Stem Cell Reports, Volume *13*

Supplemental Information

SOX21 Ensures Rostral Forebrain Identity by Suppression of WNT8B during Neural Regionalization of Human Embryonic Stem Cells

Zhuoqing Fang, Xinyuan Liu, Jing Wen, Fan Tang, Yang Zhou, Naihe Jing, and Ying Jin

Inventory of Supplemental Information

1. Supplemental Figures and Legends

1) Figure S1. Related to Figure 1. Region- specific transcriptome analysis of hESC-derived neural progenitor cells.

2) Figure S2. Related to Figure 2. SOX21 participates in forebrain regionalization.

3) Figure S3. Related to Figure 4. Genome-wide analysis of SOX21 binding peaks.

4) Figure S4. Relative to Figure 5. Restoration of the rostral forebrain fate in SOX21 knockout cells by DKK1 overexpression.

5) Figure S5. Related to Figure 6. WNT8B knockout largely rescues defects caused by SOX21 deletion.

2. Supplemental Tables

1) Table S1. Related to Figure S1. Modules classified by WGCNA.

2) Table S2. Related to Figure 4. The list of SOX21 positively regulated putative target genes.

3) Table S3. Related to Figure 4. The list of SOX21 negatively regulated putative target genes.

4) Table S4. Related to Figure 6. The list of potential SOX21 interacting proteins identified by mass spectrometric analysis.

5) Table S5. The list of real-time qPCR primers used in this study.

6) Table S6. The list of antibodies used in this study.

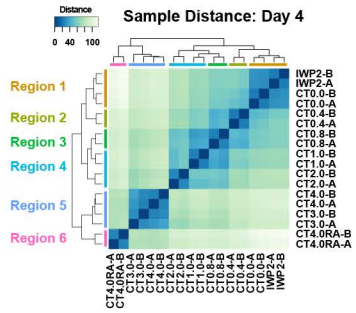
7) Table S7. The list of ChIP-qPCR primers used in this study.

