

Smad2 mediates Activin/Nodal signaling in mesendoderm differentiation of mouse embryonic stem cells

Teng Fei¹, Shanshan Zhu², Kai Xia², Jianping Zhang¹, Zhongwei Li¹, Jing-Dong J Han^{2,3}, Ye-Guang Chen¹

¹The State Key Laboratory of Biomembrane and Membrane Biotechnology, School of Life Sciences, Tsinghua University, Beijing 100084, China; ²CAS Key Laboratory of Molecular Developmental Biology and Center for Molecular Systems Biology, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China; ³CAS Key Laboratory for Computational Biology, Chinese Academy of Sciences-Max Planck Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yueyang Road, Shanghai 200031, China

Although Activin/Nodal signaling regulates pluripotency of human embryonic stem (ES) cells, how this signaling acts in mouse ES cells remains largely unclear. To investigate this, we confirmed that mouse ES cells possess active Smad2-mediated Activin/Nodal signaling and found that Smad2-mediated Activin/Nodal signaling is dispensable for self-renewal maintenance but is required for proper differentiation toward the mesendoderm lineage. To gain insights into the underlying mechanisms, Smad2-associated genes were identified by genome-wide chromatin immunoprecipitation-chip analysis. The results showed that there is a transcriptional correlation between Smad2 binding and Activin/Nodal signaling modulation, and that the development-related genes were enriched among the Smad2-bound targets. We further identified *Tapbp* as a key player in mesendoderm differentiation of mouse ES cells acting downstream of the Activin/Nodal-Smad2 pathway. Taken together, our findings suggest that Smad2-mediated Activin/Nodal signaling orchestrates mesendoderm lineage commitment of mouse ES cells through direct modulation of corresponding developmental regulator expression.

Keywords: embryonic stem cell; Activin; Nodal; Smad2; ChIP-chip

Cell Research (2010) 20:1306-1318. doi:10.1038/cr.2010.158; published online 16 November 2010

Introduction

Embryonic stem (ES) cells derived from blastocyst-stage embryos can be maintained in a self-renewal state or induced to differentiate into all the cell lineages of the body [1-3]. Extrinsic stimuli of cytokines produced cell autonomously or imposed exogenously play crucial roles in ES cell fate determination by triggering distinct signaling cascades (reviewed in [4]). Unlike the conserved pluripotency maintenance roles of core transcription factors like Nanog, Oct4 and Sox2 etc., mouse and human ES cells respond differently to extrinsic signals to main-

tain their self-renewal status.

Among those signals, members of the transforming growth factor- β (TGF- β) superfamily were recognized as the crucial ones for both mouse and human ES cell fate choices and were representative to show the distinct requirements of the two species [5-7]. The TGF- β superfamily comprises two main groups: the TGF- β /Activin/Nodal group and the bone morphogenetic protein (BMP)/growth and differentiation factor (GDF) group. These extracellular cytokines transduce signals via binding to their cognate pairs of receptors (type I and type II receptors), which in turn activate R-Smad (receptor-regulated Smad) proteins via phosphorylation. Activated R-Smads then form a heterocomplex with the common mediator Smad4 to regulate gene expression in the nucleus with help of other cofactors or transcription factors [8-12]. The two groups of TGF- β superfamily members employ different sets of R-Smads to transduce their signals: TGF- β /Activin/Nodal mainly activate Smad2 and Smad3, whereas BMP/GDF generally utilize Smad1, Smad5 and

Correspondence: Ye-Guang Chen^a, Jing-Dong J Han^b

^aTel: 86-10-62795184; Fax: 86-10-62794376

E-mail: ygchen@tsinghua.edu.cn

^bTel: 86-10-64845843; Fax: 86-10-64845797

E-mail: jdhan@genetics.ac.cn

Received 6 September 2010; revised 16 September 2010; accepted 20 September 2010; published online 16 November 2010

Smad8. Genetic and biochemical studies have well documented that Smad2, but not Smad3, plays a critical role in mediating Activin/Nodal signaling to regulate mouse embryo development [13-15], indicating the potential significance of Smad2 in ES cell fate determination.

In human ES cells, TGF- β /Activin/Nodal signaling was regarded as one of the most crucial pathways for self-renewal maintenance, while BMP signaling can drive ES cells to trophoblast differentiation [16-19]. Distinct from human ES cells, mouse ES cells employ BMP signaling to promote self-renewal in combination with leukemia inhibitory factor (LIF) signaling mainly by inhibiting neural differentiation [20, 21]. TGF- β signaling was recognized as being inert in undifferentiated mouse ES cells probably due to low expression of the TGF- β type II receptor [22, 23]. Activin/Nodal signaling has been reported to promote mouse ES cell proliferation [22]. A recent report also suggested that endogenous Nodal signaling indirectly modulates pluripotency via negative regulation of BMP signaling [24]. When ES cells are induced to differentiate, both branches of the TGF- β family signals are implicated in various lineage commitment, serving as either inducers or repressors in a context-dependent manner [5-7]. However, the exact role of the Smad-mediated Activin/Nodal signaling in self-renewal and differentiation is still largely unclear.

In the present study, we found that the Smad2-mediated Activin/Nodal signaling is dispensable for self-renewal maintenance, but is essential for proper commitment of several lineages, especially of mesendoderm. Through mapping Smad2-associated genes by the genome-wide chromatin immunoprecipitation (ChIP)-chip approach, we found that Smad2-bound genes mainly encode development-related regulators, which may in turn orchestrate mouse ES cell fate choices downstream of Activin/Nodal-Smad2 signaling. Finally, we exemplified the above point by identifying a Smad2 target gene – *Tapbp*, as a new key player mediating Activin/Nodal-Smad2 functions during mesendoderm differentiation.

Results

Smad2-mediated Activin/Nodal signaling is active in mouse ES cells

Before investigating the roles of Activin/Nodal signaling in mouse ES cells, we first examined the endogenous Activin/Nodal signaling activity in R1 ES cells and to what extent those cells respond to the exogenous ligand stimuli. By employing the Smad2/3-responsive TGF- β /Activin/Nodal signaling luciferase reporter constructs – CAGA-Luciferase and 3TP-Luciferase, we determined the reporter activities in serum-free cultured ES cells

after treatment with 25 ng/ml Activin, 10 μ M SB431542 (specific small molecule inhibitor of TGF- β /Activin/Nodal signaling type I receptors – Alk4, Alk5 and Alk7) and both together for 24 h. As shown in Figure 1A and 1B, exogenous Activin dramatically upregulated both CAGA-Luciferase and 3TP-Luciferase activities, indicating that R1 ES cells possess a well-responsive signaling machinery for Activin/Nodal signaling. Furthermore, SB431542 significantly repressed the expression of these reporters at the basal level, reflecting the existence of endogenous Activin/Nodal signaling activity in these cells. Consistently, SB431542 also inhibited the reporter activities induced by Activin.

To confirm that Smad2 mediates Activin/Nodal signaling transduction in mouse ES cells, we then monitored the phospho-Smad2 levels. SB431542 caused a significant decrease of the basal phospho-Smad2 levels, whereas Activin induced ~3-fold increase of phospho-Smad2 (Figure 1C). Notably, TGF- β slightly enhanced phospho-Smad2 level, which is consistent with the previous reports that TGF- β responsiveness is weak in mouse ES cells. These data clearly showed that R1 mouse ES cells possess endogenous Smad-mediated Activin/Nodal signaling and respond to the extrinsic stimuli.

Smad2 is not required for self-renewal of mouse ES cells

To further determine how Smad2-mediated Activin/Nodal signaling exerts its biological functions in mouse ES cells, we generated a stable Smad2 knockdown ES cell line with short hairpin RNA (shRNA) (Figure 2A). The Smad2 knockdown ES cells cultured in the standard conditions do not show discriminable difference in morphology (data not shown). To further clarify whether endogenous Smad2-mediated Activin/Nodal signaling affects mouse ES cell self-renewal, we examined the mRNA levels of the self-renewal markers *Nanog*, *Oct4* and *Zfp42/Rex1*. As shown in Figure 2B, all of them showed no difference between control and Smad2 knockdown cells. This result was further confirmed by examination of the protein levels of *Nanog* and *Oct4* (Figure 2C), suggesting that Smad2-mediated Activin/Nodal signaling may not play an essential role in mouse ES cell pluripotency maintenance. In addition, Smad2 knockdown had no effect on cell proliferation (Supplementary information, Figure S1). Consistently, a long-term treatment of SB431542 (2 days) did not change the expression of *Nanog*, *Oct4* and *Rex1* either (Figure 2D). As a control, *Lefty2*, a well-characterized target gene of Activin/Nodal, was repressed significantly by SB431542. These data together indicate that endogenous Smad2-mediated Activin/Nodal signaling is not essential for mouse ES cell self-renewal.

