

Comprehensive profiling reveals mechanisms of SOX2-mediated cell fate specification in human ESCs and NPCs

Chenlin Zhou^{1,4,*}, Xiaoqin Yang^{2,*}, Yiyang Sun^{1,4}, Hongyao Yu³, Yong Zhang², Ying Jin^{1,3,4}

¹Key Laboratory of Stem Cell Biology, CAS Center for Excellence in Molecular Cell Science, Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences/Shanghai JiaoTong University School of Medicine, Shanghai 200031, China; ²School of Life Sciences and Technology, Tongji University, Shanghai 200092, China; ³Laboratory of Molecular Developmental Biology, Shanghai JiaoTong University School of Medicine, Shanghai 200025, China; ⁴University of Chinese Academy of Sciences, Beijing 100049, China

SOX2 is a key regulator of multiple types of stem cells, especially embryonic stem cells (ESCs) and neural progenitor cells (NPCs). Understanding the mechanism underlying the function of SOX2 is of great importance for realizing the full potential of ESCs and NPCs. Here, through genome-wide comparative studies, we show that SOX2 executes its distinct functions in human ESCs (hESCs) and hESC-derived NPCs (hNPCs) through cell type- and stage-dependent transcription programs. Importantly, SOX2 suppresses non-neural lineages in hESCs and regulates neurogenesis from hNPCs by inhibiting canonical Wnt signaling. In hESCs, SOX2 achieves such inhibition by direct transcriptional regulation of important Wnt signaling modulators, WLS and SFRP2. Moreover, SOX2 ensures pluripotent epigenetic landscapes via interacting with histone variant H2A.Z and recruiting polycomb repressor complex 2 to poise developmental genes in hESCs. Together, our results advance our understanding of the mechanism by which cell type-specific transcription factors control lineage-specific gene expression programs and specify cell fate.

Keywords: SOX2; hESCs; hNPCs; Wnt; H2A.Z

Cell Research (2016) 26:171-189. doi:10.1038/cr.2016.15; published online 26 January 2016

Introduction

Cell type-specific master transcription factors play central roles in establishing and maintaining cell identity through directing specific gene regulatory networks. The same transcription factor may exert distinct roles in different cell types and at different developmental stages. However, how they regulate lineage-specific developmental programs is not fully elucidated. As a unique pioneer factor, SOX2 exerts critical roles in embryonic

stem cells (ESCs) and progenitors of multiple lineages, especially neural progenitor cells (NPCs) [1, 2]. To date, most studies on SOX2 were conducted in mouse cells or animal models. The role of SOX2 in human ESCs (hESCs) and human NPCs (hNPCs) is poorly defined; controversies remain regarding differentiated lineages after silencing SOX2 in hESCs [3, 4]. Recently, Wang *et al.* [5] reported that *SOX2* knockdown (KD) caused a compensational increase in *SOX3* expression and that simultaneous depletion of *SOX2* and *SOX3* led to primitive streak induction in hESCs. Also, *SOX2* KD was reported to affect hNPCs proliferation and neurogenesis rather than apoptosis through a *SOX2*-*Lin28/let7* pathway [6]. Apart from these studies, how SOX2 exerts cell type-dependent roles in these two distinct and developmentally related cell types is largely unknown. Comparative studies of Sox2-binding motifs in mouse ESCs (mESCs) and mouse NPCs (mNPCs) have led to the identification of Pou3f2 as an mNPC-specific Pou family member that is

*These two authors contributed equally to this work.

Correspondence: Ying Jin^a, Yong Zhang^b

^aTel: 86-21-54923342

E-mail: yjin@sibs.ac.cn

^bTel: 86-21-65981196

E-mail: yzhang@tongji.edu.cn

Received 10 September 2015; revised 20 November 2015; accepted 24 November 2015; published online 26 January 2016

extensively colocalized with Sox2 [7]. However, a recent study challenges the functional association between the two factors in mNPCs [8]. Moreover, these studies mainly focused on the relationship between Sox2 and POU family proteins in mouse cells. The entire SOX2 protein interactome as well as the authentic targets that constitute the transcription programs underlying SOX2's versatile functions remain to be identified, especially in hESCs and hNPCs.

The crosstalk between core pluripotency factors and key signaling pathways is important for the maintenance of self-renewal and pluripotency in hESCs. Canonical Wnt signaling is needed for self-renewal of mESCs as well as proliferation and neurogenesis of mNPCs [9, 10]. In contrast, reports concerning the role of canonical Wnt signaling in hESCs were controversial. Earlier studies showed that Wnt signaling facilitates hESCs self-renewal [11, 12], while recent studies demonstrated that Wnt signaling induces hESCs to differentiate [13], and that it functions downstream of BMP4 or by switching Smad2/3 signaling in the TGF β pathway from maintaining self-renewal to promoting lineage specification [14, 15]. The role of canonical Wnt signaling in hNPCs is less studied. Recently, Wnt signaling was reported to be dispensable for hNPC generation and proliferation; instead, it contributes to the specification of the anterior and posterior identity [16], and hinders neuronal differentiation from hNPCs [17]. These findings suggest a difference in the role of Wnt signaling between human and mouse cells. On the other hand, although the relationship between SOX2 and Wnt signaling has been studied in various contexts, including neural development in animal models [18], few targets have been identified in mediating either positive or negative regulation between SOX2 and Wnt signaling. Till now, the relationship between SOX2 and Wnt signaling in hESCs and hNPCs remains elusive.

In addition, a growing body of evidence points to the influence of unique chromatin state on transcription programs in cell fate specification [19, 20]. Interestingly, PRC2, responsible for the generation of repressive histone mark H3K27me₃, and core pluripotency factors (OCT4, SOX2 and NANOG) were shown to co-occupy a significant proportion of developmental genes in hESCs [21], suggesting a link between the binding of pluripotency factors and the deposition of H3K27me₃. Moreover, histone variant H2A.Z was also shown to enrich at PRC2 target genes in ESCs and to function in organizing the pluripotent chromatin state [22, 24-26]. However, how pluripotency factors participate in the preservation of genomic plasticity and coordinate with chromatin regulators to repress developmentally poised genes in hESCs has not been clearly answered.

In this study, we conducted a comprehensive study of SOX2 in hESCs and their progenies along neural differentiation. We compared SOX2 global binding profiles, SOX2-regulated genes and SOX2 protein interactome between hESCs and hNPCs, and performed functional assays to test several important regulatory activities. Our results clarify SOX2's roles in hESCs at various stages during neural differentiation and provide mechanistic insights into SOX2's action. SOX2 controls distinct signaling pathways and interacts with proteins associated with chromatin regulation, thus exerting the role of a master transcription factor in the specification of cell fates in multiple contexts.

Results

Genome-wide DNA-binding characteristics of SOX2 in hESCs and hNPCs

To answer the question of how SOX2 could function in a cell type-specific manner, we first generated hNPCs from hESCs and conducted full characterization as previously described [27-29]. We then examined the global binding profiles of SOX2 and Pol II in the genetically identical hESCs and hNPCs by ChIP-seq. Altogether, 17 992 and 67 021 highly confident SOX2-binding sites were detected in hESCs and hNPCs, respectively. The quality of our ChIP-seq data was verified by ChIP-qPCR examination of selected loci (Supplementary information, Figure S1A). About 50% of the peaks in hESCs (8 744) overlapped with those in hNPCs (Figure 1A). The promoter proximal sites only accounted for less than 10% of total peaks, and the majority of SOX2 peaks located in promoter distal sites in both cell types (Figure 1B).

We next compiled our SOX2 ChIP-seq data and Pol II ChIP-seq data as well as published ChIP-seq data of histone modifications in hESCs and hNPCs [30], and classified SOX2-bound sites into promoter peaks and enhancer peaks, which were further categorized into active, poised and repressed peaks. The number of repressed promoter peaks marked by H3K4me₃-/H3K27me₃+ was very small; therefore, the promoter and enhancer peaks were divided into 4 categories: A, poised promoter by H3K4me₃+/H3K27me₃+; B, active promoter by H3K4me₃+/H3K27me₃-; C, active enhancer by H3K4me₁+/H3K27ac+; D, poised enhancer by H3K4me₁+/H3K27ac- and H3K4me₁+/H3K27me₃+ (Figure 1C).

Promoter-bound genes (PBGs) are presumably directly activated or repressed by transcription factors and over 85% of SOX2 PBGs were marked by active histone modifications in both hESCs and hNPCs, indicating a