3. Supplemental Experimental Procedures

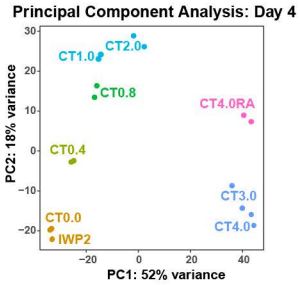
4. Supplemental References

Figure S1

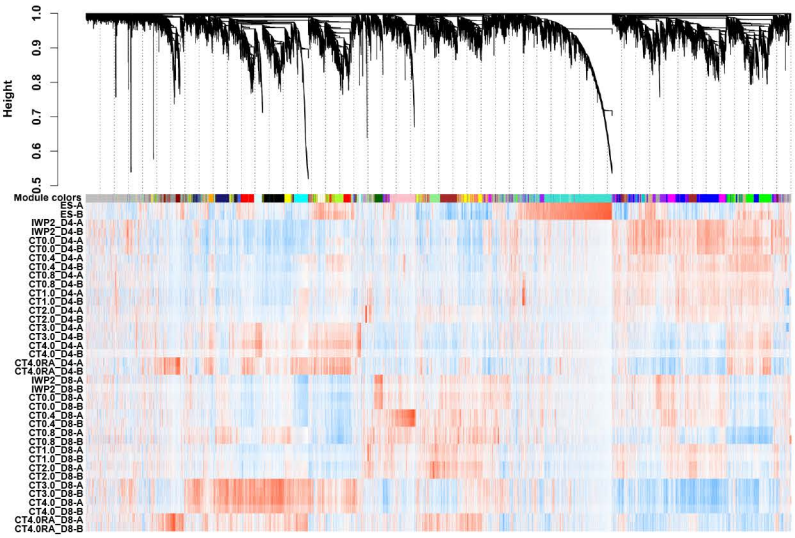
A



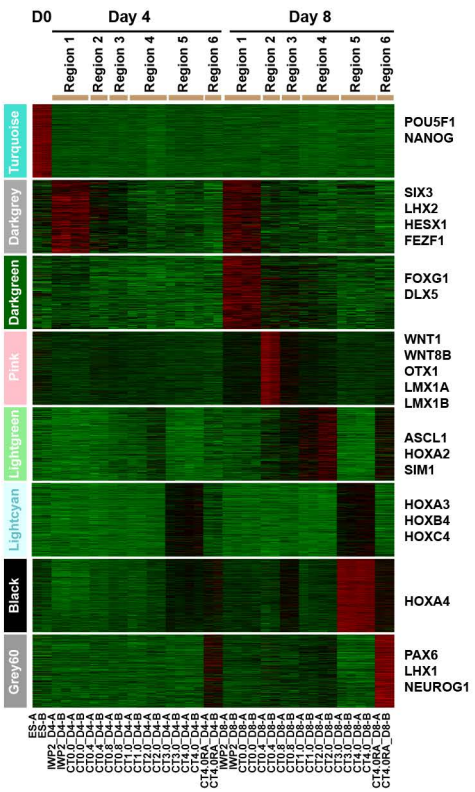
B



C



D



E

	Biological Process	Combined Score
Turquoise	response to interferon-gamma (GO:0034341)	34.3
	cellular response to cytokine stimulus (GO:0071345)	28.6
	cytokine-mediated signaling pathway (GO:0019221)	28.5
	cofactor metabolic process (GO:0051186)	27.3
	cellular response to interferon-gamma (GO:0071346)	26.2
Darkgrey	axis specification (GO:0009798)	20.5
	embryonic axis specification (GO:0000578)	19.5
	pattern specification process (GO:0007389)	19.4
	mitotic sister chromatid segregation (GO:0000070)	19.1
	anterior/posterior axis specification (GO:0009948)	19.1
Darkgreen	extracellular matrix organization (GO:0030198)	52.5
	extracellular structure organization (GO:0043062)	52.3
	glycosaminoglycan metabolic process (GO:0030203)	34.7
	aminoglycan metabolic process (GO:0006022)	33.5
	keratan sulfate catabolic process (GO:0042340)	33.0
Pink	axon guidance (GO:0007411)	49.9
	neuron projection guidance (GO:0097485)	49.8
	regulation of osteoblast differentiation (GO:0045667)	48.4
	midbrain development (GO:0030901)	47.6
	extracellular matrix organization (GO:0030198)	47.6
Lightgreen	neuron differentiation (GO:0030182)	42.6
	negative regulation of nervous system development (GO:0051961)	36.1
	regulation of neuron differentiation (GO:0045664)	34.0
	central nervous system neuron differentiation (GO:0021953)	33.7
	neuroblast proliferation (GO:0007405)	32.8
Lightcyan	anterior/posterior pattern specification (GO:0009952)	88.4
	regionalization (GO:0003002)	81.0
	pattern specification process (GO:0007389)	60.0
	organ morphogenesis (GO:0009887)	49.4
	embryonic morphogenesis (GO:0048598)	40.9
Black	regulation of translation (GO:0006417)	27.3
	posttranscriptional regulation of gene expression (GO:0010608)	27.3
	alcohol biosynthetic process (GO:0046165)	26.9
	positive regulation of translation (GO:0045727)	22.9
	cholesterol biosynthetic process (GO:0006695)	22.4
Grey60	pattern specification process (GO:0007389)	28.3
	regulation of neuron differentiation (GO:0045664)	26.6
	embryonic morphogenesis (GO:0048598)	24.3
	cellular response to growth hormone stimulus (GO:0071378)	19.4
	nervous system development (GO:0007399)	19.3

Figure S2

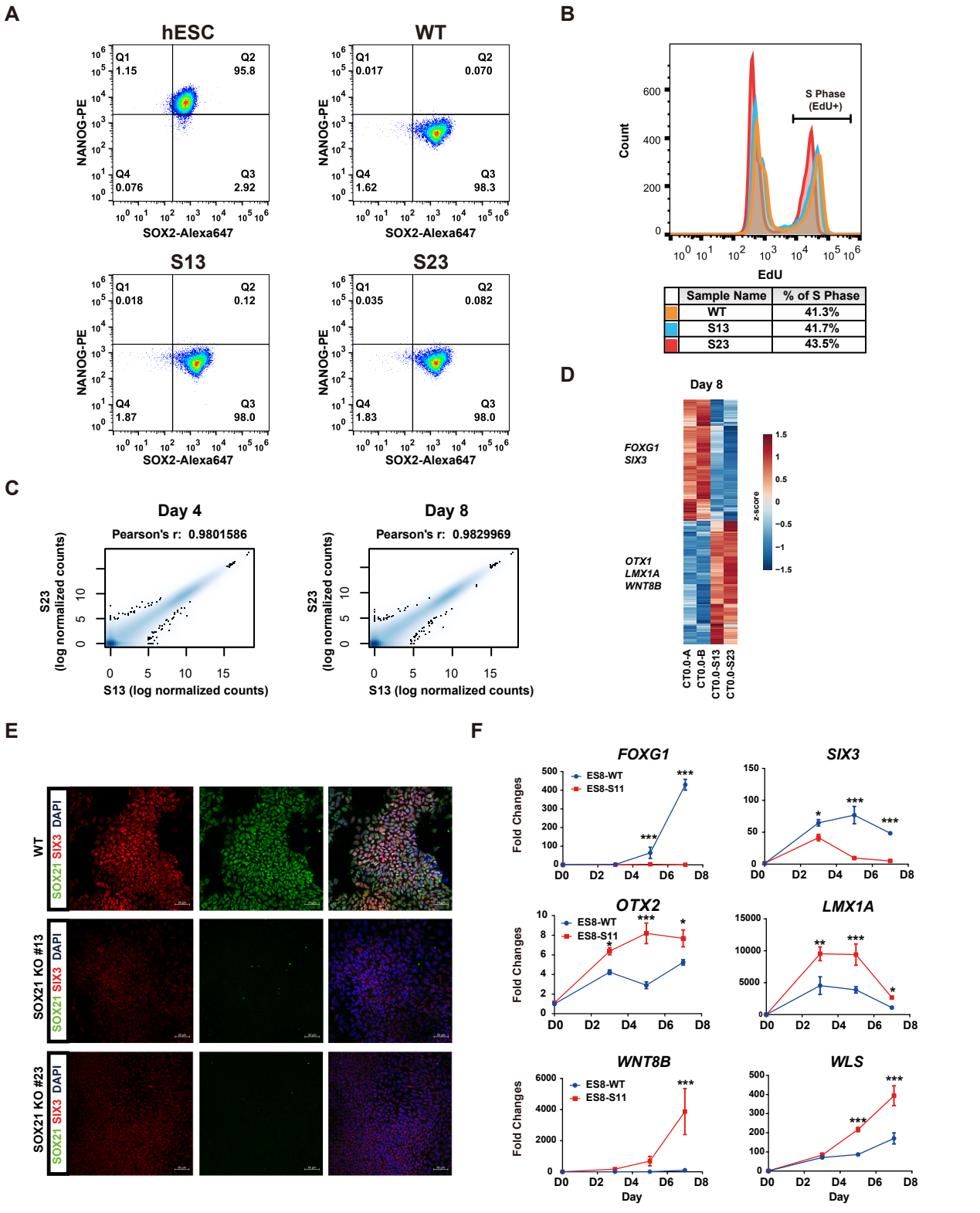

A
hESC
Q1 1.15
Q2 95.8
Q4 0.076
Q3 2.92
WT
Q1 0.017
Q2 0.070
Q4 1.62
Q3 98.3
S13
Q1 0.018
Q2 0.12
Q4 1.87
Q3 98.0
S23
Q1 0.035
Q2 0.082
Q4 1.83
Q3 98.0
NANOG-PE
SOX2-Alexa647
B
S Phase (EdU+)
Count
EdU
Sample Name | % of S Phase
WT | 41.3%
S13 | 41.7%
S23 | 43.5%
C
Day 4
Pearson's r: 0.9801586
Day 8
Pearson's r: 0.9829969
S23 (log normalized counts)
S13 (log normalized counts)
D
Day 8
FOXG1
SIX3
OTX1
LMX1A
WNT8B
z-score
CT0.0-A
CT0.0-B
CT0.0-S13
CT0.0-S23
E
WT
SOX21 KO #13
SOX21 KO #23
SOX21 SIX3 DAPI
F
FOXG1
SIX3
OTX2
LMX1A
WNT8B
WLS
ES8-WT
ES8-S11
Fold Changes
Day
D0 D2 D4 D6 D8