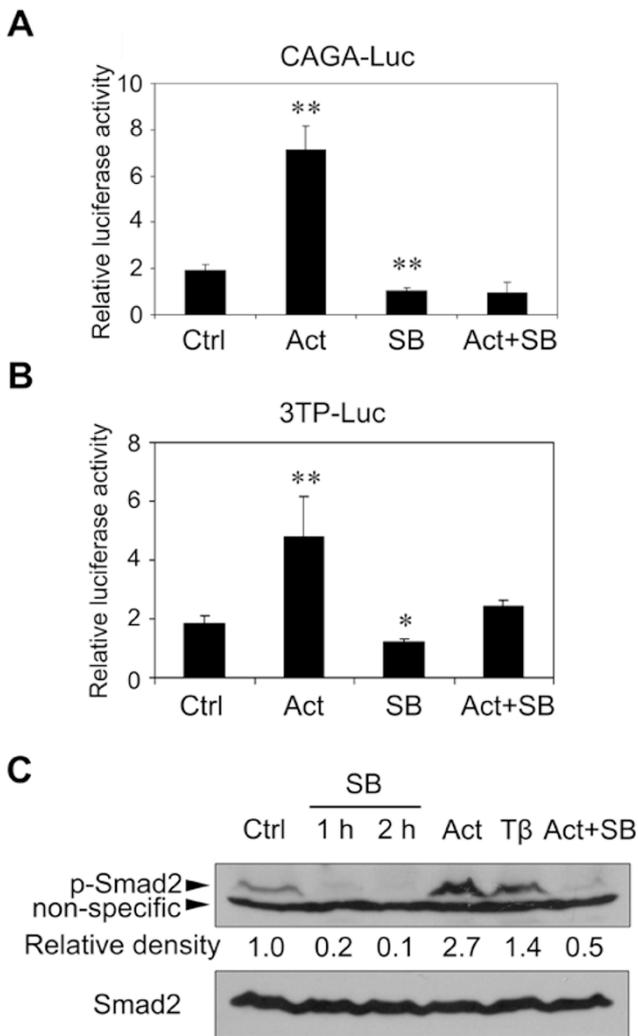


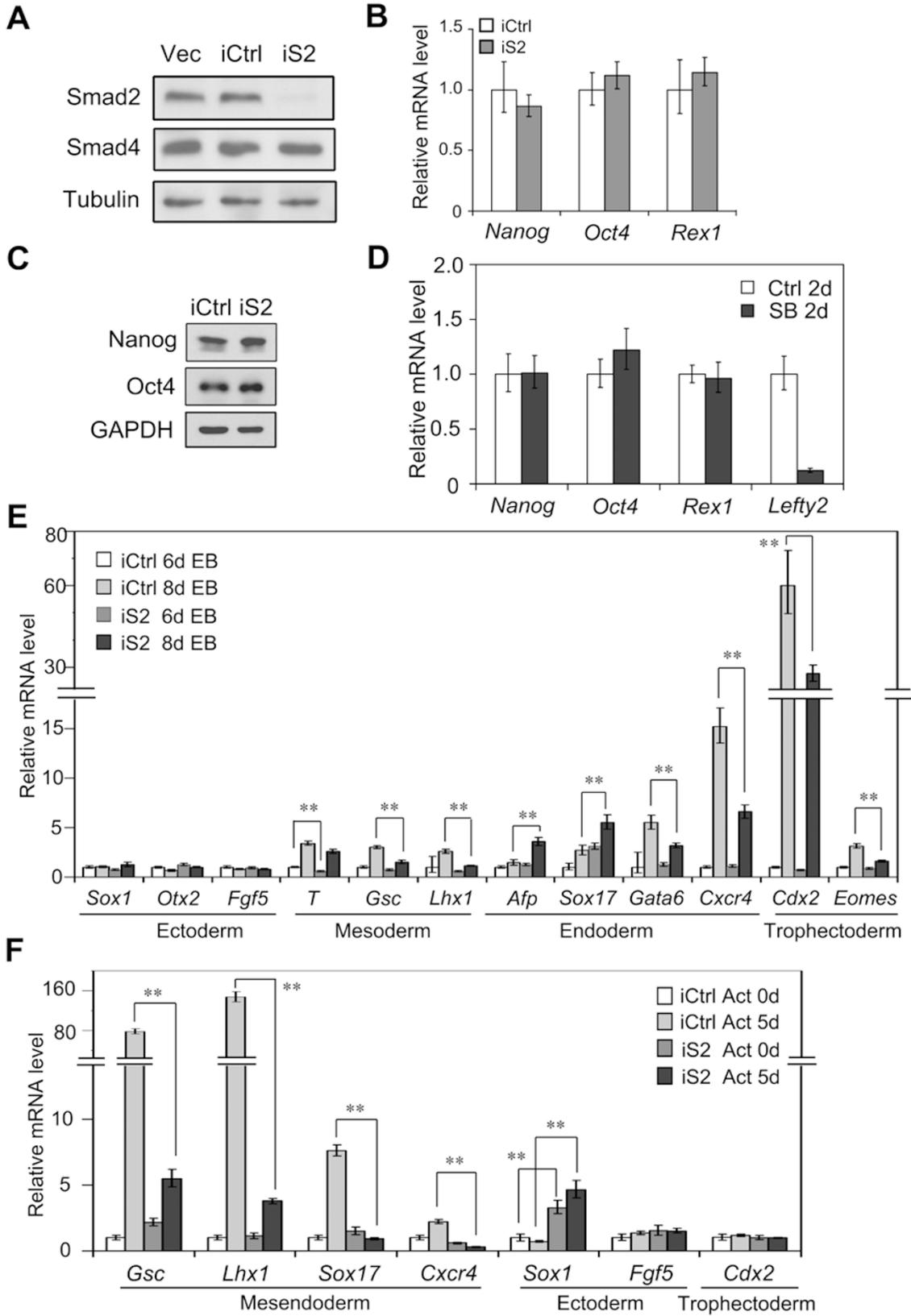
Figure 1 Mouse ES cells possess active Activin/Nodal-Smad2 signaling. **(A, B)** R1 cells transfected with CAGA-luciferase **(A)** or 3TP-luciferase **(B)** were treated with DMSO (Ctrl), 25 ng/ml Activin (Act), 10 μM SB431542 (SB) or both for 24 h, respectively. Cells were then harvested for luciferase assay. The experiment was performed in triplicate, and the data are presented as mean ± SEM of three independent experiments after normalization to Renilla activity (** $P < 0.01$, * $P < 0.05$). **(C)** R1 cells were treated with 10 μM SB431542 (SB), 25 ng/ml Activin (Act), 100 pM TGF-β (Tβ) or both 25 ng/ml Activin and 10 μM SB431542 for 1 h (except one lane of SB431542 for 2 h), respectively. Phospho-Smad2 level was determined by immunoblotting. Anti-Smad2 antibody was used to detect total smad2 level. The numbers below the phospho-Smad2 bands indicate the relative density of phospho-Smad2 compared with corresponding total Smad2.