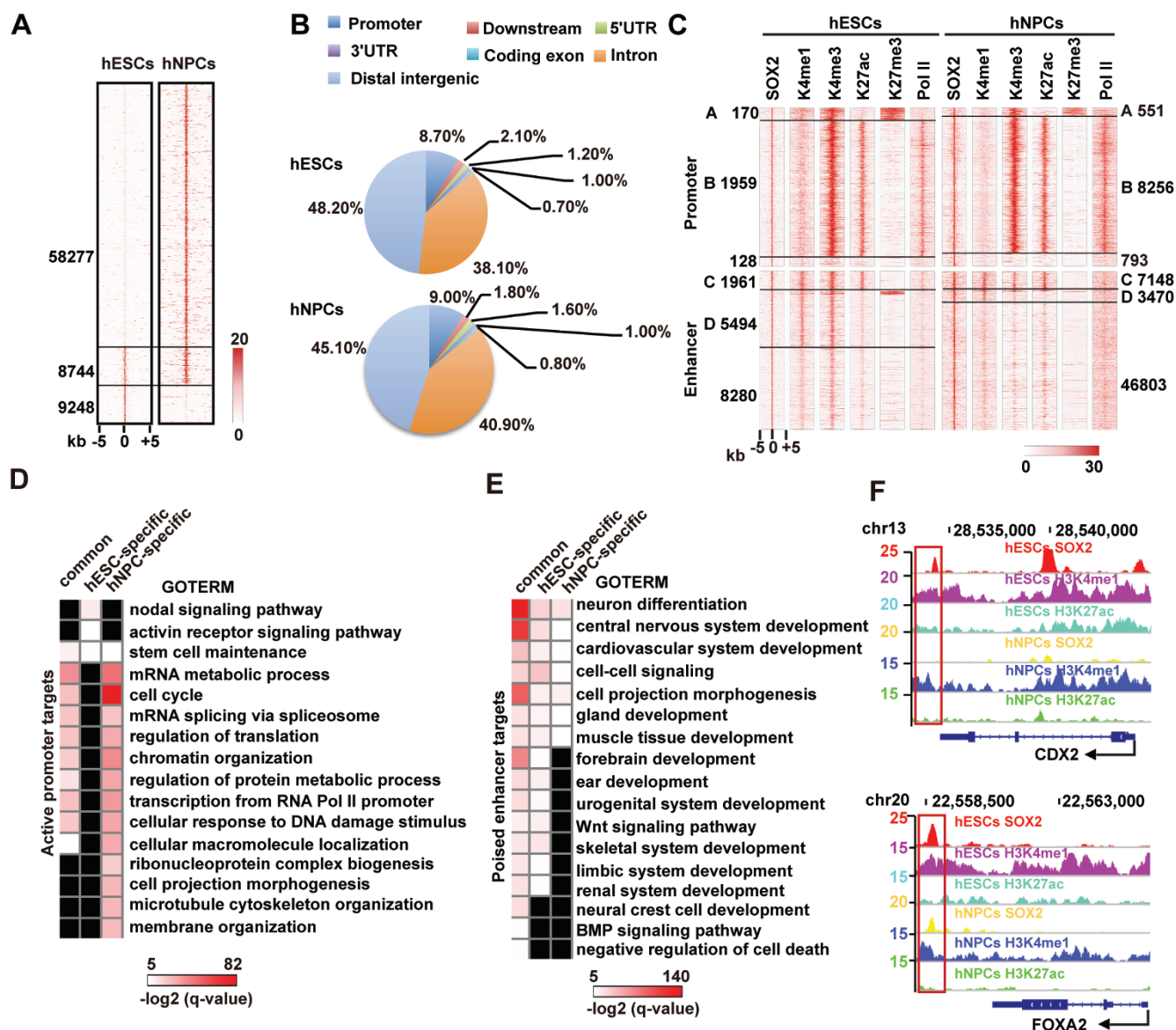


Figure 1 SOX2-binding sites in hESCs and hNPCs characterized by ChIP-seq data of SOX2, Pol II and major histone modifications. **(A)** Heatmap of SOX2 and Pol II enrichments in hESCs and hNPCs centered on peaks of enrichment and extended 5 kb in each direction. Numbers of common and specific peaks are indicated. **(B)** The genomic distribution of SOX2 peaks in hESCs and hNPCs. **(C)** Heatmaps depicting Pol II, H3K4me3, H3K4me1, H3K27me3 and H3K27ac at regions spanning all SOX2-binding sites (± 5 kb) in hESCs and hNPCs. SOX2 peaks are divided into 4 categories (A, poised promoter by H3K4me3+/H3K27me3+; B, active promoter by H3K4me3+/H3K27me3-; C, active enhancer by H3K4me1+/H3K27ac+; D, poised enhancer by H3K4me1+/H3K27ac- and H3K4me1+/H3K27me3+). Numbers of peaks in each category are indicated. **(D, E)** GO analyses of active promoter-bound genes (**D**) and poised enhancer-bound genes (**E**) in hESCs and hNPCs. Benjamini-Hochberg method is applied to adjust the *P*-values in order to account for multiple testing. Enrichment levels of selected GO terms are marked by $-\log_2(q\text{-value})$. Missing values are shown as black. **(F)** Integrative Genomics Viewer screenshots showing the hESC-specific SOX2 peak on *CDX2* enhancer and the SOX2 peak on *FOXA2* enhancer in both hESCs and hNPCs.

positive role for promoter binding by SOX2. Amongst the active hESC-PBGs, 84% of them overlapped with those in hNPCs (Supplementary information, Figure

S1B). Gene Ontology (GO) analyses revealed that these common PBGs, including *MED12*, *MCM4* and *HDAC2* (Supplementary information, Figure S1C), participate in

fundamental biological processes such as cell proliferation, cell cycle, regulation of transcription and translation that safeguard self-renewal and stem cell maintenance (Figure 1D). Moreover, SOX2 PBGs common in hESCs and hNPCs are also involved in mRNA metabolic processes, RNA splicing and chromatin organization (Figure 1D). The remaining hESC-specific PBGs (16%) included genes in the Activin/Nodal pathway known to be important for hESCs self-renewal, while cell projection and cytoskeleton organization were significantly overrepresented in the remaining hNPC-specific PBGs, reflecting the cell type-specific functions of SOX2 (Figure 1D).

In contrast to the high percentage of active promoters, different portions of enhancer-bound peaks were poised in hESCs (74%) and hNPCs (33%) (Figure 1C). Since tissue-specific transcription factors usually bind to enhancers to poise developmental genes [31], we focused on SOX2-bound poised enhancers. Only 5.5% of poised enhancer peaks and 38% of poised enhancer-bound genes (EBGs) in hESCs overlapped with those in hNPCs (Supplementary information, Figure S1B), suggesting the cell type-specific utilization of enhancers and different developmental programs controlled by SOX2 through enhancer binding. Both common and cell type-specific EBGs showed significant enrichment of terms such as neuron differentiation and central nervous system (CNS) development (Figure 1E), probably reflecting SOX2's role as a master regulator in priming neural lineage differentiation from hESCs and CNS development from hNPCs onward. However, common EBGs were also associated with non-neural lineage differentiation, especially mesoderm- and endoderm-derived organogenesis such as cardiovascular system development, skeletal system development and the Wnt signaling pathway (Figure 1E). In addition, SOX2 might reinforce the control of these developmental processes through the remaining 62% hESC-specific EBGs, while hNPC-specific EBGs were largely confined to the development of CNS (Figure 1E), indicating a restriction in the scope of regulation from pluripotency to multipotency. Examples of specific and common EBGs included *CDX2* and *FOXA2* (Figure 1F), master regulators of the trophectoderm and endoderm, respectively.

Collectively, comparative analysis of SOX2-binding profiles combined with epigenetic landscapes reveals that SOX2 controls self-renewal and developmental potential through promoter binding and enhancer binding, respectively. Fundamental self-renewal-related processes are positively regulated through a large group of common PBGs, while developmental potential is poised by SOX2 through common and cell type-specific EBGs, emphasizing the role of enhancers in cell fate determination.

SOX2 occupies a poised proneural transcription program in hESCs

Despite the well-established role of SOX2 as a master regulator of neural development, so far, few genes have been identified as SOX2's authentic targets in promoting neural differentiation. Since CNS development was among the most significantly enriched terms in SOX2 EBGs for hESCs, we hypothesized that SOX2 may occupy a group of proneural genes, which are inactive in hESCs and can be switched on in response to neural induction cues through changes in chromatin landscapes. In search for these proneural genes, we compared histone modification turnover between hESCs and hNPCs and found 640 genes that switched from being poised to being active, including both PBGs and EBGs (Figure 2A and 2B). GO enrichment analysis indicated that the function of these genes was closely associated with CNS development (Figure 2C). Notch signaling ligand *DLL1* and downstream bHLH factor *HES5* were among these genes (Figure 2D). Time course ChIP-qPCR results showed that SOX2 bound promoters of *DLL1* and *HES5* in hESCs with increasing intensity along neural differentiation (Figure 2E). In addition, SOX2 began to bind promoters of Notch ligand *JAG1* and downstream bHLH factor *HES3*, but not *HES1*, as soon as neural differentiation initiated on day 6 (Figure 2E). Notably, the peak expression of these Notch pathway components coincided with SOX2 binding on day 6 of neural differentiation (Figure 2F). In hNPCs, SOX2 continued to bind to these promoters (Figure 2E) and SOX2 KD led to a decrease in the expression of these Notch pathway components (Figure 2G). These findings place SOX2 upstream of Notch signaling during neural differentiation and the maintenance of hNPCs, revealing a previously underappreciated regulation of the Notch pathway by SOX2 at various stages of neural development. Together, our analyses identify a group of potential proneural targets of SOX2, which may allow SOX2 to function as a master regulator of neural development.

SOX2 safeguards stemness and directs specific developmental paths in hESCs and hNPCs through controlling distinct transcription programs

To uncover the functional targets of SOX2, we examined the global transcriptome in SOX2-depleted hESCs and hNPCs using RNA-seq. First, we verified the compensation effect of SOX3 on SOX2 depletion reported by Wang *et al.* [5] in two lines of hESCs, H9 [32] and SHhES2 [33] (Supplementary information, Figure S2A and S2B). Such compensation effect was not seen in hNPCs (Supplementary information, Figure S2C and S2D). Two sets of siRNA oligos targeting both *SOX2* and *SOX3*

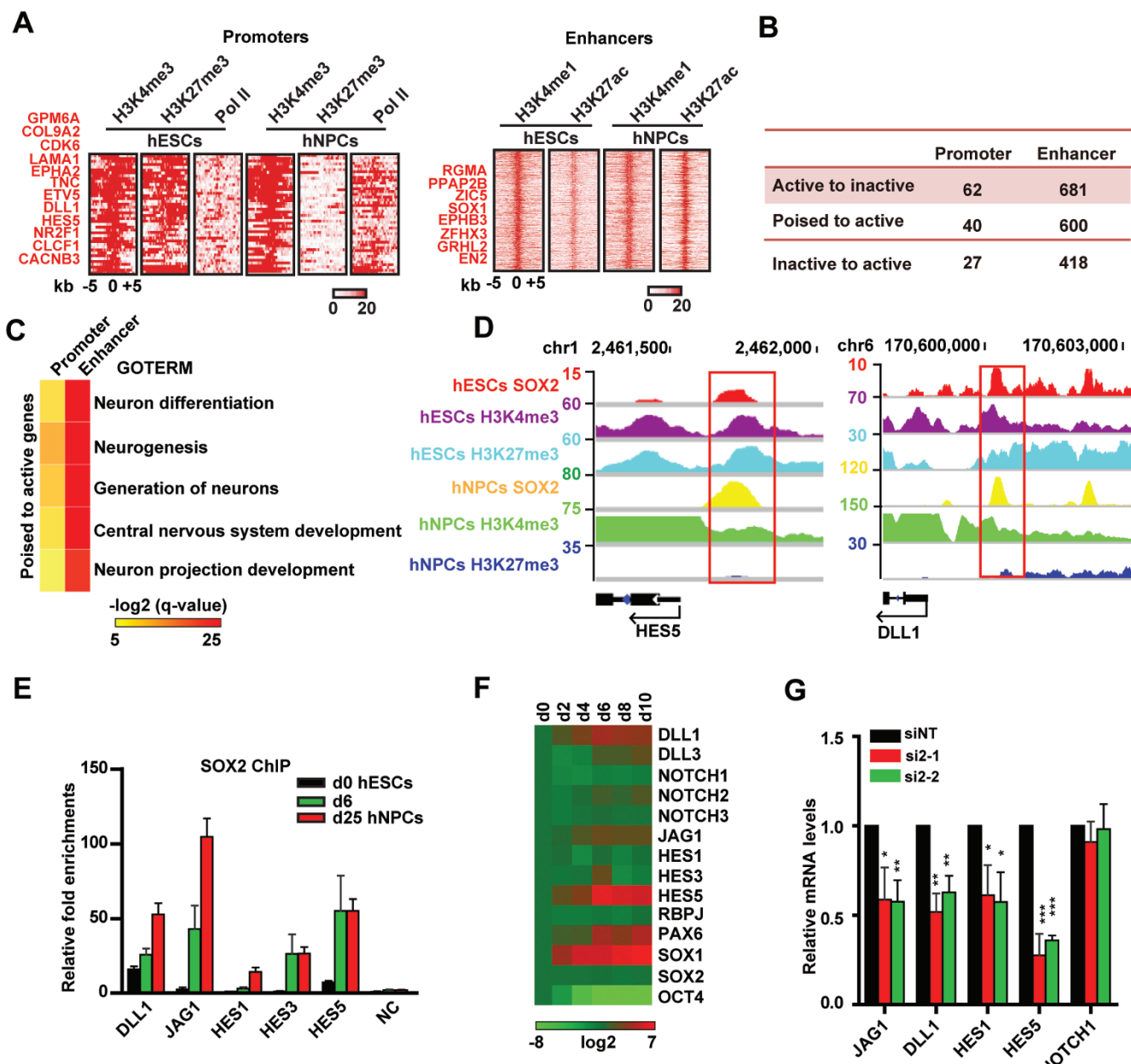


Figure 2 SOX2 promotes differentiation toward neural lineage through occupying proneural targets in hESCs. **(A)** A heatmap showing epigenetic activation of a group of promoters and enhancers. Loss of H3K27me3 and gain of H3K27ac indicate a switch from poised to active promoters (left panel) and enhancers (right panel). A cluster of important neural development genes nearest to these peaks is listed. **(B)** A chart showing histone modification turnover of SOX2's promoter-bound genes and enhance-bound genes in hESCs to hNPCs. Repressed and unmarked histone modifications are defined as inactive ones. **(C)** GO analysis of genes from **A** whose promoter or enhancer switches from the poised state to active state. Benjamini-Hochberg method is applied to adjust the P -values in order to account for multiple testing. Enrichment levels of selected GO terms are marked by $-\log_2(q\text{-value})$. **(D)** Integrative Genomics Viewer screenshots showing the density of SOX2, H3K4me3 and H3K27me3 on promoters of *HES5* and *DLL1* in hESCs and hNPCs. **(E)** ChIP-qPCR results of SOX2 enrichments at regulatory elements of Notch pathway components during neural differentiation at day 0 (hESCs), day 6 and day 25 (hNPCs). **(F)** Expression patterns of Notch pathway components during neural differentiation from day 0 (hESCs) to day 10. *PAX6*, *SOX1* and *OCT4* are used as controls. Scale bar represents \log_2 of expression levels. **(G)** RT-qPCR results for Notch pathway components and proneuronal genes upon SOX2 KD in hNPCs. NT, non-targeting control; si2-1 and si2-2 are two siRNA oligos targeting SOX2.

