4 in both ESCs and XEN cells, the majority of these were occupied at the same specific location within the 8 kb region interrogated. However, it is interesting to note that there were instances of cell-specific binding sites on a small number of these promoters such as those observed on the *Sall4* promoter (Figure 2B).

The surprisingly large proportion (>75%) of Sall4-bound genes that are cell type specific strongly suggest that Sall4 participates in the regulation of distinct transcriptional programs in ESCs and XEN cells. In particular, we found Sall4 to be present on the promoters of many lineage-associated genes in a cell-specific manner. For example, in a mutually exclusive manner, Sall4 occupied the promoters of *Oct4*, *Sox2*, and *Tcf3* in ESCs, but that of *Gata4*, *Gata6*, *Sox17*, and other XEN-associated genes in XEN cells

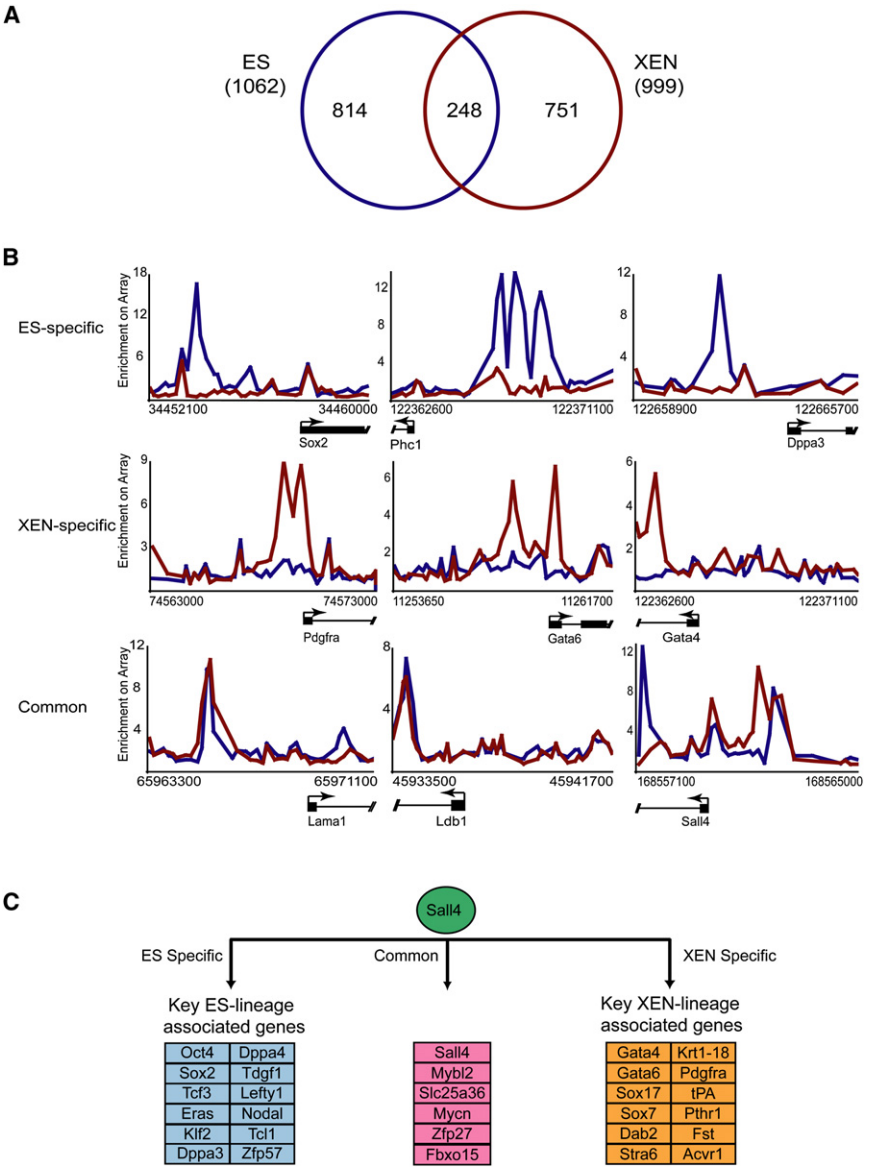


**Figure 2. ChIP-Chip Analysis of Genome-wide Promoter Occupancy Reveals that Sall4 Regulates Specific and Common Target Genes in ESCs and XEN Cells**

(A) Diagram showing overlap of genes bound by Sall4 in ESCs and XEN cells. Numbers in parentheses indicate total number of genes bound in each cell type.

(B) Chromosomal location and enrichment ratio of Sall4 for all probes of each indicated gene. Blue graph indicate Sall4 enrichment profile in ESCs. Red graph represent Sall4 enrichment profile in XEN cells.

(C) Examples of selected Sall4 target genes that are cell type specific as well as common to both ESCs and XEN.

(Figure 2C). These data suggest that Sall4 may regulate the expression of critical lineage-specifying factors in the two cell types.

To gain insights into how Sall4 may be recruited to different promoters in the two blastocyst-derived cell lines, we investigated the histone modification features that are present at Sall4-occupied regions. In a global analysis, we examined the distribution of H3K4me3 and H3K27me3 modifications at Sall4-bound regions in ESCs by intersecting Sall4-bound loci identified in our ChIP-chip studies with the recently described epigenome maps of these two histone marks (Mikkelsen et al., 2007). We found that >66% of Sall4-bound regions bear the H3K4me3 mark, while 14% contained both H3K4me3 and H3K27me3 modifications (Figure 3A). The observation that Sall4-bound regions in ESCs are enriched for the activating H3K4me3 mark indicates an association of Sall4 with nonrepressed genes in ESCs. In contrast, intersection of the histone marks present in ESCs at XEN-specific Sall4-binding loci showed significantly less enrichment of the H3K4me3 marks at these regions (Figure 3A). Instead, more than 60% of these regions displayed chromatin features frequently associated with gene repression, such as presence of the H3K27me3 mark or the absence of both H3K4me3 and H3K27me3 modifications. These results suggest that, in ESCs, Sall4 is predominantly present at the promoters of active target genes but absent at repressed promoters.

We then sought to examine if Sall4 occupancy is similarly correlated with target gene activity in XEN cells. We performed ChIP-qPCR analysis to investigate these histone modifications on the Sall4-binding sites of selected XEN-specific Sall4 target genes, as well as ESC-specific and common genes, in these cells. In XEN cells, we observed that the repressive H3K27me3 mark was highly enriched at genes such as *Nanog*, *Oct4*, *Sox2*, and *Dppa5* where Sall4 was absent (Figure 3B). In contrast, H3K27me3 was absent at Sall4-bound target genes such as *Gata6*, *Gata4*, and *Sox17*, which carried the active H3K4me3 mark instead. These results suggest that Sall4 promoter occupancy in XEN cells is also closely correlated with gene activity. As a comparison, we performed the ChIP-qPCR analysis in ESCs and observed significant enrichment of H3K27me3 at all the repressed XEN-specific Sall4 target genes (Figure 3B), consistent with the findings of the genome-wide correlation analysis. Taken together, these results indicated that Sall4 occupancy in each cell type is correlated with active histone modification signatures.