Smad2-mediated Activin/Nodal signaling regulates distinct fate choices of ES cells

We then asked whether Smad2 knockdown may inter-

fer with the proper differentiation processes of mouse ES cells using embryoid body (EB), in which ES cells autonomously undergo multilineage differentiation after LIF deprivation. We monitored the expression of multiple lineage markers in control and Smad2 knockdown cells at both days 6 and 8 time point of EB formation using quantitative RT-PCR. Significant downregulation of mesodermal and endodermal markers (*T*, *Gsc*, *Lhx1*, *Gata6* and *Cxcr4*) was observed (Figure 2E). Trophectodermal markers (*Cdx2* and *Eomes*) were also repressed in Smad2 knockdown cells, whereas ectodermal markers (*Sox1*, *Otx2* and *Fgf5*) remained unchanged. Intriguingly, in contrast to *Gata6* and *Cxcr4*, the endodermal markers *Afp* and *Sox17* exhibited upregulation in Smad2 knockdown cells, indicating that there is a differential regulation pattern for endoderm differentiation (Figure 2E). We speculated that this paradox may result from the differential regulatory mechanisms of these distinct lineage markers by endogenous Smad2 signaling and the Smad2 effect on differentiation is lineage-specific but not the whole germ layer-specific. Furthermore, the dynamic regulation depends on the time and differentiation stages, and certain bias of this differentiation system cannot be excluded. Nevertheless, these data strongly indicate that endogenous Smad2-mediated Activin/Nodal signaling plays important roles during mouse ES cells fate determination, especially toward mesendodermal and trophectodermal lineages.

To obtain a more precise picture of the roles of Smad2-mediated Activin/Nodal signaling in mouse ES cell differentiation, we adopted a more direct differentiation system, in which ES cells were cultured as monolayer in a defined N2B27-supplemented medium without feeder, serum or serum replacement [21]. The cells were differentiated toward mesendoderm with Activin treatment for 4-5 days; otherwise the cells underwent neuroectoderm differentiation in the absence of Activin (data not shown). We then compared the mesendoderm differentiation status of control and Smad2 knockdown cells before and after Activin treatment for 5 days. As shown in Figure 2F, mesendodermal markers – *Gsc*, *Lhx1* and *Sox17* – were dramatically induced, and knockdown of Smad2 significantly attenuated their induction. Distinct from the EB differentiation system, *Cxcr4* seemed not to respond well to Activin in this system, but Smad2 knockdown could still reduce its induction. The ectoderm markers *Sox1* and *Fgf5* were not induced by Activin, but knockdown of Smad2 led to a higher expression of *Sox1*, which is also a marker of early neural lineages, consistent with that Activin/Nodal signaling has a repressive role in early neural differentiation [25]. The trophectodermal marker *Cdx2* was not effectively induced



by Activin, nor repressed by Smad2 knockdown (Figure 2F). These data together indicate that Smad2-mediated Activin/Nodal signaling regulates distinct mouse ES cell lineage commitment, especially promoting mesendoderm differentiation.

Genome-wide identification of Smad2-bound genes in ES cells

To systematically identify direct Smad2 target genes downstream of Activin/Nodal signaling, we employed ChIP coupled high-throughput promoter array technique (ChIP-chip) to map the Smad2 binding sites in R1 ES cells. Undifferentiated R1 cells cultured in serum replacement medium on gelatin-coated feeder-free plates were harvested for ChIP with the antibodies specifically recognizing Smad2 in R1 cells (Supplementary information, Figure S2). Genomic DNA immunoprecipitated by Smad2 was then amplified, labeled with fluorophores and finally detected on the Agilent mouse promoter arrays, consisting of ~60-base pair (bp) probes at ~200-bp intervals covering from upstream -5.5 kilobases (kb) to downstream +2.5 kb promoter regions relative to the transcriptional start sites (TSS) for ~17 000 annotated genes of the mouse genome. Binding sites were defined by the method described previously [21], which was shown to obtain a highly credible data set with a false-positive rate of less than 20% as validated by ChIP-PCR and other criteria. Seven hundred ninety-two genes harboring Smad2 binding sites were identified (Supplementary information, Table S1). By comparing Smad2 binding sites with the reported Smad4 binding sites in the mouse genome [21], we found that Smad2 and Smad4 have a significant co-occupation in the genome of R1 cells with empirical $P < 0.001$ (Figure 3A). Binding peak plots of *Casc4* and *Fbn2* exemplified the co-occupation of Smad2 and Smad4 (Figure 3B). Representative ChIP-

PCR validation of several Smad2 binding events is shown in Figure 3C. Further examination of Smad2 binding sites distribution along the promoter regions showed that Smad2 binding is mostly concentrated in the region near the TSS (Figure 3D).

Smad2 binding correlates with Activin/Nodal signaling-modulated transcriptional regulation

To determine whether Smad2-DNA association in undifferentiated ES cells accounts for the transcriptional outputs of Activin/Nodal signaling, we compared our Smad2-DNA binding data with the Activin- or SB431542-treated mouse ES cell expression array data obtained by Galvin *et al.* [24] by Parametric Analysis of Gene Set Enrichment (PAGE) analysis (Supplementary information, Table S2). PAGE is a statistical method to detect whether a prior defined gene set shows statistically significant differences between two samples or states [26]. As shown in Figure 4A and Supplementary information, Figure S3, neither Smad2- nor Smad4-binding genes showed significant correlation with expression level change upon Activin or SB431542 stimuli, whereas those genes co-occupied by both Smad2 and Smad4 exhibited a significant upregulation (judged by the Z-score) as evidenced by the P -value of 0.001 in SB431542-treated group, indicating that genes bound by both Smad2 and Smad4 tend to show transcriptional regulation upon SB431542 treatment in undifferentiated ES cells. We then selected several genes from Smad2-associated targets and examined their expression upon the treatment of Activin or SB431542 for 4 h or in the Smad2 knockdown cells. Although the well-characterized TGF- β /Activin/Nodal target *Lefty2* was not found in our Smad2 binding target list, which might result from the relative moderate binding strength under experimental conditions or stringent cutoff algorithms, another classical

Figure 2 Smad2-mediated Activin/Nodal signaling regulates mouse ES cell fate determination. **(A)** The lysates from R1 cells stably transfected with different shRNA producing p13.7-puro plasmids – empty vector (Vec), luciferase control shRNA (iCtrl) and Smad2 shRNA (iS2) – were subjected to immunoblotting using anti-Smad2 and anti-Smad4 antibodies. Tubulin served as a loading control. **(B)** Control and Smad2 knockdown cells were cultured in the medium with serum replacement plus LIF without feeders, and RNA was isolated to determine the mRNA levels of *Nanog*, *Oct4* and *Zfp42/Rex1* by qRT-PCR. **(C)** Protein levels of Nanog and Oct4 in control and Smad2 knockdown cells revealed by immunoblotting. GAPDH served as a loading control. **(D)** Inhibition of endogenous Activin/Nodal signaling does not affect the expression of self-renewal markers. R1 ES cells were treated with DMSO (Ctrl) or 10 μ M SB431542 (SB) for 2 days, and mRNA level of indicated genes were determined by qRT-PCR. **(E)** Smad2 knockdown differentially influences distinct lineage marker expression during EB differentiation of R1 cells. EBs derived from 6 day and 8 day differentiation of control and Smad2 knockdown (iCtrl and iS2, respectively) were harvested for qRT-PCR analysis of distinct germ layer markers. The significance in expression was analyzed by Student's t -test, and data are presented as mean \pm SEM ($n = 3$, $**P < 0.01$). **(F)** Smad2 knockdown impairs mesendoderm marker expression in a monolayer culture-based Activin-induced differentiation system. Cells cultured in a defined N2B27 medium were treated with 25 ng/ml Activin for 5 days and the mRNA levels of indicated lineage markers were determined by qRT-PCR. The significance of expression was analyzed by Student's t -test and data are presented as mean \pm SEM ($n = 3$, $**P < 0.01$).

TGF- β /Activin/Nodal target gene – *Smad7*, was in the list (Supplementary information, Table S1). Their expression was differentially induced by Activin to various extents and significantly repressed by SB431542 or Smad2 knockdown (Figure 4B). In contrast, *Id2* and *Tmem109* were upregulated by SB431542 or Smad2 knockdown and remained unchanged upon Activin treatment, indicating that endogenous Smad2-mediated Activin/Nodal signaling plays key roles in their expression regulation, possibly due to the high endogenous Activin activity or to other Smad2-mediated Activin-independent signaling (for instance, Nodal signaling). Interestingly, another two genes – *Pmp22* and *Tapbp* – are representative of the

gene set, which exhibited better response to exogenous Activin/Nodal signaling than the endogenous signaling (Figure 4B). Therefore, Smad2 association may mediate different types of gene regulation with regard to the signaling intensity thresholds.