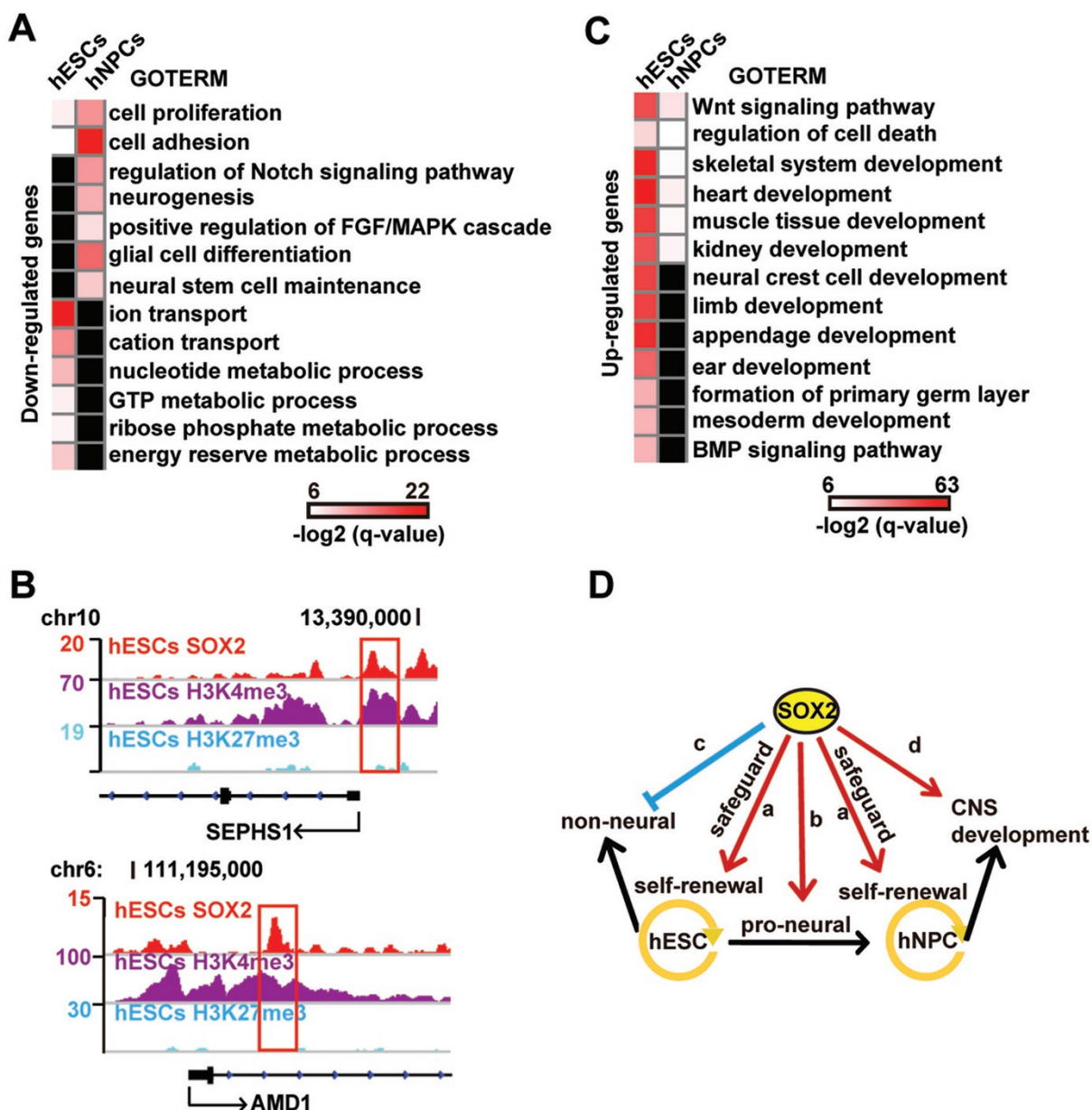


Figure 3 SOX2 ensures self-renewal and differentiation potential of hESCs and hNPCs through dynamic control of distinct transcriptional programs. **(A)** GO analyses of common and specific features of downregulated genes in hESCs and hNPCs. si2/3-1 and si2/3-2 are two sets of oligos used for SOX2/3 KD in hESCs; si2-1 and si2-2 are for SOX2 KD in hNPCs. Cells were collected on day 3 of KD for RNA-seq. Analysis was done using genes altered by both sets of oligos in each cell type. Benjamini-Hochberg method is used to adjust the *P*-values in order to account for multiple testing. Enrichment levels of selected GO terms are marked by $-\log_2(q\text{-value})$. Missing values are marked in black. **(B)** Integrative Genomics Viewer screenshots showing SOX2 peaks on the active promoters of *SEPHS1* and *AMD1*. **(C)** GO analyses of common and specific features of upregulated genes in hESCs and hNPCs. Benjamini-Hochberg method is applied to adjust the *P*-values in order to account for multiple testing. Enrichment levels of selected GO terms are marked by $-\log_2(q\text{-value})$. Missing values are marked in black. **(D)** An illustrative diagram showing SOX2-regulated cell type- and stage-dependent transcriptional programs: a, common in both cell types to safeguard the stem cell identity; b, proneural differentiation in hESCs; c, inhibition of non-neural lineages in hESCs; d, regulation of neurogenesis in hNPCs.

(si2/3-1 and si2/3-2) were used to KD *SOX2* in hESCs and two sets of siRNA oligos targeting *SOX2* (si2-1 and

si2-2) were used in hNPCs. Analyses of RNA-seq data identified around 4 000 altered genes for each oligo set in

hESCs and 900 in hNPCs. Altogether 1 323 direct functional targets of SOX2 in hESCs and 295 in hNPCs were detected by combining results of ChIP-seq and RNA-seq.

Downregulated genes in both cell types were enriched in cell proliferation, an important aspect of self-renewal (Figure 3A). Indeed, *SOX2/3* KD in hESCs and *SOX2* KD in hNPCs both led to a decrease in the expression of stemness markers (Supplementary information, Figure S2B and S2D) and a drastic decrease in the cell number as a result of increased apoptosis and proliferation defects (Supplementary information, Figure S2E-S2K), indicating impaired self-renewal abilities in both cell types. In addition, two critical pathways for self-renewal in NPCs, Notch and FGF/MAPK, were specifically over-represented in hNPC downregulated genes (Figure 3A). The finding was further verified by decreased expression of core Notch components and phosphorylated ERK1/2 (Figure 2G and Supplementary information, Figure S2E). Moreover, analysis of hNPC downregulated genes also indicated impaired differentiation potential toward neurons and glia (Figure 3A). Interestingly, hESC-specific downregulated genes were highly enriched in ion/cation transport and metabolism (Figure 3A). Metabolism was recently proved to be a critical aspect of the unique energetic and biosynthetic programs associated with ESCs' self-renewal [34, 35]. Indeed, many actively expressed metabolic genes with reported or unknown roles in ESC maintenance were targeted by SOX2. For example, *AMD1* encodes a critical enzyme for polyamine biosynthesis and was reported to be required for mESC self-renewal, while *SEPHS1* is a shared signature gene of human mature oocytes and ESCs, although its role in hESC maintenance is unknown [36, 37] (Figure 3B). These results indicate a potential role of SOX2 in positively regulating hESC-specific metabolic pathways essential for the maintenance of hESC self-renewal.

On the other hand, analyses of upregulated genes in hESCs and hNPCs revealed that SOX2 might control differentiation potential in a cell type- and differentiation stage-dependent manner. Although common and cell type-specific EBGs were involved in the development of all three germ layers in both cell types (Figure 1E), only upregulation of genes associated with mesoderm- and endoderm-derived organogenesis such as the skeletal system, heart and kidney was seen upon *SOX2* depletion (Figure 3C). Such induction was significantly more intense in hESCs than in hNPCs (Figure 3C), consistent with our RT-qPCR examination of lineage markers in hESCs (Supplementary information, Figure S2B). These results reinforce the notion that SOX2 directly represses a specific non-neural differentiation program in hESCs, while its function becomes restricted to CNS develop-

ment from NPCs onward. Intriguingly, genes in the Wnt signaling pathway are upregulated in both hESCs and hNPCs (Figure 3C), demonstrating a repressive role of SOX2 in regulating Wnt signaling.

Collectively, we have identified SOX2-regulated direct targets in hESCs and hNPCs for safeguarding self-renewal characteristics and differentiation potentials. Our integrative analyses for ChIP-seq and RNA-seq data reveal that SOX2 may control at least the following four cell type- and differentiation stage-dependent transcription programs: (a) a common program safeguarding the self-renewal of hESCs and hNPCs; (b) a proneural program in hESCs; (c) a hESC-specific non-neural differentiation program, and (d) a CNS differentiation program in hNPCs (Figure 3D). The direct and functional targets associated with these programs most likely contribute to the function of SOX2 in hESCs and hNPCs.

SOX2 represses non-neural lineage differentiation through inhibition of canonical Wnt signaling in hESCs

To examine how regulation of Wnt signaling by SOX2 contributes to its dynamic functions in both hESCs and hNPCs, we determined the role of Wnt signaling in the repression of non-neural lineages by SOX2 in hESCs. First, significantly increased levels of β -catenin and p-LRP6 proteins indicated an increase in the canonical Wnt pathway activity as soon as 48 h post *SOX2/3* KD (Figure 4A and 4B). The marked increase in the nuclear β -catenin level, TOP/FOP luciferase reporter activity and mRNA levels of various Wnt ligands, receptors and downstream targets further support the notion (Figure 4C-4E). Second, forced activation of Wnt signaling by either GSK3 β inhibitor ChIR99021 or increasing dosages of Wnt3A for 48 h mimicked *SOX2/3* KD phenotypes in differentiation toward non-neural lineages and down-regulation of the pluripotency markers (Figure 4F and 4G, Supplementary information, Figure S3A). It is worth noting that BMP signaling was also elevated after *SOX2/3* KD as shown by increased p-SMAD1/5 levels (Supplementary information, Figure S3B). However, a time course analysis showed that the BMP4 treatment failed to induce significant increases of endoderm markers *FOXA2* and *GATA4*. The treatment only induced a transient and weak increase in the expression of mesoderm markers *WNT3*, *MIXL1* and a strong increase in the level of *EOMEs*, which quickly dropped after 24 h. In contrast, trophoderm markers continued to increase after BMP4 treatment for 24 h (Supplementary information, Figure S3A). These results indicate that the developmental path induced by BMP4 signaling is different from that induced by Wnt signaling or *SOX2/3* KD in hESCs. Third, canonical Wnt pathway inhibitor IWR1-e