## Sall4 Regulates Differential Transcriptional Programs in ESCs and XEN Cells

To examine how Sall4 may transcriptionally regulate target gene expression in ESCs and XEN cells, we profiled perturbations to the global gene expression profiles induced by the loss of *Sall4*. We then examined how many of the genes that exhibited

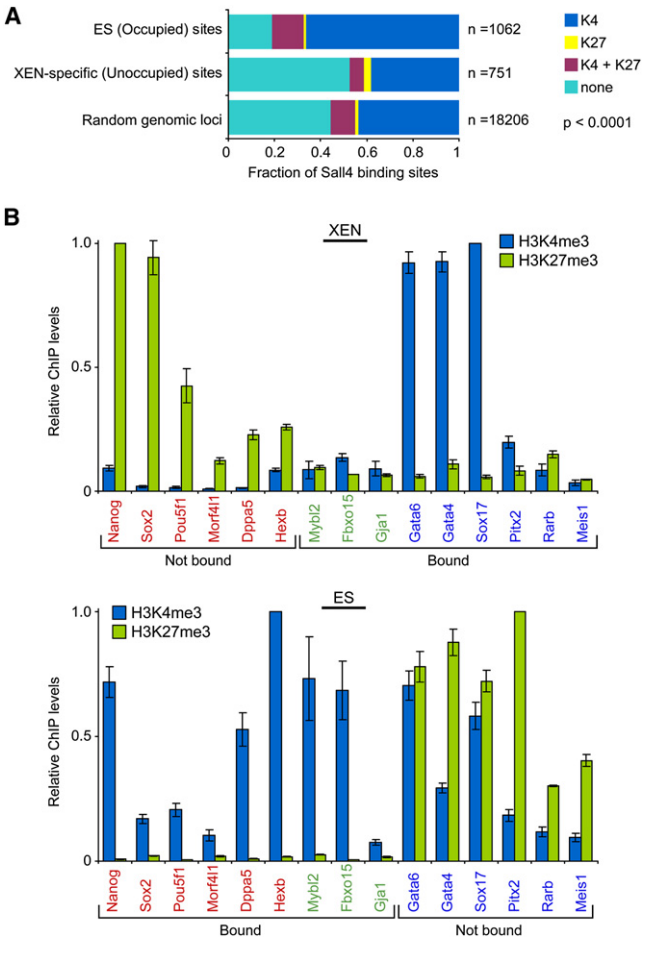


**Figure 3. Lineage-Specific Sall4 Promoter Occupancy Is Correlated with Specific Histone Modifications**

(A) Correlation of H3K4me3 and H3K27me3 histone modifications with Sall4 occupancy in ESCs was examined by intersecting epigenome maps (Mikkelsen et al., 2007) with Sall4-binding sites. Sall4-bound regions in ESCs are denoted as "Occupied." Sall4-occupied regions uniquely identified in XEN cells are denoted as "Unoccupied." Correlation of occupied and unoccupied sites with different chromatin states were found to be statistically significant at $p < 0.0001$ (z-test for two proportions), compared to a control set of randomly selected promoter regions.

(B) Levels of H3K4me3 and H3K27me3 on ESC-specific (red labels), XEN-specific (blue labels), and common (green labels) Sall4-binding sites in ESCs and XEN cells were examined by ChIP-qPCR. For comparison of different modifications that give different enrichment values, the data are plotted relative to the maximum enrichment value in each cell line for each ChIP. In ESCs, H3K4me3 and H3K27me3 levels for each locus were normalized to that for Hexb and Pitx2, respectively. In XEN cells, H3K4me3 and H3K27me3 levels for each locus were normalized to that for Nanog and Sox17, respectively. Error bars indicate the SEM of three independent experiments.

>1.75-fold change in transcript levels upon *Sall4* depletion had promoters that were occupied by Sall4. In ESCs, ~15% of the downregulated and ~9% of the upregulated genes were Sall4-bound targets, suggesting that these genes are potentially activated or repressed transcriptionally by Sall4 in the undifferentiated state (Figure 4A; Table S5). In XEN cells, Sall4 appears to regulate the expression of ~15% of the total number of genes significantly up- or downregulated upon the loss of *Sall4* (Figure 4A; Table S6). Therefore, these data indicate that Sall4 functionally regulates a subset of the genes whose expressions were altered in a Sall4-dependent manner.

To further evaluate the correlation between Sall4 promoter occupancy and transcriptional control of its targets, we analyzed the expression profiles of Sall4-bound genes upon *Sall4* knockdown in ESCs and XEN cells. In *Sall4*-depleted ESCs, expression of more than 60% of Sall4-bound target genes exhibited changes of >1.25-fold compared to control (Figure 4B). These results suggest that Sall4 directly regulates the expression of 60% of its target genes in ESCs. Consistent with the finding that most of these genes are ESC-specific Sall4 targets, the majority of these genes did not show similar expression changes in *Sall4*-knockdown XEN cells (Figure 4B). In XEN cells, ~75% of Sall4 target genes exhibited expression changes upon the loss of Sall4 (Figure 4B). As most of these genes were XEN-specific targets of Sall4, their expression were not altered in a similar manner in ESCs. Thus, these results suggest that Sall4 is recruited to, and appears to regulate, the expression of distinct sets of genes in a cell-type-specific manner.

To demonstrate that Sall4 directly controls the expression of its target genes, we performed reporter assays on 15 target genes, which were downregulated upon *Sall4* depletion in ESCs or XEN cells. Promoter regions encompassing the Sall4-binding sites of these genes were cloned upstream of a luciferase reporter and cotransfected into HEK293T cells with either a Sall4 overexpression plasmid or an empty vector. In these experiments, all promoters tested were activated by Sall4 compared to vector control (Figure S4). These results illustrate the ability of Sall4 to directly activate the expression of its target genes. The observation that certain genes were upregulated upon *Sall4* depletion from our microarray analyses suggests that Sall4 could function to repress these target genes in ESCs or XEN cells. However, we did not observe Sall4-mediated repression on three of these genes tested in the reporter assays (Figure S4). This could suggest that transcriptional repression mediated by Sall4 in ESCs and XEN cells may involve more complex mechanisms that are likely to require the presence of additional factors, which may be absent in HEK293T cells.

## Sall4 Is an Integral Component of Core Circuitries in Stem Cells

To gather additional insights into the comparative functional roles of Sall4 in ESCs and XEN cells, we first performed Gene Ontology (GO) analysis on the sets of ESC-specific, XEN-specific, and shared target genes bound by Sall4. Consistent with its role as a stem cell transcription factor, all three sets of target genes were highly enriched for genes involved in developmental processes and transcriptional regulation (Table S7).