Notably, there was a large population of Smad2-associated genes that did not exhibit transcriptional changes upon Activin/Nodal signaling modulation in undifferentiated ES cells. We speculated that Smad2 association with DNA not only is for real-time transcription events but also can act as a pre-occupation status ready to meet proper stimuli to initiate transcription. To test this possibility, we monitored the transcriptional responsiveness

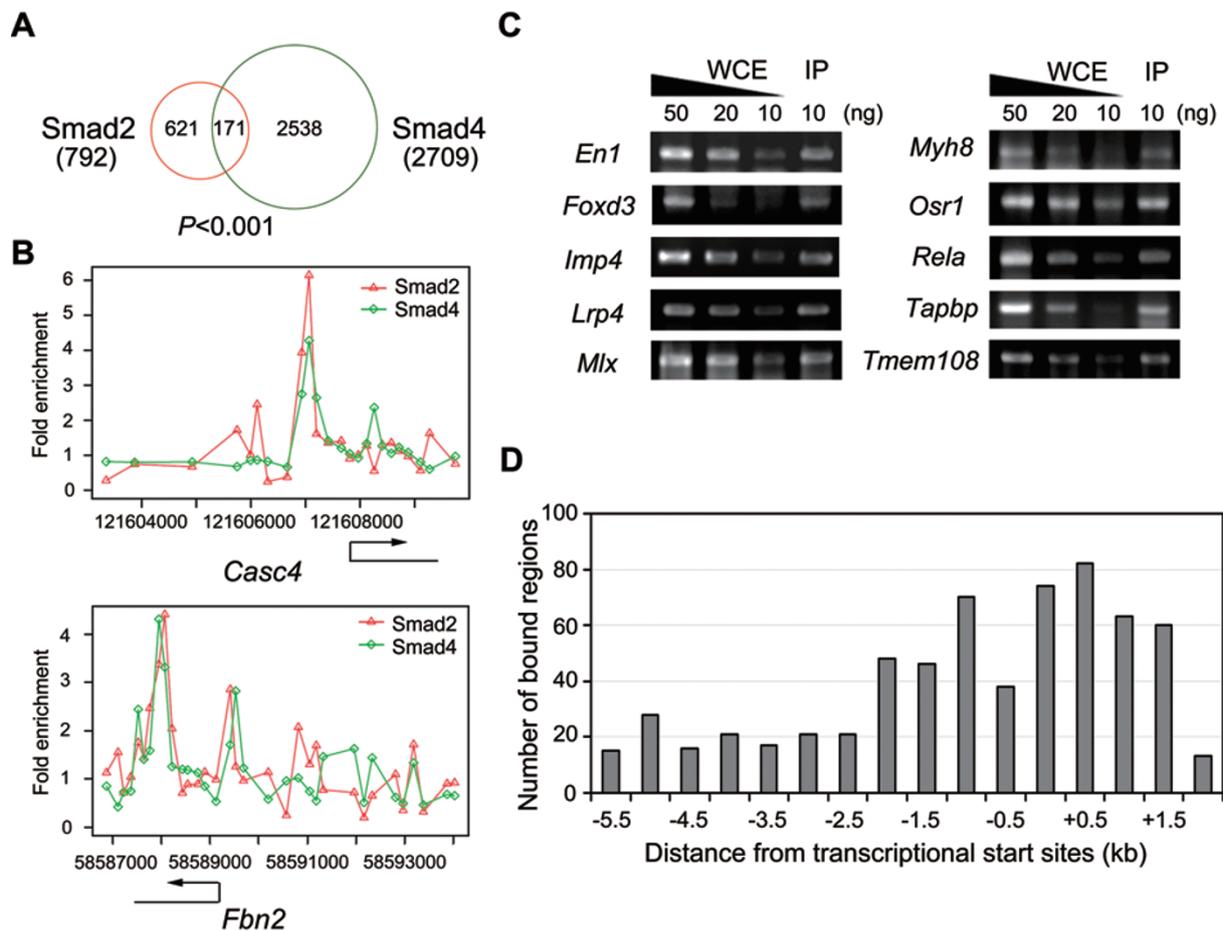


Figure 3 Genome-wide mapping of Smad2-associated genes in R1 cells. **(A)** Venn diagram showing the overlap among genes bound by Smad2 and Smad4. Smad2-bound genes were compared with previously reported Smad4-bound ones identified in the same conditions. **(B)** A representative view of co-occupancy by Smad2 and Smad4. Plots display unprocessed ChIP-enriched ratios within the chromosomal region indicated in the X axis. The corresponding gene was depicted below the plot with transcriptional direction denoted by arrows. **(C)** Representative results of Smad2-DNA association validation by ChIP-site-specific PCR. Anti-Smad2 immuno-enriched (IP) DNA (10 ng) from R1 cells and a range of unenriched whole cell extract (WCE) DNA amounts (10, 20, 50 ng of DNA) were used for each primer pair designed according to the predicted bound regions. **(D)** Distribution of Smad-bound regions along promoter regions.

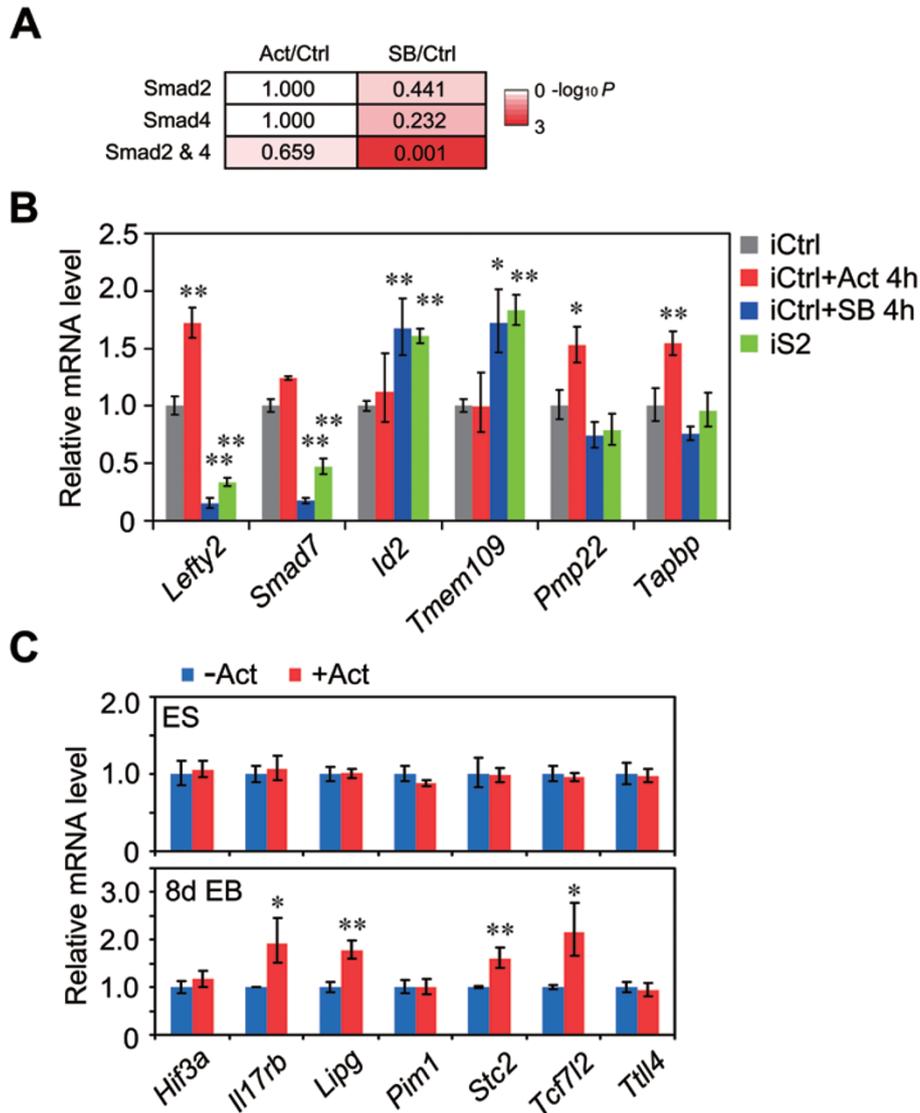


Figure 4 Smad2-associated genes transcriptionally respond to Activin/Nodal in ES cells. **(A)** The significance of correlation between Smad2-bound targets with Activin and SB responsive genes analyzed by PAGE. Fold change for each gene, under Activin or SB treatment, was used in this analysis to detect if a target gene set listed in the row headers was significantly changed (a high positive z-score indicates upregulation and a low negative z-score indicates downregulation). The significance of *P* values derived from the z-scores is indicated inside each box which is colored according to the color legend. **(B)** Gene expression validation of several Smad2-associated genes displays distinct regulatory patterns by Activin/Nodal signaling. Smad2 knockdown (iS2) or control (iCtrl) cells were treated with 25 ng/ml Activin (Act) or 10 μ M SB431542 (SB) for 4 h, then total RNA was extracted and gene expression was determined by qRT-PCR. **(C)** Smad-associated genes display context-dependent regulation in ES cells upon Activin treatment. ES cells and 8-day EB were treated with 25 ng/ml Activin for 4 h. Total RNA was extracted and gene expression was analyzed by qRT-PCR. The significance in expression was analyzed by Student's *t*-test, and data are presented as mean \pm SEM ($n = 3$, ** $P < 0.01$, * $P < 0.05$).

