SOX15 regulates stem cell pluripotency and promotes neural fate during differentiation by activating *Hes5*

Eun-Bee Choi[1,2,3,†], Munender Vodnala[1,2,3,†], Prince Saini[1,2,3], Madeleine Zerbato[1,2,3], Jaclyn J. Ho[4,5], Sharath Anugula[1,2,3], Shannan J. Ho Sui[6], Joon Yoon[6], Carla Inouye[4,5], Yick W. Fong[1,2,3,*]

Affiliations

[1]Brigham Regenerative Medicine Center, Brigham and Women's Hospital, Boston, MA, USA.

[2]Department of Medicine, Cardiovascular Medicine Division, Harvard Medical School, Boston, MA, USA.

[3]Harvard Stem Cell Institute, Cambridge, MA, USA

[4]Department of Molecular and Cell Biology, Li Ka Shing Center for Biomedical and Health Sciences, California Institute for Regenerative Medicine Center of Excellence, University of California at Berkeley, Berkeley, CA, USA.

[5]Howard Hughes Medical Institute, Berkeley, CA, USA.

[6]Harvard Chan Bioinformatics Core, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

[†]The authors wish it to be known that, in their opinion, the first 2 authors should be regarded as joint First Authors

[*]To whom correspondence should be addressed. Email: yfong@bwh.harvard.edu

Abstract

SOX2 and SOX15 are Sox family transcription factors enriched in embryonic stem cells (ESCs). The role of SOX2 in activating gene expression programs essential for stem cell self-renewal and acquisition of pluripotency during somatic cell reprogramming is well-documented. However, the contribution of SOX15 to these processes is unclear and often presumed redundant with SOX2 because both transcription factors display similar DNA binding specificity. Here, we show that SOX15 contributes to stem cell maintenance by cooperating with ESC-enriched transcriptional coactivators to ensure optimal expression of pluripotency-associated genes. We demonstrate that SOX15 depletion compromises reprogramming of fibroblasts to pluripotency which cannot be compensated by SOX2. Ectopic expression of SOX15 promotes the reversion of a post-implantation, epiblast stem cell state back to a pre-implantation, ESC-like identity even though SOX2 is expressed in both cell states. We also uncover a role of SOX15 in lineage specification, by showing that loss of SOX15 leads to defects in commitment of ESCs to neural fates. SOX15 promotes neural differentiation by binding to and activating a previously uncharacterized distal enhancer of a key neurogenic regulator, *Hes5*. Together, these findings identify a multifaceted role of SOX15 in induction and maintenance of pluripotency and neural differentiation.

Introduction

Sox family transcription factors (TFs) play essential roles in development (1, 2). Sox TFs are characterized by their high-mobility group (HMG) DNA-binding domain that recognizes highly conserved DNA sequences (3). Despite having similar sequence specificity, Sox TFs are able to choreograph complex and divergent developmental processes by driving specific gene expression programs. In addition to their tissue-specific expression patterns, Sox proteins regulate cell type-specific transcription by leveraging their ability to partner with different TFs and recruit cell-specific transcriptional coactivators (4–6).

In embryonic stem cells (ESCs), SOX2 along with TFs OCT4 and NANOG activates a core regulatory network that is essential for stem cell self-renewal and pluripotency (7–11). SOX2 activates transcription by recruiting OCT4 and stem cell-enriched coactivators to gene enhancers containing sox-oct sequence motifs (12–15). The central role of SOX2 in reprogramming somatic cells to induced pluripotent stem cells (iPSCs) further highlights the importance of SOX2 in pluripotency (16–19). It has been shown that defects in stem cell self-renewal and iPSC generation in the absence of SOX2 can be rescued by ectopic expression of SOX1, SOX3, and SOX15 with varying degree of efficiency (20, 21). However, the physiological relevance of the apparent functional redundancy is unclear, given that these SOX proteins were overexpressed in the rescue experiments, and that, except for SOX15, they are not naturally expressed in ESCs (22). In vitro binding assays indicate that SOX2 and SOX15 display essentially identical binding sequence preferences (22). Furthermore, OCT4 can pair with SOX2 or SOX15 and assemble onto an oct-sox enhancer element in vitro with comparable affinity (23). Another study showed by chromatin immunoprecipitation (ChIP) that several SOX proteins including SOX15 occupy oct-sox enhancers at select pluripotency genes that are also targeted by SOX2 (24). However, it is not known what the relative occupancy of SOX15 vs SOX2 is across the genome and, more

importantly, whether SOX15's interaction with endogenous enhancers is functionally important for transcriptional activation. Therefore, the mechanism by which SOX15 contributes to the pluripotency gene network remains unclear.

When ESCs exit from pluripotency and undergo differentiation, collapse of the pluripotency gene network is accompanied by activation of lineage-specific differentiation programs. Pluripotency-associated TFs have been shown to play an active role in orchestrating this critical transcriptional switch (25). For example, SOX2 is redirected to neurogenic genes to promote neural fate specification (26, 27). The role of SOX15 in lineage specification is unknown.

In this study, we find that SOX2 depletion in ESCs leads to loss of SOX15 expression, raising important questions about whether defects in pluripotency gene expression and self-renewal found in SOX2-deficient ESCs could be due to simultaneous loss of SOX15. We therefore address the function of SOX15 in stem cell self-renewal, induction of pluripotency, and lineage specification. Using complementary in vitro biochemical approaches and genome-wide analyses, we show that SOX15 can partner with OCT4 to directly activate transcription of key pluripotency genes such as *NANOG*, by recruiting ESC-enriched transcriptional coactivators that are also targeted by SOX2. We show that ESCs depleted of SOX15 display defects in self-renewal and expression of pluripotency-associated genes despite having normal levels of SOX2. Likewise, we find that SOX15 depletion blocks reacquisition of pluripotency during cellular reprogramming which cannot be compensated by SOX2. Finally, we demonstrate that in differentiating ESCs, SOX15 is required for neural fate specification via direct binding and activation of a previously uncharacterized enhancer upstream of a key neurogenic gene, *Hes5*. Our findings identify

common mechanisms of transactivation by SOX15 and SOX2 and reveal new non-redundant functions of SOX15 in induction of pluripotency and neural differentiation.

Material and Methods

DNA constructs and antibodies

Complementary DNAs (cDNAs) for human OCT4, SOX2, and SOX15 were obtained from cDNA libraries generated from human embryonic carcinoma cell line NTERA-2 (NT2). Intronless human SOX1 and SOX11 cDNAs were amplified from human genomic DNA. C-terminal FLAG-tagged OCT4 and various SOX TFs were generated by subcloning cDNAs into pFLAG-CMV5a mammalian expression vector (Sigma-Aldrich). For stable expression of mCherry and SOX15 in OEC-2 cells, mCherry and SOX15 cDNA were cloned into lentiviral overexpression vector pHAGE-IRES-Neo (12). Untagged human DKC1 cDNA was subcloned into pHAGE-IRES-Neo. Mammalian expression constructs for untagged XPC-RAD23B-CETN2 complex (pHAGE-XPC.com) (12) and V5-tagged ABCF1 (pFLAG-CMV5a-ABCF1-V5) (15) were described. Note that the V5 tag followed by a translation stop codon was inserted upstream of the built-in C-terminal FLAG tag in pFLAG-CMV5a vector. pLKO lentiviral vectors expressing non-targeting shRNA and shRNAs targeting SOX2 and SOX15 were purchased from Sigma-Aldrich.

Polyclonal antibodies against SOX2 were purchased from EMD Millipore (AB5603); XPC (A301-122A) from Bethyl Laboratories; DKC1 (H-300) and OCT4 (N-19) from Santa Cruz Biotechnology; SOX15 (25415-1-AP) from ProteinTech. Monoclonal antibodies against hemagglutinin (HA) tag (C29F4) for ChIP and western blotting were purchased from Cell Signaling Technology; ACTB (66009-1) from ProteinTech; FLAG tag (M2) and α-tubulin (T5168) from Sigma-Aldrich; βIII-tubulin from Biolegend (801201).

Cell culture

NT2 cell line was obtained from the American Type Culture Collection (ATCC). NT2 and 293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) high glucose with GlutaMAX (2 mM; Gibco) supplemented with 10% fetal bovine serum (FBS, Gibco). Large scale culture of NT2 cells to generate P1M fraction was described (12).

D3 mouse ESC line was purchased from ATCC and adapted to feeder-free condition as described (12). Medium was changed daily. Cell cultures were passaged by trypsin. For ChIP-seq experiments, SOX2-HA and SOX15-HA knockin mouse ESC lines were first adapted to 2i/LIF culture condition as described (15). Differentiation of ESCs by embryoid body (EB) formation was performed as described (28). EBs were formed by using the hanging drop method. Briefly, hanging drops (30 μl) containing ~300 cells were pooled after 2 days and transferred to low attachment plates. EBs were collected for analysis after 8 days of differentiation. Differentiation of ESCs into neuroectodermal precursors was performed as described (29). Differentiation medium was changed every other day for 6 days. EpiSC line OEC-2 was kindly provided by Dr. Austin Smith laboratory. OEC-2 was cultured on human fibronectin-coated plates (15 μg/ml; EMD Millipore) in serum-free N2B27 medium (DMEM/F12 [Gibco] and Neurobasal [Gibco] 1:1 ratio, 1:100 N2 Supplement [Gibco], 1:50 B27 Supplement [Gibco], 2 mM Glutamax [Gibco], beta-mercaptoethanol [0.1 mM, Sigma-Aldrich], and 1x penicillin and streptomycin [100 U/ml, Gibco]) supplemented with 20% KnockOut Serum Replacement (Gibco), bovine serum albumin (0.01%, Sigma-Aldrich), Activin A (20 ng/ml, R&D Systems), and FGF2 (12 ng/ml, R&D Systems). For in vitro differentiation of EpiSC-L cells. ESCs cultured on fibronectin-coated plates in 2i/LIF medium were exchanged with serum-free N2B27 with Activin A and FGF2 as described above with the addition of 1% KnockOut Serum Replacement (Gibco). EpiSC-L cells are collected for analysis after 10 days of differentiation.

In vitro transcription assay

The human *NANOG* transcription template, purification of activators OCT4 and SOX TFs, general transcription factors, RNA polymerase II, recombinant XPC complex and DKC1 complexes reconstituted from insect Sf9 cells (12, 13), and recombinant ABCF1 purified from *Escherichia coli* (15) were described. Partially-purified phosphocelluose 1M (P1M) nuclear cell extracts from NT2 cells were described (12). In vitro transcription and primer extension assays were described (12).

Coimmunoprecipitation assay

pHAGE-XPC.com plasmid expressing XPC and its subunits RAD23B and CETN2, pHAGE-DKC1-IRES-Neo, pFLAG-CMV5a-ABCF1-V5 were co-transfected with control empty pFLAG-CMV5a vector or pFLAG-CMV5a-SOX15 into 293T using polyethylenimine (PEI, Polysciences) at a DNA:PEI (μg) ratio of 1:2.5. 40 hours after transfection, cells on 10-cm dishes were lysed with 1 ml of lysis buffer [50 mM HEPES pH 7.6, 0.12 M NaCl, 1% NP-40, 1 mM EDTA, and 10% glycerol] containing 1 mM DTT, 0.25 mM PMSF, and complete protease inhibitor cocktail (Sigma-Aldrich). Cell lysates were cleared by centrifugation at 15,000 rpm for 30 min at 4°C. Cleared lysates were then incubated with anti-FLAG M2 agarose for 3 to 4 hours at 4°C. Bound proteins were washed extensively with lysis buffer followed by FLAG peptide elution (0.4 mg/ml) in elution buffer [50 mM HEPES pH 7.6, 0.15 M NaCl, 0.5% NP-40, 10% glycerol) containing 0.5 mM DTT, 1 mM benzamidine, and 0.5 mM PMSF.

Lentiviral production and transduction

For lentivirus production, nontargeting control and pLKO plasmids targeting mouse SOX15 (Sigma-Aldrich), or pHAGE plasmids for overexpression of SOX15 or control mCherry were cotransfected with packaging vectors (psPAX2 and pMD2.G; Addgene) into 293T cells using PEI methods. Supernatants were collected at 48 hours and again at 72 hours. Virus preparation was performed as described (12). Functional viral titer was determined by transduction of limited dilution of concentrated viruses into HeLa or NIH 3T3 cells followed by counting antibiotic-resistant colonies to obtain colony-forming unit per milliliter. Cells were infected with viruses in the presence of polybrene (8 μg/ml). For knockdown experiments, mouse ESCs transduced with pLKO viruses (multiplicity of infection [MOI] of 5) were selected with puromycin (1.5 μg/ml). Stable expression of SOX15 or mCherry in transduced OEC-2 cells was maintained in serum-free N2B27 medium containing geneticin (500 μg/ml).

Cellular reprogramming and flow cytometry

For generation of iPSCs from fibroblasts, CF-1 MEFs (Charles River Laboratories) were transduced with inducible STEMCCA and rtTA lentivirus–containing supernatants overnight in polybrene (8 μg/ml; Sigma-Aldrich). Cells were expanded and transduced with pLKO viruses expressing control shRNA or shRNAs targeting SOX15. Doxycycline (2 μg/ml; Sigma-Aldrich) was supplemented to ESC medium to induce expression of OCT4, KLF4, SOX2, and c-MYC. Reprogramming was assayed by AP staining (EMD Millipore) or by flow cytometry analysis using anti-SSEA1, anti-THY1, and anti-EpCAM (BioLegend) on BD LSRFortessa, performed according to the manufacturers' protocols.

For EpiSC to ESC conversion, ~$1x10^6$ stable OEC-2 cells expressing SOX15 or mCherry (MOI of 3) were plated in OEC-2 culture medium. After 24 hours, medium was changed to 2i/LIF medium. 2i/LIF medium was changed every other day for 14 days.

ChIP-seq analysis

HA-ChIP assays were performed in SOX15-HA knockin mouse ESCs essentially as described (13) except cells were cross-linked with formaldehyde (1%) for 5 min and cross-linked chromatin immunoprecipitated using an anti-HA rabbit monoclonal antibody (Cell Signaling Technology). Knock-in ESCs were first adapted to serum-free 2i/LIF condition as described before ChIP (15). ChIP-seq DNA libraries were constructed using the KAPA HyperPlus DNA Library Prep Kit for Illumina (Roche) according to the manufacturer instructions. Purified libraries were sequenced on NextSeq Illumina sequencer using NextSeq.

ChIP-seq data was processed using the ChIP-seq pipeline implemented in bcbio-nextgen v1.2.0.(https://bcbio-nextgen.readthedocs.org/). Briefly, raw sequence quality was evaluated using FASTQC (30). Reads were filtered and trimmed with Atropos (31) and mapped to the mouse genome build mm10 using Bowtie2 (32) with default parameters. Peaks were called using MACS2 (33) with default parameters after excluding multi-mapping reads and removing duplicates. Quality control was performed with ChIPQC (34) and DeepTools (35) to assess peak quality, coverage and reproducibility of the peaks across replicates. Likely peak artifacts were filtered out using the ENCODE blacklist (36). Overlapping peaks were identified using Bedtools (37). ChIP-seq peaks were annotated and functional enrichment analysis was performed using ChIPseeker

(38). Known and de novo motifs were identified with HOMER (39). Data visualization was performed using the Integrated Genomics Viewer (IGV) (40).

Comparisons to Pol II-S5p and H3K27me3 binding in mouse ESCs were performed using GEO data sets GSM2474111 and GSM2474113, respectively.

RT-qPCR and RNA-seq analysis

Total RNA was extracted and purified from cells directly lysed in TRIzol reagent (Life Technologies). For RT-qPCR, 1 μg of purified RNAs were treated with deoxyribonuclease I (Life Technologies) before subjected to first strand cDNA synthesis using iScript cDNA Synthesis Kit (Bio-Rad) or SuperScript IV VILO Master Mix (Thermo Fisher Scientific) for low abundance transcripts. RT-PCR analysis was carried out with iTaq Universal SYBR Green (Bio-Rad) and gene-specific primers using the CFX Touch Real-Time PCR Detection System (Bio-Rad). Results were normalized to β-actin. Primer sequences are shown in Supplementary table S1.