Figure S3

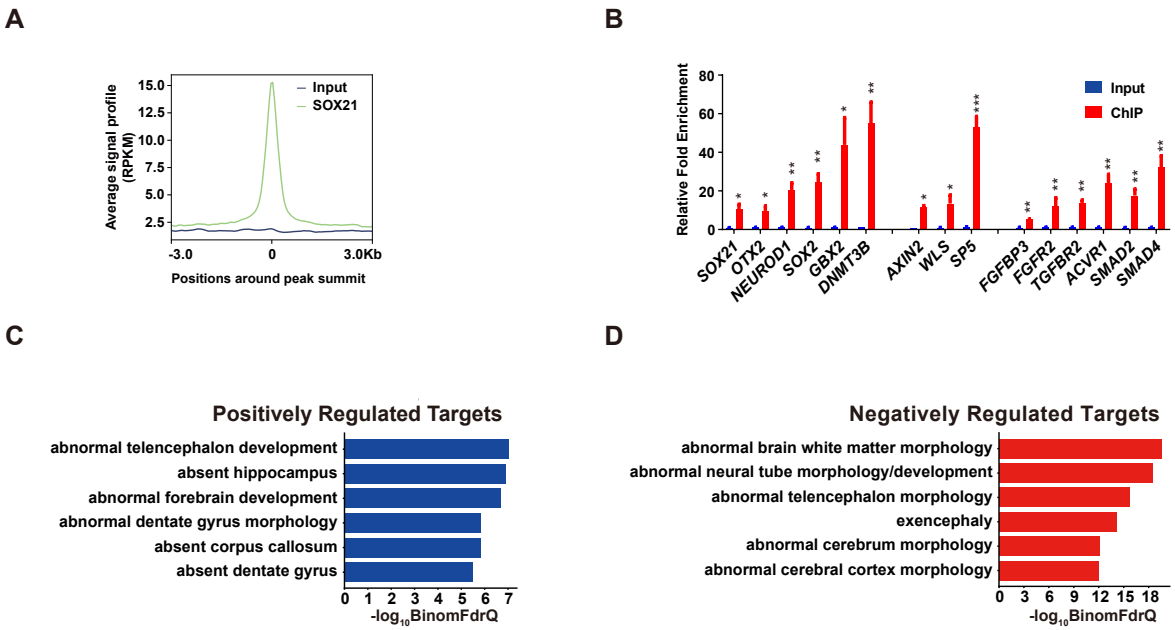

A
Average signal profile (RPKM)
15.0
12.5
10.0
7.5
5.0
2.5
Input
SOX21
-3.0
0
3.0Kb
Positions around peak summit
B
Relative Fold Enrichment
80
60
40
20
0
Input
ChIP
SOX21
OTX2
NEUROD1
SOX2
GBX2
DNMT3B
AXIN2
WLS
SP5
FGFBP3
FGFR2
TGFBR2
ACVR1
SMAD2
SMAD4
C
Positively Regulated Targets
abnormal telencephalon development
absent hippocampus
abnormal forebrain development
abnormal dentate gyrus morphology
absent corpus callosum
absent dentate gyrus
0 1 2 3 4 5 6 7
-log10BinomFdrQ
D
Negatively Regulated Targets
abnormal brain white matter morphology
abnormal neural tube morphology/development
abnormal telencephalon morphology
exencephaly
abnormal cerebrum morphology
abnormal cerebral cortex morphology
0 3 6 9 12 15 18
-log10BinomFdrQ

Figure S4

A

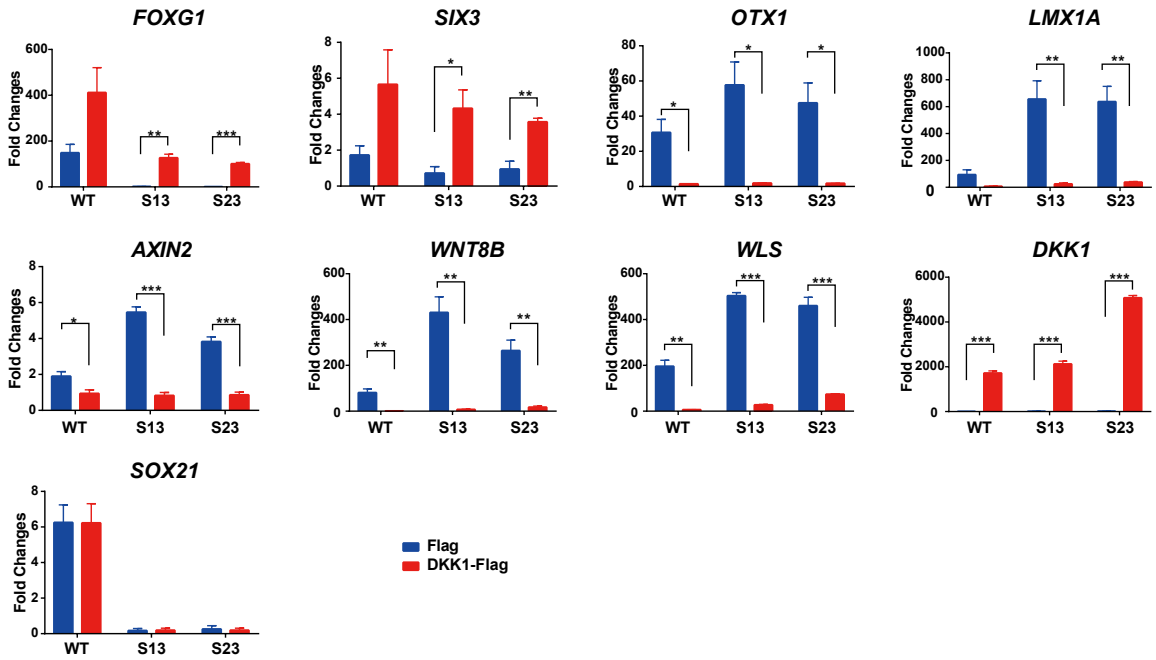

FOXG1
SIX3
OTX1
LMX1A
AXIN2
WNT8B
WLS
DKK1
SOX21
Fold Changes
WT
S13
S23
Flag
DKK1-Flag