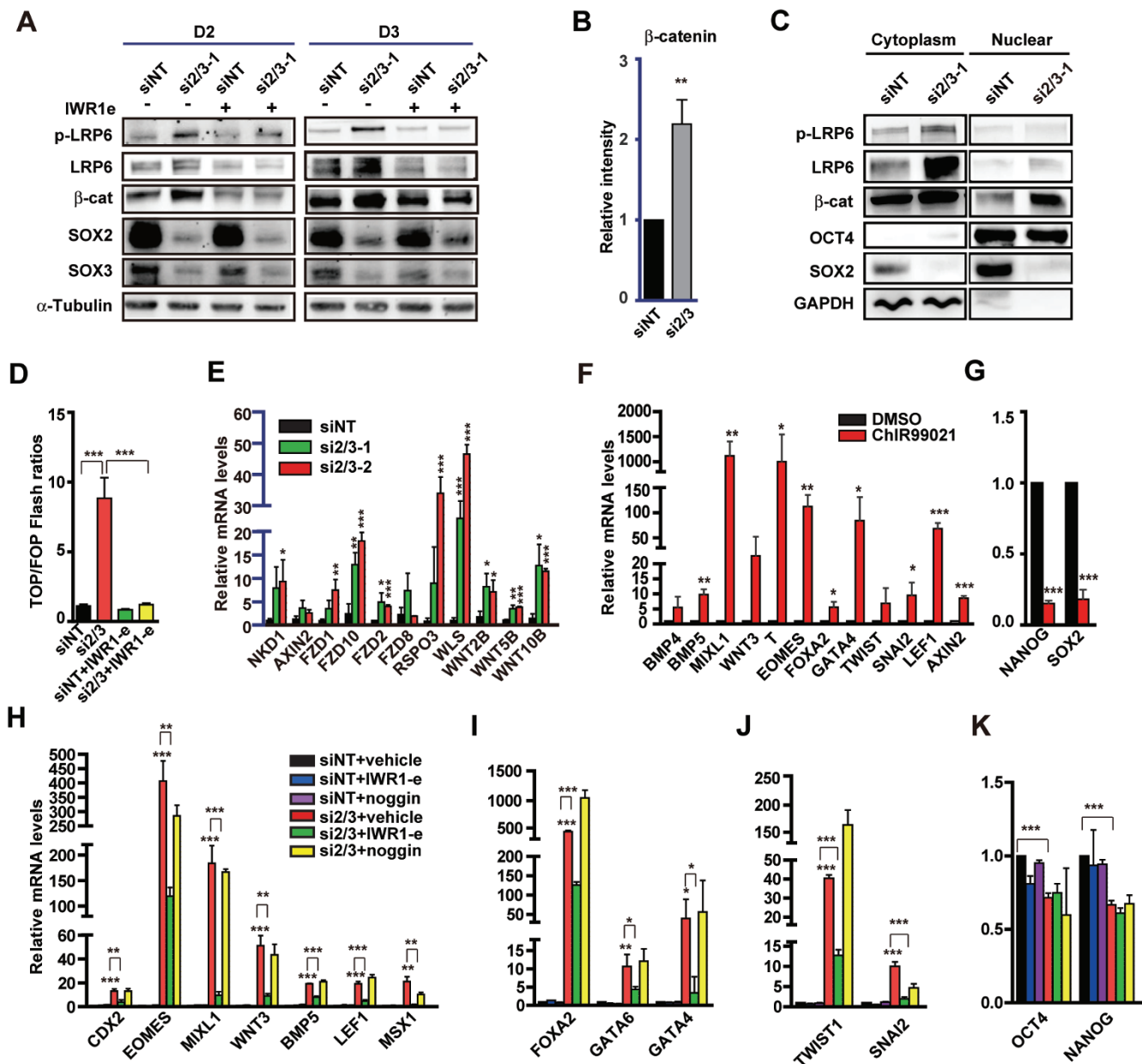


Figure 4 SOX2 KD leads to enhanced non-neural lineage differentiation in hESCs due to activation of canonical Wnt signaling. (A) Western blot results of p-LRP6, LRP6, total β -catenin in H9 hESCs after SOX2/3 KD for 2 and 3 days with or without IWR1-e (10 μ M). (B) Quantification of total β -catenin proteins on day 2 of SOX2/3 KD in hESCs shown in A by densitometry. (C) Western blot results of β -catenin, p-LRP6 and LRP6 in cytoplasm and nuclear extracts of hESCs on day 2 of SOX2/3 KD. OCT4 and GAPDH are loading controls for nuclear and cytoplasmic proteins, respectively. (D) Results of the 8 \times TOP/FOP luciferase reporter assay in hESCs on day 2 of SOX2/3 KD with or without IWR1-e (10 μ M). (E) RT-qPCR results of Wnt pathway components in hESCs on day 3 of SOX2/3 KD. (F, G) RT-qPCR results of germ layer markers and Wnt targets (F) as well as pluripotency markers (G) on day 3 of ChIR99021 treatment (3 μ M). (H-K) RT-qPCR results of mesoderm and trophectoderm markers (H), endoderm markers (I), EMT markers (J) and pluripotency markers (K) on day 3 of SOX2/3 KD in hESCs in the presence of vehicle, or noggin (200 ng/ml) or IWR1-e (10 μ M).

but not BMP inhibitor Noggin significantly abrogated the increase in the expression of almost all lineage markers caused by SOX2/3 KD and restored the morphology of

hESCs (Figure 4H-4J, Supplementary information, Figure S3C), indicating that the activation of Wnt signaling was primarily responsible for these changes and is not

downstream of BMP signaling. These results were reproducible with two additional Wnt pathway inhibitors (IWP2 and DKK1) and in two lines of hESCs (H9 and SHhES2) (Supplementary information, Figure S3D and S3E).

Interestingly, defects in proliferation and the increased apoptosis following *SOX2/3* KD were not rescued by IWR1-e (Supplementary information, Figure S2H), neither was the expression of *SOX2*, *SOX3*, *OCT4* and *NANOG* (Figure 4A and 4K), suggesting that *SOX2* might protect self-renewal of hESCs through both Wnt

signaling-dependent and Wnt signaling-independent mechanisms. Since studies in mESCs have suggested that *Sox2* governs pluripotency mainly through transcriptional regulation of *Oct4* [38] and that *OCT4* was reported to repress Wnt signaling in hESCs [13], we examined whether *SOX2* inhibited Wnt signaling through *OCT4*. Our results showed that activation of the Wnt pathway took place on day 2 post *SOX2/3* KD (Figure 4A), while the decrease in *OCT4* levels and the most distinct differentiation phenotype occurred on day 3 (Supplementary information, Figure S3C, S3F and S3G). Therefore, it is

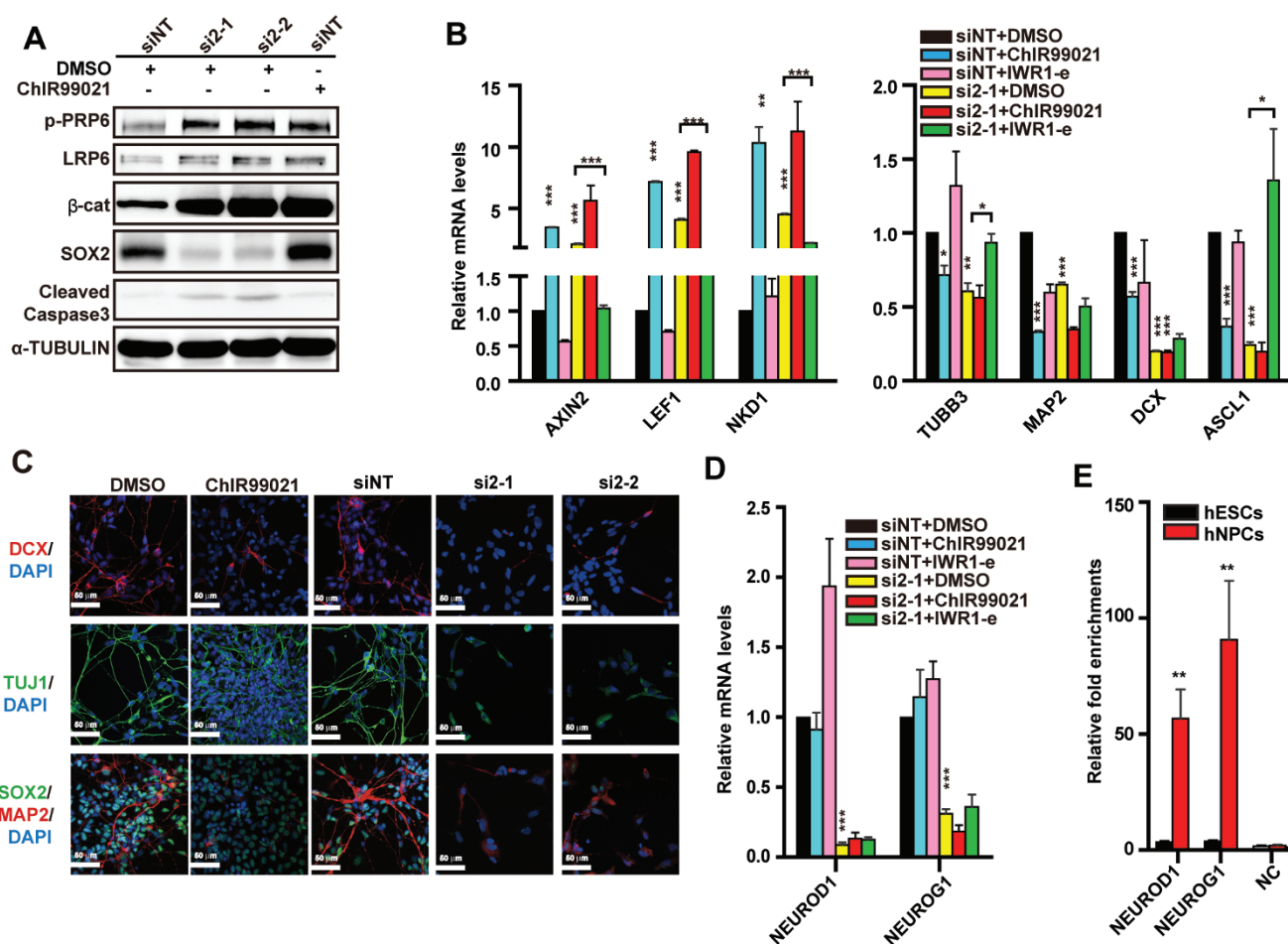


Figure 5 *SOX2* KD leads to impaired neuronal differentiation due to the increased activity of canonical Wnt signaling and reduced expression of pro-neuronal genes. **(A)** Western blot for p-LRP6, LRP6 and active caspase3 on day 3 of neuronal differentiation from hNPCs. *SOX2* KD was conducted on day 1. ChIR99021 (3 μ M) was used as a positive control. **(B)** RT-qPCR results of WNT pathway genes (left panel), and neuronal markers (right panel) on day 3 of *SOX2* KD under neuronal differentiation conditions in the presence of DMSO, or IWR-1e (10 μ M), or ChIR99021 (3 μ M). The values in the control (siNT+DMSO) are set as 1. Statistical analyses are made by comparing control and experiment groups. **(C)** Immunostaining of MAP2, TUJ1 and DCX upon ChIR99021 treatment or *SOX2* KD under neuronal differentiation conditions. *SOX2* KD was carried out in the beginning of neuronal differentiation (day 1). IWR1-e and ChIR99021 were added from day 2 to day 7. Immunostaining was performed on day 7. **(D)** RT-qPCR results of pro-neuronal genes (*NEUROD1* and *NEUROG1*) under the conditions in **B**. **(E)** ChIP-qPCR results of *SOX2* enrichments at promoters of *NEUROD1* and *NEUROG1* in hESCs and hNPCs.

unlikely that SOX2 inhibits the Wnt signaling pathway through regulating OCT4 levels.

Taken together, these results illustrate how a core pluripotency factor suppresses alternative lineages to prime differentiation toward its preferred direction. In this case, SOX2 suppresses non-neural lineages via the control of canonical Wnt signaling, thus acting as a neural master regulator. More importantly, such inhibition is not dependent on the level of OCT4, suggesting a unique role of SOX2 in hESCs.