For RNA-seq, total RNA was extracted from cells using TRizol Reagent (Invitrogen). Ribosomal RNAs depletion and libraries preparation were conducted using KAPA HyperPrep kit with Ribo-Erase (Roche) according to manufacturer instructions. Sequencing was performed using Nextseq Illumina sequencer using a NextSeq High-Output 75-cycle kit.

RNA-seq samples were processed using the RNA-seq pipeline implemented in bcbio-nextgen v1.1.3 (https://bcbio-nextgen.readthedocs.org). Raw reads were examined for quality issues using FastQC (30) and aligned the UCSC build mm10 of the mouse genome using STAR (41).

Alignments were checked for evenness of coverage, rRNA content, exonic and intronic mapping rates, complexity and other quality metrics using FastQC, Qualimap (42), MultiQC (43) and custom tools. RNA-seq counts were generated using Salmon (44). DESeq2 (45) was used to identify differentially expressed genes with a BH adjusted *P* value cutoff of 0.01; no fold-change cutoff was applied.

Micrococcal nuclease ChIP-qPCR analysis

MNase-ChIP was performed essentially as described (15) except that ESCs were cross-linked with formaldehyde for 5 min. Cross-linked chromatin were immunoprecipitated with rabbit anti-HA monoclonal antibodies (Cell Signaling Technology). Purified DNA was quantified by real-time PCR with KAPA SYBR FAST qPCR Master Mix (KAPA Biosystems) and gene-specific primers using the CFX Touch Real-Time PCR Detection System (Bio-Rad). Primer sequences are shown in Supplementary table S2.

CRISPR/Cas9-mediated genome editing

For generation of HA-tagged SOX2 and SOX15 knockin ESC lines, single-guide RNA (sgRNA) targeting genomic region immediately upstream of the translation stop codon of *Sox2* or *Sox15* was cloned into LentiCRISPRv2 vector (Addgene). Single-stranded (ss) donor oligonucleotides containing a flexible linker GSGT followed by HA tag sequence and translation stop codon, which is flanked by left and right homology arms of about 60-75 base pairs (bp), were synthesized [Integrated DNA Technologies (IDT)]. Both the LentiCRISPRv2-sgRNA plasmid and the ss donor oligonucleotides were transfected into D3 mouse ESC line using Lipofectamine 3000 (Invitrogen). Transfected cells were selected with puromycin (1.5 μg/ml) for 3 days. Cells were then expanded in the absence of puromycin. Single clones were plated into 96-well plates. Positive clones were

identified by PCR and confirmed by sequencing and western blotting. Clones selected for further analysis were confirmed to be puromycin-sensitive and express similar levels of key pluripotency genes such as *Nanog*, *Oct4*, and *Sox2*. See Supplementary table S3 for sgRNA and ss donor oligonucleotide sequences.

For generation of *Sox15* knockout ESC lines, sgRNA targeting immediately downstream of the translation start codon or sgRNAs flanking the entire *Sox15* gene locus were cloned into LentiCRISPRv2 vector. See Supplementary table S4 for sgRNA sequences.

To delete the -17 kb region upstream of the *Hes5* gene locus, sgRNAs flanking a ~1 kb region containing the SOX15 ChIP-seq peaks were cloned into LentiCRISPRv2 vector. Transfection and selection of knockout clones were as described for SOX2 and SOX15 HA-knockin lines. Homozygous deletion clones were confirmed by genomic PCR and sequencing. See Supplementary table S5 and S6 for sgRNA sequences and primer sequences for genomic PCR, respectively.

Colony formation assay

200 cells of WT and *Sox15* KO clones (KO1 and KO2) were plated on 24-well plates either under medium containing serum and LIF or serum-free 2i/LIF medium. After 6 days, cells were fixed and stained for AP activity (EMD Millipore). AP-positive cells were counted and analyzed.

Immunofluorescence microscopy

WT and *Sox15* KO clones (KO1 and KO2) ESCs were plated on cover glasses and induced to undergo neural differentiation by changing ESC medium to serum-free N2B27 medium. At day 6, cells were fixed in 4% paraformaldehyde for 10 min at room temperature (RT) and permeabilized with 0.1% triton X-100 in PBS for 10 min at RT. After three washes in PBS for 5 min each, cells were blocked with 5% bovine serum albumin (BSA) and 5% goat serum for 1hr at RT and incubated with primary antibodies (anti-βIII-tubulin antibody [BioLegend]) overnight at 4°C. Following three washes, cells were incubated with secondary antibodies (Goat anti-Rabbit IgG Alexa Fluor-488, Life Technologies, 1:500) for 1 hr at RT. Cover glasses were mounted on slides with mountant with 4',6-diamidino-2-phenylindole (DAPI) (Life Technologies) and imaged with a confocal microscope.

Statistical analysis

To determine statistical significance, *P* values were calculated by using unpaired two-sided Student's *t* test. *P* values less than 0.05 (<0.05) were considered as statistically significant, and they were indicated with * (*$p < 0.05$). All data represent the mean ± SEM (error bars).

Results

OCT4 can cooperate with SOX2 or SOX15 to directly activate *NANOG* transcription

To address whether SOX15 and other related SOX proteins can cooperate with OCT4 to activate transcription, we took advantage of our fully reconstituted in vitro transcription assay that recapitulates stem cell-specific transcriptional activation by OCT4 and SOX2 (12). We expressed and purified a panel of SOX proteins (SOX1, SOX2, SOX11, and SOX15) and asked whether

they can activate a well-characterized OCT4/SOX2-target, the human *NANOG* gene promoter (Figure 1 and Supplementary figure 1A) (46, 47). In the presence of a partially purified pluripotent cell nuclear fraction (phosphocellulose 1M [P1M]) previously shown to contain transcriptional coactivators for OCT4/SOX2 (12), we found that, in addition to SOX2, SOX15 was able to robustly activate the *NANOG* gene with OCT4 (Figure 1A). Although it has been shown that SOX11 binds the enhancer of the *Nanog* gene in mouse ESCs (24), our result indicates that it is likely inactive for transcriptional activation, thus highlighting the selective nature of SOX-OCT4 protein complexes for stem cell-specific transcriptional activation. To confirm that SOX15 indeed binds sox-oct enhancers of the endogenous *Nanog* promoter and core pluripotency genes *Oct4* and *Sox2*, we used CRISPR/Cas9-mediated homologous recombination to knock-in an HA-epitope tag at the C-terminus of *Sox15* or *Sox2* gene locus in mouse ESCs (Figure 1B and Supplementary figure 1B). Binding of SOX15 at the sox-oct enhancers at *Oct4, Sox2*, and *Nanog* was confirmed by performing micrococcal nuclease (MNase)-chromatin immunoprecipitation (ChIP) using a highly specific anti-HA antibody (Figure 1C and Supplementary figure 1C).

We have previously shown that transcriptional activation by SOX2 and OCT4 requires three stem cell coactivators (SCCs) present in P1M fraction, namely, the XPC DNA repair complex, the dyskerin (DKC1) ribonucleoprotein complex, and the ATP-binding cassette subfamily F member 1 (ABCF1) (12, 13, 15). We therefore asked whether SOX15 also potentiates transcription by cooperating with SCCs. We found that optimal transcriptional activation of *NANOG* requires all three SCCs (Figure 1D and Supplementary figure 1D). The co-dependent nature of SCCs on transcriptional activation by SOX15 is highly reminiscent of our previous findings with SOX2 (13, 15), suggesting that SOX2 and SOX15 employ similar mechanisms to potentiate *NANOG* transcription in ESCs. Indeed, using transient transfection and coimmunoprecipitation experiments in 293T cells, we showed that SOX15, like SOX2, binds XPC, DKC1, and ABCF1

(Figure 1E) (13, 15, 48, 49). Together, we conclude that SOX15 activates stem cell-specific transcription by cooperating with SCCs through direct recruitment.

SOX15 contributes to stem cell self-renewal and maintenance

To examine the functional relationship between SOX2 and SOX15 in ESCs, we performed acute depletion of SOX2 or SOX15 in mouse ESCs by targeted small hairpin RNAs (shRNAs) and assessed their effect on pluripotency gene expression. While knockdown of SOX15 has no discernible effect on SOX2 expression, we found that partial knockdown of SOX2 reduced SOX15 both at the mRNA and protein level (Figure 2A and 2B). We do not believe SOX2 directly regulates *Sox15* gene expression because we did not detect SOX2 binding at the vicinity of the *Sox15* locus using our own SOX2 ChIP-seq data (*see below*) or published ChIP-seq results (11). These observations also emphasize the importance of studying SOX15 function in ESCs directly, because defects in transcription and pluripotency previously shown in SOX2-deficient ESCs may be attributed to loss of SOX15 (24, 50, 51). Consistent with previous observations, both SOX2 and SOX15 are enriched in the pluripotent state (Supplementary figure 2A) (22). Even though SOX15 mRNA and protein levels are significantly lower than that of SOX2 (Figure 1B and Supplementary figure 2A), we found that knockdown of SOX15 was nevertheless able to reduce the expression of several key pluripotency genes such as *Nanog* and *Klf4*, although to a lesser degree than SOX2-knockdown (Supplementary Figure 2B). These data suggest that both SOX2 and SOX15 contribute to the expression of these genes in ESCs. To determine whether incomplete nature of shRNA-mediated knockdown might obscure the impact of SOX15 depletion in ESCs, we employed CRISPR/Cas9 to knockout *Sox15* in mouse ESCs. In one knockout ESC line (KO1), frameshift indels were introduced near the 5' end of *Sox15* coding region, leading to a premature translation stop at amino acid 72 (Figure 2C). Because the *Sox15* gene is only ~1.4 kb long, we also deleted *Sox15* by using two sgRNAs flanking the entire gene locus (KO2).

Western blot analysis confirmed the absence of SOX15 in both KO ESC lines while SOX2 and OCT4 remained largely unchanged (Figure 2D). The fact that we could establish SOX15-null ESC lines indicated that self-renewal can occur in the absence SOX15. However, we noticed that when these SOX15-deficient ESCs were challenged with single cell dissociation, they were less efficient in forming pluripotent, alkaline phosphatase (AP)-positive cell colonies (Figure 2E). Consistent with their reduced self-renewal capacity, we found that the expression levels of several key pluripotency genes such as *Nanog*, *Klf4*, *Prdm14,* and *Esrrb* are compromised (Figure 2F). Therefore, our data suggested that optimal expression of these pluripotency genes and stem cell maintenance require both SOX15 and SOX2.

ESCs are typically cultured in medium containing serum and the self-renewal cytokine leukemia inhibitory factor (LIF). Under this condition, ESCs exhibit a high degree of heterogeneity, reflected in their mosaic expression of several pluripotency-associated genes such as *Nanog* (52) and *Esrrb* (53). Cell populations expressing lower levels of these genes have reduced self-renewal capacity and are more prone to irreversible exit from pluripotency. However, this metastable pluripotent state can be stabilized in serum-free medium containing LIF and inhibitors (2i) of mitogen-activated protein kinase signaling and glycogen synthase kinase-3 (GSK3) (54, 55), which blocks signals that promote spontaneous differentiation. 2i/LIF conditions have been shown to bypass the requirement of TFs such as ESRRB and PRDM14 for self-renewal in serum/LIF (56, 57), suggesting that these TFs function to buttress the pluripotency gene network to counteract differentiation signals. Indeed, we found that 2i/LIF also overcame defects in self-renewal and restored expression of *Nanog* and *Prdm14* in *Sox15* KO ESCs (Supplementary figure 2C and 2D), indicating some SOX15-independent pathways enforced by 2i/LIF signaling may be sufficient to sustain the expression of these genes. Our results thus demonstrate that *Sox15* KO ESCs can be maintained in a stable pluripotent state under specific conditions, and that SOX15

stabilizes stem cell self-renewal in serum/LIF conditions by promoting optimal expression of key pluripotency genes such as *Nanog*, *Esrrb, and Prdm14*.

Non-redundant function SOX15 in cellular reprogramming

Barriers to reacquisition of pluripotency during somatic cell reprogramming are high, which include the faithful reactivation of the pluripotency gene network in somatic cells (58–60). We reasoned that such processes may be more sensitive to perturbation of pluripotency-associated TFs. We therefore asked whether induced pluripotent stem cell (iPSC) generation requires SOX15 (18, 19). To address this question, we transduced mouse embryonic fibroblasts (MEFs) with lentiviruses expressing non-targeting control shRNA or two independent shRNAs specific for SOX15. We next initiated reprogramming of these MEFs by doxycycline-induced expression of OCT4, KLF4, SOX2, and c-MYC (61). Knockdown of SOX15 led to a marked decrease in the number of AP-positive iPSC colonies formed (Figure 3A). Flow cytometry analysis showed that loss of SOX15 interfered with the downregulation of fibroblast-associated cell surface marker THY1, and compromised the expression of iPSC early marker SSEA1 and the late stage marker EpCAM (Figure 3B) (62), which has been shown to correlate with the reactivation of *Nanog* and endogenous *Oct4* (63). Our results show that SOX15 is necessary for efficient suppression of somatic cell identity and reactivation of pluripotency-associated genes during reprogramming, and cannot be replaced by exogenously overexpressed SOX2.

It is possible that the observed reprogramming block in SOX15 knockdown MEFs is due to defects in events prior to pluripotency gene network reactivation. To test more directly the role of SOX15 in induced pluripotency, we chose a different reprogramming paradigm, in which we assessed the ability of SOX15 to regulate the conversion between two related pluripotent states found in

mouse ESCs and epiblast stem cells (EpiSCs). Unlike ESCs which are derived from pre-implantation blastocysts, EpiSCs are isolated from post-implantation embryos (64, 65). Similar to ESCs, EpiSCs express core TFs OCT4 and SOX2 and can self-renew and differentiate into three embryonic germ layers. However, EpiSCs require fibroblast growth factor (FGF) and Activin/Nodal signaling instead of LIF for self-renewal, and display restricted cell fates in part due to downregulation of several key pluripotency genes such as *Klf4* and *Nanog* and reactivation of lineage-specific genes representing the three germ layers (65, 66). EpiSCs are poised to undergo differentiation and thus represent a more advanced, "primed" pluripotent state as opposed to the more pristine, "naive" state found in ESCs. We found that SOX15 expression is restricted to ESCs when compared to an established EpiSC line, OEC-2 (67) while SOX2 protein level remains largely the same in both cell lines (Supplementary figure 3A and 3B). To ensure that the observed differences in SOX15 expression pattern is not a cell line specific phenomenon, we directly converted mouse ESCs to an EpiSC-like (EpiSC-L) state by culturing ESCs in medium containing FGF2 and Activin (67, 68) and confirmed that SOX15 expression was also markedly reduced in EpiSC-L cells (Supplementary figure 3C), concomitant with upregulation of lineage-specific genes such as *Fgf5* and *T* (Supplementary figure 3D). These observations prompted us to ask whether SOX15 could play a role in promoting the acquisition of naive pluripotency in EpiSCs.