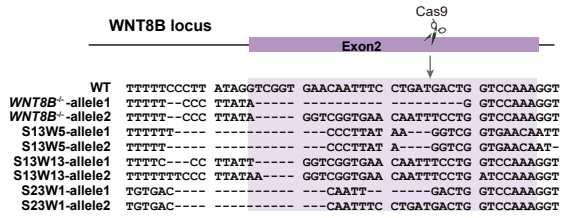
Figure S5

A

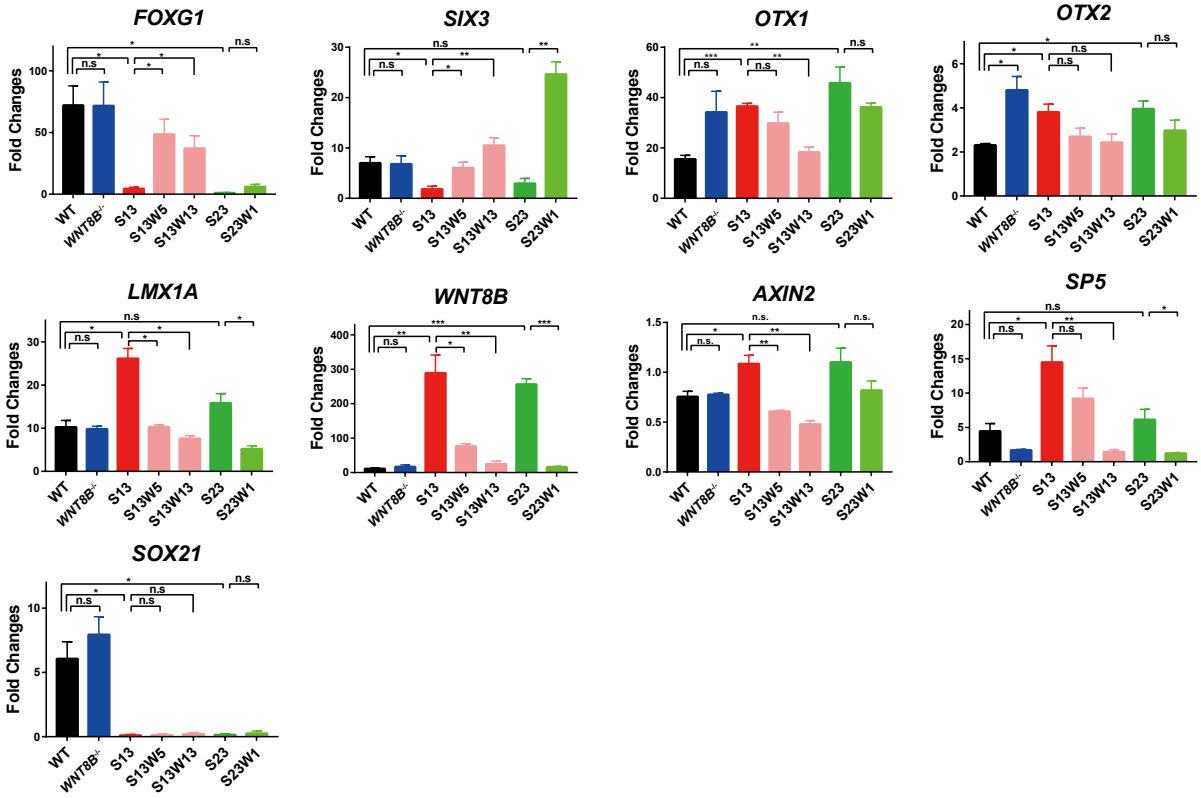
Top 10 Upregulated Putative Target Genes at Day 8

Gene Name	log2FoldChange	FDR
WNT8B	8.35	6.88E-60
HES3	5.68	3.05E-19
FOXB1	5.00	9.96E-21
WLS	4.87	2.46E-61
MTUS1	4.24	4.41E-08
NR5A1	4.09	6.64E-04
RSPO2	4.01	1.64E-07
DMRT3	3.89	1.02E-03
FGFR2	3.79	1.33E-81
TMEM125	3.50	2.21E-01

B



C



Supplemental Information

Supplemental Figure legends

Figure S1. Related to Figure 1. Region- specific transcriptome analysis of hESC-derived neural progenitor cells (NPCs).

(A) Unsupervised hierarchical clustering of 18 samples at day 4.

(B) Principal component analysis of 18 samples at day 4.

(C) The gene dendrogram and co-expression modules were identified by the weighted gene co-expression network analysis (WGCNA). The first color band underneath the dendrogram is modules determined by dynamic tree cut. Remaining color bands reveal highly correlated genes for the particular sample.

(D) The heatmap shows the relative gene expression in 7 representative region- specific modules and 1 module of undifferentiated WT hESCs. Representative genes belonging to each module are shown on the right side.

(E) Top 5 gene ontology terms for each module are shown.

See also Table S1.

Figure S2. Related to Figure 2. SOX21 participates in forebrain regionalization

(A) The representative result from flow cytometric analysis of $NANOG^{-}SOX2^{+}$ cells in WT and *SOX21* KO cells at day 7 of neural differentiation. $NANOG^{+}SOX2^{+}$ undifferentiated WT hESCs were used as a control.

(B) The representative result from cell proliferation analysis by EdU labeling assays in WT and *SOX21* KO cells at day 7. The percentage of EdU positive cells is indicated by the horizontal bar and summarized in the table at the bottom.

(C) The correlation analysis of RNA-seq data between two *SOX21* KO clones at day 4 and 8, respectively.

(D) A heatmap of 730 differentially expressed genes (DEGs) between WT (CT0.0-A and CT0.0-B) and *SOX21* KO cells (S13 and S23) (fold changes > 2, FDR < 0.05) at day 8 (CT = 0.0 μM). Some known markers involved in neural development are listed on the left. Each row of the heatmap represents the Z-score transformed and normalized read count values of DEGs across all samples.

(E) Representative result of immunofluorescence staining using antibodies against SIX3 and SOX21 at day 7. Scale bars: 50 μm.

