Supporting Information

Material and Methods

Reprogramming MEFs into iPS cells

iPS cell reprogramming protocol was followed as described (1). Briefly, lentiviral particles (OSKM and M2rtTA from Addgene) were packaged in 293T cells. After incubation of 293T cells, infected medium containing virus particles was filtered (0.45 μm) and then transferred to plates growing either WT or KO MEFs (less than 3 passages). The following day, medium was changed into ES cell medium as described above. Doxycycline (Stemgent) was added at 2 μg/ml. The above medium including doxycycline was replaced at least every other day for approximately 20 or more days. ES cell clones were hand–picked, then propagated and tested for ES cell characteristics.

CG methylation analysis

For whole genome bisulfite sequencing of *Lsh-/-* and WT primary murine embryonic fibroblasts the protocol was applied as previously described (2). Single-read MethylC-Seq sequences were processed and aligned to MM9, and an additional filter was added to remove any mapped reads in which a read-C base was aligned to a reference-T base. The sequence data was filtered for methyl-cytosines with ≥ 5 sequence reads. The methylation ratio presents the fraction of methyl-cytosine to the number of sequence reads. For Bisulfite sequencing at single loci genomic DNA (0.5μg) was bisulfite-converted, PCR amplified and fragments subcloned. 10 clones for each sample were sequenced to identify the methylation status of cytosines. The

methylation result is expressed as CG methylation ratio (fraction of methylated cytosine over total cytosine) or as percentage of methylated cytosine.

Immunofluorescence staining

Cells in chamber slides were fixed in 4% paraformaldehyde for 10 min at 25 °C, washed twice with PBS, then permeabilized for 15 min with 0.2% TritonX-100 in PBS, and blocked for 20 min with 5% goat serum in PBS. After incubation with primary antibodies against Tubb3 (Abcam, ab18207); Nestin (Abcam, ab11306); NeuN (Millipore,Mab377); Olig2(Abcam,ab42453); Map2 (Sigma,11M4848); SSEA1 (Stemgent, 09-0085); Oct4 (Stemgent, 09-0023); Nanog (Stemgent, 09-0020); and Sox2 (Millipore, MAB4343) for 2 hrs in 1% BSA in PBS, cells were washed three times with PBS and incubated with fluorophore-labeled appropriate secondary antibodies purchased from Millipore. The images were acquired with a Zeiss confocal microscope.

Enhancer reporter assay

Predicted enhancer sequences were selected for validation in reporter assays. The chromosome location and primers are listed in Supplementary Table S1. Cloning and reporter assays were carried out as previously reported (3), Fragments were subcloned into the pGL3 promoter plasmids (Promega E1761) with BamHI and SalI. The constructs were transfected into Lsh-/- MEFs, HeLa S3 and Neuro-2a cells (ATCC) with lipofectamine 2000 (Invitrogen) and GenJet™ DNA *In Vitro* Tranfection Reagent (SignaGen Laboratories) separately. Fragments were designated as active if their relative luciferase value was significantly higher than control in HeLa S3 or Neuro-2A cells (p-value < 0.05).

Teratoma formation assay

Three sets of iPS clones were resuspended in PBS, kept on ice and drawn into 1-mL syringe immediately before injection. Approximately 1×10^6 cells in 100 µl/injection site were used. Nude mice were injected in the dorso-lateral area into the subcutaneous space on both sides. After 4 weeks, teratoma were collected for size measurement, immunostaining, FACS, and gene expression analysis.

Flow cytometry analysis

Teratoma were sliced into small pieces, the undesirable tissue such as fat or necrotic material was removed, and incubated with 1mg/ml collagenase D at 37^{o} C for 20 min. The cell suspension was passed through a nylon mesh to remove undigested fragments. After centrifugation of the supernatants at 100g for 3 min, the pellet was resuspended in cell culture medium. Dissociated cells were fixed in 1% paraformaldehyde for 10 min, and the standard immunostaining protocol was performed. For surface marker staining, 10^5-10^6 cells were resuspended in 100 µl blocking buffer for 30 min at 4^{o} C. Directly conjugated primary antibodies (APC-CXCR4 or APC-FLK1, Biolegend) were added to the cells at a final dilution of 1:200 and incubated for 30 min at 4^{o} C. Stained cells were washed twice in PBS-FBS buffer. Isotype-matched mouse or rabbit IgG1 (Millipore) were used as negative controls to set the threshold for non-specific binding. FACS Calibur (BD Biosciences) instruments were used for analysis (Flow cytometry core facility of NCI/NIH).