The finding that Sall4 is recruited to distinct groups of genes in ESCs and XEN cells suggests that Sall4 governs important transcription programs required for the establishment or maintenance of the two distinct stem cell lineages. In ESCs, the contribution of Sall4 is critical, as evident through its regulation of *Oct4*, and interactions with Nanog to co-occupy and potentially coregulate a number of downstream targets, suggesting that Sall4-regulated target genes could potentially feed into the Oct4/Sox2/Nanog circuitry. To evaluate the connectivity of Sall4, Oct4, Sox2, and Nanog in the ESC transcription network,

Please cite this article in press as: Lim et al., Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages, Cell Stem Cell (2008), doi:10.1016/j.stem.2008.08.004



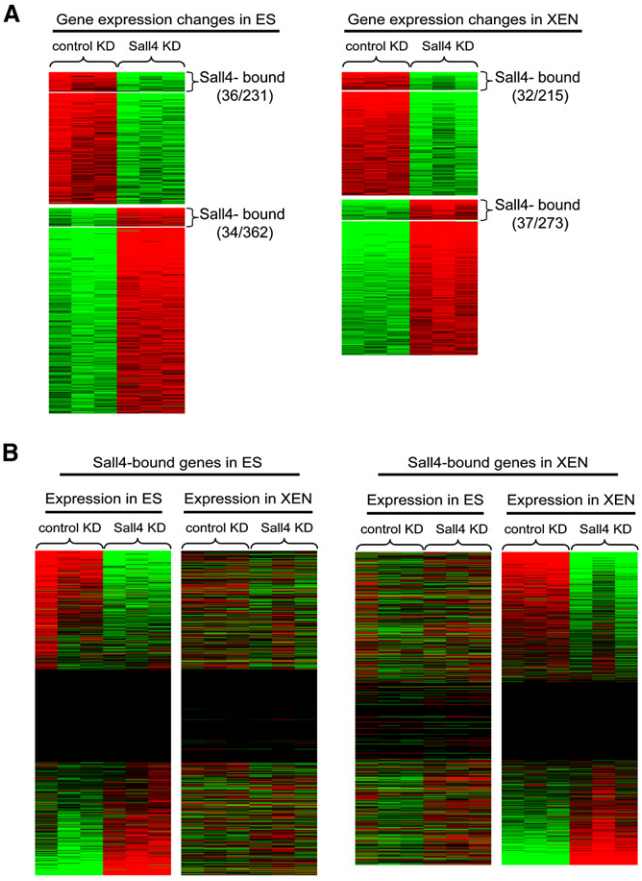


**Figure 4. Sall4 Transcriptionally Regulates a Subset of Bound Target Genes**

(A) Differential gene expression profile upon depletion of Sall4 in ESCs and XEN cells. Triplicate expression profiling was performed on ESCs and XEN cells transfected with Sall4 or control shRNA and cultured under puromycin selection for 4 (ESC) or 5 (XEN) days. KD, knockdown. Differential gene expression was calculated based on the mean of the $\log_2$ fold difference of Sall4 knockdown in ESCs or XEN cells to the corresponding controls. Heatmaps represent genes with greater than 1.75-fold difference sorted based on mean fold differences.

(B) Gene expression changes of all Sall4-bound genes were calculated based on the mean of the $\log_2$ fold difference of Sall4 knockdown in ESCs or XEN cells to the corresponding controls. Heatmaps were generated using hierarchical clustering with average linkage, and genes were sorted based on the mean fold changes.

we intersected the target genes of Sall4 with those of Oct4, Sox2, and Nanog identified in a genome-wide ChIP-Seq study (Chen et al., 2008). We found that Sall4 co-occupies the promoters of many genes common to Oct4, Sox2, and Nanog, including a set of 305 genes that were co-occupied by all four factors (Table S8). Interestingly, we observed that Sall4 participates in an interconnected autoregulatory circuit with Oct4, Sox2, and Nanog where each of the four factors appears to regulate its own expression as well as that of the others (Figure 5A). We further examined the expression profiles of the 305 common target genes in undifferentiated and differentiating ESCs (Hirst et al., 2006; Palmqvist et al., 2005). We found that >60% of these genes were differentially expressed in differentiating compared to undifferentiated ESCs (Figure 5B). These results indicate that the genes co-occupied by Sall4, Oct4, Sox2, and Nanog exhibit ESC-specific expression, which suggests that these genes may be functionally important for maintaining the undifferentiated ESCs. Thus, Sall4 is an integral part of the transcriptional network in ESCs and is involved in the coregulation of functionally important target genes, together with Oct4, Sox2, and Nanog.

We then sought to extend the repertoire of functionally relevant genes in the ESC circuitry by examining the roles of some Sall4 target genes in ESCs. We performed RNAi knockdown of three selected Sall4 target genes, *Slc25a36*, *Zfp27*, and *Ezh2*, whose expression were downregulated upon depletion of Sall4 in ESCs but had not been previously implicated in maintaining pluripotency of ESCs. Depletion of *Slc25a36* by >50% led to the differentiation of ESCs with the appearance of giant trophoblast-like cells when cultured in ESC growth conditions, partially phenocopying *Sall4* knockdown (Figures 5C and S5A). When *Slc25a36* knockdown in ESCs was performed in culture conditions that support the growth of trophoblast stem cells, cells that immunostained positive for Cdx2 were observed (Figure S6). Furthermore, *Slc25a36* depletion in ESCs led to the downregulation of pluripotency markers *Oct4*, *Sox2*, and *Utf1*, while trophoblast-associated markers *Hand1*, *Cdx2*, *Eomes*, *Esx1*, and *Psx1* were uniformly upregulated (Figure 5D). In contrast, consistent changes in the levels of marker genes associated with the endoderm, ectoderm, and mesoderm lineages were not observed (Figure S5A). Thus the loss of *Slc25a36* in ESCs resulted in a phenotype that partially resembled the loss of *Sall4*, suggesting that *Slc25a36* is an important Sall4 target gene that may function to maintain the undifferentiated state of ESCs.

RNAi knockdown of *Ezh2*, a member of the PRC2 polycomb repressor complex, and *Zfp27*, a zinc-finger protein, resulted in less distinct morphological changes in cells cultured in ESC media (Figure 5C). However, we also observed the emergence of Cdx2-expressing cells in the *Ezh2* and *Zfp27* knockdowns that were cultured in TS culture conditions (Figure S6). Lineage marker gene analysis in *Zfp27* knockdown ESCs revealed significant upregulation of the trophectoderm-associated markers *Hand1*, *Cdx2*, *Eomes*, *Esx1*, and *Psx1*, while other lineage-associated markers were generally unchanged or did not exhibit consistent patterns (Figures 5D and S5B). In *Ezh2* knockdown cells, we observed substantial increases in the expression of genes associated with the trophectoderm, mesoderm, endoderm, and ectoderm lineages (Figures 5D and S5C). These observations are consistent with previous reports that ESCs deficient in *Suz12* and *Eed*, which are components of the PRC2 complex, aberrantly expressed a variety of lineage-specific markers (Boyer et al., 2005; Chamberlain et al., 2008; Pasini et al., 2007). Hence, we have uncovered additional Sall4-regulated genes that have functional roles in the maintenance of the ESC identity, providing further evidence that the transcriptional program controlled by Sall4 is an essential component of the ESC circuitry.