(F) Results from real-time qPCR analysis of DEGs in another *SOX21* KO hESC line (SHhES8). Data are represented as fold changes relative to undifferentiated WT hESCs (day 0) and shown as mean ± SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (n=3).

Figure S3. Related to Figure 4. Genome-wide analysis of SOX21 binding peaks.

(A) The average signal profile of SOX21 relative to the summit of SOX21 ChIP-seq peaks.

(B) Results of ChIP-qPCR assays to validate the recruitment of SOX21 to the selected loci. Data are shown as mean ± SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (n=3).

(C-D) Disease ontology analysis of positively and negatively regulated targets by SOX21, respectively. BinomFdrQ: the Binomial FDR q-value.

Figure S4. Relative to Figure 5. Restoration of the rostral forebrain fate in SOX21 knockout cells by DKK1 overexpression

(A) The gene expression of regional markers and Wnt signaling genes was quantified by real-time qPCR at day 5 of WT and *SOX21* KO NPCs with *DKK1* or Flag control overexpression. Data are represented as fold changes relative to undifferentiated WT hESCs (day 0) and shown as mean ± SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (n=3).

Figure S5. Related to Figure 6. WNT8B knockout largely rescues defects caused by SOX21 deletion

(A) The top 10 upregulated targets in *SOX21* KO cells at day 8. SOX21 targets were identified by integrated analysis of ChIP-seq and RNA-seq data and ranked by the log2FoldChange.

(B) The strategy to knockout *WNT8B* using the wild type CRISPR/Cas9. A guide RNA was targeted to the *WNT8B* exon 2, which is showed in the purple. Sanger sequencing for 2 mutated alleles (with suffix -1 and -2, respectively) of *WNT8B* gene in WT and *SOX21* KO hESCs, as well as for a normal allele in WT hESCs. WT, wild-type; double KO clones: S13W5 (*SOX21*$^{-/-}$#13 *WNT8B*$^{-/-}$#5), S13W13 (*SOX21*$^{-/-}$#13 *WNT8B*$^{-/-}$#13), S23W1 (*SOX21*$^{-/-}$#23 *WNT8B*$^{-/-}$#1).

(C) Results from real-time qPCR analysis of regional markers and Wnt signaling genes in WT, *SOX21*$^{-/-}$, *WNT8B*$^{-/-}$, *SOX21*$^{-/-}$*WNT8B*$^{-/-}$ NPCs at day 5. Data are represented as fold changes relative to undifferentiated WT hESCs (day 0) and shown as mean ± SEM. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (n=3); n.s, not significant.

Table S1. Related to Figure S1. Modules classified by WGCNA

Table S2. Related to Figure 4. The list of SOX21 positively regulated putative target genes

Table S3. Related to Figure 4. The list of SOX21 negatively regulated putative target genes

Table S4. Related to Figure 6. The list of potential SOX21 interacting proteins identified by mass spectrometric analysis

Table S5. The list of real-time qPCR primers used in this study

Human genes		
Gene	**Forward 5' to 3'**	**Reverse 5' to 3'**
ACTB	TGAAGTGTGACGTGGACATC	GGAGGAGCAATGATCTTGAT
AXIN2	ACTTCTGGTTTGCCTGCAATGGA	GTGGCAGGCTTCAGCTGCTT
FOXG1	TGCAATGTGGGGAGAATACA	CAGGTTTGAATGAAATGGCA
LHX2	GAAGGGGCGGCCGAGGAAAC	GCTGGTCACGGTCCAGGTGC
LMX1A	GCTCAGATCCCTTCCGACAG	GCCAGTGACCCCTCAAAGAA
OTX1	CACTAACTGGCGTGTTTCTGC	GGCGTGGAGCAAAATCG
DKK1	ATAGCACCTTGGATGGGTATTCC	CTGATGACCGGAGACAAACAG
OTX2	ACAAGTGGCCAATTCACTCC	GAGGTGGACAAGGGATCTGA
IRX3	GCGGAACAGATCGCTGTAGT	GAGAGCCGATAAGACCAGGG
SIX3	ACCGGCCTCACTCCCACACA	CGCTCGGTCCAATGGCCTGG
SOX21	CCGAGTGGAAACTGCTCACA	TTCTTGAGGAGCGTCTTGGG
SP5	CTTCGGGTGTCCATGCCTC	GTGCGGTCCTGGAGAAAGG
WLS	TGCCATGAAGACCTTCCTTACG	CCAGTCAAACCCGATGGAAAAC
WNT8B	TTGTCGATGCCCTGGAAACA	TTGAGTGCTGCGTGGTACTT

Table S6. The list of antibodies used in this study

Antibody	Company
SOX21	R&D, AF3538
GAPDH	Bioworld, 20301707-1
FOXG1	Abcam, ab18259
OTX1/2	Abcam, ab25985
SIX3	Abcam, ab221750
LMX1A	Abcam, ab139726
p-LRP6	CST, #2568P
LRP6	CST, #3395P
DVL2	CST, #3216
DVL3	CST, #3218
TCF4	CST, #2569
Active-β-catenin	CST, #8814
β-catenin	BD Biosciences, 610154
α-Tubulin	Sigma, T9026
GBX2	Sigma, HPA067809
WNT8B	Lifespan, LS-C117181-50
SOX2-Alexa647	BD Biosciences, 51-9006407
NANOG-PE	BD Biosciences, 560483
Isotype-Alexa647	BD Biosciences, 557783
Isotype-PE	BD Biosciences, 555749