By directly culturing EpiSCs in 2i/LIF medium, it has been shown that only a small fraction of EpiSCs are converted to ESC-like cells (rESCs) (67). However, this inefficient process can be enhanced by ectopic expression of naive pluripotency-enriched transcription factors such as KLF4 (67), ESRRB (56, 69), PRDM14 (70) and NANOG (71). Because we showed that expression of these genes is partially dependent on SOX15 in ESCs (Figure 2F), we next tested whether ectopic expression of SOX15 could also promote rESC generation. We transduced OEC-2 cells with lentiviruses expressing SOX15 or RFP as control (Supplementary figure 3E). We found that

ectopic expression of SOX15 significantly increased reprogramming efficiency of EpiSC to rESC (Figure 3C). We further noticed that rESC colonies obtained from SOX15-overexpressing cells are tightly packed with more uniform AP staining, characteristics of *bona fide* ESCs (Figure 3C). By contrast, RFP-rESC colonies tended to be more loosely packed with large, elongated cells and variegated AP staining intensity, indicating that reprogramming may be less faithful than SOX15-rESCs. qPCR analysis confirmed SOX15-overexpressing rESCs restored the expression of key naive pluripotency-associated genes (*Nanog* and *Klf4*) and suppressed lineage-specific gene *Fgf5*, to levels comparable to ESCs (Figure 3D). Note that SOX2 level remains high in OEC-2 cells that is similar to ESCs (Supplementary figure 3B). Our results point to a role of SOX15 distinct from SOX2 in promoting the reacquisition of naive pluripotency in EpiSCs.

SOX15 and SOX2 binding sites partially overlap in ESCs

To identify transcriptional networks regulated by SOX15 and SOX2 in ESCs, we performed ChIP-seq in HA-tag knockin mouse ESC lines. We reasoned that using the same, highly specific anti-HA antibody for ChIP would offer a more accurate comparison of genomic binding sites for both TFs. By performing anti-HA ChIP-seq analysis on respective homozygous knockin ESC lines, we identified 2378 SOX2 binding sites (n=2 biological replicates) and 170 SOX15 binding sites (n=3 biological replicates) (Supplementary figure 4). To validate the SOX15 ChIP-seq results, we performed MNase ChIP-qPCR and confirmed one of the strongest SOX15 binding sites upstream of the *Perc* gene in ESCs (Supplementary Figure 4C). Consistent with previous SOX2 ChIP-seq results (11), about 16% of the peaks are located in the promoters of RefSeq genes (0-3 kb), while the majority lies in distal intergenic and intronic regions (Figure 4A). On the other hand, we found that the majority of SOX15 peaks (50%) are located within 1 kb from transcription start site (TSS) of the closest RefSeq gene, compared to 8.5% for SOX2. These results indicate that the strongest SOX15 binding sites are found primarily at promoter proximal regions. Importantly, almost all 170

SOX15 peaks overlap with active RNA polymerase II (phosphoserine 5) and are excluded from histone H3 lysine 27 trimethylation (H3K27me3) repressive histone marks (72) (Figure 4B). These results are consistent with SOX15 acting as a transcriptional activator in ESCs. De novo motif discovery within DNA sequences that are recognized by both SOX15 and SOX2 confirmed Sox sequences as one of the enriched motifs, likely reflecting their intrinsic sequence specificity (Figure 4C). However, we found that more than half of SOX15 binding sites (100/170) did not overlap with SOX2, supporting the idea that SOX15 may regulate genes distinct from SOX2 in ESCs. Analysis of sequences enriched in these SOX15-unique binding regions identified motifs including those recognized by OCT4 and an unknown element previously shown to be enriched in a NANOG ChIP-seq dataset (11) (Figure 4D). While the precise role of these SOX15-unique targets in ESCs is unclear, these observations suggest a potential additional mechanism by which SOX15 contributes to the pluripotency gene network, by partnering with NANOG and OCT4 at their cognate sequences instead of the canonical sox motifs.

Sox15 **deletion compromises stem cell pluripotency**

To gain a more comprehensive view of the transcriptional programs regulated by SOX15, we performed RNA-seq analysis on three biological replicates each of wild-type (WT) and *Sox15* KO1 ESCs. ~7% of protein-coding genes in mouse ESCs are either up (830 genes) or down (589) regulated (Figure 5A). We performed unbiased gene ontology (GO) analysis on these transcripts and observed a significant overrepresentation of categories related to stem cell maintenance and blastocyst formation, consistent with a role of SOX15 in pluripotency (Figure 5B). Interestingly, GO analysis also identified genes involved in cell migration, trophectodermal specification, and neural stem cell and neuronal differentiation. In agreement with our qPCR results (Figure 2F), key pluripotency genes such as *Nanog*, *Klf4*, and *Esrrb* are downregulated in *Sox15* KO ESCs (Figure 5C). However, genes associated with trophectoderm and neuronal differentiation were also

decreased in *Sox15* KO ESCs. This result was somewhat unexpected given that *Sox15* KO ESCs are more prone to spontaneous exit from pluripotency, and thus one might anticipate that lineage-specific genes should be upregulated compared to WT ESCs. However, SOX15 has previously been implicated in trophectoderm differentiation by cooperating with TF HAND1 (73). While the role of SOX15 in neurogenesis has not been reported, we found that neurogenic transcription factors such as *Sall3* (74, 75) and *Nfib* (76), and Slit/Robo signaling, previously shown to modulate neurogenesis (77, 78), are downregulated in *Sox15* KO ESCs. Reduced levels of these transcripts in ESCs may reflect that, as *Sox15* KO ESCs transiently exit from pluripotency under serum/LIF condition, these cells may have a reduced capacity to commit towards the neural cell fates by upregulating neurogenic gene expression programs. These observations suggest that *Sox15* KO ESCs may have restricted potential to differentiate into the three embryonic germ layers.

SOX15 is required for *Hes5* expression and neural differentiation

Our ChIP-seq analysis revealed that SOX15 occupies a region ~17 kb upstream of the *Hes5* gene locus (Figure 6A). *Hes5* belongs to the *Hairy and Enhancer of Split* (*Hes*) family transcriptional repressors that include *Hes1* and *Hes3* (79). *Hes1* and *Hes5* are expressed in the ventricular zone of the developing brain, in which neural stem cells (NSCs) reside (80, 81). In mice, it has been shown that *Hes1* and *Hes5* display some degree of functional redundancy in promoting self-renewal and maintenance of NSCs (81, 82). However, these repressors appear to also regulate the timing of cell fate specification of neural progenitor cells derived from NSCs (83–85). Therefore, the *Hes* family transcriptional repressors play complex roles in controlling both the maintenance of NSCs and their differentiation during development. SOX2 has been implicated in regulating *Hes5* expression in neural progenitor cells by binding at promoter proximal regions (86, 87). However, the role of SOX15 in neural differentiation is unknown.

We first confirmed our ChIP-seq results by performing MNase ChIP-qPCR and demonstrated specific enrichment of SOX15 at -17 kb but not -2 kb or TSS of the *Hes5* locus in undifferentiated ESCs (Figure 6B). To examine the potential role of SOX15 in regulating *Hes5* expression during neural differentiation, we directly differentiated ESCs in serum-free, pro-neurogenic medium (29). *Hes5* expression was induced in SOX15-HA knockin ESCs 2 days after differentiation (Supplementary figure 5A). Even though SOX15 mRNA and particularly protein levels were downregulated at day 2 (Supplementary figure 5B), we found that SOX15 remained engaged at the -17 kb region, with no change in relative enrichment at -2 kb or TSS (Figure 6C). These results suggest that SOX15 is already bound at the -17 kb region prior to *Hes*5 induction during neural differentiation.

To examine whether SOX15 is required for *Hes5* expression, we compared its expression in WT and *Sox15* KO ESCs during neural differentiation. We observed robust induction of *Hes5* in WT ESCs over the course of 6-day differentiation process, wherein a population of neuroectodermal precursors such as NSCs and neural progenitors are formed (Figure 6D) (29). However, both *Sox15* KO ESC lines completely failed to induce *Hes5* expression. Consistent with the function of HES5 in NSC maintenance and multipotency (84), we also found that expression of NSC marker such as *Nestin* (*Nes*) and transcription factors important for neural commitment, *Mash1* and *Pax6*, were significantly compromised in *Sox15* KO ESCs, while *Sox1* was moderately reduced at day 6 (Figure 6E). Interestingly, even though *Hes1* expression was not affected in *Sox15* KO ESCs, we observed a pronounced defect of KO ESCs in expressing neural marker βIII-tubulin (Tuj1) (Figure 6F) (29), indicating that loss of *Hes5* expression cannot be compensated by *Hes1*. Given that SOX2 levels remained the same in WT and *Sox15* KO ESCs and during the early stages of differentiation (Supplementary figure 5B), our results uncovered a critical function of SOX15 in regulating *Hes5* expression and neural differentiation in vitro.

The -17 kb region bound by SOX15 displays distal enhancer signatures such as DNase I hypersensitivity and enrichment of H3K27 mono-acetylation (H3K27Ac) (ENCODE). This region is also surrounded by several genes including *Tnfrsf14*, *Pank4*, and *Plch2* (Supplementary figure 6A). While there are no reports that directly implicate these genes in early neural fate specification, we tested whether expression of these genes is also dependent on SOX15. We found that expression of *Tnfrsf14*, *Pank4*, and *Plch2* in undifferentiated (D0) and differentiated ESCs (D6) was not affected by SOX15 depletion (Supplementary figure 6B). Therefore, it appears that the -17 kb region occupied by SOX15 may be specifically required for *Hes5* expression. It has previously been shown that RNA polymerase II (Pol II) recruitment to enhancers correlate with enhancer activities and gene activation (88). Therefore, to evaluate the activity of this putative distal enhancer in WT and *Sox15* KO ESCs at the onset of neural differentiation, we examined Pol II occupancy at the -17 kb region at day 2 of differentiation. We found striking reduction of Pol II occupancy at the -17 kb region in KO ESCs compared to WT, as well as at the -2 kb region and TSS (Figure 7A). These results are consistent with a defect in activating the -17 kb enhancer and transcription of *Hes5* in *Sox15* KO ESCs during neural differentiation. Importantly, SOX2 binding at the *Hes5* locus is not significantly affected by the absence of SOX15 in KO ESCs (Supplementary figure 6C). Together, our data point to a specific requirement of SOX15 in regulating enhancer and promoter activities of *Hes5* during early stages of neural differentiation.

To more firmly establish the -17 kb region as an important enhancer that drives *Hes5* expression during neural differentiation, we employed CRISPR/Cas9 gene editing to delete this putative enhancer. Two independent homozygous deletion ESC lines (Δenh1 and 2) were generated and confirmed by genomic DNA PCR and sequencing (Figure 7B). Deletion of this enhancer region did not affect the expression of core pluripotency genes (*Nanog*, *Oct4*, and *Sox2*), genes in the vicinity of the -17 kb region (i.e., *Tnfrsf14*, *Pank4*, and *Plch2*), or basal expression of *Hes5* in

undifferentiated ESCs (Figure 7C and Supplementary figures 6D and 6E). However, compared to control ESCs expressing a non-targeting sgRNA, induction of *Hes5* in ∆enh1/2 was significantly compromised at day 6 of differentiation while *Hes1* mRNA levels were unaffected (Figures 7C and 7D). *Mash1*, *Pax6*, and *Sox1* were also reduced in ∆enh1/2 compared to control ESCs, indicating that neural differentiation was compromised (Figure 7D). These results are reminiscent of the differentiation defect observed in *Sox15* KO ESCs (Figure 6F), suggesting that activation of *Hes5* by SOX15 is an important early event in neural differentiation in vitro. It has recently been reported that *Hes5* is a direct target of the TF GLI3 in the Hedgehog (Hh) pathway, which plays an important role in neural differentiation during development (89). We found that *Gli3* mRNAs were elevated in WT ESCs during neural differentiation (Supplementary figure 6F). However, no significant differences in *Gli3* levels were observed between WT and KOs at day 4 of differentiation (Supplementary figure 6G), even though *Hes5* expression was already reduced in *Sox15* KO cells (Figure 6D). Therefore, *Gli3* alone does not appear to be sufficient to induce *Hes5* expression. Using in vitro neural differentiation system as a paradigm for studying the role of Sox TFs in regulating exit from pluripotency and lineage specification, we uncovered an important function of SOX15 in the early neural fate commitment that appears to be distinct from SOX2.

Discussion

SOX15 function in maintenance and reacquisition of pluripotency

In this study, we show that SOX15 and SOX2 activate transcription of pluripotency genes due to their ability to recruit a common set of transcriptional coactivators enriched in ESCs. Outside the highly conserved HMG domain (~80% sequence identity), SOX15 and SOX2 protein sequences diverge significantly, raising questions on how SOX15 and SOX2 (but not other Sox family TFs) are able to stimulate transcription using the same transcription factor ensemble. We have recently

shown that one of the stem cell-enriched coactivators, ABCF1, activates ESC-specific transcription by selectively interacting with the HMG domain of SOX2 (but not SOX11, for example) (15). We find here that ABCF1 also binds SOX15, thus revealing a role of cell-specific coactivators in discriminating various SOX-OCT4 complexes that may assemble at a gene promoter in ESCs, likely by recognizing small differences in the HMG box sequence. Take together with our previous observation that OCT4 alone with SCCs is not sufficient to activate *NANOG* transcription (12), our results indicate that the ability of SOX15 to bind ABCF1 is a critical determinant for transactivation. However, considering that SOX15 abundance in ESCs is significantly lower than that of SOX2, disruption of a SOX15-ABCF1 interaction alone may not be sufficient to account for the degree with which genes targeted by both SOX15 and SOX2 are downregulated in *Sox15* KO ESCs. These observations suggest non-redundant mechanisms by which SOX15 contributes to pluripotency gene activation. Sox TFs have been shown to interact with various chromatin remodeling and modifying complexes to drive cell type-specific transcription (87, 90–92). It is possible that the unique C-terminal region outside of the HMG box also play a role, potentially by mediating SOX15-specific interactions that are conducive to gene activation in the context chromatin in ESCs. Furthermore, studies have shown that the unique region can modulate DNA binding specificity of the HMG domain (93). We speculate that the unique C-terminal region of SOX15 may function to direct SOX15 to genomic sites that are distinct from SOX2. Our ChIP-seq results suggest that core pluripotency TFs NANOG and OCT4 may "hijack" SOX15 to their respective cognate binding sites, a phenomenon that has previously been reported for OCT4 in ESCs (94). Future studies will be required to examine how these SOX15-dependent interactions may contribute to stem cell self-renewal and pluripotency gene expression.

Sox15-null mice are viable, suggesting no overt defects in the formation of pluripotent cells in the inner cell mass (ICM) of the blastocyst (22). Because maternal *Sox15* transcripts are enriched in metaphase II oocytes (95–97), maternally supplied SOX15 proteins in KO embryos may be sufficient to compensate for the absence of functional zygotic *Sox15* gene products. Alternatively, given the shared mechanisms by which SOX15 and SOX2 regulate the expression of core pluripotency genes, SOX2 alone may be sufficient to initiate pluripotency in the ICM, likely because the pluripotent state in ICM is relatively transient during development. By contrast, our data suggest that long-term maintenance of ESCs in culture requires the combined action of SOX2 and SOX15, likely by counteracting cell-intrinsic and extrinsic signals that promote ESCs to exit from pluripotency. Our study also reveals a critical role of SOX15 in overcoming barriers to reacquisition of pluripotency during reprogramming. It is known that several key primordial germ cell (PGC) TFs including PRDM14 promote iPSC generation (70, 98–100). Coupled with recent reports implicating a role of SOX15 in PGC fate specification (101, 102) and our observation that *Prdm14* is downregulated in *Sox15* KO ESCs, we speculate that SOX15 may regulate reprogramming by promoting the expression of critical PGC-associated TFs. It is also possible that SOX15, together with SOX2, ensures optimal reactivation of the pluripotency gene network during reprogramming. Paradoxically, ectopic expression of SOX15 in place of SOX2 fails to generate iPSCs from fibroblasts (103). However, the same study demonstrated that exogenous SOX15 can promote iPSC derivation as efficiently as SOX2 from a cell type that is considered more permissive to TF-induced reprogramming (e.g., menstrual blood stromal cells (MnSCs)). These observations indicate that SOX15 has an intrinsic capacity to induce pluripotency, consistent with our data showing that it can activate the transcription of key pluripotency genes by partnering with OCT4 and SCCs. However, it appears that this activity is cell-context dependent. We surmise that the epigenetic environment and transcription factor repertoire in MnSCs that allow SOX15 to efficiently target and activate pluripotency genes may be less favorable or underdeveloped in skin-derived fibroblasts. This may also explain why endogenous

SOX15 is still required for iPSC generation from fibroblasts when SOX2 is already overexpressed. As fibroblasts traverse through various reprogramming intermediate stages, the accompanied changes in chromatin landscape and coactivator expression levels may become more permissive to activation of SOX15 target genes important for pluripotency reacquisition (104). Similarly, we posit that during EpiSC to ESC conversion, the more relaxed, “open” chromatin environment associated with pluripotent EpiSCs (compared to fibroblasts) may facilitate transcriptional activation of naive pluripotency-associated genes by exogenous SOX15 (105). Human ESCs have been shown to share features with mouse EpiSCs including their restricted differentiation potential, which in turn limit the use of human ESCs for regenerative medicine and human disease modeling (106, 107). The ability of SOX15 to promote an ESC-like state in mouse EpiSCs may inform new strategies to promote naive pluripotency in human ESCs.