In contrast to the extensive knowledge about key transcriptional networks governing ESCs, the core genetic circuitry regulating the extraembryonic endoderm and its XEN cell derivative is much less understood. Using Sall4 as the starting point, we sought to map the core circuitry of critical factors in XEN cells. From our studies, we observed that Sall4 directly activates many extraembryonic endoderm lineage-associated genes,

Please cite this article in press as: Lim et al., Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages, Cell Stem Cell (2008), doi:10.1016/j.stem.2008.08.004



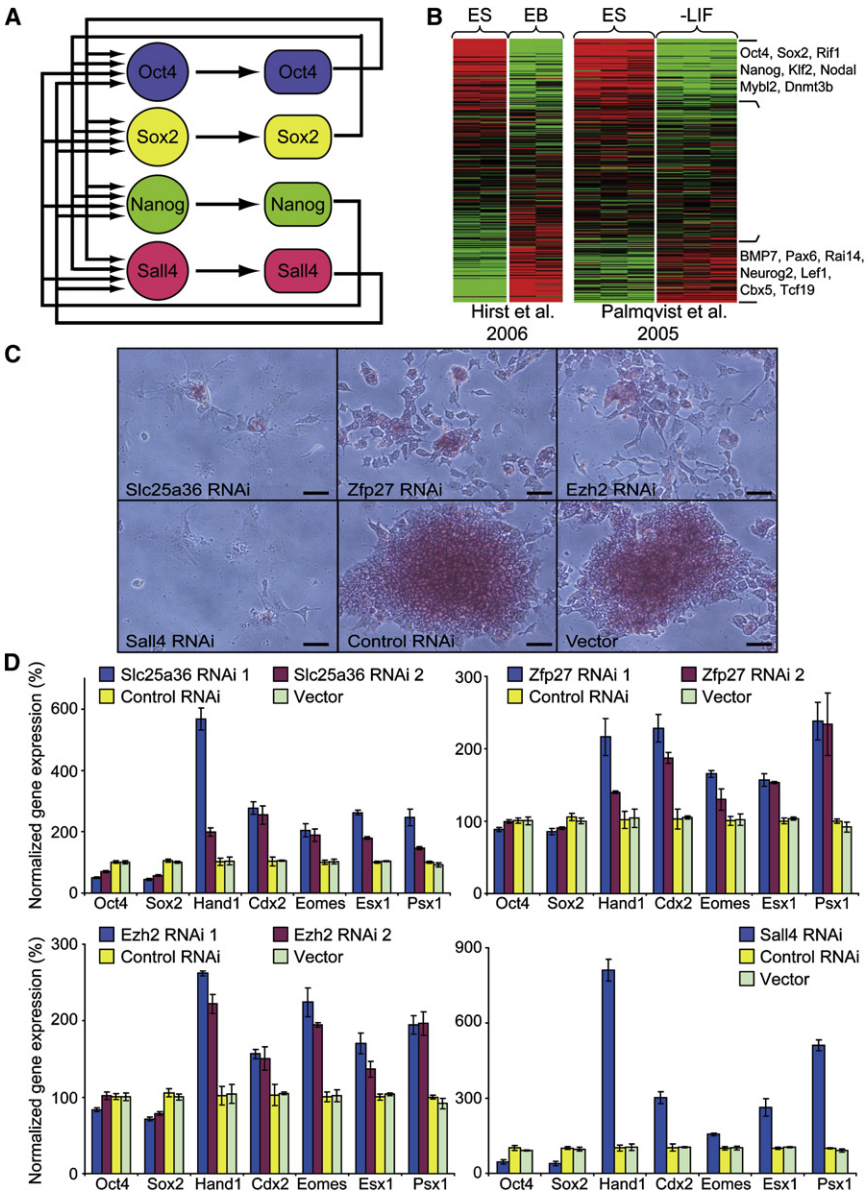


**Figure 5. Sall4 Is an Integral Part of the ESC Transcription Network and Regulates Functionally Important Genes**

(A) Circuit diagram showing the interconnected autoregulatory, positive feedback network formed by Sall4, Oct4, Sox2, and Nanog.

(B) Heatmap representation of the expression profiles of the 305 common target genes in undifferentiated and differentiating ESCs.

(C) Alkaline phosphatase staining of ESCs transfected with *Slc25a36*, *Zfp27*, *Ezh2*, or control shRNA and cultured for 4 days under puromycin selection. Scale bars, 100 μm.

(D) Quantitative RT-PCR analysis of marker gene expression in ESCs 4 days after transfection with gene-specific shRNA knockdown or control vector. All values were normalized to levels of β-actin and plotted relative to the expression levels in control shRNA-treated cells. Error bars indicate standard error of three replicates.

such as *Gata6*, *Sox7*, and *Sox17*. Although Gata6 has been implicated in primitive endoderm development during embryogenesis, its corresponding function in XEN cells has not been tested. Hence, we sought to determine if these lineage-associated genes are required for maintaining the undifferentiated state of XEN cells. When we depleted the levels of *Gata6*, *Sox17*, and *Sox7* individually, by RNAi, we observed morphological differentiation of the shRNA-treated XEN cells, which resembled the phenotype of *Sall4* knockdown XEN cells (Figure 6A). These morphological changes were accompanied by the downregulation of the extraembryonic endoderm lineage-associated genes by 20%–50% (Figure 6B). When we reduced the levels of *Gata6*, *Sox17*, and *Sox7* simultaneously, we observed very rapid induction of cellular differentiation and a marked reduction in cell numbers. The transcript levels of *Gata4* and *Sall4* were also drastically decreased by >70% relative to the controls. These results confirm that these downstream targets of Sall4 are critical factors for XEN cell function, as the loss of one or more of these factors is detrimental to XEN cells and could potentially affect primitive endoderm formation in the embryo. Interestingly, we observed that depletion of each of these factors led to significant downregulation of all four extraembryonic endoderm lineage-associated genes, *Gata4*, *Gata6*, *Sox17*, and *Sox7* as well as *Sall4*. This suggests that Gata6, Sox17, Sox7, and Sall4 might participate in an autoregulatory feedback circuit in XEN cells, reminiscent of the interconnected regulatory circuit of master regulators present in ESCs.

## DISCUSSION

The molecular mediators and mechanism for specification of germ layers and extraembryonic lineages during early embryonic development are an important area that is poorly understood. With the emergence of genome-wide transcription factor localization analyses, precise and detailed data can be obtained to reveal comprehensive views of the molecular circuitry controlling cell-fate maintenance and commitment, best exemplified in several recent studies of ESCs. In particular, much effort has been focused on how different transcription factors act singly or collectively to regulate extensive networks in ESCs (Boyer et al., 2005; Chen et al., 2008; Kim et al., 2008; Loh et al., 2006). However, there is virtually no information on how a single transcription factor might behave in different stem cells. This is largely due to the paucity of identified factors that are present in more than one stem cell type, and yet retain distinctive functional relevance in each cell lineage (Fortunel et al., 2003).