Table S7. The list of ChIP-qPCR primers used in this study

Gene	Forward 5' to 3'	Reverse 5' to 3'
NegCtrl	GGGGGATCAGATGACAGTAAA	AATGCCAGCATGGGAAATA
ACVR1	TTTGAACGCTGCTTGCATGG	AGGCTCTTGGTCACATCTGC
AXIN2	TGACCAAGCAGACGACGAAG	TTGCGTTTGGGCAAGGTACT
DNMT3B	TTGACTTGGTGATTGGCGGA	CCAGGAACCGTGAGATGTCC
FGFBP3	AATAAGGCTTTGGCGCCTCT	CAGCTTCGGAGGAGTCATGC
FGFR2	CCTGCGGAGACAGGTAACAG	GGTGTCTGCCGTTGAAGAGA
GBX2	GCGGTCAGGCTTAATAGGATCA	CCGCCTTCGTCCTAAAGGG
NEUROD1	CGCTTTGCAAGGGCTTATCC	AGGCGACTGGTAGGAGTAGG
OTX2	TGCATTCCTATCCCTACATTTGC	CTGCAAATGGCCCCAATCAA
SMAD2	AGCCAATGGCAAGTGAAGGA	CACCAAGGATGCAGCCACTA
SMAD4	GTGACACCACCCTCCTAAGTG	AGAGCTCTGAATGATCCAGCC
SOX2	CGAGGCTTTGTTTGACTCCG	ATAGGTAGGCGCTCAATGCG
SOX21	TTCTTGGCCGGTAAACCTATTCA	TGACAAAACGGTGAAAGGGGAA
SP5	GGCCCCCTTTGATCAGGAAA	AGTTTGCCGCTACCCAATCA
TGFBR2	GCTCTGGTGCTCTGGGAAAT	CCAGCACTCAGTCAACGTCT
WLS	ACTCCTGGGGAATAGGGACCA	TCCACAAAAGATTATGGGGCACCT
WNT8B	CCGGCCTTTCTTCCCTTCAA	CGGGCTAGATGTGTGTGTGT

Supplemental Experimental Procedures

Culture and differentiation of human ESCs

Human ESCs of H9 (Karyotype, XX) and SHhES8 (Karyotype, XX) lines were cultured on Matrigel (BD, #354277) coated dishes with a daily change of the mTeSR1 medium (STEMCELL). For neural differentiation, human ESC colonies were detached with dispase (1 mg/ml, Gibco) and suspended for aggregate formation in the DMEM/F12:Neurobasal (1:1, Gibco), N2 supplement (1:100, Gibco), and B27 supplement without vitamin A (1:50, Gibco). Rock inhibitor (Y-27632, 10 μM, Selleck) was present from day 0 to day 2. On day 4, aggregates were plated onto 6-well plates coated with Matrigel, and grew in the same medium. From day 0 to day 8, SB431542 (5 μM, Stemgent) and LDN193189 (50 nM, Stemgent) were present in the medium, with or without IWP-2 (2 μM, Millipore), CHIR99021 (0.0 - 4.0 μM, STEMCELL), and RA (1μM, Sigma, #R2625) for region- specific neural differentiation.

Targeting plasmid construction

For donor-based *SOX21* knockout (KO), the single exon of *SOX21* was replaced with a PGK-neo cassette. A 0.8 kb short homology arm upstream the translation initiation site and a 1.5 kb long homology arm downstream the stop codon of *SOX21* were generated by PCR from H9 ESC genomic DNA using primers that had restriction sites on the ends of each amplicon and cloned into the PGK-NEO-DTA targeting vector. SalI/Hind III were used for the short arm and XbaI/NotI for the long arm: SOX21-short_fw 5'-acgcgtcgacagccgggagaacttcctct-3', SOX21-short_rev 5'-cccaagcttggcccggaggaaatcaatgt-3'; SOX21-long_fw 5'-gctctagaataggtgccaggtagaggca-3', SOX21-long_rev 5'-

ataagaatgcggccgcaggcacgtaagggcaattca-3'. Two paired gRNAs flanking the coding sequencing of *SOX21* were designed using the online tool from Zhang-lab (http://tools.genome-engineering.org). The gRNAs were cloned into px335 (Addgene plasmid, #48873) using the following gRNA pairs: for Cas9n-1, using g1A 5'-agccggtggaccacgtcaag-3', and g1B 5'-gacatgctctcgccctgccg-3'; for Cas9n-2, using g2A 5'-atgtataggtacgagcgctg-3', and g2B 5'-gcacaccggtcctcgcgagg-3'. For donor-free gene targeting, we designed 2 different sets of gRNAs, which were cloned into px459 (Addgene plasmid, #48139). The sequence of gRNAs for human *SOX21*: g19 5'-cttgacgtggtccaccggct-3', g69 5'- gcgggctcagcggcgcaaga-3', g14 5' -gctgccgcgctatgaccccg-3', gb 5'- atgtataggtacgagcgctg-3'. For *SOX21* knockdown experiments, 2 oligonucleotide sequences (3# 5'-tcgcaattttatcgaagatta-3'; 5# 5'-tccctgtttgtactatttgaa-3') targeting *SOX21* were designed to construct the *SOX21* shRNA inserts, respectively, according to a published protocol (Chang et al., 2013), and the AAVS1-TRE3G-EGFP plasmid (Qian et al., 2014) was used as a vector to generate *SOX21* shRNA plasmids. We also designed a nontarget (NT) control oligonucleotide sequence (5'-aggaattataatgcttatcta-3') to make the NT control shRNA plasmid. The sequence of gRNA for *WNT8B* KO is 5'-ttacctttggaccagtcatc-3'.

Generation of genetically modified hESC lines

Human ESCs were treated with Rho kinase inhibitor (Y27632, 10 μM) for 1 hr before electroporation. Cells were digested into single cells by Accutase (Innovative Cell Techonologies) for 7 min, and 5 x 10^6 cells were electroporated with plasmids in 200 μl of the electroporation buffer (5 mM KCl, 5 mM $MgCl_2$, 15 mM HEPES, 102.94 mM Na_2HPO_4,