SOX15 function in neural differentiation

We noticed that SOX15 is already engaged at the -17 kb distal enhancer of the *Hes5* gene in undifferentiated ESCs. This observation suggests that SOX15 is necessary but not sufficient for *Hes5* induction. Because SOX2 recruitment to the *Hes5* locus does not appear to be affected by the absence of SOX15 in KO ESCs, we propose that transcriptional activation of *Hes5* by SOX15 likely requires other TFs and coactivators that are induced at the onset of neural differentiation, in conjunction with epigenetic remodeling at the *Hes5* gene locus (108). The pre-binding of SOX15 at the *Hes5* locus may represent a common strategy by which developmentally poised genes are regulated upon differentiation (109).

Sox15 knockout mice appear phenotypically normal, indicating no gross defect in neurogenesis (22, 110). Interestingly, we find that *Hes1* expression in ESCs is not affected by the absence of SOX15. Given that multiple *Hes* genes appear to be functionally redundant in development (80,

111, 112), it is conceivable that normal *Hes1* expression levels in *Sox15* KO mice may be sufficient to support neurogenesis. While the precise mechanism of mammalian neurogenesis remains unclear (113), at least in vitro, our results suggest a hierarchical organization of HES proteins, in which *Hes5* appears to be play an important role in promoting the commitment of ESCs to NSC fate and neural lineage. This is in agreement with previous observations wherein *Hes5* expression correlates with NSC identity and multipotency better than *Hes1* (84). However, other neurogenic TFs and signaling pathways that are downregulated in *Sox15* KO ESCs (Figure 5C) may also contribute to this process. We show that SOX15 levels decrease sharply shortly after neural differentiation is initiated. We hypothesize that this may allow other Sox TFs including SOX2 to precisely tune *Hes5* expression levels in a spatiotemporal manner to control NSC maintenance vs neural cell fate specification (86, 114, 115). In summary, our findings reveal a multifaceted role of SOX15 in regulating the pluripotent stem cell fate and execution of neural differentiation programs.

Data Availability

The data that support the findings of this study are available within the manuscript and associated supplementary data files. The ChIP-seq and RNA-seq data have been deposited in NCBI's Gene Expression Omnibus and can be accessed through GEO series accession number GSE200688. Go to https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE200688.

Acknowledgments

We thank G. Dailey for help with expression constructs and S. Zheng for tissue culture assistance; We thank S. Agarwal, Z. Zhang, C. Cattoglio and members of the Fong lab for valuable discussion.

Funding

This work was supported by the National Institute of Health (NIH) [R01HL125527]; the Harvard Stem Cell Institute (HSCI); Boston Biomedical Innovation Center; Charles H. Hood Foundation; Brigham Research Institute; Blavatnik Therapeutics Challenge Awards; and Brigham and Women's Hospital HVC Junior Faculty Research Awards to Y.W.F. Work by S.H.S and J.Y was funded in part by the HSCI Center for Bioinformatics and Harvard Catalyst and the NIH [UL 1TR002541].

Conflict of Interest Disclosure

Y.W.F. is a consultant to Rejuveron Life Sciences.

References

1. Sarkar,A. and Hochedlinger,K. (2013) The Sox family of transcription factors: Versatile regulators of stem and progenitor cell fate. *Cell Stem Cell*, **12**, 15–30.

2. Carl,S.H. and Russell,S. (2015) Common binding by redundant group B Sox proteins is evolutionarily conserved in Drosophila. *BMC Genomics*, **16**, 292.

3. Wegner,M. (1999) From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.*, **27**, 1409–1420.

4. Fong,Y.W., Cattoglio,C., Yamaguchi,T. and Tjian,R. (2012) Transcriptional regulation by coactivators in embryonic stem cells. *Trends Cell Biol.*, **22**, 292–298.

5. Kamachi,Y. and Kondoh,H. (2013) Sox proteins: regulators of cell fate specification and differentiation. *Development*, **140**, 4129–4144.

6. Vodnala,M., Choi,E.-B. and Fong,Y.W. (2021) Low complexity domains, condensates, and stem cell pluripotency. *World J. Stem Cells*, **13**, 416–438.

7. Dunn,S.J., Martello,G., Yordanov,B., Emmott,S. and Smith,A.G. (2014) Defining an essential transcription factor program for naïve pluripotency. *Science (80-.).*, **344**, 1156–1160.

8. Yeo,J.-C. and Ng,H.-H. (2013) The transcriptional regulation of pluripotency. *Cell Res.*, **23**, 20–32.

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

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

11. Marson,A., Levine,S.S., Cole,M.F., Frampton,G.M., Brambrink,T., Johnstone,S., Guenther,M.G., Johnston,W.K., Wernig,M., Newman,J., *et al.* (2008) Connecting microRNA Genes to the Core Transcriptional Regulatory Circuitry of Embryonic Stem Cells. *Cell*, **134**, 521–533.

12. Fong,Y.W., Inouye,C., Yamaguchi,T., Cattoglio,C., Grubisic,I. and Tjian,R. (2011) A DNA repair complex functions as an Oct4/Sox2 coactivator in embryonic stem cells. *Cell*, **147**, 120–131.

13. Fong,Y.W., Ho,J.J., Inouye,C. and Tjian,R. (2014) The dyskerin ribonucleoprotein complex as an OCT4/SOX2 coactivator in embryonic stem cells. *Elife*, **3**, 1–30.

14. Kagey,M.H., Newman,J.J., Bilodeau,S., Zhan,Y., Orlando,D.A., van Berkum,N.L., Ebmeier,C.C., Goossens,J., Rahl,P.B., Levine,S.S., *et al.* (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**, 430–435.

15. Choi,E.B., Vodnala,M., Zerbato,M., Wang,J., Ho,J.J., Inouye,C., Ding,L. and Fong,Y.W. (2021) ATP-binding cassette protein ABCF1 couples transcription and genome surveillance in embryonic stem cells through low-complexity domain. *Sci. Adv.*, **7**, 1–19.

16. Takahashi,K. and Yamanaka,S. (2016) A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.*, **17**, 183–193.

17. Buganim,Y., Faddah,D.A., Cheng,A.W., Itskovich,E., Markoulaki,S., Ganz,K., Klemm,S.L., van Oudenaarden,A. and Jaenisch,R. (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell*, **150**, 1209–1222.

18. Takahashi,K. and Yamanaka,S. (2006) Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, **126**, 663–676.

19. Yu,J., Vodyanik,M.A., Smuga-Otto,K., Antosiewicz-Bourget,J., Frane,J.L., Tian,S., Nie,J., Jonsdottir,G.A., Ruotti,V., Stewart,R., *et al.* (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science*, **318**, 1917–1920.

20. Niwa,H., Nakamura,A., Urata,M., Shirae-Kurabayashi,M., Kuraku,S., Russell,S. and Ohtsuka,S. (2016) The evolutionally-conserved function of group B1 Sox family members confers the unique role of Sox2 in mouse ES cells. *BMC Evol. Biol.*, **16**, 173.

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

22. Maruyama,M., Ichisaka,T., Nakagawa,M. and Yamanaka,S. (2005) Differential roles for Sox15 and Sox2 in transcriptional control in mouse embryonic stem cells. *J. Biol. Chem.*, **280**, 24371–24379.

23. Ng,C.K.L., Li,N.X., Chee,S., Prabhakar,S., Kolatkar,P.R. and Jauch,R. (2012) Deciphering the Sox-Oct partner code by quantitative cooperativity measurements. *Nucleic Acids Res.*, **40**, 4933–4941.

24. Masui,S., Nakatake,Y., Toyooka,Y., Shimosato,D., Yagi,R., Takahashi,K., Okochi,H.,

Okuda,A., Matoba,R., Sharov,A.A., *et al.* (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.*, **9**, 625–635.

25. Thomson,M., Liu,S.J., Zou,L.-N., Smith,Z., Meissner,A. and Ramanathan,S. (2011) Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*, **145**, 875–889.

26. Lodato,M.A., Ng,C.W., Wamstad,J.A., Cheng,A.W., Thai,K.K., Fraenkel,E., Jaenisch,R. and Boyer,L.A. (2013) SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. *PLoS Genet.*, **9**, e1003288.

27. Alejandro,A.-A., Flavio,C., Chun-Teng,H., Rebecca,W., Susan,L., H.,G.F. and V.,T.A. (2015) SOX2 primes the epigenetic landscape in neural precursors enabling proper gene activation during hippocampal neurogenesis. *Proc. Natl. Acad. Sci.*, **112**, E1936–E1945.

28. Samuelson,L.C. and Metzger,J.M. (2006) Differentiation of embryonic stem (ES) cells using the hanging drop method. *Cold Spring Harb. Protoc.*, **2006**, pdb-prot4485.

29. Ying,Q.-L., Stavridis,M., Griffiths,D., Li,M. and Smith,A. (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat. Biotechnol.*, **21**, 183–186.

30. Andrews,S. (2010) FastQC: a quality control tool for high throughput sequence data. *http://www.bioinformatics.babraham.ac.uk/projects/fastqc*.

31. Didion,J.P., Martin,M. and Collins,F.S. (2017) Atropos: specific, sensitive, and speedy trimming of sequencing reads. *PeerJ*, **5**, e3720.

32. Langmead,B., Trapnell,C., Pop,M. and Salzberg,S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.*, **10**, R25.

33. Zhang,Y., Liu,T., Meyer,C.A., Eeckhoute,J., Johnson,D.S., Bernstein,B.E., Nusbaum,C., Myers,R.M., Brown,M., Li,W., *et al.* (2008) Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.*, **9**, R137.

34. Carroll,T.S., Liang,Z., Salama,R., Stark,R. and de Santiago,I. (2014) Impact of artifact removal on ChIP quality metrics in ChIP-seq and ChIP-exo data . *Front. Genet.* , **5**.

35. Ramírez,F., Ryan,D.P., Grüning,B., Bhardwaj,V., Kilpert,F., Richter,A.S., Heyne,S., Dündar,F. and Manke,T. (2016) deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.*, **44**, W160-5.

36. Amemiya,H.M., Kundaje,A. and Boyle,A.P. (2019) The ENCODE Blacklist: Identification of Problematic Regions of the Genome. *Sci. Rep.*, **9**, 9354.

37. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

38. Yu,G., Wang,L.-G. and He,Q.-Y. (2015) ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, **31**, 2382–2383.

39. Heinz,S., Benner,C., Spann,N., Bertolino,E., Lin,Y.C., Laslo,P., Cheng,J.X., Murre,C., Singh,H. and Glass,C.K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell*, **38**, 576–589.

40. Robinson,J.T., Thorvaldsdóttir,H., Winckler,W., Guttman,M., Lander,E.S., Getz,G. and

Mesirov,J.P. (2011) Integrative genomics viewer. *Nat. Biotechnol.*, **29**, 24–26.

41. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

42. García-Alcalde,F., Okonechnikov,K., Carbonell,J., Cruz,L.M., Götz,S., Tarazona,S., Dopazo,J., Meyer,T.F. and Conesa,A. (2012) Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics*, **28**, 2678–2679.

43. Ewels,P., Magnusson,M., Lundin,S. and Käller,M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, **32**, 3047–3048.

44. Patro,R., Duggal,G., Love,M.I., Irizarry,R.A. and Kingsford,C. (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods*, **14**, 417–419.

45. Love,M.I., Huber,W. and Anders,S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.

46. Rodda,D.J., Chew,J.L., Lim,L.H., Loh,Y.H., Wang,B., Ng,H.H. and Robson,P. (2005) Transcriptional regulation of Nanog by OCT4 and SOX2. *J. Biol. Chem.*, **280**, 24731–24737.

47. Kuroda,T., Tada,M., Kubota,H., Kimura,H., Hatano,S., Suemori,H., Nakatsuji,N. and Tada,T. (2005) Octamer and Sox elements are required for transcriptional cis regulation of Nanog gene expression. *Mol. Cell. Biol.*, **25**, 2475–2485.

48. Cattoglio,C., Zhang,E.T., Grubisic,I., Chiba,K., Fong,Y.W. and Tjian,R. (2015) Functional and mechanistic studies of XPC DNA-repair complex as transcriptional coactivator in embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, E2317--E2326.

49. Zhang,E.T., He,Y., Grob,P., Fong,Y.W., Nogales,E. and Tjian,R. (2015) Architecture of the human XPC DNA repair and stem cell coactivator complex. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, 14817–14822.

50. Wang,Z.-X., Teh,C.H.-L., Kueh,J.L.L., Lufkin,T., Robson,P. and Stanton,L.W. (2007) Oct4 and Sox2 Directly Regulate Expression of Another Pluripotency Transcription Factor, Zfp206, in Embryonic Stem Cells*. *J. Biol. Chem.*, **282**, 12822–12830.

51. Fong,H., Hohenstein,K.A. and Donovan,P.J. (2008) Regulation of self-renewal and pluripotency by Sox2 in human embryonic stem cells. *Stem Cells*, **26**, 1931–1938.

52. Kalmar,T., Lim,C., Hayward,P., Muñoz-Descalzo,S., Nichols,J., Garcia-Ojalvo,J. and Martinez Arias,A. (2009) Regulated Fluctuations in Nanog Expression Mediate Cell Fate Decisions in Embryonic Stem Cells. *PLOS Biol.*, **7**, e1000149.

53. van den Berg,D.L.C., Zhang,W., Yates,A., Engelen,E., Takacs,K., Bezstarosti,K., Demmers,J., Chambers,I. and Poot,R.A. (2008) Estrogen-related receptor beta interacts with Oct4 to positively regulate Nanog gene expression. *Mol. Cell. Biol.*, **28**, 5986–5995.

54. Silva,J., Barrandon,O., Nichols,J., Kawaguchi,J., Theunissen,T.W. and Smith,A. (2008) Promotion of Reprogramming to Ground State Pluripotency by Signal Inhibition. *PLOS Biol.*, **6**, e253.

55. Sim,Y.-J., Kim,M.-S., Nayfeh,A., Yun,Y.-J., Kim,S.-J., Park,K.-T., Kim,C.-H. and Kim,K.-S. (2017) 2i Maintains a Naive Ground State in ESCs through Two Distinct Epigenetic

Mechanisms. *Stem cell reports*, **8**, 1312–1328.

56. Festuccia,N., Osorno,R., Halbritter,F., Karwacki-Neisius,V., Navarro,P., Colby,D., Wong,F., Yates,A., Tomlinson,S.R. and Chambers,I. (2012) Esrrb is a direct Nanog target gene that can substitute for Nanog function in pluripotent cells. *Cell Stem Cell*, **11**, 477–490.