ChIP-Seq analysis

Biologic replica were evaluated for correlation as previously described (4). In brief, reads were sorted in bins of 1kb. Samples were normalized based on total aligned reads, and input signals were subtracted from sample signals. For H3K4me1 ChIP-Seq analysis the correlation of wt1 and wt2 was 0.998 and ko1 to ko2 was 0.943 for 2,464,195 windows. Reads from biological replicates for H3K4me1 (two WT and two KO MEF samples derived from litter mates) were merged. Peak detection (genomic bins that contained statistically significant ChIP-Seq enrichment) was performed using Partek Genomics Suite at an FDR<0.05 for autosomes only. Each read was extended toward the interior of the sequenced fragment (about 135-150bp), based on the strand of the alignment. H3K4me1peaks were determined at a window of 150bp and selected at an FDR<0.0001 over their respective input. The following criteria was applied to term H3K4me1 modifications as predictive enhancers (4): H3K4me1 peaks were selected that had no overlap with any H3K4me3 peaks (either in WT or KO samples) and that were at least 2kb distant of transcriptional start sites (2kb upstream or downstream of TSS). Differentially enriched H3K4me1 peaks in KO MEFs over WT MEFs (or vice versa) was determined using an FDR <0.00001. In addition, raw sequence reads are presented in Fig1A (lower panel), supplement Fig S1B, and FigS4 to give highest resolution. The total read number in KO samples is as follows: KO#1 Input 21,874,452; KO#1 H3K4me1 21,579,939, KO#2 H3K4me1 22,601,971; WT#1 Input 23,387,459; WT#1 H3K4me1 28,134,577; WT#2 H3K4me1 30,242,670. Motif analysis was performed in Partek using the Jaspar Core Transcription Factor Database with a threshold p value<1e-250 with the exception of MafB p value<1.5e-58 (Dataset 3). KO-H3K4me1 peaks were defined as overlapping with cis-regulatory enhancer sequences if peaks

were less than 1 kb distant from the center base pair of the enhancer (4). Genes are considered as 'marked' with KO-H3K4me1 (WT-H3K4me1) if a KO-H3K4me1 (WT-H3K4me1) peak was localized within the 5' and 3' transcription unit including 5kb of flanking regions. For H3K4me1 association profiles at KO-H3K4me1 and WT-H3K4me1 sites normalized ChIP-Seq reads were tabulated in 100 bp windows including the center of the peak and 5 kb of flanking regions (Fig 2A, B). For CG methylation profiles CG methylation ratios were tabulated in 10 bp windows including the center of the peak and 5 kb of flanking regions (Fig 2C,D). For gene ontology (GO) pathway analysis we used Functional Annotation Pathway analysis in David (Dataset 1). The p-value represents the EASE score, a modified Fisher exact p-value as described (http://david.abcc.ncifcrf.gov/home.jsp). For GO term analysis of the Biologic Process Partek software was used and biologic process terms with respect to development selected. The raw ChIPs-seq data set of H3K4me1 and H3K27ac ChIPs derived from adult murine cortex, adult murine cerebellum and fetal murine brain used in supplemental Fig S4 is downloaded from Geo omnibus (GSM722632, GSM722664, GSM851281, GSM851271, GSM851272, GSM851284).

ChIP-qPCR and gene expression analysis

ChIP (chromatin immunoprecipitation) was performed under a combination of Young (5) and McMahon lab's protocol (Harvard University) with slight modifications. The following Chip-grade antibodies were used: H3K4me1 (Abcam, ab8895), H3K27ac (Abcam, ab4729), YY1 (Abcam, ab12132), Ets1 (Santa Cruz, sc-350), MafB (Santa Cruz, sc-22830 X), Zic1 (Santa Cruz, sc-28149X), Setd7 (Abcam, ab14820), Lsd1 (Santa Cruz, sc-271720X), MLL5 (Santa Cruz, sc-

292360 X) and normal rabbit/Mouse IgG1. Precipitated DNA was re-suspended in 60 µl of Nuclease-Free water (Invitrogen) and analyzed by qPCR using the specific primers shown in Supplementary Table S1 (primers are listed for each gene in sequence according to Figures 1B). The normalization method for ChIP analysis is percent of input. In addition, the relative values were normalized to one site of each gene in WT sample. Every ChIPs results represents three independent experiments (mean and s.d. of n=3). For p-value computation the student t-test was applied. Anti-rabbit IgG served as negative control. All RNA samples were extracted by RNeasy Mini Kit (Qiagen) from 1 to $5x10^6$ cells, according to the manufacturer's instructions, and treated with DNase I (Roche). The cDNA was prepared with Superscript III First Strand cDNA Synthesis Kit (Invitrogen) from total RNA using oligo dT primer, followed by qPCR with the specific primers shown in Supplementary Table S1. All quantifications were normalized to an endogenous Gapdh and Actb control. The relative quantification value for each target gene compared to the calibrator for that target is expressed as $2^{-(Ct-Cc)}$ (Ct and Cc are the mean threshold cycle differences after normalizing to Gapdh). For all qPCR, iQ SYBR Green Supermix was used in iCycler instrument (Bio-rad) according to the manufacturer's instructions.