of several genes to Activin under two distinct cellular conditions, i.e., the undifferentiated ES cell state and the differentiated 8-day EB state. As shown in Figure 4C, all the tested genes did not respond to Activin in undifferentiated ES cells, whereas some of them like *Il17rb*, *Lipg*,

Stc2 and *Tcf7l2* apparently were induced by Activin in differentiated cells.

De novo prediction of significantly enriched Smad2-associated DNA motifs revealed canonical TGF- β /Activin/Nodal-responsive CAGA motifs (Supplementary

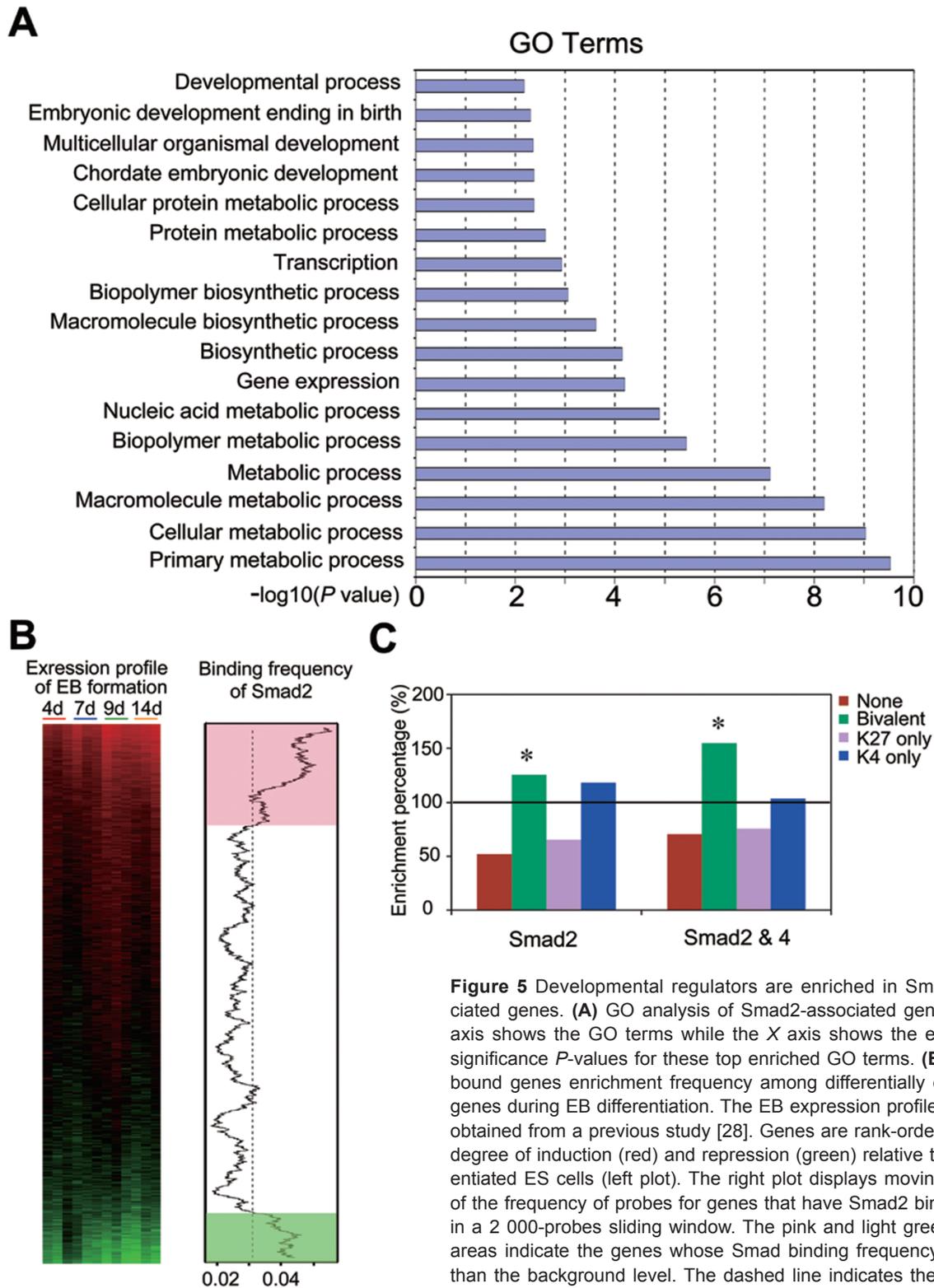


Figure 5 Developmental regulators are enriched in Smad2-associated genes. **(A)** GO analysis of Smad2-associated genes. The Y axis shows the GO terms while the X axis shows the enrichment significance P -values for these top enriched GO terms. **(B)** Smad2-bound genes enrichment frequency among differentially expressed genes during EB differentiation. The EB expression profile data was obtained from a previous study [28]. Genes are rank-ordered by the degree of induction (red) and repression (green) relative to undifferentiated ES cells (left plot). The right plot displays moving average of the frequency of probes for genes that have Smad2 binding sites in a 2 000-probes sliding window. The pink and light green shaded areas indicate the genes whose Smad binding frequency is higher than the background level. The dashed line indicates the expected average (background level or the ratio of the number of probes for Smad2-bound targets over the total number of interrogated gene probes). **(C)** Percentage of different histone modification marks within Smad-bound genes was compared with that over all promoters. Asterisk indicates Fisher's exact test $P < 0.005$. "None" stands for the genes that are not modified by either H3K27me3 or H3K4me3.

information, Figure S4). Although the major form of Smad2 with an insert in its MH1 domain prevents direct DNA binding, an alternatively spliced form of Smad2 without the insert can directly bind to DNA [27]. It is also possible that Smad2 is brought to DNA by other factors. Indeed, the Smad2-associated DNA elements are closely related to the consensus binding elements of Stat1, Ets family members and Smad3 and Smad4 (Supplementary information, Figure S4).

Development-related factors are enriched among Smad2 target genes

To gain insights into the biological processes where Smad2-associated genes are highly involved, we carried out the gene ontology (GO) analysis. As shown in Figure 5A, metabolic- and developmental-related processes were the most enriched functional categories of Smad2-binding target genes, indicating that Smad2-mediated Activin/Nodal signaling orchestrates ES cell fate decisions through direct regulation of metabolic and developmental regulators.

Then we asked whether the Smad2-associated genes are dynamically regulated during the differentiation pro-

cess of ES cells to EB by analyzing a time course expression array data reported by Hailesellasse Sene *et al.* [28]. Differentially expressed genes were ranked according to their expression levels with the upregulated ones (red) on the top and downregulated ones (green) on the bottom of the left panel, and the frequency of Smad2-association was plotted in a sliding window on the right panel (Figure 5B). Obviously, the differentially regulated genes with Smad2 association are highly enriched in both ends where the genes exhibit dramatic expression changes during this differentiation process.