SOX2 inhibits canonical Wnt signaling and activates pro-neuronal genes for the initiation of neuronal differentiation in hNPCs

We next examined the role of canonical Wnt signaling in *SOX2* KD-caused impairment of differentiation potential in hNPCs. An increase in canonical Wnt signaling activity was detected upon *SOX2* KD under neuronal differentiation conditions. Protein levels of p-LRP6, total LRP6 and β -catenin as well as mRNA levels of Wnt downstream targets *AXIN2*, *LEF1* and *NKDI* quickly increased to the levels comparable to those under conditions of Wnt signaling activation by ChIR99021 (Figure 5A and 5B). In addition, ChIR99021 treatment mimicked *SOX2* KD in reducing expression of pan-neuronal markers *TUBB3*, *MAP2*, *DCX* and *ASCL1* (Figure 5B). Conversely, inhibition of Wnt signaling by IWR1-e not only attenuated the *SOX2* KD-mediated increase in Wnt signaling activity, but also restored the expression of *ASCL1* and *TUBB3* significantly (Figure 5B). Moreover, in line with the gene expression patterns, *SOX2* KD and ChIR99021 treatment both impaired neuronal differentiation, especially cellular projection as shown by immunostaining of *DCX*, *MAP2* and *TUJ1* (encoded by *TUBB3*) (Figure 5C). These results indicate that enhanced canonical Wnt signaling may be partially responsible for neuronal differentiation defects induced by *SOX2* KD and that canonical Wnt signaling needs to be silenced when neuronal differentiation begins in hNPCs. The inhibitory role of canonical Wnt signaling in neuronal differentiation in hNPCs is in agreement with the finding by Bengoa *et al.* [17], but is in contrast to the inductive role of Wnt signaling in mouse adult neurogenesis [10], possibly reflecting a species-specific function of the pathway in neuronal differentiation. However, the reduced cell number due to massive cell death, as indicated by increased levels of active caspase3 (Figure 5A), was only seen after *SOX2* KD but not after ChIR99021 treatment, suggesting the presence of additional mechanism underlying *SOX2*'s role in supporting cell survival during neuronal differentiation. We found that ChIR99021 failed to mimic *SOX2* KD in downregulating pro-neuronal genes

NEUROD1 and *NEUROG1* (Figure 5D). ChIP-qPCR results revealed that SOX2 directly bound *NEUROD1* and *NEUROG1* promoters in hNPCs (Figure 5E). Therefore, SOX2 may directly control neuronal differentiation of hNPCs through both Wnt signaling-dependent and Wnt signaling-independent mechanisms.

SOX2 inhibits the Wnt pathway by transcriptional control of Wnt modulators SFRP2 and WLS in hESCs

To pinpoint the molecular mechanism by which SOX2 represses the Wnt pathway, we searched for SOX2's direct targets that regulate the Wnt pathway. This led to the identification of two Wnt signaling modulators, SFRP2 and WLS, whose promoter regions are occupied by SOX2 in both hESCs and hNPCs (Figure 6A). SFRP2 belongs to the SFRP family of Wnt antagonists [39], while WLS is a Wnt-trafficking protein responsible for transporting Wnt proteins to the cell membrane for subsequent release [40]. The mRNA levels and protein levels of SFRP2 and WLS changed quickly upon *SOX2/3* KD in hESCs, with SFRP2 decreasing on day 1 and WLS increasing on day 2 (Figure 6B and 6C). The increase in the *WLS* level was the most significant among all Wnt pathway components examined after *SOX2/3* KD (Figure 6B). In addition, IWR1-e abrogated the effect of *SOX2/3* KD on the expression of *WLS* but not *SFRP2* (Figure 6C), suggesting that *WLS*, but not *SFRP2*, was also subject to the regulation of Wnt signaling. On the other hand, *SOX2* KD in hNPCs reduced the expression of *SFRP2* but did not affect the expression of *WLS* (Supplementary information, Figure S4A). These results suggest a direct transcriptional control of *SFRP2* in both hESCs and hNPCs and a hESCs-specific functional control of *WLS*. ChIP-qPCR verified the existence of a SOX2 peak on the *SFRP2* promoter and two SOX2 peaks on the *WLS* promoter (Supplementary information, Figure S4B and S4C), which are 1 kb apart and henceforth named WLS-1 and WLS-2 (Figure 6A).

We next sought to compare the contribution of SFRP2 and WLS in *SOX2/3* KD-caused differentiation of hESCs. First, we clarified the relationship between SFRP2 and the Wnt pathway in hESCs, since SFRP2 was reported to be an agonist of the Wnt pathway under certain conditions [41, 42]. Inhibition of SFRP2 by either siRNA or a neutralizing antibody markedly enhanced WNT3A-caused upregulation of mesendoderm markers (Supplementary information, Figure S4D and S4E), indicating an antagonistic role of SFRP2 in Wnt signaling in hESCs. Second, we investigated their rescue effect upon *SOX2/3* KD in hESCs. *WLS* KD more effectively repressed the upregulation of markers caused by *SOX2/3* KD, including *SNAI2*, *TWIST1*, *CDX2*, *EOMES*, *GATA4*

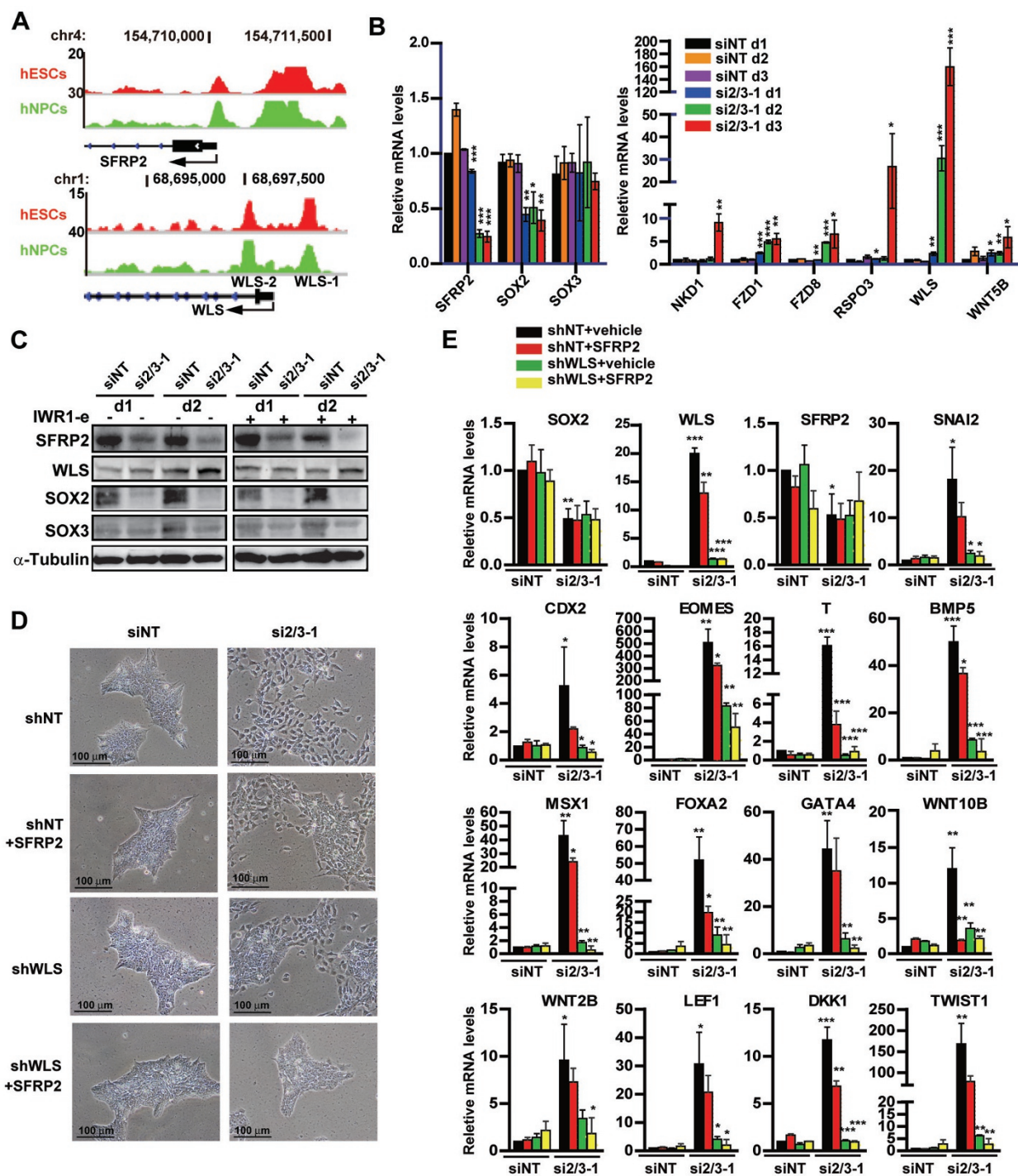


Figure 6 WLS and SFRP2 are major transcriptional targets of SOX2, and mediate SOX2's function in the control of Wnt signaling in hESCs. **(A)** Integrative Genomics Viewer screenshots showing common SOX2 peaks on the promoters of *SFRP2* and *WLS* in hESCs and hNPCs. Two WLS peaks are named WLS-1 and WLS-2. **(B)** RT-qPCR results of *SFRP2* and *WLS* along with other Wnt pathway components at day 1, 2 and 3 after *SOX2/3* KD in hESCs. **(C)** Western blot results of *SFRP2* and *WLS* at day 1 and day 2 after *SOX2/3* KD in hESCs with or without IWR1-e (10 μ M). **(D)** Morphology of H9 hESCs infected with lentiviral *shRNA* (*shNT* or *shWLS*) in combination with NT or *SOX2/3* KD using siRNA (*siNT* or *si2/3-1*). Recombinant *SFRP2* was supplemented at 200 ng/ml. **(E)** RT-qPCR results of germ layer markers and Wnt signaling components in cells from **D**. Statistical analysis is performed as follows: the 3 rescue approaches within the *si2/3-1* group are compared with the rescue effect of the vehicle (*shNT*+vehicle); the *shNT*+vehicle in *si2/3-1* group is compared with the *shNT*+vehicle in the *siNT* group, whose values are set as 1.

and *LEF1*, while recombinant SFRP2 partially rescued the cell morphology and mitigated the increase of many mesoderm and endoderm markers such as *T*, *FOXA2*, *Wnt10B* and *MSX1* (Figure 6D and 6E). Notably, the combination of SFRP2 supplementation and *WLS* KD significantly rescued *SOX2/3* KD phenotypes in both cell morphology and marker gene expression. These findings were reproducible with two sets of siRNAs targeting *SOX2/3* and two shRNAs targeting *WLS* (Supplementary information, Figure S4F). Thus, we conclude that SFRP2 and *WLS* are two major downstream targets of SOX2 and they mediate SOX2's inhibitory function in non-neural lineage specification in hESCs.

SOX2 poises developmental genes together with H2A.Z and PRC2 in hESCs

To deepen our understanding of mechanisms underlying SOX2's function and particularly address the question of how SOX2 poises the expression of developmental genes such as *WLS*, we performed immunoprecipitation of endogenous SOX2-containing protein complexes in both hESCs and hNPCs, and mass spectrometric analyses. Altogether, we detected 68 and 191 potential SOX2-interacting partners in hESCs and hNPCs, respectively (Figure 7A and Supplementary information, Tables S1 and S2). Twenty-eight proteins were shared between the two cell types, and many of them are important splicing factors, including pre-mRNA splicing factor PRPF8 and several heterogeneous nuclear ribonucleoproteins (HnRNPs) (Supplementary information, Table S1 and S2), which were also identified in another SOX2 interactome study in human glioma [43]. Moreover, many RNA helicases were identified as SOX2 partners in hNPCs (Figure 7A). Furthermore, we detected important chromatin regulators such as BPTF, SIRT6 and HIRP3 (Supplementary information, Tables S1 and S2). Strikingly, many histone variants were identified, including H2A.Z.