Here, we have identified Sall4 to be an essential key regulator in two blastocyst-derived stem cell lines that are developmentally connected. Mapping of Sall4 targets is timely, as recent

Please cite this article in press as: Lim et al., Sall4 Regulates Distinct Transcription Circuitries in Different Blastocyst-Derived Stem Cell Lineages, Cell Stem Cell (2008), doi:10.1016/j.stem.2008.08.004



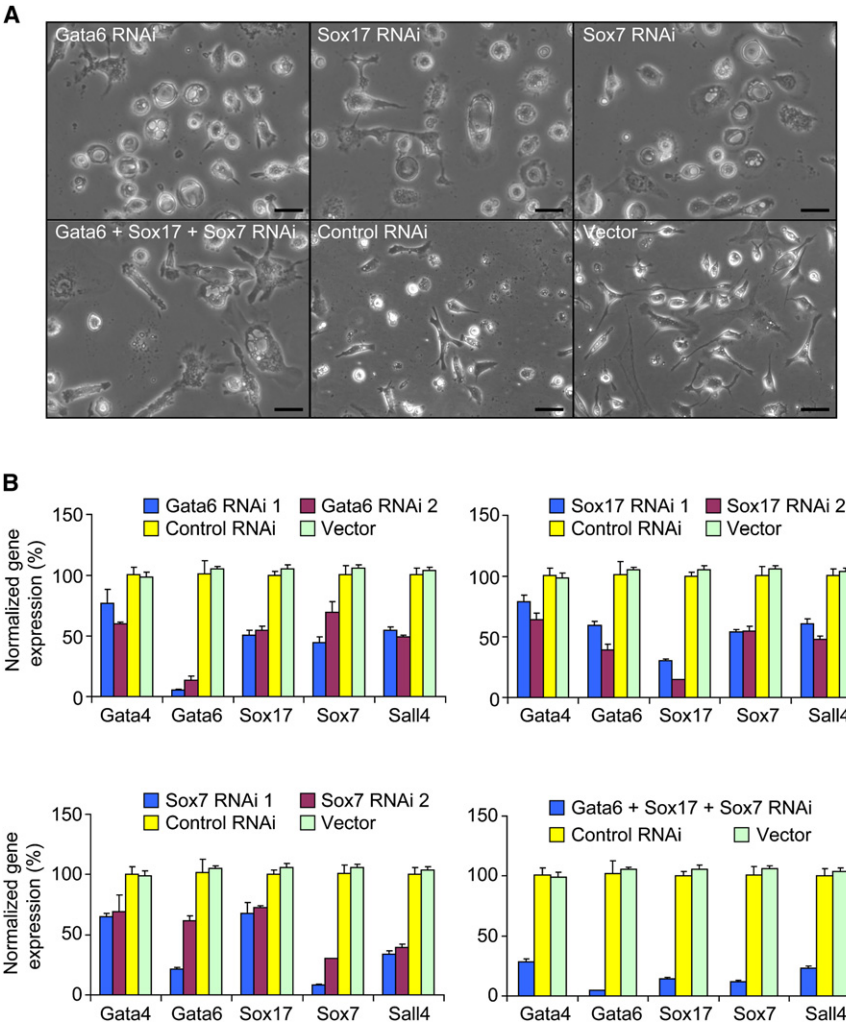


**Figure 6. Sall4 Regulates Core Circuitry Critical for Maintaining the Undifferentiated State in XEN Cells**

(A) Phase images of XEN cells treated for 4 days with Gata6, Sox17, and Sox7 shRNAs, either singly or in combination, in the presence of puromycin drug selection. Scale bars, 100 μm.

(B) Quantitative RT-PCR analysis of marker gene expression in XEN cells treated with gene-specific shRNA knockdown or control vector for 4 days. All values were normalized to levels of β-actin and plotted relative to the expression levels in control shRNA-treated cells. Error bars indicate standard error of three replicates.

mapping of multiple transcription factors has also identified Sall4 as a major node in ESC pluripotency circuitry, revealed by multifactor occupancy on the *Sall4* promoter (Chen et al., 2008; Kim et al., 2008). Our studies reveal that, in ESCs, Sall4 regulates many transcription factors critical for maintaining pluripotency and self-renewal (Figure 2C; Table S2) (Ivanova et al., 2006; Loh et al., 2006; Tam et al., 2008), as well as members of key signaling pathways such as the Wnt, Fgf, TGFβ, and MAPK pathways. Sall4 is also recruited to the promoters of *Klf2* and *Klf5*, which have been reported to have overlapping functions with *Klf4* in ESC self-renewal and somatic cell reprogramming (Jiang et al., 2008; Nakagawa et al., 2008). In addition, we have also functionally validated the contribution of additional Sall4 target genes to the maintenance of ESC identity. Thus, the integration of Sall4 into the core transcriptional network in ESCs provides an extended transcriptional circuitry that is essential for the pluripotency of ESCs. The finding that Sall4 participates in an interconnected autoregulatory circuit with Oct4, Sox2, and Nanog further suggests that Sall4 functions as an important master regulator of the ESC state.

Our conclusion that Sall4 is integral to the ESC transcriptional circuitry is strengthened by our recent ChIP-sequencing analysis of Sall4 occupancy in ESCs. Intersection of Sall4-bound loci with the ~3500 multiple transcription factor-binding loci (MTL), which are occupied by at least four key transcription factors in ESCs (Chen et al., 2008), showed that Sall4 is also present at ~1350 of the MTL (Table S9). Clustering analysis of Sall4 with the reported MTL revealed that it is most frequently colocalized with Oct4, Sox2, Nanog, and Smad1 (Figure S7A). The strong association of Sall4 with the Nanog-Oct4-Sox2 MTL cluster suggests that Sall4 is a component of the ES-cell-specific enhanceosome that drives specific gene expression programs required for maintaining pluripotency (Chen et al., 2008). With this deep Sall4 ChIP-Seq data, we attempted to define the Sall4-binding motif in ESCs. De novo sequence analysis revealed an enrichment of the sequence TCGCCATA in high-intensity Sall4-binding peaks (Figure S7B), which remains to be more definitively characterized as a bona fide Sall4 recognition motif.

In addition to defining the role of Sall4 in ESCs, we also demonstrated the importance of Sall4 function in XEN cells, which are extraembryonic endoderm-derived stem cells. We show here that the loss of *Sall4* leads to XEN cell differentiation, along with the extensive loss of expression of endoderm-associated factors. Our study reveals Sall4 as a crucial upstream activator of key lineage-defining genes in the extraembryonic endoderm. Gata6 has been identified as one of the earliest lineage-defining factors for the primitive endoderm and acts potentially as an antagonist to Nanog that is thought to specify the epiblast (Mitsui et al., 2003). However, it has been suggested that other factors are necessary for initial primitive endoderm specification, as *Gata6* mutants do not exhibit extraembryonic endoderm defects until several days after blastocyst formation. The discovery that Sall4 lies upstream of *Gata6* and *Gata4* and drives the expression of other lineage-determining genes provides a plausible explanation for why the loss of *Gata6* or *Gata4* alone does not lead to complete loss of primitive endoderm formation but only affects visceral endoderm differentiation at later stages (Koutsourakis et al., 1999; Morrisey et al., 1998; Ralston and Rossant, 2005; Soudais et al., 1995). In addition, though studies have

shown that sustained ectopic expression of GATA factors in ESCs could induce differentiation toward the extraembryonic endoderm lineage, giving rise to cells with XEN-like properties (Shimosato et al., 2007), our study is the first direct demonstration that Gata6, Sox7, and Sox17 are essential for the maintenance and propagation of the blastocyst-derived XEN cells. Given the importance of the Sall4-mediated transcriptional program in XEN cells, our data on Sall4 targets are informative and useful for future identification and characterization of additional genes involved in the progression of embryonic and extraembryonic development.