57. Yamaji,M., Ueda,J., Hayashi,K., Ohta,H., Yabuta,Y., Kurimoto,K., Nakato,R., Yamada,Y., Shirahige,K. and Saitou,M. (2013) PRDM14 Ensures Naive Pluripotency through Dual Regulation of Signaling and Epigenetic Pathways in Mouse Embryonic Stem Cells. *Cell Stem Cell*, **12**, 368–382.

58. Yamanaka,S. (2009) Elite and stochastic models for induced pluripotent stem cell generation. *Nature*, **460**, 49–52.

59. Jaenisch,R. and Young,R. (2008) Stem Cells, the Molecular Circuitry of Pluripotency and Nuclear Reprogramming. *Cell*, **132**, 567–582.

60. Plath,K. and Lowry,W.E. (2011) Progress in understanding reprogramming to the induced pluripotent state. *Nat. Rev. Genet.*, **12**, 253–265.

61. Sommer,C.A., Stadtfeld,M., Murphy,G.J., Hochedlinger,K., Kotton,D.N. and Mostoslavsky,G. (2009) Induced pluripotent stem cell generation using a single lentiviral stem cell cassette. *Stem Cells*, **27**, 543–549.

62. Polo,J.M., Anderssen,E., Walsh,R.M., Schwarz,B.A., Nefzger,C.M., Lim,S.M., Borkent,M., Apostolou,E., Alaei,S., Cloutier,J., *et al.* (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell*, **151**, 1617–1632.

63. Chen,H.-F., Chuang,C.-Y., Lee,W.-C., Huang,H.-P., Wu,H.-C., Ho,H.-N., Chen,Y.-J. and Kuo,H.-C. (2011) Surface marker epithelial cell adhesion molecule and E-cadherin facilitate the identification and selection of induced pluripotent stem cells. *Stem cell Rev. reports*, **7**, 722–735.

64. Brons,I.G.M., Smithers,L.E., Trotter,M.W.B., Rugg-Gunn,P., Sun,B., Chuva de Sousa Lopes,S.M., Howlett,S.K., Clarkson,A., Ahrlund-Richter,L., Pedersen,R.A., *et al.* (2007) Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature*, **448**, 191–195.

65. Tesar,P.J., Chenoweth,J.G., Brook,F.A., Davies,T.J., Evans,E.P., Mack,D.L., Gardner,R.L. and McKay,R.D.G. (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, **448**, 196–199.

66. Ghimire,S., Van der Jeught,M., Neupane,J., Roost,M.S., Anckaert,J., Popovic,M., Van Nieuwerburgh,F., Mestdagh,P., Vandesompele,J., Deforce,D., *et al.* (2018) Comparative analysis of naive, primed and ground state pluripotency in mouse embryonic stem cells originating from the same genetic background. *Sci. Rep.*, **8**, 5884.

67. Guo,G., Yang,J., Nichols,J., Hall,J.S., Eyres,I., Mansfield,W. and Smith,A. (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development*, **136**, 1063–1069.

68. Tosolini,M. and Jouneau,A. (2016) From Naive to Primed Pluripotency: In Vitro Conversion of Mouse Embryonic Stem Cells in Epiblast Stem Cells. *Methods Mol. Biol.*, **1341**, 209–216.

69. Adachi,K., Kopp,W., Wu,G., Heising,S., Greber,B., Stehling,M., Araúzo-Bravo,M.J.,

Boerno,S.T., Timmermann,B., Vingron,M., *et al.* (2018) Esrrb Unlocks Silenced Enhancers for Reprogramming to Naive Pluripotency. *Cell Stem Cell*, **23**, 266--275.e6.

70. Gillich,A., Bao,S., Grabole,N., Hayashi,K., Trotter,M.W.B., Pasque,V., Magnúsdóttir,E. and Surani,M.A. (2012) Epiblast stem cell-based system reveals reprogramming synergy of germline factors. *Cell Stem Cell*, **10**, 425–439.

71. Silva,J., Nichols,J., Theunissen,T.W., Guo,G., van Oosten,A.L., Barrandon,O., Wray,J., Yamanaka,S., Chambers,I. and Smith,A. (2009) Nanog is the gateway to the pluripotent ground state. *Cell*, **138**, 722–737.

72. Ferrai,C., Torlai Triglia,E., Risner-Janiczek,J.R., Rito,T., Rackham,O.J., de Santiago,I., Kukalev,A., Nicodemi,M., Akalin,A., Li,M., *et al.* (2017) RNA polymerase II primes Polycomb-repressed developmental genes throughout terminal neuronal differentiation. *Mol. Syst. Biol.*, **13**, 946.

73. Yamada,K., Kanda,H., Tanaka,S., Takamatsu,N., Shiba,T. and Ito,M. (2006) Sox15 enhances trophoblast giant cell differentiation induced by Hand1 in mouse placenta. *Differentiation.*, **74**, 212–221.

74. Kuroda,T., Yasuda,S., Tachi,S., Matsuyama,S., Kusakawa,S., Tano,K., Miura,T., Matsuyama,A. and Sato,Y. (2019) SALL3 expression balance underlies lineage biases in human induced pluripotent stem cell differentiation. *Nat. Commun.*, **10**, 2175.

75. Harrison,S.J., Parrish,M. and Monaghan,A.P. (2008) Sall3 is required for the terminal maturation of olfactory glomerular interneurons. *J. Comp. Neurol.*, **507**, 1780–1794.

76. Betancourt,J., Katzman,S. and Chen,B. (2014) Nuclear factor one B regulates neural stem cell differentiation and axonal projection of corticofugal neurons. *J. Comp. Neurol.*, **522**, 6–35.

77. Gonda,Y., Namba,T. and Hanashima,C. (2020) Beyond Axon Guidance: Roles of Slit-Robo Signaling in Neocortical Formation . *Front. Cell Dev. Biol.* , **8**.

78. Yeh,M.L., Gonda,Y., Mommersteeg,M.T.M., Barber,M., Ypsilanti,A.R., Hanashima,C., Parnavelas,J.G. and Andrews,W.D. (2014) Robo1 Modulates Proliferation and Neurogenesis in the Developing Neocortex. *J. Neurosci.*, **34**, 5717–5731.

79. Katoh,M. and Katoh,M. (2007) Notch signaling in gastrointestinal tract (review). *Int. J. Oncol.*, **30**, 247–251.

80. Sueda,R., Imayoshi,I., Harima,Y. and Kageyama,R. (2019) High Hes1 expression and resultant Ascl1 suppression regulate quiescent vs. active neural stem cells in the adult mouse brain. *Genes Dev.*, **33**, 511–523.

81. Ohtsuka,T., Sakamoto,M., Guillemot,F. and Kageyama,R. (2001) Roles of the basic helix-loop-helix genes Hes1 and Hes5 in expansion of neural stem cells of the developing brain. *J. Biol. Chem.*, **276**, 30467–30474.

82. Ohtsuka,T., Ishibashi,M., Gradwohl,G., Nakanishi,S., Guillemot,F. and Kageyama,R. (1999) Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation. *EMBO J.*, **18**, 2196–2207.

83. Kageyama,R., Ohtsuka,T. and Kobayashi,T. (2007) The Hes gene family: repressors and oscillators that orchestrate embryogenesis. *Development*, **134**, 1243–1251.

84. Basak,O. and Taylor,V. (2007) Identification of self-replicating multipotent progenitors in the embryonic nervous system by high Notch activity and Hes5 expression. *Eur. J. Neurosci.*, **25**, 1006–1022.

85. Bansod,S., Kageyama,R. and Ohtsuka,T. (2017) Hes5 regulates the transition timing of neurogenesis and gliogenesis in mammalian neocortical development. *Development*, **144**, 3156–3167.

86. Panaliappan,T.K., Wittmann,W., Jidigam,V.K., Mercurio,S., Bertolini,J.A., Sghari,S., Bose,R., Patthey,C., Nicolis,S.K. and Gunhaga,L. (2018) Sox2 is required for olfactory pit formation and olfactory neurogenesis through BMP restriction and Hes5 upregulation. *Development*, **145**, dev153791.

87. Engelen,E., Akinci,U., Bryne,J.C., Hou,J., Gontan,C., Moen,M., Szumska,D., Kockx,C., van Ijcken,W., Dekkers,D.H.W., *et al.* (2011) Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. *Nat. Genet.*, **43**, 607–611.

88. Bonn,S., Zinzen,R.P., Girardot,C., Gustafson,E.H., Perez-Gonzalez,A., Delhomme,N., Ghavi-Helm,Y., Wilczyński,B., Riddell,A. and Furlong,E.E.M. (2012) Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.*, **44**, 148–156.

89. Fleck,J.S., Jansen,S.M.J., Wollny,D., Seimiya,M., Zenk,F., Santel,M., He,Z., Gray Camp,J. and Treutlein,B. (2021) Inferring and perturbing cell fate regulomes in human cerebral organoids. *bioRxiv*, 10.1101/2021.08.24.457460.

90. Furumatsu,T., Tsuda,M., Yoshida,K., Taniguchi,N., Ito,T., Hashimoto,M., Ito,T. and Asahara,H. (2005) Sox9 and p300 cooperatively regulate chromatin-mediated transcription. *J. Biol. Chem.*, **280**, 35203–35208.

91. Marathe,H.G., Watkins-Chow,D.E., Weider,M., Hoffmann,A., Mehta,G., Trivedi,A., Aras,S., Basuroy,T., Mehrotra,A., Bennett,D.C., *et al.* (2017) BRG1 interacts with SOX10 to establish the melanocyte lineage and to promote differentiation. *Nucleic Acids Res.*, **45**, 6442–6458.

92. Nowling,T., Bernadt,C., Johnson,L., Desler,M. and Rizzino,A. (2003) The co-activator p300 associates physically with and can mediate the action of the distal enhancer of the FGF-4 gene. *J. Biol. Chem.*, **278**, 13696–13705.

93. Chen,J., Zhang,Z., Li,L., Chen,B.-C., Revyakin,A., Hajj,B., Legant,W., Dahan,M., Lionnet,T., Betzig,E., *et al.* (2014) Single-Molecule Dynamics of Enhanceosome Assembly in Embryonic Stem Cells. *Cell*, **156**, 1274–1285.

94. Mullen,A.C., Orlando,D.A., Newman,J.J., Lovén,J., Kumar,R.M., Bilodeau,S., Reddy,J., Guenther,M.G., DeKoter,R.P. and Young,R.A. (2011) Master transcription factors determine cell-type-specific responses to TGF-β signaling. *Cell*, **147**, 565–576.

95. Liu,X., Wang,C., Liu,W., Li,J., Li,C., Kou,X., Chen,J., Zhao,Y., Gao,H., Wang,H., *et al.* (2016) Distinct features of H3K4me3 and H3K27me3 chromatin domains in pre-implantation embryos. *Nature*, **537**, 558–562.

96. Assou,S., Cerecedo,D., Tondeur,S., Pantesco,V., Hovatta,O., Klein,B., Hamamah,S. and De Vos,J. (2009) A gene expression signature shared by human mature oocytes and embryonic stem cells. *BMC Genomics*, **10**, 10.

97. Awe,J.P. and Byrne,J.A. (2013) Identifying candidate oocyte reprogramming factors using

cross-species global transcriptional analysis. *Cell. Reprogram.*, **15**, 126–133.

98. Sybirna,A., Tang,W.W.C., Pierson Smela,M., Dietmann,S., Gruhn,W.H., Brosh,R. and Surani,M.A. (2020) A critical role of PRDM14 in human primordial germ cell fate revealed by inducible degrons. *Nat. Commun.*, **11**, 1282.

99. Chan,Y.-S., Göke,J., Lu,X., Venkatesan,N., Feng,B., Su,I.-H. and Ng,H.-H. (2013) A PRC2-dependent repressive role of PRDM14 in human embryonic stem cells and induced pluripotent stem cell reprogramming. *Stem Cells*, **31**, 682–692.

100. Nagamatsu,G., Kosaka,T., Kawasumi,M., Kinoshita,T., Takubo,K., Akiyama,H., Sudo,T., Kobayashi,T., Oya,M. and Suda,T. (2011) A germ cell-specific gene, Prmt5, works in somatic cell reprogramming. *J. Biol. Chem.*, **286**, 10641–10648.

101. Wang,X., Veerapandian,V., Yang,X., Song,K., Xu,X., Cui,M., Yuan,W., Huang,Y., Xia,X., Yao,Z., *et al.* (2021) The chromatin accessibility landscape reveals distinct transcriptional regulation in the induction of human primordial germ cell-like cells from pluripotent stem cells. *Stem cell reports*, **16**, 1245–1261.

102. Pierson Smela,M., Sybirna,A., Wong,F.C.K. and Surani,M.A. (2019) Testing the role of SOX15 in human primordial germ cell fate. *Wellcome open Res.*, **4**, 122.

103. Lopez-Caraballo,L., Martorell-Marugan,J., Carmona-Saez,P. and Gonzalez-Muñoz,E. (2020) Analysis of Menstrual Blood Stromal Cells Reveals SOX15 Triggers Oocyte-Based Human Cell Reprogramming. *iScience*, **23**, 101376.

104. Apostolou,E. and Hochedlinger,K. (2013) Chromatin dynamics during cellular reprogramming. *Nature*, **502**, 462–471.

105. Orkin,S.H. and Hochedlinger,K. (2011) Chromatin connections to pluripotency and cellular reprogramming. *Cell*, **145**, 835–850.

106. Bernemann,C., Greber,B., Ko,K., Sterneckert,J., Han,D.W., Araúzo-Bravo,M.J. and Schöler,H.R. (2011) Distinct developmental ground states of epiblast stem cell lines determine different pluripotency features. *Stem Cells*, **29**, 1496–1503.

107. Bock,C., Kiskinis,E., Verstappen,G., Gu,H., Boulting,G., Smith,Z.D., Ziller,M., Croft,G.F., Amoroso,M.W., Oakley,D.H., *et al.* (2011) Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*, **144**, 439–452.

108. Hitoshi,S., Ishino,Y., Kumar,A., Jasmine,S., Tanaka,K.F., Kondo,T., Kato,S., Hosoya,T., Hotta,Y. and Ikenaka,K. (2011) Mammalian Gcm genes induce Hes5 expression by active DNA demethylation and induce neural stem cells. *Nat. Neurosci.*, **14**, 957–964.

109. Crispatzu,G., Rehimi,R., Pachano,T., Bleckwehl,T., Cruz-Molina,S., Xiao,C., Mahabir,E., Bazzi,H. and Rada-Iglesias,A. (2021) The chromatin, topological and regulatory properties of pluripotency-associated poised enhancers are conserved in vivo. *Nat. Commun.*, **12**, 4344.

110. Lee,H.-J., Göring,W., Ochs,M., Mühlfeld,C., Steding,G., Paprotta,I., Engel,W. and Adham,I.M. (2004) Sox15 is required for skeletal muscle regeneration. *Mol. Cell. Biol.*, **24**, 8428–8436.

111. Hatakeyama,J., Bessho,Y., Katoh,K., Ookawara,S., Fujioka,M., Guillemot,F. and Kageyama,R. (2004) Hes genes regulate size, shape and histogenesis of the nervous

system by control of the timing of neural stem cell differentiation. *Development*, **131**, 5539–5550.

112. Imayoshi,I., Isomura,A., Harima,Y., Kawaguchi,K., Kori,H., Miyachi,H., Fujiwara,T., Ishidate,F. and Kageyama,R. (2013) Oscillatory control of factors determining multipotency and fate in mouse neural progenitors. *Science*, **342**, 1203–1208.

113. Galiakberova,A.A. and Dashinimaev,E.B. (2020) Neural Stem Cells and Methods for Their Generation From Induced Pluripotent Stem Cells in vitro . *Front. Cell Dev. Biol.* , **8**.

114. Braccioli,L., Vervoort,S.J., Puma,G., Nijboer,C.H. and Coffer,P.J. (2018) SOX4 inhibits oligodendrocyte differentiation of embryonic neural stem cells in vitro by inducing Hes5 expression. *Stem Cell Res.*, **33**, 110–119.