Our SOX2 protein interactome reveals important aspects of SOX2's participation in the regulation of gene expression programs, cell proliferation and survival. In this study we chose to focus on the interaction between SOX2 and H2A.Z in hESCs, since both were reported to co-localize extensively with PRC2 on developmental genes. Specific interaction between endogenous SOX2 and H2A.Z was validated by reciprocal co-immunoprecipitation assays (Figure 7B). In addition, we found that H2A.Z but not SOX2 interacted with PRC2 components EZH2, JARID2 and SUZ12 in hESCs (Figure 7B). To learn how the interaction between SOX2 and H2A.Z might affect the outcome of gene transcription, we examined whether H2A.Z would be enriched on the SOX2 targets and whether *SOX2/3* KD would affect the

occupancy of H2A.Z on these sites. ChIP-qPCR detected a selective enrichment of H2A.Z on *WLS-2*, but not on *WLS-1*, and the enrichment decreased upon *SOX2/3* KD (Figure 7C). Moreover, H2A.Z and SOX2 co-localized on the promoter of another SOX2 target *Wnt5B* and the H2A.Z occupancy was dependent on the expression of SOX2 (Figure 7C and Supplementary information, Figure S5A). Thus, sequence-specific SOX2 binding to developmental genes probably occurs prior to H2A.Z incorporation into chromatin.

Given the interaction and co-occupancy between H2A.Z and PRC2, we anticipated that *SOX2/3* KD could also affect recruitment of PRC2 to the SOX2-H2A.Z-occupying developmental genes. Indeed, *SOX2/3* KD reduced the recruitment of EZH2 to the SOX2-H2A.Z co-occupied sites within the promoters of *WLS* and *Wnt5B* (Figure 7D), with a concomitant drastic decrease in the enrichment of the repressive histone marker H3K27me3 and a significantly elevated ratio of H3K4me3/H3K27me3 on their originally bivalently modified promoters (Figure 7E and 7F, Supplementary information, Figure S5B and S5C). In contrast, histone acetylation levels were low and unaffected at these sites. Notably, such reconstruction of the epigenetic signature on the *WLS* promoter following *SOX2/3* KD was not confined to the *WLS-2* site, but extended to the adjacent *WLS-1* site (Supplementary information, Figure S5D and S5E). For a negative control, the *SFRP2* promoter was solely occupied by H3K4me3 and unaffected by *SOX2/3* KD (Supplementary information, Figure S5F).

To investigate the relationship between SOX2 and H2A.Z on a genome-wide scale, we integrated a published H2A.Z ChIP-seq data in hESCs [26] with our SOX2 ChIP-seq data, and found that 12% of SOX2-bound sites showed significant enrichment of H2A.Z (Figure 7G). Comparing with other SOX2-binding sites, these sites were preferentially localized in promoter regions (Figure 7H). Of all the co-occupied genes, 156 genes were downregulated and 205 were upregulated upon *SOX2/3* KD. The upregulated genes, including *WLS*, were involved in important developmental processes such as brain, muscle and cardiovascular system development (Figure 7I), while the biological functions of the downregulated genes were not enriched in any specific biological processes. A close inspection of ChIP-seq data revealed co-localization of SOX2 and H2A.Z on *WLS-2* but not *WLS-1* (Figure 7J) and the co-localized region was bivalently modified by both H3K4me3 and H3K27me3, supporting the finding in our ChIP-qPCR assays. Such co-localization was also seen on other important developmental genes such as *PPP1R9B*, *GAS7*, *ATP6V1B1* and *EYAI* (Supplementary information, Fig-

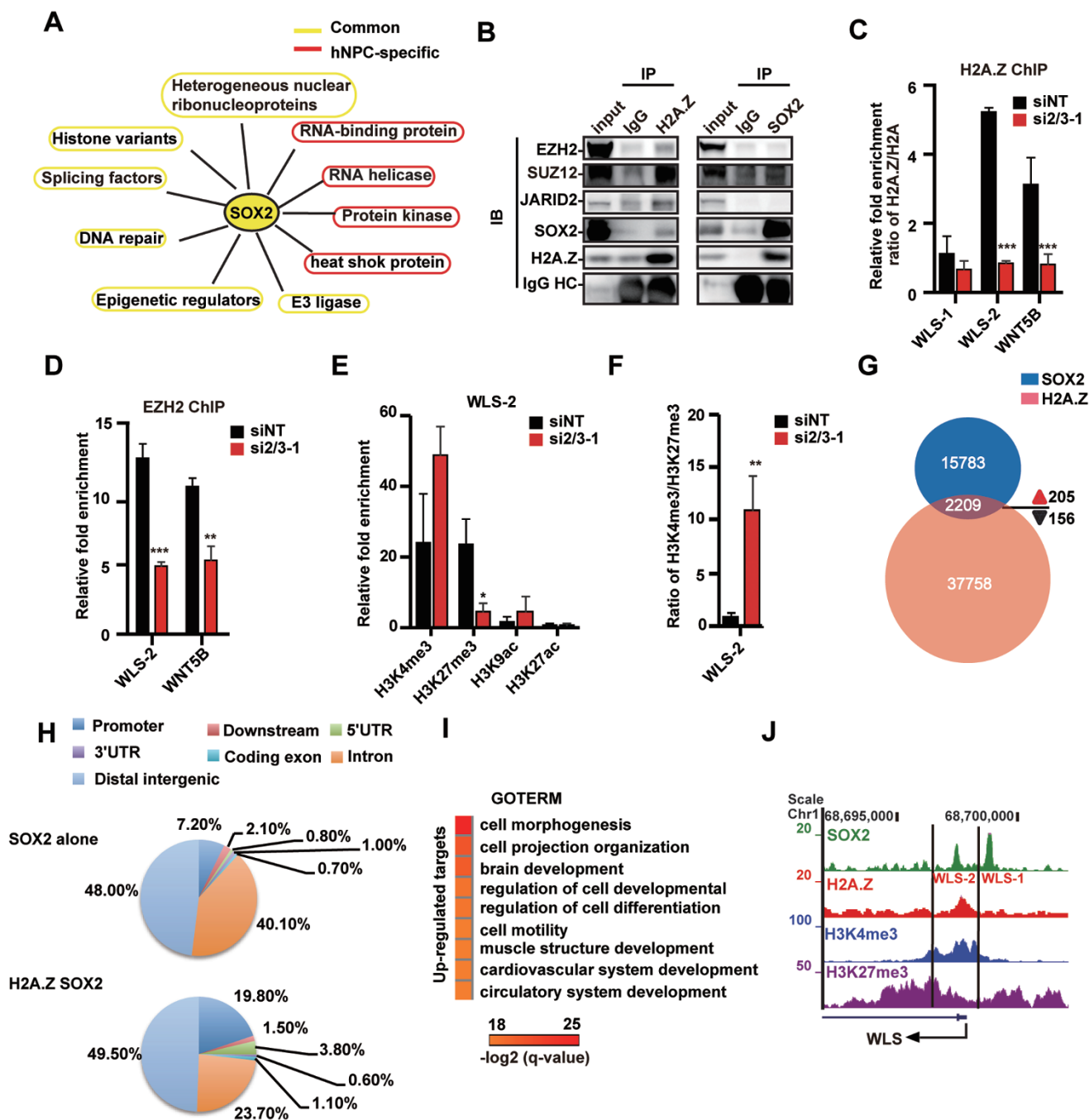


Figure 7 SOX2 interacts with H2A.Z to recruit PRC2 and poise developmental genes in hESCs. **(A)** SOX2 protein partners in hESCs and hNPCs are classified into 10 categories. Protein partners common in the two cell types or specific to one of the cell types are shown. **(B)** Co-immunoprecipitation results of the interaction between H2A.Z, SOX2 and PRC2 components (EZH2, JARID2 and SUZ12) in hESCs. IB, immunoblotting; HC, immunoglobulin heavy chain. **(C, D)** ChIP-qPCR for H2A.Z **(C)** and EZH2 **(D)** binding at promoters of *WLS* and *WNT5B* in NT and *SOX2/3* KD H9 hESCs. **(E, F)** ChIP-qPCR analysis of histone modification enrichments **(E)** and the ratio of H3K4me3/H3K27me3 at the *WLS-2* site **(F)** in NT and *SOX2/3* KD hESCs. **(G)** A Venn diagram showing the overlap between target regions of SOX2 and H2A.Z in hESCs. Numbers of indicated peaks and altered genes upon *SOX2/3* KD are shown. **(H)** The genomic distribution of SOX2 alone sites and SOX2-H2A.Z co-localized sites. **(I)** GO analysis of common targets of SOX2 and H2A.Z that are upregulated upon *SOX2/3* KD in hESCs. Benjamini-Hochberg method is applied to adjust the *P*-values in order to account for multiple testing. Enrichment levels of selected GO terms are marked by $-\log_2(q\text{-value})$. **(J)** Integrative Genomics Viewer screenshots showing enrichments of H3K4me3, H3K27me3, H2A.Z and SOX2 on the promoter of *WLS*.

ure S5G). Based on these data we suggest that SOX2 exerts its regulatory function by recruiting PRC2 through H2A.Z as an intermediary.

Conserved and divergent transcription programs governed by SOX2 in human and mouse

We next used published Sox2 ChIP-seq data from mESCs and mNPCs [7] and data generated from hESCs and hNPCs in this study to compare the regulatory network of SOX2 between mouse and human and found some species-specific features. For example, common SOX2 PBGs for hESCs and hNPCs were most significantly enriched in RNA metabolic processes, RNA splicing as well as chromatin organization (Figure 1D). Examples of such common PBGs included many HNRNPs, known to affect various aspects of mRNA processing such as alternative splicing, metabolism and mRNA transport, and chromatin regulators such as HELLS (Figure 8A). Our SOX2 protein interactome in hESCs and hNPCs further supported the participation of SOX2 in RNA splicing and chromatin organization in both hESCs and hNPCs. In contrast, the study in mouse cells found RNA splicing to be a mNPC-specific feature of Sox2 PBGs, but not for mESCs, and chromatin organization to be a mESC-specific feature, but not for mNPCs [7]. In addition, we also found differences in the regulation of signaling pathways. For example, Wnt signaling was highly enriched in poised SOX2 EBGs and upregulated genes after SOX2 KD in hESCs and hNPCs (Figures 1E and 3C). Examples include *Wnt10B*, an hESC-specific EBG and *FZD10*, a common EBG for hESCs and hNPCs (Figure 8B). Their expression increased upon *SOX2/3* KD in hESCs (Figure 4E). However, many SOX2 EBGs involved in Wnt signaling were found to be active in mESCs and mNPCs [7], revealing the dynamic relationship between SOX2 and Wnt signaling between species.

To further compare the species-common and species-specific transcription programs, we defined the species-common peaks with 10% overlap as retained peaks, peaks shifted within the range of 10 kb as turnover peaks, and the rest as unique peaks (Figure 8C). Interestingly, hESCs and hNPCs had considerably more unique peaks whereas mESCs and mNPCs contained a higher percentage of retained peaks. Nevertheless, the majority of peaks were unique in both species (Figure 8C), reflecting the low conservation in SOX2-binding profiles between the two species. Moreover, genomic distribution analyses revealed that the percentage of promoter peaks was highest in the retained group especially for NPCs (Supplementary information, Figure S6A). Also, DNA sequences of the retained peaks were the most conserved (Supplementary information, Figure S6B), indicating a relatively

conserved promoter-binding profile of SOX2 in the same cell type between species. In contrast, almost all unique peaks, preferentially located to the gene bodies and intergenic regions. Such genomic distribution for retained and unique SOX2 peaks likely reflects the conservation of regulatory mechanism through promoter binding in maintaining stem cell identity and species-specific mechanisms utilizing distal regulatory elements to regulate developmental programs. This notion is consistent with our findings that promoter binding by SOX2 is associated with fundamental biological processes of protecting the stem cell identity and that promoter distal binding correlates with tissue-specific gene regulation.