Specific state of histone modification plays crucial roles in ES cell fate determination, and the genes harboring both histone H3K4 trimethylation (me3) and H3K27me3 modification — so called “bivalent” genes — are regarded to largely encode important developmental regulators in ES cells [29-31]. Thus, we examined the status of bivalent and other related modifications of our Smad2-associated genes by comparing with previously reported histone modification data of ES cells [32]. As shown in Figure 5C, bivalent modification is most enriched in the Smad2-associated genes. In addition, when we extended this analysis to the genes co-occupied by

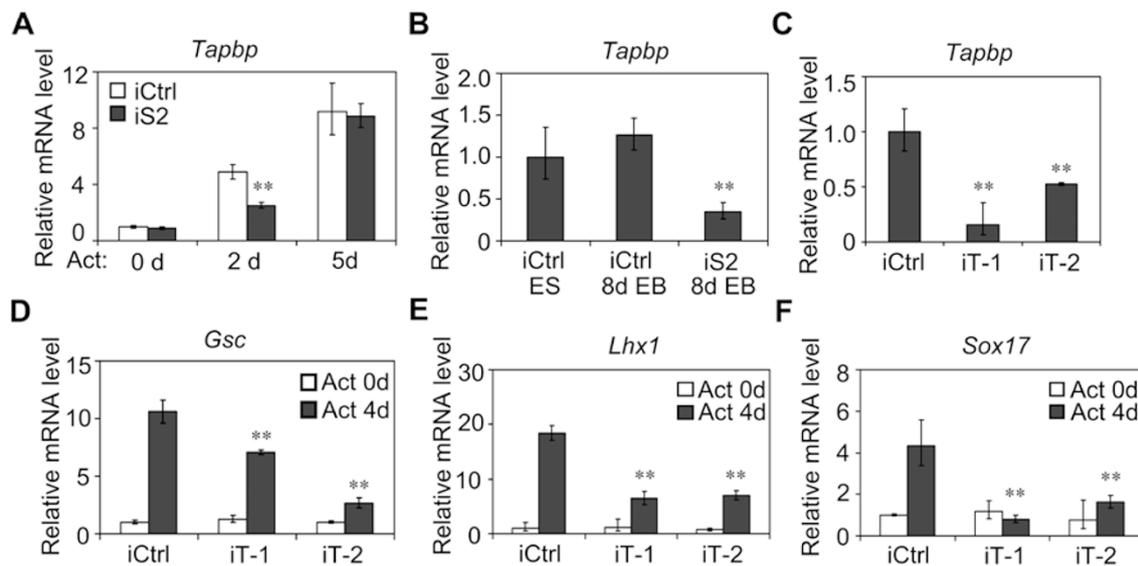


Figure 6 *Tapbp* is required for proper mesendoderm differentiation and acts downstream of Activin/Nodal-Smad2 signaling. **(A)** Expression pattern of *Tapbp* during Activin-induced mesendoderm differentiation. Control and Smad2 knockdown cells were treated with 25 ng/ml Activin for the indicated time, and RNA was harvested to determine the mRNA level of *Tapbp* by qRT-PCR. **(B)** Samples of control ES cells, 8-day control EB and 8-day Smad2 knockdown EB were harvested to determine the mRNA level of *Tapbp* by qRT-PCR. **(C)** Knockdown of *Tapbp* in R1 ES cells. R1 cells stably transfected with control shRNA and two distinct *Tapbp* shRNA (iT-1 and iT-2, respectively) plasmids were harvested to determine the mRNA level of *Tapbp* by qRT-PCR. **(D, E and F)** *Tapbp* knockdown reduces several mesendoderm markers (*Gsc*, *Lhx1* and *Sox17*) induced by Activin. Cells stably expressing control shRNA and two distinct *Tapbp* shRNA were treated with 25 ng/ml Activin for 4 days to induce mesendoderm differentiation. Expression of markers was determined by qRT-PCR. The significance in expression was analyzed by Student's *t*-test, and data are presented as mean \pm SEM ($n = 3$, $**P < 0.01$).

both Smad2 and Smad4, bivalent modification becomes more enriched, which echoes to the significant transcriptional correlation described above (Figure 4A). These results indicate that the Smad2-associated genes are largely bivalently modified and encode developmental regulators, which exert functions during ES cell fate determination.

Tapbp, a downstream target of Activin/Nodal-Smad2 signaling, is required for proper mesendoderm differentiation

As the above results suggested that the Smad2-mediated Activin/Nodal signaling promotes mesendoderm differentiation, we attempted to identify specific undiscovered genes implicated in this process from the Smad2-bound gene list. *Tapbp* encodes the transmembrane glycoprotein tapasin, which is involved in endoplasmic reticulum (ER) quality control of major histocompatibility complex (MHC) class I antigen presentation [33, 34]. The data in Figures 3C and 4B showed that Smad2 binds to the *Tapbp* promoter and Activin can upregulate *Tapbp* expression. Therefore, we examined the importance of *Tapbp* in mesendoderm differentiation. As shown in Figure 6A, *Tapbp* expression was dramatically increased as mesendoderm differentiation proceeded, and was significantly downregulated by Smad2 knockdown in the initial stages (2 day) of differentiation, indicating the regulation of *Tapbp* by the Activin/Nodal-Smad2 axis and the potential functional relevance at the initiation of mesendoderm differentiation. To determine whether the acute induction of *Tapbp* is Activin induced and not the result of differentiation, we compared *Tapbp* mRNA levels between ES cells and 8-day differentiated EB. The data in Figure 6B showed that the mRNA level of *Tapbp* in differentiated EB is comparable with that of ES cells, supporting that the expression elevation in Figure 6A is indeed a direct effect of Activin. Moreover, Smad2 knockdown in 8-day EB significantly reduced *Tapbp* expression, further supporting the transcriptional regulation of *Tapbp* by Smad2-mediated Activin signaling.

To determine whether *Tapbp* expression is important for Activin-induced mesendoderm differentiation, we knocked down *Tapbp* in ES cells by using two independent *Tapbp*-targeted shRNA (iT-1 and iT-2) (Figure 6C). After Activin-induced mesendoderm differentiation, the mesendoderm lineage markers *Gsc*, *Lhx1* and *Sox17* exhibited differential but significant downregulation in two lines of *Tapbp* knockdown cells (Figure 6D-6F). These data indicate that *Tapbp*, acting downstream of Activin/Nodal-Smad2 signaling, is an essential player in Activin-triggered commitment of specific mesendoderm lineages of ES cells.

Discussion

Activin/Nodal signaling has been established as a pivotal regulatory cue to instruct ES cell fate determination [5-7], however, the underlying molecular basis is not fully understood. In this study, we found that Smad2-dependent Activin/Nodal signaling is dispensable for mouse ES cell self-renewal maintenance but required for various lineage differentiation, especially toward the mesendoderm direction. CHIP-chip identification of Smad2-associated genes indicated that Activin/Nodal-Smad signaling may function by directly regulating expression of a cohort of developmental regulators. Among them, *Tapbp* was for the first time identified as a direct target gene and plays an important role in mediating Activin/Nodal-Smad2 signaling to induce mesendoderm differentiation.

Due to the poor response to TGF- β of mouse ES cells, few systematic studies were carried out to investigate the significance of Activin/Nodal signaling. By employing Smad2 shRNA and the TGF- β /Activin/Nodal type I receptor inhibitor SB431542, we demonstrated that Smad2-mediated Activin/Nodal signaling has no apparent regulatory effect on mouse ES cell self-renewal (Figure 2 and Supplementary information, Figure S1). A previous study reported that Activin/Nodal signaling promotes mouse ES cell propagation [22], but Smad2 knockdown did not lead to aberrant cell proliferation in our study. This may be attributed to incomplete deprivation of Smad2 expression by RNA interference, or Smad3 but not Smad2 plays a dominant role in this process. A recent study showed that inhibition of Activin/Nodal signaling by SB431542 can increase BMP signaling and thus reinforce mouse ES cell pluripotency [24], consistent with our notion that endogenous Activin/Nodal signaling is not essential for self-renewal maintenance.