115. Matsuda,S., Kuwako,K., Okano,H.J., Tsutsumi,S., Aburatani,H., Saga,Y., Matsuzaki,Y., Akaike,A., Sugimoto,H. and Okano,H. (2012) Sox21 promotes hippocampal adult neurogenesis via the transcriptional repression of the Hes5 gene. *J. Neurosci.*, **32**, 12543–12557.

Figure 1

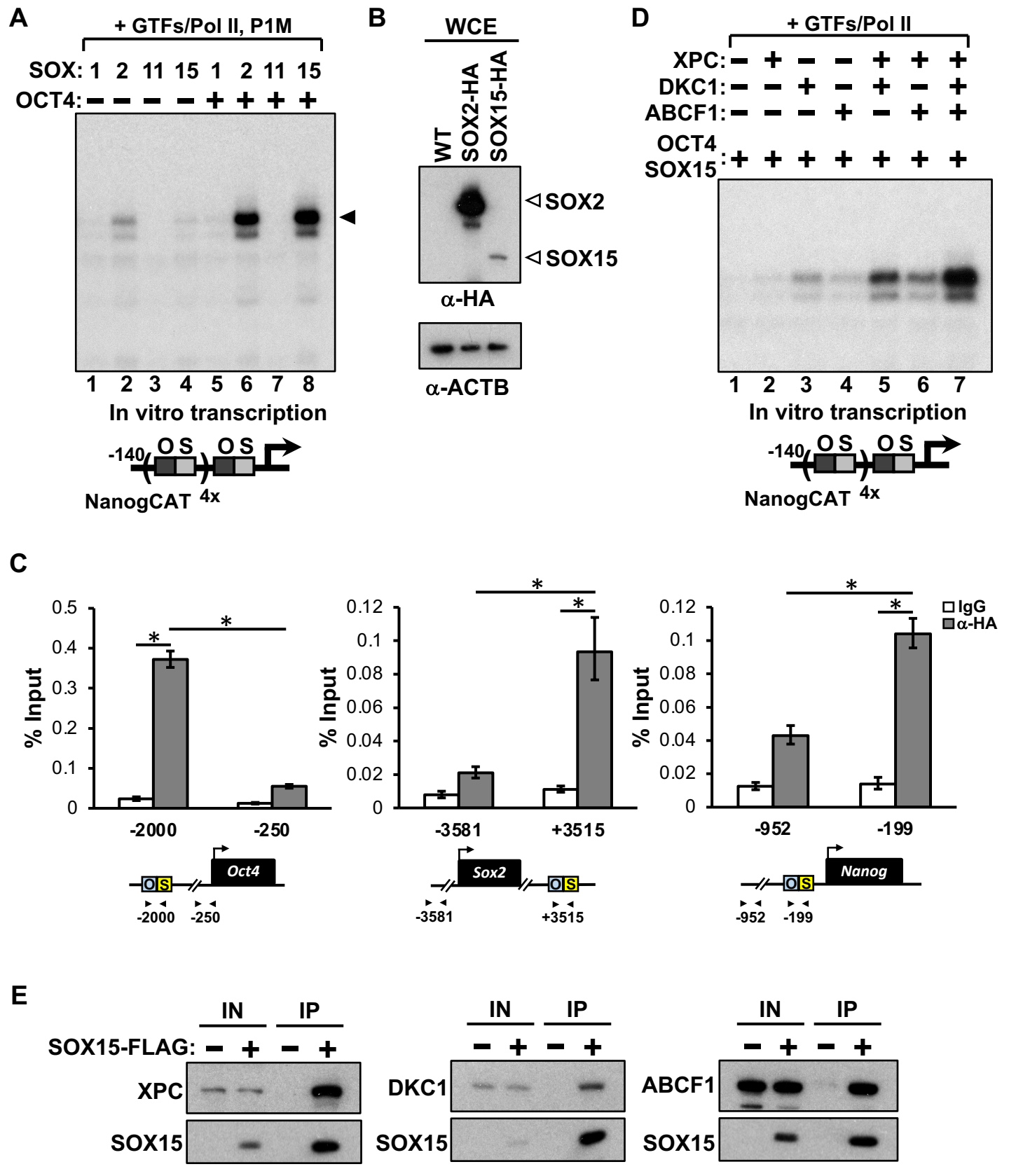

A
+ GTFs/Pol II, P1M
SOX: 1 2 11 15 1 2 11 15
OCT4: − − − − + + + +
1 2 3 4 5 6 7 8
In vitro transcription
-140
O S
O S
NanogCAT 4x
B
WCE
WT
SOX2-HA
SOX15-HA
SOX2
SOX15
α-HA
α-ACTB
D
+ GTFs/Pol II
XPC: − + − − + + +
DKC1: − − + − + − +
ABCF1: − − − + − + +
OCT4 SOX15: + + + + + + +
1 2 3 4 5 6 7
In vitro transcription
-140
O S
O S
NanogCAT 4x
C
% Input
IgG
α-HA
-2000
-250
Oct4
-3581
+3515
Sox2
-952
-199
Nanog
E
IN
IP
SOX15-FLAG: − + − +
XPC
SOX15
DKC1
ABCF1

Figure 1. Transcriptional activation of *NANOG* by SOX2 and SOX15 requires Stem Cell Coactivators (SCCs). (A) Equal molar of indicated SOX proteins are added to in vitro transcription reactions in the absence (-) or presence (+) of OCT4 and tested for their ability to stimulate transcription from the human *NANOG* promoter template engineered with four extra copies of the oct-sox enhancer element (bottom). All reactions contain purified general transcription factors (GTFs), RNA polymerase II (Pol II), and a partially purified nuclear fraction (phosphocellulose 1M, [P1M]) containing stem cell-enriched coactivators (SCCs) essential for SOX2/OCT4-dependent transcriptional activation. Transcribed RNA products are subjected to primer extension and visualized by autoradiography. Filled arrowhead indicates major transcription products. **(B)** Expression of SOX2-HA and SOX15-HA in whole-cell extracts (WCEs) of knockin mouse ESC lines is analyzed by western blotting using an antibody against HA tag. β-actin (ACTB) is used as loading control. **(C)** Micrococcal nuclease (MNase) ChIP analysis of SOX15 occupancy on control and enhancer regions of *Oct4*, *Sox2*, and *Nanog* gene loci in SOX15-HA knockin mouse ESCs. Representative data showing the enrichment of SOX15-HA (gray bars) compared to control IgGs (white bars) are analyzed by qPCR and expressed as percentage of input chromatin. Schematic diagrams of known oct-sox binding sites of each gene and the relative positions of the amplicons are shown at the bottom. **(D)** Transcriptional activation of *NANOG* by SOX15 and OCT4 in the presence of purified recombinant SCCs (XPC, DKC1, and ABCF1). **(E)** WCEs from 293T cells cotransfected with plasmids expressing FLAG-tagged SOX15 together with XPC, DKC1, or ABCF1 are immunoprecipitated with anti-FLAG antibody. Proteins coimmunoprecipitated with SOX15 are detected by western blotting using antibodies against XPC, DKC1, FLAG tag for SOX15-FLAG, and V5 tag for ABCF1-V5.

Figure 2

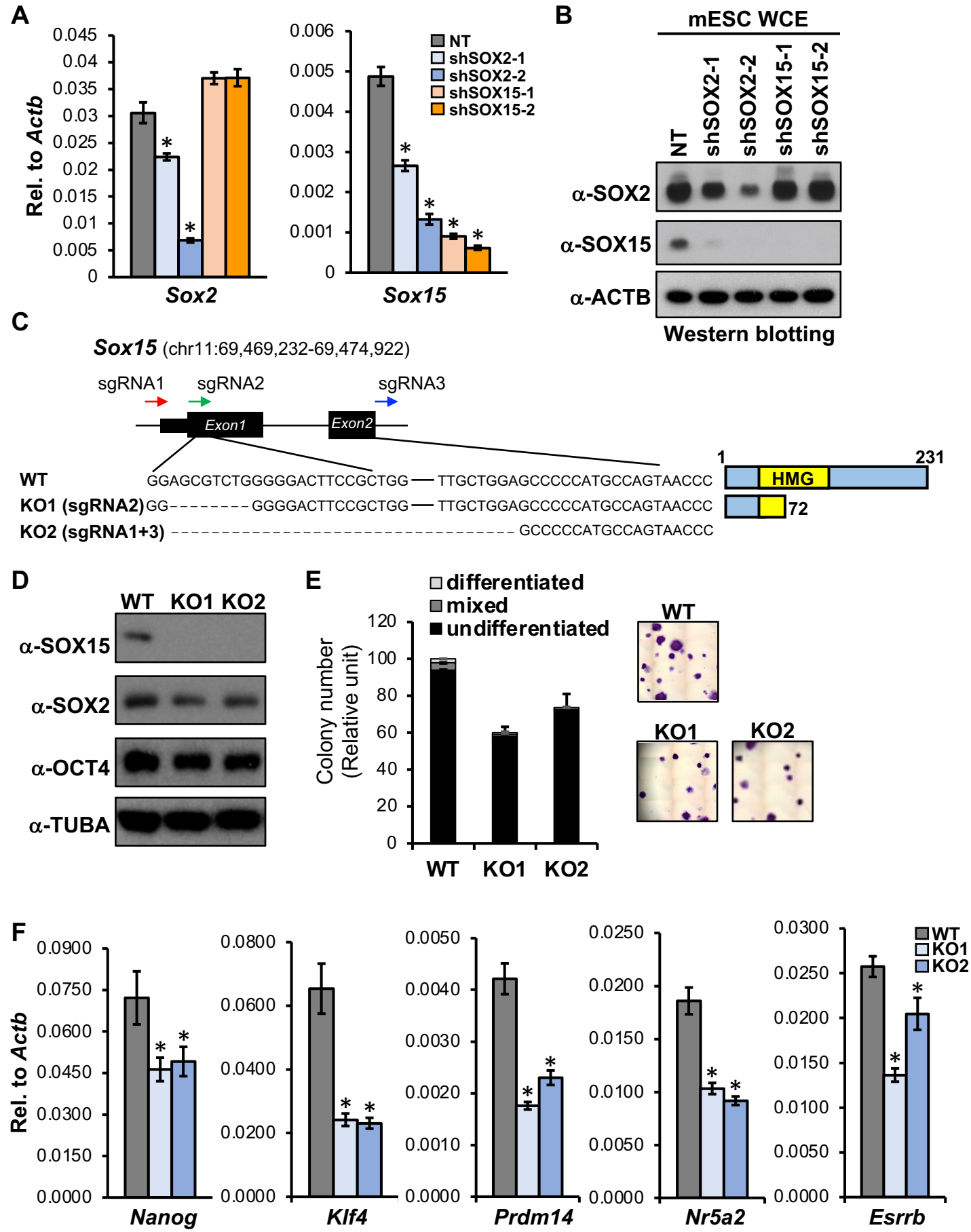

A
Rel. to Actb
0.045
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
Sox2
0.006
0.005
0.004
0.003
0.002
0.001
0
Sox15
NT
shSOX2-1
shSOX2-2
shSOX15-1
shSOX15-2
B
mESC WCE
NT
shSOX2-1
shSOX2-2
shSOX15-1
shSOX15-2
α-SOX2
α-SOX15
α-ACTB
Western blotting
C
Sox15 (chr11:69,469,232-69,474,922)
sgRNA1
sgRNA2
sgRNA3
Exon1
Exon2
WT GGAGCGTCTGGGGGACTTCCGCTGG — TTGCTGGAGCCCCCATGCCAGTAACCC
KO1 (sgRNA2) GG-------- GGGGACTTCCGCTGG — TTGCTGGAGCCCCCATGCCAGTAACCC
KO2 (sgRNA1+3) ---------- GCCCCCATGCCAGTAACCC
1
231
HMG
72
D
WT KO1 KO2
α-SOX15
α-SOX2
α-OCT4
α-TUBA
E
differentiated
mixed
undifferentiated
Colony number (Relative unit)
120
100
80
60
40
20
0
WT
KO1
KO2
F
Rel. to Actb
0.0900
0.0750
0.0600
0.0450
0.0300
0.0150
0.0000
Nanog
0.0800
0.0600
0.0400
0.0200
0.0000
Klf4
0.0050
0.0040
0.0030
0.0020
0.0010
0.0000
Prdm14
0.0250
0.0200
0.0150
0.0100
0.0050
0.0000
Nr5a2
0.0300
0.0250
0.0200
0.0150
0.0100
0.0050
0.0000
Esrrb
WT
KO1
KO2

Figure 2. SOX15 deficiency compromises stem cell self-renewal and pluripotency gene expression. (A) Partial depletion of SOX2 by targeted shRNAs compromises SOX15 expression. Quantification of *Sox2* and *Sox15* mRNA levels in SOX2 or SOX15-knockdown ESCs are analyzed by qPCR and normalized to *Actb*. **(B)** SOX2 and SOX15 protein levels in WCEs of SOX2- and SOX15-knockdown ESCs are analyzed by western blotting, ACTB as loading control. **(C)** Left: Schematic diagram of *Sox15* deletion strategies using CRISPR/Cas9 system. KO1 contains an 8 bp deletion induced by sgRNA2, leading to frameshift translation stop at amino acid (aa) 72 (full-length SOX15 contains 231 aa's). KO2 is generated by sgRNA1 and 3-mediated excision of the entire exon 1 and 2 of the *Sox15* gene. Right: Schematic representation of full-length SOX15 in WT ESCs and predicted SOX15 fragment (1-72 aa) in KO1. Relative position of HMG box is highlighted in yellow. **(D)** Western blotting of ESC WCEs confirms the absence of SOX15 in KO1 and KO2, while SOX2, OCT4, and α-tubulin (TUBA) levels remain largely the same as wild-type (WT) ESCs. **(E)** Colony formation assays of WT and *Sox15* KO ESCs. Differentiation status is evaluated on the basis of AP staining intensity and colony morphology. Representative images of AP staining of control and KO ESCs are shown (right). **(F)** Loss of SOX15 in ESCs compromises pluripotency gene expression. Quantification of mRNA levels of indicated pluripotency-associated genes by qPCR, normalized to *Actb*.