Our observation of a substantial induction of apoptosis upon the reduction of Sall4 in XEN cells suggests that Sall4 may also regulate apoptotic pathways. Indeed, we have observed in our GO analysis of Sall4 targets that Sall4 occupies the promoter of several proapoptotic genes, such as *Cul1*and *Stk17b*, as well as apoptosis-inhibiting genes, such as *Cish*, *Mybl2*, and *Bcl2l1*, whose expression levels are altered in a Sall4-dependent manner. This is worth noting in the light of recent reports showing that Sall4 functions as a regulator of cell survival in human acute promyelocytic leukemic cells (Yang et al., 2008) by targeting and regulating a wide range of genes in both proapoptotic and antiapoptotic pathways. These and our findings, therefore, implicate Sall4 as an important regulator not only of the stemness state and survival of several types of normal stem cells, but also the survival and expansion of cancer cells and possibly cancer stem cells.

The observation that a common transcription factor is differentially recruited to very distinct genes in different progenitor or stem cell types poses interesting mechanistic questions. It is possible that Sall4 recognizes and binds to different sequence motifs in the two cell types; however, the derivation of high-confidence Sall4 recognition sequences by de novo motif search of Sall4-bound regions proved elusive. One explanation could be attributed to a highly degenerate Sall4 recognition motif, which would make identification by de novo sequence analysis difficult. A second mechanism by which lineage-specific recruitment of Sall4 to different target genes could be mediated is via its interactions with other cell-type-specific cofactors. One such lineage-restricted cofactor is Nanog, which has been shown to directly interact and colocalize with Sall4 at many promoters in ESCs (Wu et al., 2006). Although specific cofactors of Sall4 have not been identified in XEN cells, the extraembryonic endoderm lineage factors such as Gata6, Sox7, and Sox17, which are regulated by Sall4, could be potential XEN-specific partners of Sall4. In addition, recent reports of how transcription factor recruitment can be modulated by epigenetic modifications to the chromatin (Guccione et al., 2006; Lupien et al., 2008) suggest a third mechanism by which cell-type-specific Sall4 promoter occupancy could be facilitated. It has been shown recently that lineage-specific recruitment of FoxA1 transcription factor to different target sites in LNCaP prostate cancer cells and MCF7 breast cancer cells was positively determined by the distribution of H3K4 methylation at the DNA recognition sites, which have the same sequence in both cell types (Lupien et al., 2008). Our findings that Sall4-occupied regions in ESCs and XEN cells are correlated with active histone marks, while promoters of target genes not bound by Sall4 are frequently associated with repressive H3K27me3 modifications may point to a role for the chromatin landscape in the lineage-restricted selection of Sall4-binding sites.

In conclusion, our study establishes the importance of Sall4 throughout early embryonic development (Figure 7A). As early as in the oocyte, Sall4 is present together with Oct4 and participates in specifying the first lineage commitment decision of forming either the ICM or the trophectoderm of the blastocyst. Subsequently, Sall4 interacts and regulates the key lineage-defining selectors, Gata6 and Nanog, which help specify the primitive endoderm from the epiblast, respectively. At the mechanistic level, Sall4 forms a key overarching component of the distinct core transcription circuitries present in the blastocyst-derived stem cell lines. Interestingly, it appears that, while the recruitment of Sall4 is needed, in association with other factors, to coordinate the expression of a subset of genes in establishing a specific lineage program in one cell type, it is equally important that the promoters of the same subset of genes become inaccessible to Sall4 in another stem cell type to ensure the complete absence of Sall4 regulatory input of those genes (Figure 7B). This study illustrates and provides basis for future investigations into how other similar factors could have unique roles in more than one precursor cell type and further our understanding on the genetic and epigenetic bases for self-renewal and lineage potency in stem cells.

## EXPERIMENTAL PROCEDURES

### Cell Culture and Transfection

All cell cultures were maintained at 37°C with 5% $CO_2$. The culture of mouse E14 ESCs was described previously (Zhang et al., 2006). XEN cells were derived from E4.5 blastocysts as described (Kunath et al., 2005) and maintained in RPMI1640 supplemented with 20% FBS, L-glutamine, β-mecaptoethanol, sodium pyruvate, and penicillin/streptomycin. In vitro differentiation of XEN cells was induced by culturing the cells at low density, in the absence of gelatin for 4 days. HEK293T cells were maintained in DMEM supplemented with 10% FBS and penicillin/streptomycin. Transfection of plasmids into mouse ESCs, XEN cells, and HEK293T cells was performed using Lipofectamine 2000 (Invitrogen).

### Cell Viability and Proliferation Assays by Flow Cytometry

XEN cells were transfected with control or Sall4 shRNA and selected with puromycin for 3 days prior to seeding into 6-well plates at a density of 10,000 cells per well. At various time points, cells in each well were trypsinized, harvested, and washed with 1% FBS-containing PBS and fixed in 90% ethanol. The cells were then washed with PBS followed by treatment with RNaseA and resuspended in 0.5 ml PBS containing 10 μg/ml propidium iodide. Cells were analyzed on a FACS Calibur for 20 s on high mode. Data were analyzed using the CellQuest program (BD Biosciences). The proportion of apoptotic cells in control or Sall4 shRNA-treated cells was determined using Annexin V-FITC Apoptosis Kit (BD Biosciences) by fluorescence-activated cell sorting (FACS) performed on a FACS Calibur (BD Biosciences).

### Plasmid Construction

For RNA interference (RNAi) design and construction of plasmids for shRNA synthesis, 19 base-pair gene-specific regions were designed based on the algorithm of Reynolds et al. (2004). Oligonucleotides were cloned into pSuper.puro (Oligoengine; see the Supplemental Data for sequences). At least two shRNAs were designed to target each gene for functional validation in ESCs and XEN cells. All sequences were analyzed by BLAST to ensure specificity.

### RNA Isolation, Reverse Transcription, and Real-Time PCR Analysis

Cells were rinsed twice in ice-cold PBS. Total RNA was extracted using Trizol (Invitrogen) and column purified with RNeasy kits (QIAGEN). cDNA synthesis