In addition, downregulated retained genes upon *SOX2/3* KD in hESCs were associated with negative regulation of gene expression and differentiation, while the downregulated unique genes in hESCs were enriched for terms of spindle assembly, membrane potential, transport system and energy metabolism (Figure 8D). On the other hand, while the upregulated retained genes contained conserved targets in regulating trophoblast differentiation and morphogenesis of mesoderm- and endoderm-derived organs, the upregulated unique genes contained considerably more hESC-specific targets in controlling the morphogenesis of the same organs (Figure 8E). These results are in agreement with SOX2-regulated hESC-specific transcription programs generated through our ChIP-seq and RNA-seq data analyses. Notably, upregulated genes in both retained and unique groups were enriched for Wnt signaling, indicating common and specific targets between species in regulating the same process. Similar species-specific transcriptional targets were also seen in the regulation of Notch signaling. In mouse retinal progenitor cells, Sox2 was reported to positively regulate Notch pathway through *Notch1* [44]. Such regulation of *NOTCH1* was not observed in hESCs and hNPCs. Instead, SOX2 controls the Notch pathway through multiple other core components of the pathway (Figure 2G). Most of the overrepresented terms listed above were not enriched by the target genes of turnover peaks. Together, these results demonstrate the species-dependent functions of SOX2, arguing for the importance of studying SOX2's function and underlying mechanisms in human cells.

Discussion

In this study, we combined functional studies with analyses of genome-wide data produced through RNA-seq, ChIP-seq and protein complex identification, to probe the mechanisms underlying SOX2's distinct functions in stem cell maintenance and neural develop-

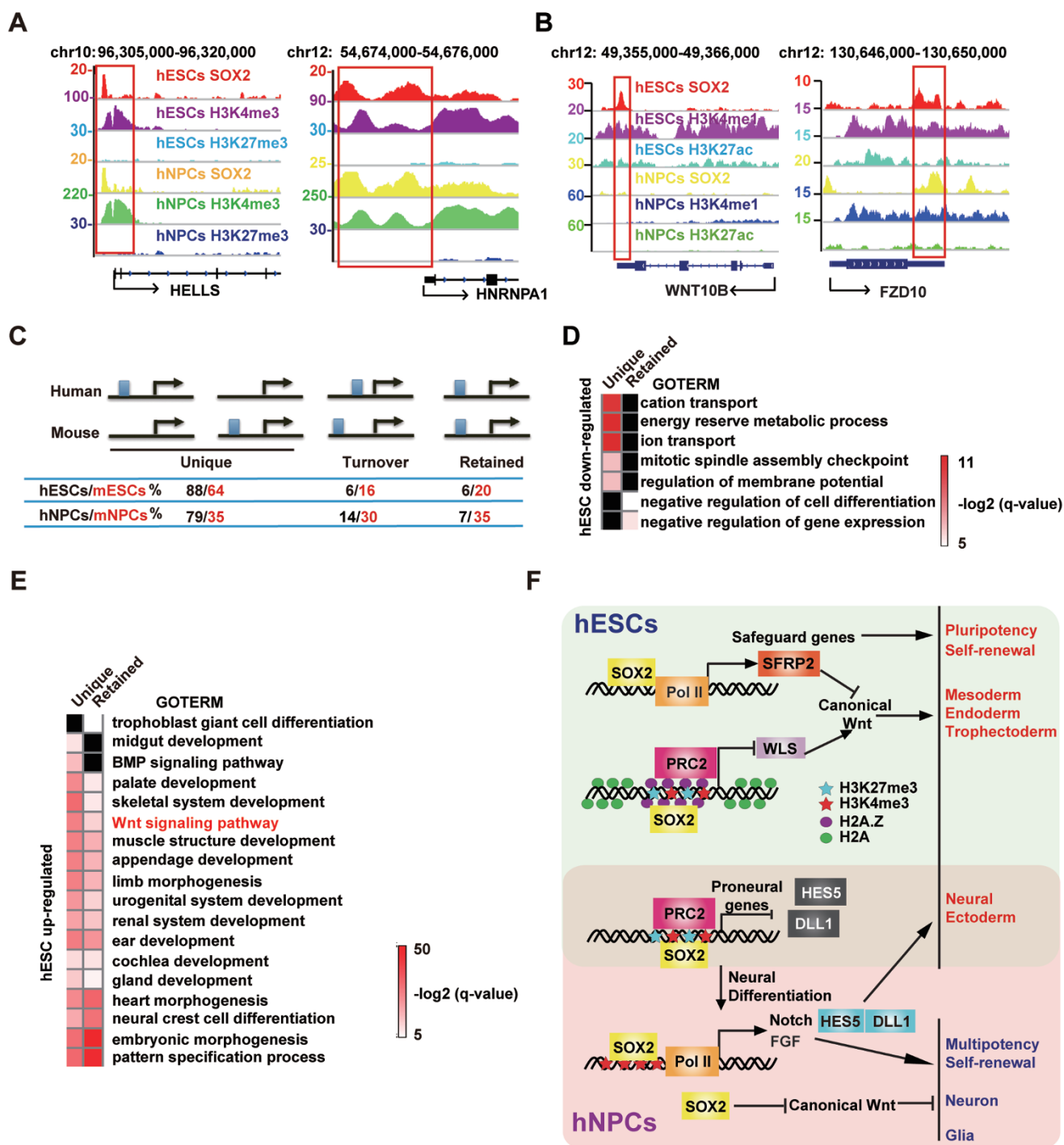


Figure 8 Comparative analysis between mouse and human SOX2 ChIP-seq reveals conserved and divergent regulation. **(A)** Integrative Genomics Viewer screenshots of common SOX2 peaks on the active promoters of *HELLS* and *HNRNPA1*. **(B)** Integrative Genomics Viewer screenshots of common SOX2 peaks on the poised enhancer of *FZD10* and the hESC-specific peak on the poised enhancer of *Wnt10B*. **(C)** A schematic illustration of the classification of retained, turned over and unique SOX2 peaks in ESCs and NPCs of the human and mouse origin. The percentages of corresponding peaks are indicated. **(D, E)** GO analyses of genes downregulated **(D)** and upregulated **(E)** upon *SOX2/3* KD in the retained and unique groups. Benjamini-Hochberg method is applied to adjust the *P*-values in order to account for multiple testing. Enrichment levels of selected GO terms are marked by $-\log_2(q\text{-value})$. Missing values are marked in black. **(F)** A working model showing the cell type- and stage-dependent functions of SOX2.

ment. Our findings provide experimental evidence for the notion that master transcription factors specify cell fates through controlling distinct signaling pathways and recruiting histone variants to modify the chromatin state.

SOX2 suppresses the non-neural lineage through inhibiting the Wnt pathway in hESCs. WLS and SFRP2, two important regulators of Wnt signaling, are novel direct targets of SOX2 and mediate SOX2's function in hESCs. In fact, both Wls and Sfrp2 are needed for the axis formation in mouse, although functional redundancy between SFRP family members exists [45, 46]. In human, both WLS and SFRP2 have been implicated in various cancers. In particular, concomitant epigenetic silencing of *SOX1* and *SFRPs* was reported in hepatocellular carcinoma [47-50]. It would be interesting to test whether SOX2 contributes to tumor development through transcriptionally controlling *WLS* and *SFRP2* expression. The function of these two proteins in cell fate commitment of hESC was previously unknown. We show that the expression of *SFRP2* is abundant in hESCs and decreases quickly upon differentiation, suggesting its important role in maintaining hESC self-renewal. Moreover, the reduction of *SFRP2* levels or increase of *WLS* levels is able to induce differentiation of hESCs into non-neural lineages. Thus, WLS and SFRP2 are important developmental regulators that can specify the cell fate of hESCs, consistent with the *in vivo* evidence.

Apart from transcriptional control, the role of pluripotency factors in shaping pluripotent epigenetic landscapes and chromatin organization is particularly intriguing. In the search for protein partners of SOX2 in hESCs, we uncovered a specific interaction between SOX2 and H2A.Z. H2A.Z was initially reported to co-occupy promoters of developmentally important genes with PRC2 component Suz12 in mESCs and repress the expression of these genes [51]. Moreover, H2A.Z was reported to be required for efficient binding of Oct4 to its targets in mESCs [24] and to provide a binding platform for pioneer factor Foxa2 to bind the promoter of genes activated during endoderm differentiation [22, 52]. Another histone variant H2A.X was also showed to occupy Cdx2-binding sites on extraembryonic genes and to repress their induction in mESCs and mouse induced pluripotent stem cells [53]. These results suggest that histone variants could either facilitate or inhibit transcription factor binding to DNA. In contrast, our results lead to a different conclusion that the occupancy of H2A.Z, the subsequent binding of PRC2 and generation of H3K27me3 on important developmental genes all depend on the binding of SOX2 to its specific targets. This model could also be applied to other pioneer transcription factors. Our data raise the following two hypotheses: i) SOX2 may facilitate H2A.

Z incorporation into specific chromatin positions of developmental genes; ii) H2A.Z may serve as a bridge to link SOX2 and PRC2. However, we do not exclude the possibility of the interplay between H2A.Z and pioneer factors in generating an active epigenetic state and open chromatin structures in a context-dependent manner. Recently, H2A.Z has been reported to interact with Nanog, another key pluripotency factor, and to stabilize Nanog protein levels in mESCs. Moreover, KD of *H2A.Z* induces differentiation of mESCs [54]. Further studies of pioneer factors, histone variants and histone modifications will help illustrate the mechanistic interplay between them.

Our ChIP-seq analyses detected considerably more peaks in hNPCs than in mNPCs. Another group who conducted SOX2 ChIP-seq in hNPCs using the same antibody obtained a similar result of 55 771 peaks; in contrast, a small number of altered genes upon *SOX2* KD were detected in hNPCs [55]. This could be due to the functional redundancy of SoxB1 family members. On the other hand, our study detected vast differences between mouse and human in ESCs and NPCs, including the species-specific targets of SOX2 and the dynamic role of Wnt signaling in stem cell maintenance. Given that mouse and human ESCs are thought to represent the naive state and prime state respectively, it is interesting to speculate that these differences may also represent SOX2's stage-dependent regulation and interplay with important signaling pathways.

Collectively, the comprehensive and systematic study of SOX2 in this study produces the following working model (Figure 8F): SOX2 binds and positively regulates genes required for the maintenance of the identity of hESCs. Simultaneously, it associates with histone variant H2A.Z and PRC2 on a set of important developmental genes including *WLS* to repress Wnt signaling and non-neural lineage genes through coordination of chromatin structures in hESCs. Moreover, SOX2 occupies proneural genes such as Notch signaling components in hESCs and turns them on in response to neural induction cues. In hNPCs, SOX2 promotes expression of FGF and Notch signaling components to sustain the status of NPCs. At last, SOX2 is also required to turn down the canonical Wnt pathway for appropriate initiation of neuronal differentiation in the presence of neuronal induction cues.

Further analysis of our comprehensive dataset will likely reveal more underappreciated roles and molecular mechanisms of SOX2 in cell fate specification and address how the multi-level coordination between SOX2 and chromatin regulators shapes the pluripotent landscape and how SOX2 participates in the regulation of

cell type-specific RNA splicing. Considering the unique role of SOX2 in pluripotency and neural differentiation as well as its close relationship with many human diseases, we expect that the rich information obtained from this study will greatly facilitate the understanding of SOX2's molecular regulation in cell fate determination.

Materials and Methods

Cell culture

hESC lines (H9 and SHhES2) were maintained in mTeSR1 medium (Stemcell Technologies) in Matrigel (BD Biosciences)-coated dishes. hNPCs were maintained in N2B27 medium (1:1 mixture of DMEM/F12 and neurobasal medium supplemented with N2 and B27 supplement, Gibco/Life Technologies) containing 10 ng/ml bFGF. The medium was changed daily. hESCs were routinely passaged with dispase (1 mg/ml; Sigma-Aldrich) or accutase (Stemcell Technologies), and hNPCs were passaged with 0.05% trypsin. All KD and rescue experiments involving hESCs were conducted in the mTeSR1 medium.