Activin/Nodal has been employed as differentiation-priming cytokines to induce mesendoderm differentiation of ES cells, indicating their important roles in ES cell lineage commitment [35, 36]. Smad2 has also been suggested to be required for mesendoderm induction in zebrafish [37]. Here, we found that Smad2 plays an essential role in expression of several mesendodermal lineage markers (e.g., *Gsc*, *Lhx1* and *Sox17*). This is also in agreement with the roles of Activin/Nodal signaling in early embryo development [14, 15]. Notably, Smad2 knockout mice are embryonic lethal due to severe defects in germ layer specification, whereas Smad3-deficient mice are viable and fertile but display aberrant phenotypes in immunity (reviewed in [13]). Although Smad2 and Smad3 are closely related mediators of TGF- β /Activin/Nodal signaling, their expression patterns during embryo devel-

opment are very distinct with the expression of Smad2, but not Smad3, in the anterior visceral endoderm, which accounts for the essential requirement of Smad2, but not Smad3, in early development of mouse embryos [38]. Together with other studies, it was strongly suggested that Smad2, but not Smad3, plays a major role in mediating Activin/Nodal signaling during early embryo development. It is the reason why we focused on Smad2 in this study.

To gain insight into the detailed mechanisms of how Smad2-mediated Activin/Nodal signaling functions, we employed high-throughput ChIP-chip technology to identify the Smad2-associated promoters of genes. Combinatorial analysis, together with our previous identification of Smad4-associated genes [21], indicated that Smad2 may work in combination with Smad4 to regulate transcription downstream of Activin/Nodal signaling in mouse ES cells (Figures 3A, 3B, 4A and 5C). The fact that a large amount of Smad2 binding events do not display transcriptional functions in self-renewing ES cells indicated their dynamic regulation, which is confirmed by our random validation (Figure 4) and is consistent with previous observation of other transcription factors [21, 39-41]. For those genes that display transcriptional response to Activin/Nodal signaling modulation, their regulation manners depend on different modulation approaches – some more responsive to signal activation (e.g., *Pmp22* and *Tapbp*), some more sensitive to signal attenuation (e.g. *Smad7*, *Id2* and *Tmem109*) and others responsive to both (e.g. *Lefty2*) (Figure 4B). These data suggest that Smad2-bound genes may display differential regulation under different thresholds of Activin/Nodal signaling as reported recently [42].

Functional classification of Smad2-bound genes by GO analysis indicated significant enrichment of developmental- and metabolic-related processes (Figure 5A). Therefore, Smad2-mediated Activin/Nodal signaling may regulate mouse ES cell fate by directly modulating these developmental regulators. Evidence from EB gene expression profile analysis also showed significant enrichment of Smad2-bound genes in highly differentially regulated gene clusters, which tend to contain more key developmental regulators during ES cell differentiation (Figure 5B). Histone modification analysis reinforced this notion by showing that Smad2-associated genes harbor significant enrichment of H3K4me3/H3K27me3 bivalent marks (Figure 5C), a sign of gene modification of developmental regulators [31].

We identified *Tapbp* as a downstream target of Activin/Nodal-Smad2 signaling in mesendoderm induction. *Tapbp* knockdown significantly attenuated Activin-mediated induction of mesendoderm lineage markers *Gsc*,

Lhx1 and *Sox17*, demonstrating that Tapasin, the product of *Tapbp*, is an essential downstream mediator for Activin/Nodal-Smad2 signaling in promoting mesendoderm differentiation. Tapasin is known as an ER chaperone facilitating antigen presentation of MHC class I molecules [33, 34]. Tapasin is also implicated in cancer development and schizophrenia [43, 44]. Tapasin is expressed in pre-implantation embryos, but little is known about its function during embryo development [45]. *Tapbp* knock-out mice developed and bred normally with impaired immune response [46], but it cannot be excluded that *Tapbp* takes part in developmental processes at the cellular level. More detailed work is required to uncover the mechanisms of how Tapasin regulates mesendoderm differentiation. It is expected that more crucial developmental regulators mediating Activin/Nodal signaling would be found in those Smad2-associated genes.

Materials and Methods

Cell culture and antibodies

R1 mouse ES cells were cultured under typical ES cell conditions with DMEM supplemented with 15% fetal calf serum (Hyclone), LIF (Chemicon), non-essential amino acids (Gibco), GlutaMax-I (Gibco), β -mercaptoethanol (Gibco) and penicillin/streptomycin. For feeder-free culture, cells were plated on gelatinized tissue culture plates, and Knockout Serum Replacement (KO-SR) (Gibco) was used to substitute fetal calf serum. For ChIP experiments and immunoblotting, affinity-purified Smad2 antibody (generated with aa183-273) and Smad4 antibody [21] were used. Other antibodies were purchased: anti-phospho-Smad2 (Chemicon, AB3849), anti-Nanog (Bethyl Laboratories, A300-397A), anti-Oct4 (Santa Cruz Biotech, sc-8628) and anti-GAPDH (Santa Cruz Biotech, sc-20357).

EB formation and Activin-induced mesendoderm differentiation

For EB formation, monolayer undifferentiated R1 ES cells were first trypsinized into single cells and then replaced at a density of 3×10^5 cells/60 mm non-adherent dish in KO-SR containing ES cell medium devoid of LIF. Floating EB was harvested at 6 day or 8 day for analysis. For Activin-induced mesendoderm differentiation, ES cells were cultured as monolayer in gelatinized feeder-free six-well plates with the initial plating density of $\sim 1.5 \times 10^5$ cells/well and the time when 25 ng/ml Activin is added was counted as day 0. The medium was composed of 1:1 mixture of DMEM/F12 supplemented with N2 supplement (Gibco) and NeuralBasal medium supplemented with B27 supplement (Gibco) and with β -mercaptoethanol and Glutamax-I. Cells were harvested at day 4 or day 5 for analysis.

Luciferase reporter assay

Cells seeded in 24-well plates were transfected with plasmids as indicated by Lipofectamine 2000 (Invitrogen). The internal control pRenilla-TK vector was used to normalize luciferase activity. Each experiment was performed in triplicate and the data represent

the mean \pm SEM of three independent experiments.

RNA interference

Modified pII3.7 plasmid (GFP fragment was replaced by puromycin resistant gene) was employed to produce short hairpin nucleotides. Nucleotides targeting firefly *luciferase* gene, which has no homology with mouse transcriptome, were served as control shRNA. The RNAi oligonucleotides used are as follows: Control (iCtrl): 5'-GCGACCAACGCCTTGATTG-3', Smad2 (iS2): 5'-GTGATAGTGCAATCTTTGT-3', Tapbp 1# (iT-1): 5'-GGTCTCGTGAGCTAATAA-3', Tapbp 2# (iT-2): 5'-GACCCAAAGCTATACTTCA-3'. To generate knockdown cells, R1 cells were transfected separately with shRNA constructs by Lipofectamine 2000 and 1 μ g/ml puromycin was added at 48 h post-transfection. About 10 days later, stably transfected cells were expanded and then stored for further analysis.

ChIP assay

ChIP was carried out as described previously [21]. Briefly, feeder-free cultured R1 cells were first crosslinked with 1% formaldehyde solution before cells were lysed. The DNA/protein complex-containing lysates were then sonicated to obtain DNA fragments ranging ~300-500 bp on average. Samples were then centrifuged at 14 000 rpm for 10 min. The supernatant was diluted and pre-absorbed by 50 μ l Protein A beads (Zymed) and then incubated with 10 μ g anti-Smad2 antibodies overnight at 4 $^{\circ}$ C. The immunocomplex was collected with 100 μ l Protein A beads, eluted and reverse-crosslinked by overnight incubation at 65 $^{\circ}$ C. Whole cell extract (WCE) DNA (Input fraction reserved from the sonication step) was also treated for crosslinking reversal. Immunoprecipitated DNA and WCE DNA were then purified by treatment with RNaseA, proteinase K and multiple phenol:chloroform:isoamyl alcohol extraction.

RNA isolation, reverse transcription and quantitative real-time PCR

RNA was isolated using Trizol (Invitrogen). Reverse transcriptase (Toyobo) was employed for oligo dT primed first-strand cDNA synthesis. Real-time PCR was carried out on Mx3000P detection system (Stratagene) using EvaGreen dye (Biotium). $\Delta\Delta C_t$ method was used to comparatively quantify the amount of mRNA level. Primers were listed in Supplementary information, Table S3.

Bioinformatics analysis

The detailed information about promoter array design, bound regions identification, expression array analysis, motif finding etc. can be found in supplementary information, Data S1. The Smad2 ChIP-chip data have been submitted to the NCBI Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE18629.