47.06 mM NaH_2PO_4, pH 7.2) (Chen et al., 2015), using the Gene Pulser Xcell System (Bio-Rad) at 250 V, 500 μF in 0.4 cm cuvettes (Bio-Rad). For *SOX21* KO experiments, 3.125 μg of each gRNA plasmid and 12.5 μg donor plasmid was used for electroporation. For *WNT8B* KO experiments, 25 μg of the gRNA plasmid was used for electroporation. For *DKK1* overexpression experiments, the *DKK1* cDNA with the Flag sequence at its C-terminus was cloned into the AAVS1-TRE3G-EGFP (Qian et al., 2014) to replace the EGFP sequence and generate the AAVS1-TRE3G-DKK1-Flag plasmid. Fifteen μg of the AAVS1-TRE3G-DKK1-Flag plasmids or the *SOX21* shRNA plasmids, plus 6.25 μg of AAVS1 gRNA plasmids for nickase (gA 5'-gtccctagtggccccactgt-3', gB 5'-gacagaaaagccccatcctt-3'), were used for electroporation. Cells were subsequently plated onto Matrigel coated 6-well plates in the mTeSR1 containing 10 μM Y27632. Twenty-four hrs later, the medium was changed to the mTeSR1 without the Rho kinase inhibitor. Three days after electroporation, puromycin (0.2 μg/ml) or G418 (125 μg/ml) was added into the culture medium for 5 - 10 days until colonies appeared. Individual colony were then picked up and transferred into 12-well plates in the mTeSR1 medium. Positive colonies were identified by genomic PCR analysis and replated onto 6-well plates. Primers for *SOX21* KO identification are S21ko-fw 5'-ctcggccggagacactaagg-3' and S21ko-rev 5'-gtgggtcaaaacgcaacagg-3'.

RNA extraction, cDNA synthesis and quantitative real time PCR

Total RNA was extracted using TRIzol reagent (Life Technologies, #15596026) and 2 μg total RNA were reversely transcribed into cDNA using a ReverTra Ace reverse transcriptase (Toyobo, #FSK-101), according to the manuals. Real-time qPCR was performed on the ABI ViiA7 Real-Time PCR system, using the SYBR Premix Ex Taq II

(Takara, #RR820L) according to manufacture instructions. *ACTB* was used as an internal control, a list of primers is provided in Table S5.

Western blot, coimmunoprecipitation and mass spectrometric analysis

Cells were washed with PBS twice and lysed with the Co-IP buffer (50 mM Tris-HCl ,150 mM NaCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 20% Glycerol, 0.1% NP-40, 3 mM β-Mercaptoethanol; pH 7.9) combined with a protease inhibitor cocktail (Selleck, B14001). Total protein concentration was measured using the Pierce BCA Protein Assay Kit (Thermo Scientific, #23227), according to the manufacture instructions. Twenty μg total proteins (per lane) were separated by the SDS-PAGE and transferred to 0.45 μm nitrocellulose blotting membranes (GE Healthcare, #10600002). Membranes were blocked with 5% BSA diluted in TBST (19 mM Tris, 2.7 mM KCl, 137 mM NaCl, 0.1% Tween-20; pH 7.4) for 1 hr and incubated with specific primary antibodies overnight at 4 °C, followed by incubation with HRP-conjugated secondary antibodies for 1 hr. The signals were detected from blotted membranes by exposing to the SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, #34580). For SOX21 coimmunoprecipitation experiments, NPCs at differentiation day 7 were lysed in the Co-IP Buffer with 1 x protease inhibitor cocktail. Then, 10 mg NPC lysates were incubated with 10 μg of SOX21 antibodies or goat IgG (Sigma) at 4 °C overnight. Then, 100 μl protein G Magnetic Beads (Millipore, #LSKMAGG10) were used to pull down SOX21 antibody- or IgG- bound proteins for 2 hrs at 4 °C. Next, protein complexes were washed twice with the Co-IP buffer, once with the high salt buffer (50 mM Tris-HCl, 250 mM NaCl, 5 mM $MgCl_2$, 0.2 mM EDTA, 20% Glycerol, 0.1% NP-40, 3 mM β-Mercaptoethanol; pH 7.9), once with the Co-IP buffer, and finally eluted in the 2x SDS

loading buffer. The eluted mixture was separated by precast-GLgel (4-15%, Sangon Biotech, C621104-001). Proteins in the gel were stained with a fast silver stain kit (Beyotime, P0017S), and sequenced by mass spectrometry (Jiyun Biotech, Shanghai). All WB analyses were conducted in at least three independent experiments, unless otherwise indicated.

Immunofluorescence staining

For neural differentiation, aggregates were attached to Matrigel-coated coverslips (Fisherbrand, #12-545-82). On neural differentiation day 7, cells were fixed with 4% paraformaldehyde for 15 min at room temperature, washed 3 times with PBS, and incubated in the blocking buffer (10% donkey serum and 0.2% Triton X-100 in PBS) for 30 min at room temperature followed by incubation with primary antibodies overnight at 4 ^{o}C. Next day, coverslip cultures were washed 3 times with PBS, and incubated with secondary antibodies (FITC, 1:100; Cy3, 1:200) for 1 hr at room temperature. The nuclei were stained with DAPI (Sigma, #D9542). Images were captured using a Zeiss Cell Observer microscope. Antibodies used are listed in Table S6.

Luciferase reporter assays

Aggregates of NPCs formed for 4 days were dissociated with Accutase, and 3 x 10^5 cells were transfected with luciferase plasmids using Lipo2000 (Invitrogen, #11668019). For enhancer reporter assays, the *WNT8B* enhancer was cloned by genomic DNA PCR using primers having restriction sites and the reverse primer had a minimal promoter sequence (Bgl II and Hind III): WNT8B-en-F 5'-ctggcctgtgagcctaaaca-3'; WNT8B-en-R 5'-

gccaagcttctggaagtcgagcttccattatataccctctaaccctagtaaggcaggggag-3'. The wild type and mutant *WNT8B* enhancers containing a minimal promoter were inserted into pGL4.22 (Promega). Then, 500 ng of the enhancer reporter plasmid and 10 ng of the pRL-TK internal control plasmid (Promega) were used for transfection. For Wnt signaling reporter experiments, 500 ng of the 8XTOPFlash (Addgene plasmid, #12456) or 8XFOPFlash (Addgene plasmid, #12457) plasmid together with 50 ng of the pRL-TK internal control plasmid (Promega) were used for transfection. Cells were plated onto 24-well plates in the N2B27 medium supplemented with 50 nM LDN193189, 5 μM SB431542 and 10 μM Y27632. The medium was changed 24 hrs post transfection. Cells were collected at 48 hrs post transfection and luciferase activities were examined with the Dual Glow Luciferase Assay System (Promega, #E1960) according to the manufacture instructions. The luciferase intensity ratio was determined by (TOP firefly/renilla) / (FOP firefly/renilla).