Neural and neuronal induction

Differentiation toward hNPCs from hESCs was carried out as previously described [27-29]. Briefly, we used noggin (500 ng/ml, R&D Systems), SB431542 (10 μ M, Tocris Bioscience) and compound C (10 μ M, Sigma-Aldrich) to induce neural differentiation. Neural rosettes were picked up and replated onto Matrigel-coated dishes, and then dissociated into single NPCs. Neuronal induction was carried out in the neurobasal medium (Gibco/Life Technologies) supplemented with B27 (Gibco/Life Technologies, 1:100), 10 ng/ml BDNF, GDNF, and CTNF (all from R&D Systems), 50 μ M β -mercaptoethanol.

Statistical analysis

Student's *t* test was used to assess statistical significance. **P* < 0.05; ***P* < 0.01, ****P* < 0.001. In all graphs, error bars represent standard deviation (SD) values. All statistically analyzed data were obtained from at least three independent experiments.

For detailed Materials and Methods see Supplementary information, Data S1.

Accession numbers

RNA-seq and ChIP-seq data have been deposited at the Gene Expression Omnibus (GEO) database under accession numbers GSE69476 and GSE69479.

Acknowledgments

This study was supported by grants from the Chinese Academy of Science (XDA01010102), Ministry of Science and Technology of China (2011CB965101, 2013CB967101) and the National Natural Science Foundation (31471393, 91419309, 31322031, 31371288 and 31571365).

Author Contributions

CZ and YJ designed the study, interpreted data and wrote the manuscript; CZ, YS and HY performed experiments and analyzed data; the bioinformatics analysis workflow was engineered by XY

and YZ, and performed by XY.

Competing Financial Interests

The authors declare no competing financial interests.

References

- 1 Sarkar A, Hochedlinger K. The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* 2013; **12**:15-30.
- 2 Graham V, Khudyakov J, Ellis P, Pevny L. SOX2 functions to maintain neural progenitor identity. *Neuron* 2003; **39**:749-765.
- 3 Adachi K, Suemori H, Yasuda SY, Nakatsuji N, Kawase E. Role of SOX2 in maintaining pluripotency of human embryonic stem cells. *Genes Cells* 2010; **15**:455-470.
- 4 Fong H, Hohenstein, KA, Donovan PJ. Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. *Stem Cells* 2008; **26**:1931-1938.
- 5 Wang Z, Oron E, Nelson B, Razis S, Ivanova N. Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell* 2012; **10**:440-454.
- 6 Cimadamore F, Amador-Arjona A, Chen C, Huang CT, Terskikh AV. SOX2-LIN28/let-7 pathway regulates proliferation and neurogenesis in neural precursors. *Proc Natl Acad Sci USA* 2013; **110**:E3017-E3026.
- 7 Lodato MA, Ng CW, Wamstad JA, *et al.* SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. *PLoS Genet* 2013; **9**:e1003288.
- 8 Mistri TK, Devasia AG, Chu LT, *et al.* Selective influence of Sox2 on POU transcription factor binding in embryonic and neural stem cells. *EMBO Rep* 2015; **16**:1177-1191.
- 9 Davidson KC, Jamshidi P, Daly R, Hearn MT, Pera MF, Dottori M. Wnt3a regulates survival, expansion, and maintenance of neural progenitors derived from human embryonic stem cells. *Mol Cell Neurosci* 2007; **36**:408-415.
- 10 Vanderhaeghen P. Wnts blow on NeuroD1 to promote adult neuron production and diversity. *Nat Neurosci* 2009; **12**:1079-1081.
- 11 Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH. Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat Med* 2004; **10**:55-63.
- 12 Ullmann U, Gilles C, De Rycke M, Van de Velde H, Sermon K, Liebaers I. GSK-3-specific inhibitor-supplemented hESC medium prevents the epithelial-mesenchymal transition process and the up-regulation of matrix metalloproteinases in hESCs cultured in feeder-free conditions. *Mol Hum Reprod* 2008; **14**:169-179.
- 13 Davidson KC, Adams AM, Goodson JM, *et al.* Wnt/beta-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. *Proc Natl Acad Sci USA* 2012; **109**:4485-4490.
- 14 Singh AM, Reynolds D, Cliff T, *et al.* Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell Stem Cell* 2012; **10**:312-326.
- 15 Kurek D, Neagu A, Tastemel M, *et al.* Endogenous WNT sig-

- nals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Reports* 2015; **4**:114-128.
- 16 Moya N, Cutts J, Gaasterland T, Willert K, Brafman DA. Endogenous WNT signaling regulates hPSC-derived neural progenitor cell heterogeneity and specifies their regional identity. *Stem Cell Reports* 2014; **3**:1015-1028.
- 17 Bengoa-Vergniory N, Gorrono-Etxebarria I, Gonzalez-Salazar I, Kypta RM. A switch from canonical to noncanonical Wnt signaling mediates early differentiation of human neural stem cells. *Stem Cells* 2014; **32**:3196-3208.
- 18 Agathocleous M, Iordanova I, Willardsen MI, *et al.* A directional Wnt/beta-catenin-Sox2-proneural pathway regulates the transition from proliferation to differentiation in the *Xenopus* retina. *Development* 2009; **136**:3289-3299.
- 19 Mendenhall EM, Bernstein BE. Chromatin state maps: new technologies, new insights. *Curr Opin Genet Dev* 2008; **18**:109-115.
- 20 Meshorer E, Misteli T. Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* 2006; **7**:540-546.
- 21 Lee TI, Jenner RG, Boyer LA, *et al.* Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 2006; **125**:301-313.
- 22 Subramanian V, Fields PA, Boyer LA. H2A.Z: a molecular rheostat for transcriptional control. *F1000Prime Rep* 2015; **7**:01.
- 23 Chen P, Wang Y, Li G. Dynamics of histone variant H3.3 and its coregulation with H2A.Z at enhancers and promoters. *Nucleus* 2014; **5**:21-7.
- 24 Hu G, Cui K, Northrup D, *et al.* H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* 2013; **12**:180-192.
- 25 Jin C, Zang C, Wei G, *et al.* H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* 2009; **41**:941-945.
- 26 Ku M, Jaffe JD, Koche RP, *et al.* H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. *Genome Biol* 2012; **13**:R85.
- 27 Zhou J, Su P, Li D, Tsang S, Duan E, Wang F. High-efficiency induction of neural conversion in human ESCs and human induced pluripotent stem cells with a single chemical inhibitor of transforming growth factor beta superfamily receptors. *Stem Cells* 2010; **28**:1741-1750.
- 28 Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol* 2009; **27**:275-280.
- 29 Zhou Y, Jiang H, Gu J, Tang Y, Shen N, Jin Y. MicroRNA-195 targets ADP-ribosylation factor-like protein 2 to induce apoptosis in human embryonic stem cell-derived neural progenitor cells. *Cell Death Dis* 2013; **4**:e695.
- 30 Ziller MJ, Edri R, Yaffe Y, *et al.* Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* 2015; **518**:355-359.
- 31 Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet* 2011; **12**:283-293.
- 32 Thomson JA, Itskovitz-Eldor J, Shapiro SS, *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* 1998; **282**:1145-1147.
- 33 Li C, Yang Y, Lu X, *et al.* Efficient derivation of Chinese human embryonic stem cell lines from frozen embryos. *In Vitro Cell Dev Biol Anim* 2010; **46**:186-191.
- 34 Carey BW, Finley LW, Cross JR, Allis CD, Thompson CB. Intracellular alpha-ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* 2015; **518**:413-416.
- 35 Zhang J, Nuebel E, Daley GQ, Koehler CM, Teitell MA. Metabolic regulation in pluripotent stem cells during reprogramming and self-renewal. *Cell Stem Cell* 2012; **11**:589-595.
- 36 Zhang D, Zhao T, Ang HS, *et al.* AMD1 is essential for ESC self-renewal and is translationally down-regulated on differentiation to neural precursor cells. *Genes Dev* 2012; **26**:461-473.
- 37 Assou S, Cerecedo D, Tondeur S, *et al.* A gene expression signature shared by human mature oocytes and embryonic stem cells. *BMC Genomics* 2009; **10**:10.
- 38 Masui S, Nakatake Y, Toyooka Y, *et al.* Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* 2007; **9**:625-635.
- 39 Cruciat CM, Niehrs C. Secreted and transmembrane wnt inhibitors and activators. *Cold Spring Harb Perspect Biol* 2013; **5**:a015081.
- 40 Port F, Kuster M, Herr P, *et al.* Wingless secretion promotes and requires retromer-dependent cycling of Wntless. *Nat Cell Biol* 2008; **10**:178-185.
- 41 Esteve P, Sandonis A, Ibanez C, Shimono A, Guerrero I, Bovolenta P. Secreted frizzled-related proteins are required for Wnt/beta-catenin signalling activation in the vertebrate optic cup. *Development* 2011; **138**:4179-4184.
- 42 Fontenot E, Rossi E, Mumper R, *et al.* A novel monoclonal antibody to secreted frizzled-related protein 2 inhibits tumor growth. *Mol Cancer Ther* 2013; **12**:685-695.
- 43 Fang X, Yoon JG, Li L, *et al.* Landscape of the SOX2 protein-protein interactome. *Proteomics* 2011; **11**:921-934.
- 44 Taranova OV, Magness ST, Fagan BM, *et al.* SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev* 2006; **20**:1187-1202.
- 45 Fu J, Jiang M, Mirando AJ, Yu HM, Hsu W. Reciprocal regulation of Wnt and Gpr177/mouse Wntless is required for embryonic axis formation. *Proc Natl Acad Sci USA* 2009; **106**:18598-18603.
- 46 Satoh W, Gotoh T, Tsunematsu Y, Aizawa S, Shimono A. Sfrp1 and Sfrp2 regulate anteroposterior axis elongation and somite segmentation during mouse embryogenesis. *Development* 2006; **133**:989-999.
- 47 Yang PT, Anastas JN, Toroni RA, *et al.* WLS inhibits melanoma cell proliferation through the beta-catenin signalling pathway and induces spontaneous metastasis. *EMBO Mol Med* 2012; **4**:1294-1307.
- 48 Voloshanenko O, Erdmann G, Dubash TD, *et al.* Wnt secretion is required to maintain high levels of Wnt activity in colon cancer cells. *Nat Commun* 2013; **4**:2610.
- 49 Augustin I, Goidts V, Bongers A, *et al.* The Wnt secretion protein Evi/Gpr177 promotes glioma tumourigenesis. *EMBO*

- Mol Med* 2012; **4**:38-51.
- 50 Shih YL, Hsieh CB, Yan MD, *et al.* Frequent concomitant epigenetic silencing of SOX1 and secreted frizzled-related proteins (SFRPs) in human hepatocellular carcinoma. *J Gastroenterol Hepatol* 2013; **28**:551-559.
- 51 Creighton MP, Markoulaki S, Levine SS, *et al.* H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. *Cell* 2008; **135**:649-661.
- 52 Li Z, Gadue P, Chen K, *et al.* Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. *Cell* 2012; **151**:1608-1616.
- 53 Wu T, Liu Y, Wen D, *et al.* Histone variant H2A.X deposition pattern serves as a functional epigenetic mark for distinguishing the developmental potentials of iPSCs. *Cell Stem Cell* 2014; **15**:281-294.
- 54 Wang J, Qiao M, He Q, *et al.* Pluripotency activity of Nanog requires biochemical stabilization by variant histone protein H2A.Z. *Stem Cells* 2015.
- 55 Ng SY, Bogu GK, Soh BS, Stanton LW. The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis. *Mol Cell* 2013; **51**:349-359.

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