Acknowledgments

We thank Gaoyang Zhu for technical assistance. This work was supported by grants from the National Natural Science Foundation of China (30930050, 30921004), the 973 Program (2006CB943401, 2010CB833706) to YGC, and grants from the China National Science Foundation (Grant # 30890033, 30588001 and 30620120433), Chinese Ministry of Science and Technology

(Grant # 2006CB910700) to JDH.

References

- 1 Evans MJ, Kaufman MH. Establishment in culture of pluripotent cells from mouse embryos. *Nature* 1981; **292**:154-156.
- 2 Martin GR. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci USA* 1981; **78**:7634-7638.
- 3 Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *Science* 1998; **282**:1145-1147.
- 4 Pera MF, Tam PP. Extrinsic regulation of pluripotent stem cells. *Nature* 2010; **465**:713-720.
- 5 Fei T, Chen YG. Regulation of embryonic stem cell self-renewal and differentiation by TGF-beta family signaling. *Sci China Life Sci* 2010; **53**:497-503.
- 6 Seuntjens E, Umans L, Zwijsen A, et al. Transforming Growth Factor type beta and Smad family signaling in stem cell function. *Cytokine Growth Factor Rev* 2009; **20**:449-458.
- 7 Watabe T, Miyazono K. Roles of TGF-beta family signaling in stem cell renewal and differentiation. *Cell Res* 2009; **19**:103-115.
- 8 Massague J, Seoane J, Wotton D. Smad transcription factors. *Genes Dev* 2005; **19**:2783-2810.
- 9 Feng XH, Derynck R. Specificity and versatility in tgf-beta signaling through Smads. *Annu Rev Cell Dev Biol* 2005; **21**:659-693.
- 10 Datto M, Wang XF. The Smads: transcriptional regulation and mouse models. *Cytokine Growth Factor Rev* 2000; **11**:37-48.
- 11 Massague J, Chen YG. Controlling TGF-beta signaling. *Genes Dev* 2000; **14**:627-644.
- 12 ten Dijke P, Hill CS. New insights into TGF-beta-Smad signalling. *Trends Biochem Sci* 2004; **29**:265-273.
- 13 Brown KA, Pietenpol JA, Moses HL. A tale of two proteins: differential roles and regulation of Smad2 and Smad3 in TGF-beta signaling. *J Cell Biochem* 2007; **101**:9-33.
- 14 Chang H, Lau AL, Matzuk MM. Studying TGF-beta superfamily signaling by knockouts and knockins. *Mol Cell Endocrinol* 2001; **180**:39-46.
- 15 Schier AF, Shen MM. Nodal signalling in vertebrate development. *Nature* 2000; **403**:385-389.
- 16 James D, Levine AJ, Besser D, Hemmati-Brivanlou A. TGF-beta/activin/nodal signaling is necessary for the maintenance of pluripotency in human embryonic stem cells. *Development* 2005; **132**:1273-1282.
- 17 Wu Z, Zhang W, Chen G, et al. Combinatorial signals of activin/nodal and bone morphogenic protein regulate the early lineage segregation of human embryonic stem cells. *J Biol Chem* 2008; **283**:24991-25002.
- 18 Xiao L, Yuan X, Sharkis SJ. Activin A maintains self-renewal and regulates fibroblast growth factor, Wnt, and bone morphogenic protein pathways in human embryonic stem cells. *Stem Cells* 2006; **24**:1476-1486.
- 19 Xu RH, Chen X, Li DS, et al. BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol* 2002; **20**:1261-1264.
- 20 Ying QL, Nichols J, Chambers I, Smith A. BMP induction

- of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 2003; **115**:281-292.
- 21 Fei T, Xia K, Li Z, *et al.* Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. *Genome Res* 2010; **20**:36-44.
- 22 Ogawa K, Saito A, Matsui H, *et al.* Activin-Nodal signaling is involved in propagation of mouse embryonic stem cells. *J Cell Sci* 2007; **120**:55-65.
- 23 Goumans MJ, Ward-van Oostwaard D, Wianny F, *et al.* Mouse embryonic stem cells with aberrant transforming growth factor beta signalling exhibit impaired differentiation *in vitro* and *in vivo*. *Differentiation* 1998; **63**:101-113.
- 24 Galvin KE, Travis ED, Yee D, Magnuson T, Vivian JL. Nodal signaling regulates the bone morphogenic protein pluripotency pathway in mouse embryonic stem cells. *J Biol Chem* 2010; **285**:19747-19756.
- 25 Camus A, Perea-Gomez A, Moreau A, Collignon J. Absence of Nodal signaling promotes precocious neural differentiation in the mouse embryo. *Dev Biol* 2006; **295**:743-755.
- 26 Kim SY, Volsky DJ. PAGE: parametric analysis of gene set enrichment. *BMC Bioinformatics* 2005; **6**:144.
- 27 Yagi K, Goto D, Hamamoto T, *et al.* Alternatively spliced variant of Smad2 lacking exon 3. Comparison with wild-type Smad2 and Smad3. *J Biol Chem* 1999; **274**:703-709.
- 28 Hailesellasse Sene K, Porter CJ, Palidwor G, *et al.* Gene function in early mouse embryonic stem cell differentiation. *BMC Genomics* 2007; **8**:85.
- 29 Pan G, Tian S, Nie J, *et al.* Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. *Cell Stem Cell* 2007; **1**:299-312.
- 30 Zhao XD, Han X, Chew JL, *et al.* Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. *Cell Stem Cell* 2007; **1**:286-298.
- 31 Bernstein BE, Mikkelsen TS, Xie X, *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 2006; **125**:315-326.
- 32 Mikkelsen TS, Ku M, Jaffe DB, *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 2007; **448**:553-560.
- 33 Chapman DC, Williams DB. ER quality control in the biogenesis of MHC class I molecules. *Semin Cell Dev Biol* 2010; **21**:512-519.
- 34 Peaper DR, Cresswell P. Regulation of MHC class I assembly and peptide binding. *Annu Rev Cell Dev Biol* 2008; **24**:343-368.
- 35 Willems E, Leyns L. Patterning of mouse embryonic stem cell-derived pan-mesoderm by Activin A/Nodal and Bmp4 signaling requires Fibroblast Growth Factor activity. *Differentiation* 2008; **76**:745-759.
- 36 Tada S, Era T, Furusawa C, *et al.* Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development* 2005; **132**:4363-4374.
- 37 Jia S, Ren Z, Li X, Zheng Y, Meng A. smad2 and smad3 are required for mesendoderm induction by transforming growth factor-beta/nodal signals in zebrafish. *J Biol Chem* 2008; **283**:2418-2426.
- 38 Tremblay KD, Hoodless PA, Bikoff EK, Robertson EJ. Formation of the definitive endoderm in mouse is a Smad2-dependent process. *Development* 2000; **127**:3079-3090.
- 39 Zheng Y, Josefowicz SZ, Kas A, *et al.* Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature* 2007; **445**:936-940.
- 40 Marson A, Kretschmer K, Frampton GM, *et al.* Foxp3 occupancy and regulation of key target genes during T-cell stimulation. *Nature* 2007; **445**:931-935.
- 41 Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA. Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes Dev* 2008; **22**:746-755.
- 42 Guzman-Ayala M, Lee KL, Mavrikis KJ, *et al.* Graded Smad2/3 activation is converted directly into levels of target gene expression in embryonic stem cells. *PLoS One* 2009; **4**:e4268.
- 43 Fellerhoff B, Wank R. Transporter associated with antigen processing and the chaperone tapasin: are non-classical HLA genes keys to the pathogenesis of schizophrenia? *Med Hypotheses* 2009; **72**:535-538.
- 44 Deshpande A, Wheeler CM, Hunt WC, *et al.* Variation in HLA class I antigen-processing genes and susceptibility to human papillomavirus type 16-associated cervical cancer. *J Infect Dis* 2008; **197**:371-381.
- 45 Lampton PW, Goldstein CY, Warner CM. The role of tapasin in MHC class I protein trafficking in embryos and T cells. *J Reprod Immunol* 2008; **78**:28-39.
- 46 Garbi N, Tan P, Diehl AD, *et al.* Impaired immune responses and altered peptide repertoire in tapasin-deficient mice. *Nat Immunol* 2000; **1**:234-238.

(Supplementary information is linked to the online version of the paper on the *Cell Research* website.)