Flow cytometric analysis and cell proliferation assays

Cells were fixed, permeabilized (Cytofix/Cytoperm Kit, BD Biosciences) and stained using antibodies against SOX2 and NANOG. For cell proliferation analysis, cells were incubated with 10 μM EdU (BD Biosciences) for 1 hr, and labeled using a Cell-Light EdU Apollo488 *In Vitro* Flow Cytometry Kit (RIBO) according to manufacture instructions. Stained samples were acquired on an Accuri C6 flow cytometer (BD Accuri Cytometers), and data were analyzed using a FlowJo software.

ChIP-qPCR

Cells (1×10^7) were cross-linked with 1% formaldehyde (Thermo Scientific, #28906) for 10 min and quenched by 0.125 M glycine for 5 min at room temperature. Cell cultures were

suspended in the 1% SDS FA lysis buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 1% SDS) supplemented with 1x protease inhibitors (Roche, #04693132001) and rotated for 15 min at 4 °C. Cell pellets were collected by centrifugation at 15,000 g, at 4 °C for 30 min, and resuspended in 1 ml of the 0.1% SDS FA lysis buffer (50 mM HEPES-KOH, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, and 0 .1% SDS) with 1x protease inhibitors. Cell pellets were sheared by sonication (50% amplitude, 30s ON, 30s OFF, incubated in ice water) for 20 min (UibraCELL), and followed by centrifugation at 12,000 g for 10 min at 4 °C. An aliquot of 10 μl supernatant was used as an input, and the remaining supernatant was precleared and incubated with 20 μg of specific antibodies overnight at 4 °C. On the next day, the reaction mixture was incubated with 100 μl Protein G Magnetic Beads (Millipore, #LSKMAGG10) for 2 hrs at 4 °C, and followed by gentle wash with the 0.1% SDS FA lysis buffer (3 times), the high salt buffer (50 mM HEPES-KOH, pH 7.5, 350 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate and 0 .1% SDS), the ChIP wash buffer (250 mM LiCl, 0.5% NP40, 0.5% sodium deoxycholate, 1 mM EDTA, 10 mM Tris-HCl; pH8.0), and the TE buffer (10 mM Tris-HCl, 1 mM EDTA; pH8.0) sequentially. The chromatin was eluted from beads with 200 μl ChIP elution buffer (50 mM Tris-HCl, 10 mM EDTA, 1% SDS; pH7.5) for 30 min at 37 °C. The Input and eluted chromatin were decross-linked by adding 8 μl of 5 M NaCl at 65 °C for less than 16 hrs, followed by adding 2 μl of 20 mg/ml protease K (Sigma) for 2 hrs at 56 °C. The resulting DNA was purified using a QIAquick PCR Purification Kit (QIAGEN, #28106) and quantified using a Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen, P11496). The purified DNA was used for

examination of site specific enrichments. To evaluate site specific enrichments, 0.5 ng of input or enriched DNA was used for qPCR. Results were presented as the relative fold enrichment to the input. All primers used in ChIP-qPCR assays are listed in Table S7.

RNA-seq

Total RNA was isolated from cells using TRIzol Reagent (Invitrogen). Isolated RNA was enriched by Poly(A) tails and fragmented. Sequencing libraries were prepared according to Illumina manufacturer instructions. Paired-end RNA-seq of 2 x 150 base pair reads were sequenced on the Illumina Hiseq X-ten by Annoroad Gene Technology (Beijing, PR China) Co., Ltd. The Salmon (Patro et al., 2017) was used to calculate samples' TPM (Transcript Per Million) and raw counts. DESeq2 (Love et al., 2014) were then used to identify differentially expressed genes with following settings: fold changes > 2 and FDR < 0.05.

ChIP-seq

ChIP assays were performed as described in the ChIP-qPCR section, and ChIP-seq DNA libraries were constructed using the NEB Next Ultra DNA Library Prep Kit for Illumina, according to the manufacture instructions. Purified libraries were sequencing on Illumina Hiseq X-ten platforms by Annoroad Gene Technology (Beijing, PR China) Co., Ltd.

ChIP-seq data for TCF4 (GSE61475) were downloaded from the Gene Expression Omnibus (GEO), including peak files and bigwig files for visualization. For SOX21 ChIP-seq data, sequencing reads were aligned to the human genome (hg19) and uniquely mapped reads were kept for further analysis. Peaks were called using MACS2 with q < 0.01, and annotated by the GREAT(McLean et al., 2010). ChIP-seq signal profiles at the specific locus were visualized with the Integrative Genomics Viewer (IGV).

Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed by following the tutorial written by Langfelder and Horvath (Langfelder and Horvath, 2008). RNA-seq data sets of positionally patterned NPCs were used to construct a signed weighted correlation network using the WGCNA. First, we created a signed Pearson correlation matrix between 19861 genes (genes with the maximum raw read count < 30 across all samples were filtered out). In order to construct adjacency matrix, we adjusted the soft power β to 18, which is a soft-threshold of the correlation matrix. We then calculated the topological overlap matrix and performed hierarchical clustering to group genes with highly similar co-expression relationships. Modules were defined by the Dynamic Hybrid Tree Cut algorithm from the hierarchical clustering tree. The expression profile of each module can be summarized by module eigengene, which represented the first principal component of module expression profile.

Supplemental Reference

Chang, K., Marran, K., Valentine, A., and Hannon, G.J. (2013). Creating an miR30-based shRNA vector. Cold Spring Harbor protocols *2013*, 631-635.

Chen, Y., Cao, J., Xiong, M., Petersen, A.J., Dong, Y., Tao, Y., Huang, C.T., Du, Z., and Zhang, S.C. (2015). Engineering Human Stem Cell Lines with Inducible Gene Knockout using CRISPR/Cas9. Cell Stem Cell *17*, 233-244.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics *9*, 559.

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

McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejerano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. Nature biotechnology *28*, 495-501.

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.

Qian, K., Huang, C.L., Chen, H., Blackbourn, L.W.t., Chen, Y., Cao, J., Yao, L., Sauvey, C., Du, Z., and Zhang, S.C. (2014). A simple and efficient system for regulating gene expression in human pluripotent stem cells and derivatives. Stem Cells *32*, 1230-1238.