Figure 3

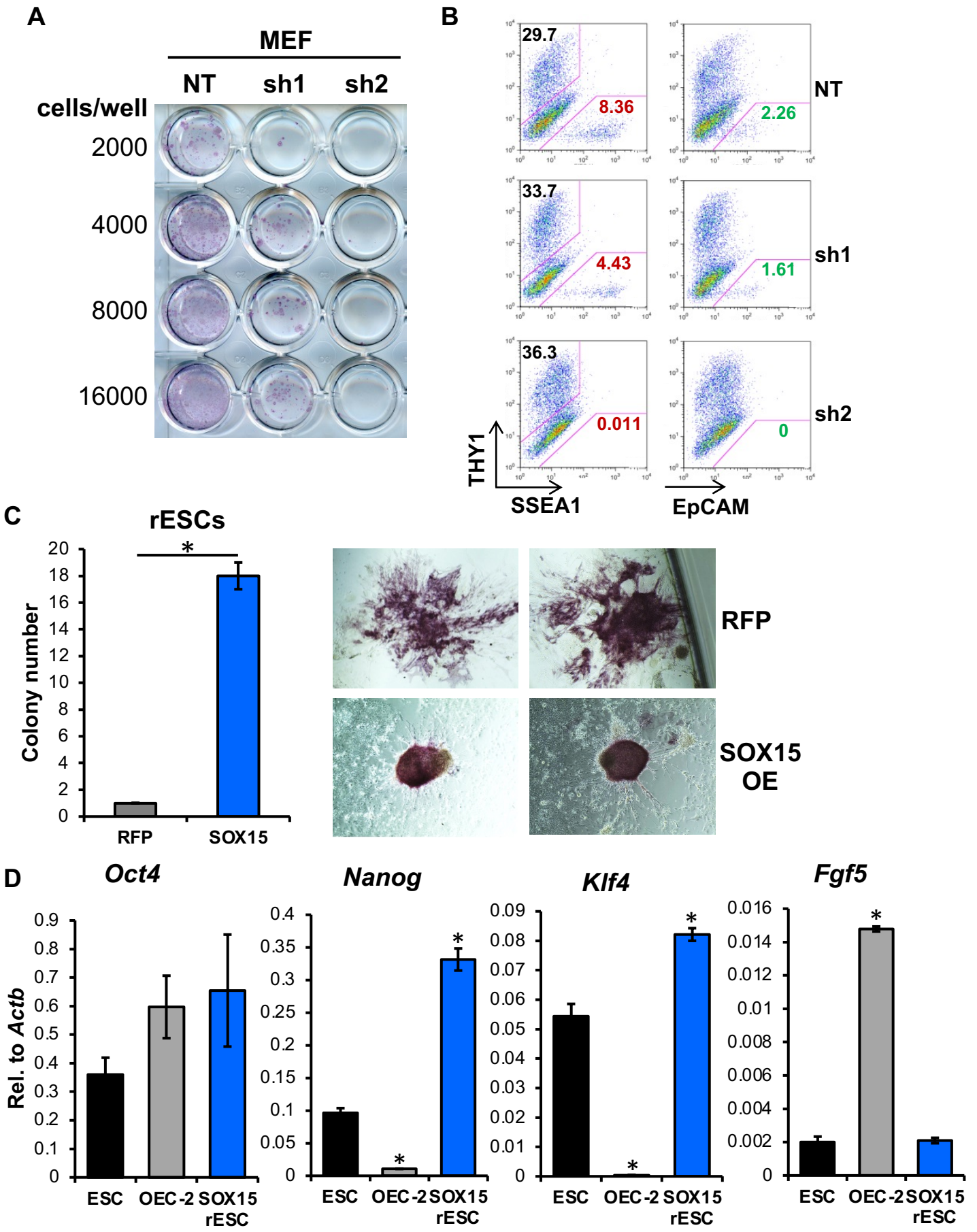
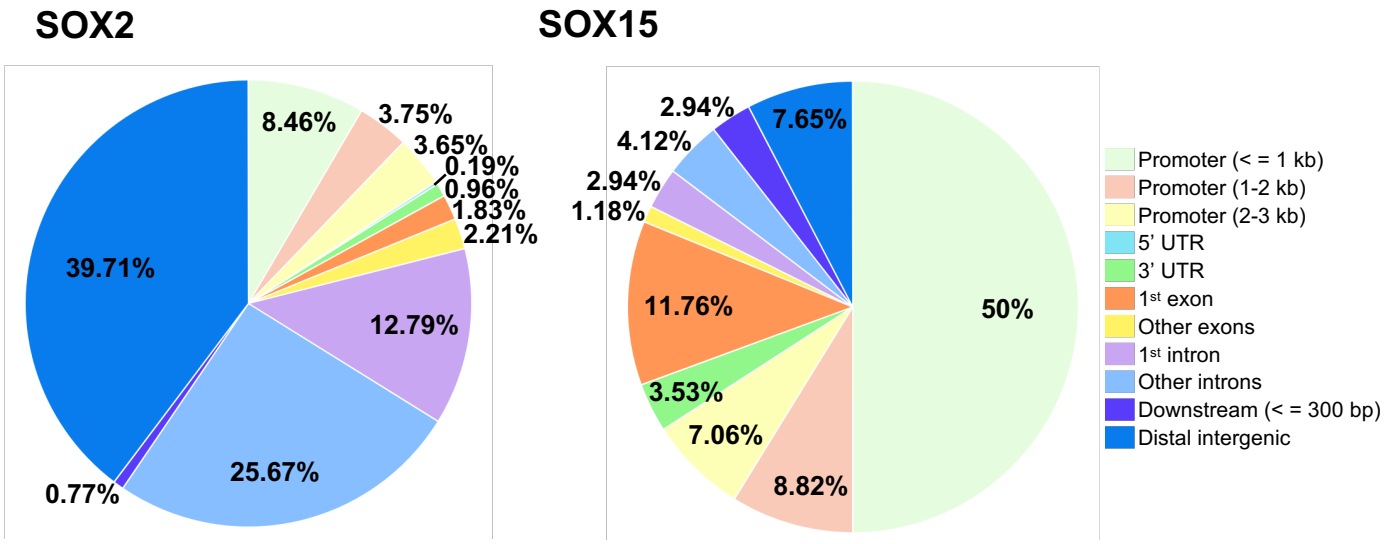

A
MEF
NT
sh1
sh2
cells/well
2000
4000
8000
16000
B
29.7
8.36
2.26
NT
33.7
4.43
1.61
sh1
36.3
0.011
0
sh2
THY1
SSEA1
EpCAM
C
rESCs
Colony number
RFP
SOX15
SOX15 OE
D
Oct4
Nanog
Klf4
Fgf5
Rel. to Actb
ESC
OEC-2
SOX15 rESC

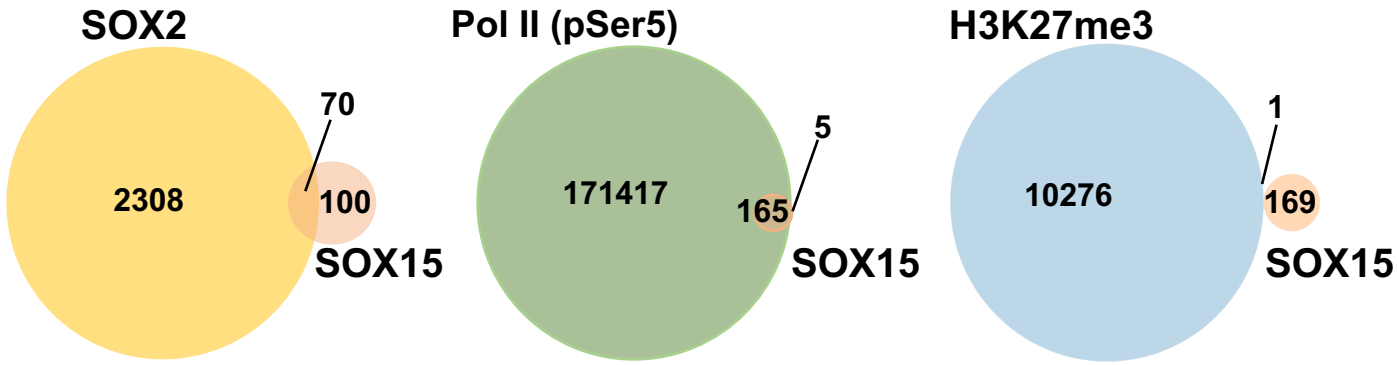
Figure 3. Regulation of cellular reprogramming by SOX15. (A) Depletion of SOX15 blocks iPSC generation from mouse embryonic fibroblasts (MEFs). Control (NT) and SOX15-knockdown (sh1 and 2) MEFs transduced with lentiviruses expressing reprogramming TFs (SOX2, OCT4, KLF4, and c-MYC) are plated at indicated cell numbers. Cells are stained for AP activity after 14 days post induction (dpi) of reprogramming TFs by doxycycline (dox) (11 days with dox followed by 3 days without dox). **(B)** Single cell suspensions of 14 dpi MEFs as described in (A) are stained with anti-THY1 together with anti-mouse SSEA1 or anti-EpCAM antibodies and analyzed by flow cytometry. Percentage of THY1+ and SSEA1+ cells are indicated. **(C)** OEC-2 epiblast stem cells (EpiSCs) are transduced with lentiviruses expressing SOX15 or RFP as control. Transduced cells are then replaced with 2i/LIF ESC medium. Colonies are stained for AP activity and counted after 14 days. Representative images of AP-staining of reprogrammed ESC (rESCs) colonies are shown (right). **(D)** mRNA levels of pluripotency-associated genes (*Oct4*, *Nanog*, and *Klf4*) and EpiSC-enriched marker *Fgf5* in ESCs, OEC-2, and SOX15-overexpressing rESCs are quantified by qPCR, normalized to *Actb*.

Figure 4

A



B



C

SOX15-SOX2 common motifs

SOX motifs

$p < 0.05$

D

SOX15-unique motifs

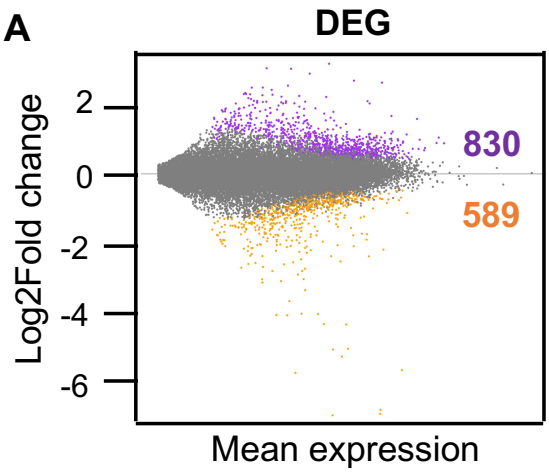
NANOG-ChIP enriched motif

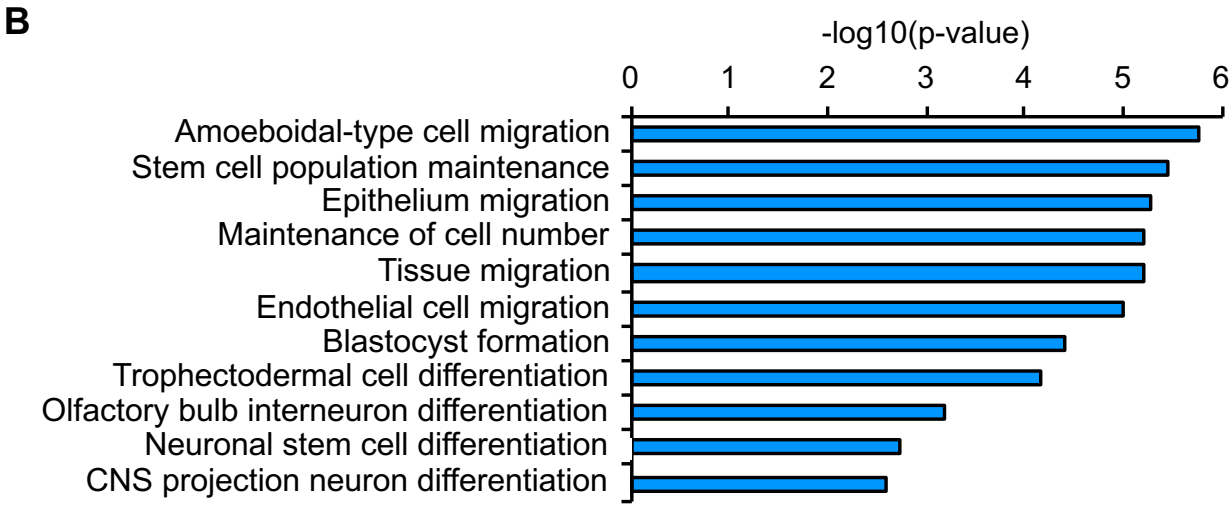
OCT4 motif

$p < 0.05$

Figure 4. SOX15 and SOX2 genomic targets in mouse ESCs partially overlap. (A) Pie charts showing distribution of SOX15 and SOX2 binding sites across the genome. Fraction of bound regions in each category is indicated. **(B)** Venn diagrams showing degree of overlap between SOX15 binding sites with SOX2, C-terminal domain serine 5 phosphorylated RNA polymerase II (Pol II-S5p), and repressive histone marks histone H3 lysine 27 trimethylation (H3K27me3). **(C)** Enrichment of TF binding motifs among binding sites common between SOX15 and SOX2, and **(D)** motifs enriched in SOX15-unique ChIP-seq peaks, analyzed by HOMER; p-values < 0.05.

Figure 5


A
DEG
Log2Fold change
2
0
-2
-4
-6
830
589
Mean expression


B
-log10(p-value)
0
1
2
3
4
5
6
Amoeboidal-type cell migration
Stem cell population maintenance
Epithelium migration
Maintenance of cell number
Tissue migration
Endothelial cell migration
Blastocyst formation
Trophectodermal cell differentiation
Olfactory bulb interneuron differentiation
Neuronal stem cell differentiation
CNS projection neuron differentiation

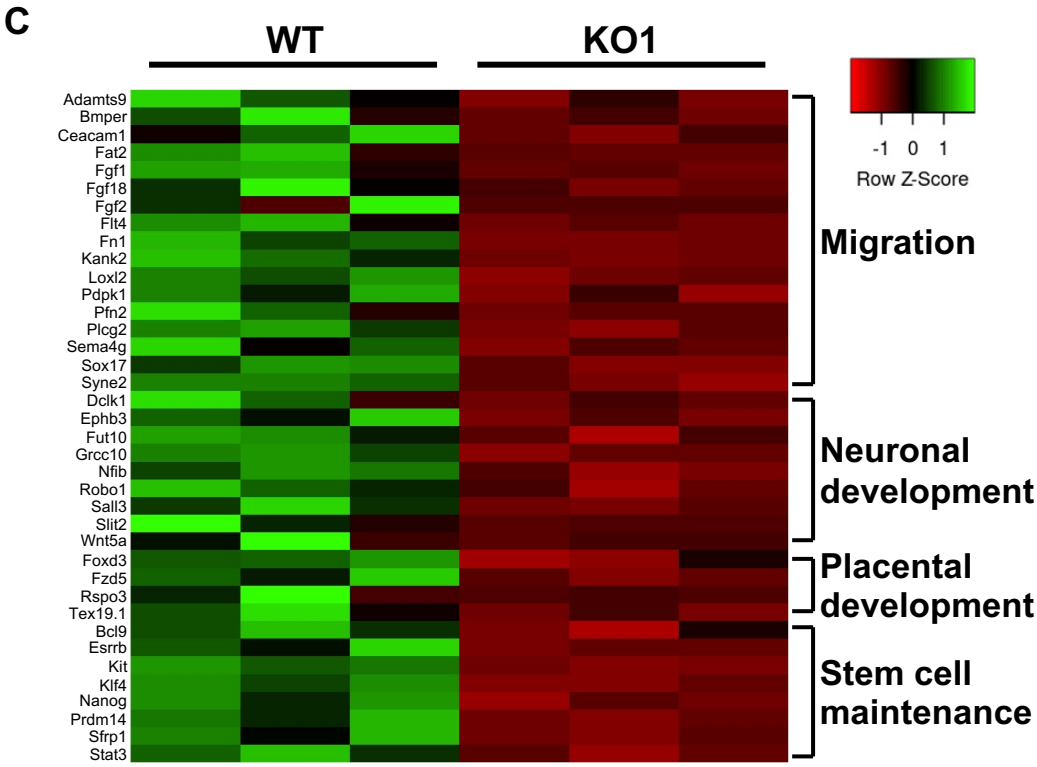
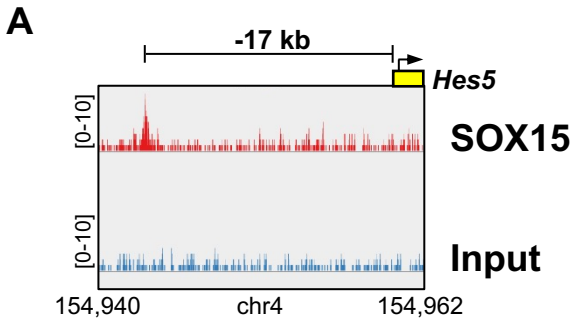
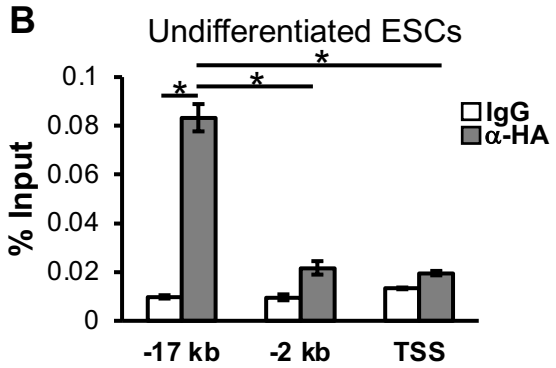
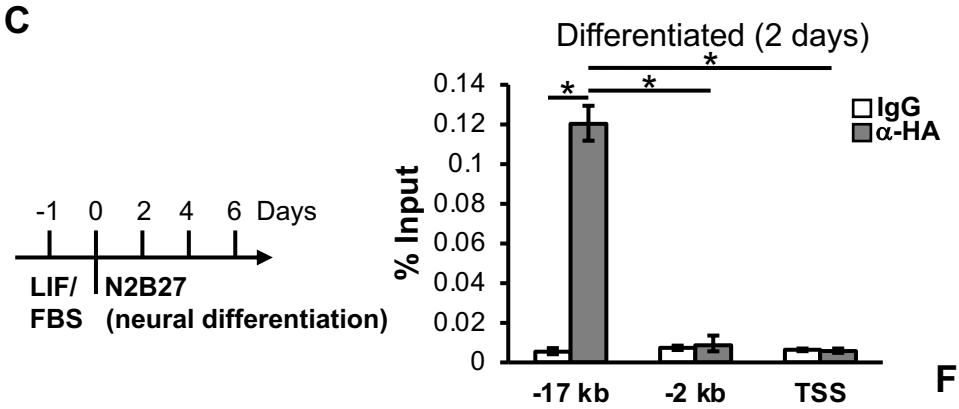