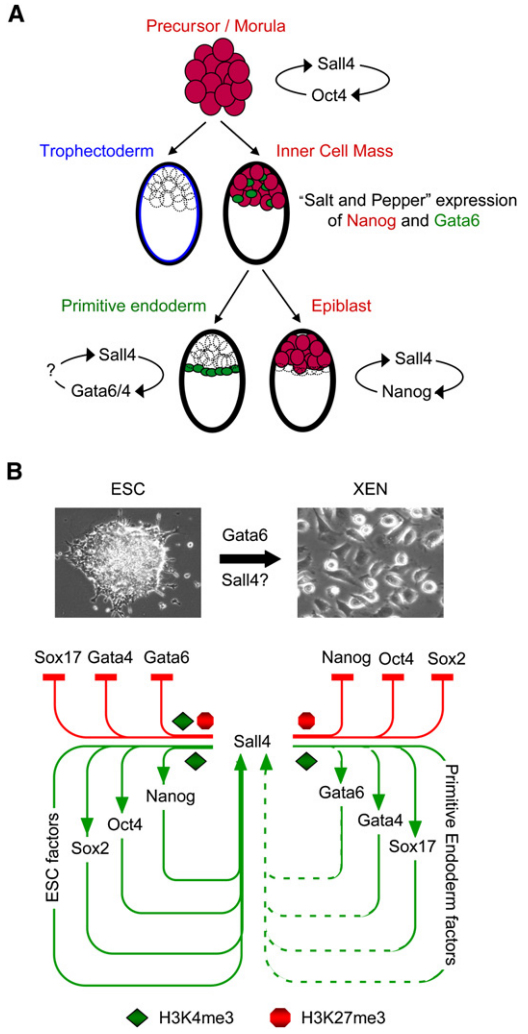


**Figure 7. Sall4 during Early Embryogenesis and Its Regulation of Distinct Core Circuitries**

(A) Sall4 is coexpressed with Oct4 from the one-cell zygote onward and acts as an activator of Oct4 for proper ICM formation (Hamatani et al., 2004; Zhang et al., 2006). In the ICM, two populations of cells mutually exclusive for Nanog and Gata6 expression can be observed interspersed in a "salt and pepper" pattern (Chazaud et al., 2006). By the late blastocyst stage, the primitive endoderm is clearly segregated from the epiblast. Sall4 regulates both the primitive endoderm selector Gata6 and the epiblast selector Nanog. Sall4 and Nanog forms an autoregulatory circuit in the epiblast, but the upstream regulation of Sall4 in the primitive endoderm is not known.

(B) Sall4 regulates distinct core circuitries in both blastocyst-derived stem cell lines. In ESCs, the pluripotency-associated factors Nanog, Oct4, and Sox2 are activated and feedback into Sall4, whereas extraembryonic the endoderm lineage-associated factors Gata6, Gata4, and Sox17 are silenced. In XEN cells, the reverse occurs, but it is unclear whether downstream Sall4 targets also regulate Sall4 directly. The maintenance of the respective undifferentiated cellular states is dependent on the proper control of lineage-determining factors, as forced expression of Gata6 can induce ESCs to form XEN-like cells. We have previously shown that forced overexpression of Sall4 in ESCs can increase endoderm markers (Zhang et al., 2006), but the importance of the precise levels of Sall4 in determining alternate cell fates remains to be tested. The activation of the circuits is, in part, due to H3K4me3 modification at the respective Sall4 targets in either cell type, whereas the H3K27me3 modification is largely associated with the repression of gene activity. Whether these epigenetic marks, and others, directly determine Sall4 recruitment to the appropriate target sites is not known.

was performed with 1 μg of total RNA at 37°C for 2 hr using the High Capacity cDNA Archive Kit (Applied Biosystems) and subsequently diluted ten times. Endogenous mRNA levels of pluripotency and differentiation markers were measured with inventoried Taqman probes using the ABI Prism 7900HT Sequence Detection System 2.2 (Applied Biosystems). Results were normalized to β-actin and analyzed using SDS 2.2 software.

### Chromatin Immunoprecipitation and DNA Microarray Analysis

ChIP assays were carried out as described previously (Zhang et al., 2006). Quantitative PCR analyses were performed in real-time using the ABI PRISM 7900 Sequence Detection System. Relative occupancy values were calculated by determining the apparent immunoprecipitation efficiency (ratios of the amount of immunoprecipitated DNA to that of the input sample) and normalized to the average levels at three control regions. ChIP combined with DNA microarray was carried out according to Agilent Mammalian ChIP-on-chip protocol (version 3) (Boyer et al., 2005). See the Supplemental Data for more detailed methodology. Antibodies used for ChIP experiments were Sall4 (Zhang et al., 2006), H3K4me3 (ab8580, Abcam), H3K27me3 (07-449, Upstate Biotechnology Inc.), and GST (sc-459, Santa Cruz Biotechnology Inc.).

### Immunofluorescence Microscopy

Cell cultures were fixed in 4% paraformaldehyde and permeabilized with 0.25% Triton X-100, followed by blocking with 1% bovine serum albumin in PBS. Cells were stained with primary antibodies, followed by the appropriate secondary antibodies conjugated with Alexa Fluor 594 (Molecular Probes). Anti-Gata6 (R&D, AF1700), anti-Gata4 (R&D, AF2606), anti-Sox7 (R&D, AF2766), and anti-Sox17 (R&D, AF1924) antibodies were used at 1:100 dilutions. Anti-Sall4 antibodies (Zhang et al., 2006) were used at 1:100 dilutions. Anti-Cdx2 (BioGenex, CDX2-88) antibodies were used at 1:200 dilutions. Images were captured with a fluorescence microscope (Axio Observer D1, Zeiss).

### Gene Expression Microarray and Data Analysis

For microarray transcriptome profiling of Sall4- and control shRNA-treated cells, cells were transfected with appropriate plasmids and cultured under puromycin selection for 3, 4, or 5 days. Cells were then harvested and washed with PBS. Total RNA was extracted using Trizol (Invitrogen) and column purified with RNeasy kits (QIAGEN). RNA amplification was performed using Ilumina TotalPrep RNA amplification kit. Expression profiling of coding genes was carried out using Illumina MouseRef-8 BeadArrays as per the manufacturer's instructions. All microarray data have been deposited with GEO repository accession number GSE12482. All data were subtracted from background intensities and were normalized using the crosscorrelation method. Differential expression was based on the mean of the $\log_2$ fold changes of Sall4 RNAi in XEN or ESCs to the corresponding scrambled controls. The cutoff mean fold change for differential expression was 1.75. All heatmaps were generated using hierarchical clustering with average linkage. In the heatmaps, $\log_2$-transformed intensities subtracted from the mean intensity were shown for each gene, and genes were sorted based on the mean fold changes. Hierarchical clustering of gene expression data was performed on different ESC lines, XEN cell lines, and MEF cell lines with two different array platforms (Illumina and Affymetrix). Prior to clustering, gene expression data were normalized using the crosscorrelation method separately for each platform (Chua et al., 2006). Normalized data were then subtracted from the mean values of XEN and ESCs and subsequently clustered based on samples. Gene Ontology analysis was performed based on biological process categories using a hypergeometric distribution. Enriched categories were obtained in comparison to the whole mouse genome. Common target genes of Oct4, Sox2, Nanog, and Sall4 were identified by colocalization of the binding loci in the neighborhood of a target gene. The colocalized binding loci of all these four factors were extracted and mapped to genes using the *C*is-regulatory *E*lement *A*nnotation *S*ystem (CEAS) (Ji et al., 2006) with the mouse genome assembly version Build 36 (mm8).

## ACCESSION NUMBERS

All microarray data have been deposited with GEO repository accession number GSE12482.

Cell Stem Cell

Sall4 Regulates Distinct Programs in Stem Cells

## SUPPLEMENTAL DATA

The Supplemental Data include Supplemental Experimental Procedures, seven figures, and seven tables and can be found with this article online at http://www.cellstemcell.com/cgi/content/full/3/5/■■■/DC1/.