C
WT
KO1
-1
0
1
Row Z-Score
Adamts9
Bmper
Ceacam1
Fat2
Fgf1
Fgf18
Fgf2
Flt4
Fn1
Kank2
Loxl2
Pdpk1
Pfn2
Plcg2
Sema4g
Sox17
Syne2
Dclk1
Ephb3
Fut10
Grcc10
Nfib
Robo1
Sall3
Slit2
Wnt5a
Foxd3
Fzd5
Rspo3
Tex19.1
Bcl9
Esrrb
Kit
Klf4
Nanog
Prdm14
Sfrp1
Stat3
Migration
Neuronal development
Placental development
Stem cell maintenance

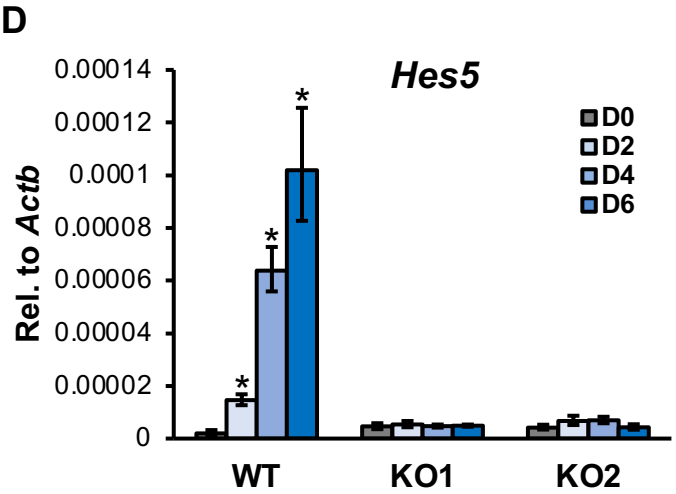
Figure 5. Gene expression profile of *Sox15*-knockout ESCs. (A) RNA-seq analysis identifies differentially expressed genes (DEGs) in *Sox15* KO1 ESCs compared to WT, plotted as a function of mean expression (fpkm) and fold changes (log2). Number of genes upregulated and downregulated are indicated. **(B)** Gene ontology (GO) analysis on DEGs in *Sox15* KO ESCs compared with all annotated genes. *p*-value per GO classification is shown (-log10(*p*-value)). **(C)** Heatmap showing the relative expression levels of genes indicated from three biological replicates each of WT and *Sox15* KO1 ESCs. Scaled values indicate relative downregulation (red) or upregulation (green) of gene expression. Genes from each category are manually curated.

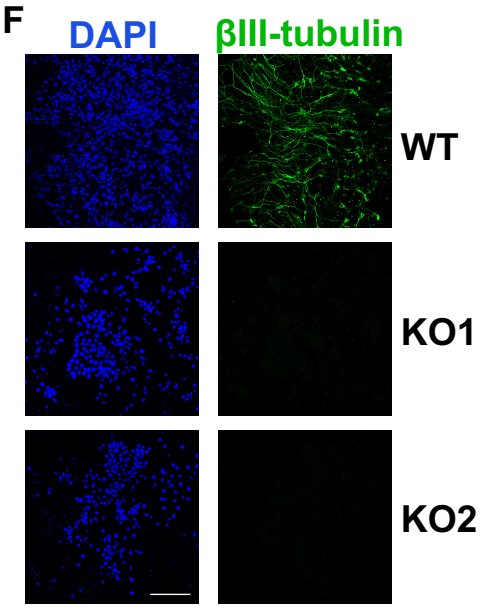
Figure 6


A
-17 kb
Hes5
[0-10]
SOX15
[0-10]
Input
154,940
chr4
154,962


B
Undifferentiated ESCs
% Input
0.1
0.08
0.06
0.04
0.02
0
IgG
α-HA
-17 kb
-2 kb
TSS


C
-1
0
2
4
6
Days
LIF/
FBS
N2B27
(neural differentiation)
Differentiated (2 days)
% Input
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
IgG
α-HA
-17 kb
-2 kb
TSS


D
Hes5
Rel. to Actb
0.00014
0.00012
0.0001
0.00008
0.00006
0.00004
0.00002
0
D0
D2
D4
D6
WT
KO1
KO2


F
DAPI
βIII-tubulin
WT
KO1
KO2

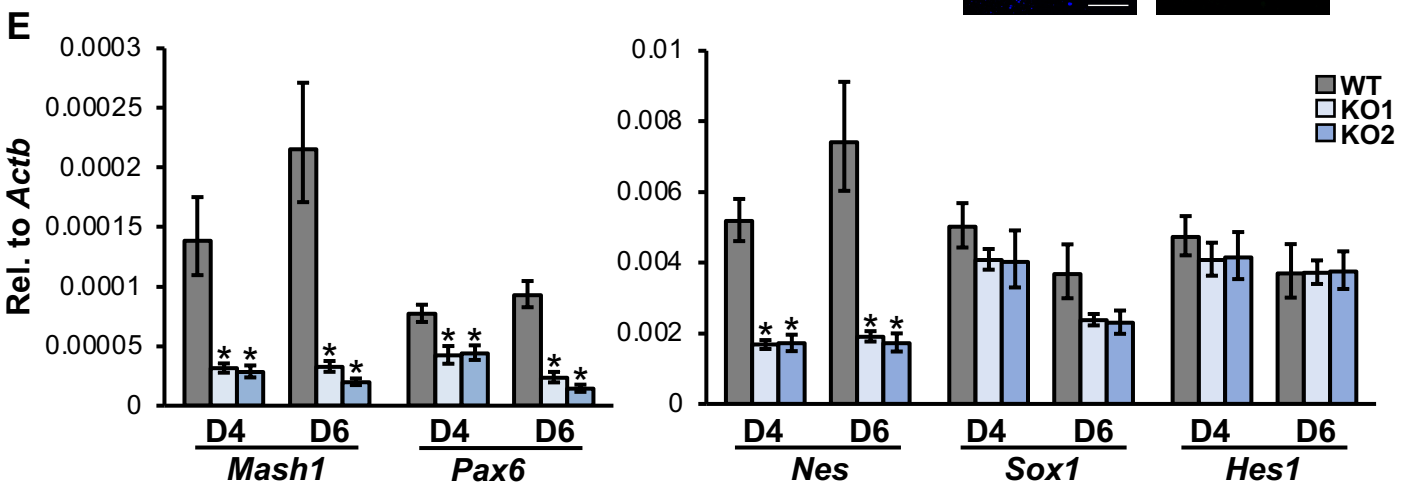

E
Rel. to Actb
0.0003
0.00025
0.0002
0.00015
0.0001
0.00005
0
D4
D6
Mash1
Pax6
0.01
0.008
0.006
0.004
0.002
0
WT
KO1
KO2
Nes
Sox1
Hes1

Figure 6. SOX15 regulates *Hes5* expression and neural differentiation. (A) Integrative Genomics Viewer (IGV)-computed ChIP-seq tracks are plotted as (number of reads) × [1,000,000/(total read count)] for the *Hes5* gene locus. Reads from SOX15-HA ChIP DNA and input chromatin are shown. Genomic coordinates are in kilobases. **(B)** MNase-ChIP analysis of SOX15 occupancy at -17 kb, -2 kb, and TSS of the *Hes5* gene in undifferentiated ESCs. Enrichment and data analysis are as described in Figure 1C. **(C)** Schematic diagram describing differentiation of ESCs into neuroectodermal precursors. ESCs cultured in LIF/fetal bovine serum (FBS) are plated 1 day before replacing with serum-free N2B27 differentiation medium. Cells are collected at 2, 4, and 6 days for analysis (left). MNase-ChIP analysis of SOX15 occupancy at the *Hes5* locus in differentiating ESCs (2 days), analyzed as described in (B). **(D)** *Hes5* mRNA levels in undifferentiated (D0) and differentiating WT, *Sox15* KO1 and KO2 ESCs collected at day 2, 4, and 6 of differentiation, analyzed by qPCR and normalized to *Actb*. **(E)** mRNA levels of indicated neural markers in differentiating WT and KO ESCs (D4 and 6) are quantified by qPCR and normalized to *Actb*. **(F)** Representative immunofluorescence images of WT and *Sox15* KO (KO1 and KO2) ESCs after 6 days of differentiation, stained with antibodies against neuronal marker βIII-tubulin. Nuclei are stained with DAPI. Scale bar, 100 μm.

Figure 7

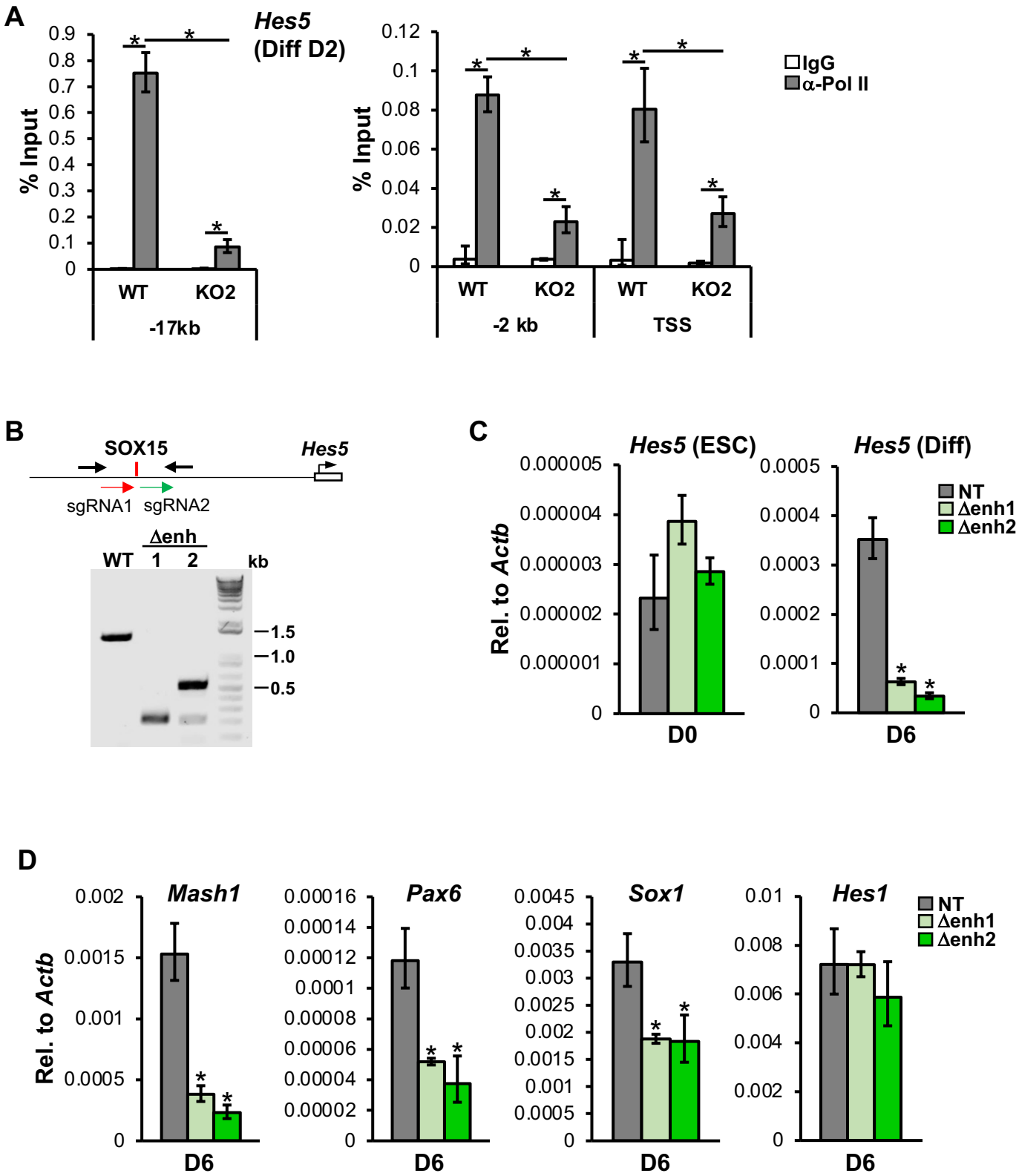

A
Hes5
(Diff D2)
% Input
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
WT
KO2
-17kb
0.12
0.1
0.08
0.06
0.04
0.02
0
-2 kb
TSS
IgG
α-Pol II
B
SOX15
Hes5
sgRNA1
sgRNA2
Δenh
WT
1
2
kb
1.5
1.0
0.5
C
Hes5 (ESC)
Hes5 (Diff)
Rel. to Actb
0.000005
0.000004
0.000003
0.000002
0.000001
0
D0
0.0005
0.0004
0.0003
0.0002
0.0001
D6
NT
Δenh1
Δenh2
D
Mash1
Pax6
Sox1
Hes1
0.002
0.0015
0.001
0.0005
0.00016
0.00014
0.00012
0.0001
0.00008
0.00006
0.00004
0.00002
0.0045
0.004
0.0035
0.003
0.0025
0.002
0.0015
0.001
0.0005
0.01
0.008
0.006
0.004
0.002

Figure 7. SOX15-bound -17 kb region controls *Hes5* gene expression during neural differentiation. (A) MNase-ChIP analysis of Pol II occupancy at indicated regions of the *Hes5* gene locus in WT and *Sox15* knockout (KO2) ESCs after 2 days of differentiation (Diff D2). Representative data showing the enrichment of Pol II (gray bars) compared to control IgGs (white bars) are analyzed as described in Figure 1C. **(B)** Schematic diagram showing CRISPR/Cas9 mediated deletion of the -17 kb region using two sgRNAs flanking a ~1 kb region (top). PCR analysis of genomic DNA from WT and two independent homozygous deletion ESC clones (Δenh1 and 2) (bottom). Primers used to amplify the deleted region indicated (black arrows). **(C)** *Hes5* mRNA levels in undifferentiated WT and Δenh ESCs and their differentiated (day 6, D6) counterparts are analyzed by qPCR, normalized to *Actb*. **(D)** mRNA levels of neural markers (*Mash1*, *Pax6*, *Sox1*, and *Hes1*) in differentiated WT and Δenh ESCs (D6) are quantified by qPCR, normalized to *Actb*.