## ACKNOWLEDGMENTS

We thank Leah Vardy for critical reading of the manuscript. We are grateful to colleagues in the Genome Institute of Singapore for helpful discussions and technical assistance. This work is supported by the Agency for Science, Technology and Research (Singapore) and Singapore Stem Cell Consortium grants (SSCC-06-003 and SSCC-06-007). The work is also partially supported by National Institutes of Health (NIH) grants to B.L. (DK47636 and AI54973).

Received: April 4, 2008
Revised: July 13, 2008
Accepted: August 15, 2008
Published online: September 18, 2008

## REFERENCES

Beddington, R.S., and Robertson, E.J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733–737.

Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell *122*, 947–956.

Chamberlain, S.J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. Stem Cells *26*, 1496–1505.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S., and 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., and 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, 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*, 1106–1117.

Chua, S.W., Vijayakumar, P., Nissom, P.M., Yam, C.Y., Wong, V.V., and Yang, H. (2006). A novel normalization method for effective removal of systematic variation in microarray data. Nucleic Acids Res. *34*, e38.

Elling, U., Klasen, C., Eisenberger, T., Anlag, K., and Treier, M. (2006). Murine inner cell mass-derived lineages depend on Sall4 function. Proc. Natl. Acad. Sci. USA *103*, 16319–16324.

Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154–156.

Fortunel, N.O., Otu, H.H., Ng, H.H., Chen, J., Mu, X., Chevassut, T., Li, X., Joseph, M., Bailey, C., Hatzfeld, J.A., et al. (2003). Comment on "'Stemness': transcriptional profiling of embryonic and adult stem cells" and "a stem cell molecular signature". Science *302*, 393.

Frei, E., Schuh, R., Baumgartner, S., Burri, M., Noll, M., Jurgens, G., Seifert, E., Nauber, U., and Jackle, H. (1988). Molecular characterization of spalt, a homeotic gene required for head and tail development in the Drosophila embryo. EMBO J. *7*, 197–204.

Guccione, E., Martinato, F., Finocchiaro, G., Luzi, L., Tizzoni, L., Dall' Olio, V., Zardo, G., Nervi, C., Bernard, L., and Amati, B. (2006). Myc-binding-site recognition in the human genome is determined by chromatin context. Nat. Cell Biol. *8*, 764–770.

Hamatani, T., Carter, M.G., Sharov, A.A., and Ko, M.S. (2004). Dynamics of global gene expression changes during mouse preimplantation development. Dev. Cell *6*, 117–131.

Hirst, C.E., Ng, E.S., Azzola, L., Voss, A.K., Thomas, T., Stanley, E.G., and Elefanty, A.G. (2006). Transcriptional profiling of mouse and human ES cells identifies SLAIN1, a novel stem cell gene. Dev. Biol. *293*, 90–103.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y., and Lemischka, I.R. (2006). Dissecting self-renewal in stem cells with RNA interference. Nature *442*, 533–538.

Ji, X., Li, W., Song, J., Wei, L., and Liu, X.S. (2006). CEAS: cis-regulatory element annotation system. Nucleic Acids Res. *34*, W551–W554.

Jiang, J., Chan, Y.S., Loh, Y.H., Cai, J., Tong, G.Q., Lim, C.A., Robson, P., Zhong, S., and Ng, H.H. (2008). A core Klf circuitry regulates self-renewal of embryonic stem cells. Nat. Cell Biol. *10*, 353–360.

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*, 1049–1061.

Koutsourakis, M., Langeveld, A., Patient, R., Beddington, R., and Grosveld, F. (1999). The transcription factor GATA6 is essential for early extraembryonic development. Development *126*, 723–732.

Kuhnlein, R.P., Frommer, G., Friedrich, M., Gonzalez-Gaitan, M., Weber, A., Wagner-Bernholz, J.F., Gehring, W.J., Jackle, H., and Schuh, R. (1994). spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the Drosophila embryo. EMBO J. *13*, 168–179.

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

Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., et al. (2006). The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. Nat. Genet. *38*, 431–440.

Lupien, M., Eeckhoute, J., Meyer, C.A., Wang, Q., Zhang, Y., Li, W., Carroll, J.S., Liu, X.S., and Brown, M. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell *132*, 958–970.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. USA *78*, 7634–7638.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553–560.

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

Morrisey, E.E., Tang, Z., Sigrist, K., Lu, M.M., Jiang, F., Ip, H.S., and Parmacek, M.S. (1998). GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. Genes Dev. *12*, 3579–3590.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., Okita, K., Mochiduki, Y., Takizawa, N., and Yamanaka, S. (2008). Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. Nat. Biotechnol. *26*, 101–106.

Palmqvist, L., Glover, C.H., Hsu, L., Lu, M., Bossen, B., Piret, J.M., Humphries, R.K., and Helgason, C.D. (2005). Correlation of murine embryonic stem cell gene expression profiles with functional measures of pluripotency. Stem Cells *23*, 663–680.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol. Cell. Biol. *27*, 3769–3779.

Ralston, A., and Rossant, J. (2005). Genetic regulation of stem cell origins in the mouse embryo. Clin. Genet. *68*, 106–112.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., and Khvorova, A. (2004). Rational siRNA design for RNA interference. Nat. Biotechnol. *22*, 326–330.

Rossant, J. (2007). Stem cells and lineage development in the mammalian blastocyst. Reprod. Fertil. Dev. *19*, 111–118.

Sakaki-Yumoto, M., Kobayashi, C., Sato, A., Fujimura, S., Matsumoto, Y., Takasato, M., Kodama, T., Aburatani, H., Asashima, M., Yoshida, N., and Nishinakamura, R. (2006). The murine homolog of SALL4, a causative gene in Okihiro syndrome, is essential for embryonic stem cell proliferation, and cooperates with Sall1 in anorectal, heart, brain and kidney development. Development *133*, 3005–3013.

Shimosato, D., Shiki, M., and 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.

Soudais, C., Bielinska, M., Heikinheimo, M., MacArthur, C.A., Narita, N., Saffitz, J.E., Simon, M.C., Leiden, J.M., and Wilson, D.B. (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.

Tam, W.L., Lim, C.Y., Han, J., Zhang, J., Ang, Y.S., Ng, H.H., Yang, H., and Lim, B. (2008). T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. Stem Cells *26*, 2019–2031.

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

Wu, Q., Chen, X., Zhang, J., Loh, Y.H., Low, T.Y., Zhang, W., Sze, S.K., Lim, B., and Ng, H.H. (2006). Sall4 interacts with Nanog and co-occupies Nanog genomic sites in embryonic stem cells. J. Biol. Chem. *281*, 24090–24094.

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

Yang, J., Chai, L., Gao, C., Fowles, T.C., Alipio, Z., Dang, H., Xu, D., Fink, L.M., Ward, D.C., and Ma, Y. (2008). SALL4 is a key regulator of survival and apoptosis in human leukemic cells. Blood *112*, 805–813.

Zhang, J., Tam, W.L., Tong, G.Q., Wu, Q., Chan, H.Y., Soh, B.S., Lou, Y., Yang, J., Ma, Y., Chai, L., et al. (2006). Sall4 modulates embryonic stem cell pluripotency and early embryonic development by the transcriptional regulation of Pou5f1. Nat. Cell Biol. *8*, 1114–1123.