Mesoderm and Endoderm differentiation

The lineage differentiation protocol was performed as previously reported (6, 7). Before differentiation, 1 million cells per well of mouse iPS cells were plated onto wells coated with growth factor-reduced Matrigel (R&D systems) in MEF-conditioned medium with 20ng/ml bFGF. The medium was changed with RPMI-B27 medium supplemented with induction reagent activin

A (100ng/ml, R&D systems) 1 day and then Bmp4 (10ng/ml, R&D systems) 3 days for mesoderm differentiation, and medium was changed with MCDB-131 medium supplemented with 1% GlutaMax, glucose (5.5mM),0.1% FAF-BSA (Invitrogen) , ITS:X (Invitrogen) , 0.25% $NaHCO_3$ and Wnt3a(20ng/ml, R&D systems)+ activin A (100ng/ml) for 1 day, then replaced with activin A (100ng/ml) for 3 days for endoderm differentiation. Cells were collected for further analysis of mesodermal and endodermal markers using real-time PCR analysis and FACS analysis. Mouse Mesenchymal Stem Cell Functional Identification Kit (Catalog # SC010) is used for chondrocyte and Osteoblast lineage differentiation.

Reference

1. **Lyssiotis CA*, et al.* (2009) Reprogramming of murine fibroblasts to induced pluripotent stem cells with chemical complementation of Klf4. (Translated from eng) *Proc Natl Acad Sci U S A* 106(22):8912-8917 (in eng).**
2. **Lister R*, et al.* (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. (Translated from eng) *Nature* 462(7271):315-322 (in eng).**
3. **Heintzman ND*, et al.* (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. (Translated from English) *Nat Genet* 39(3):311-318 (in English).**
4. **Shen Y*, et al.* (2012) A map of the cis-regulatory sequences in the mouse genome. (Translated from eng) *Nature* 488(7409):116-120 (in eng).**
5. **Odom DT*, et al.* (2004) Control of pancreas and liver gene expression by HNF transcription factors. (Translated from eng) *Science* 303(5662):1378-1381 (in eng).**
6. **Chetty S*, et al.* (2013) A simple tool to improve pluripotent stem cell differentiation. *Nature methods* 10(6):553-556.**
7. **Kattman SJ*, et al.* (2011) Stage-specific optimization of activin/nodal and BMP signaling promotes cardiac differentiation of mouse and human pluripotent stem cell lines. *Cell stem cell* 8(2):228-240.**

Supporting Figure Legends

Fig.S1 CG methylation profiles at KO-H3K4me1 peaks. (*A*) Conventional bisulfite sequencing results representing CG methylation (percentage of methylated cytosine) at sequences located within KO-H3K4me1 peaks comparing KO MEFs to WT MEFs. The black circle represents a methylated cytosine, the open circle represents an unmethylated cytosine. CG methylation is represented as percentage. (*B*) UCSC genome browser representation of raw sequence data showing input (KO Input or WT Input), and H3K4me1 enrichment in KO MEFs (KO H3K4me1) compared to WT MEFs (WT H3K4me1) at selected neuronal genes. In the lower panel CG methylation ratios for WT and KO MEFs are presented with each orange bar representing the ratio of methylated cytosine at a single C_pG site. The red line below the graph represents the area examined by conventional bisulfite sequencing (*A*) and in Fig.S2

Fig.S2 Perturbed KO-H3K4me1 modification after siDnmt1 treatment. (*A*) Dnmt1 western analysis after use of siDnmt1 in WT MEFs. (*B*) Conventional bisulfite sequencing results representing CG methylation at sequences located within KO-H3K4me1 peaks in siDnmt1 treated MEFs. For comparison are the methylation results of the same sequences analyzed for KO and WT MEFs in Fig S1. The black circle represents a methylated cytosine, the open circle an unmethylated cytosine. CG methylation is represented as percentage. (*C*) Relative enrichment of H3K4me1 after Dnmt1 reduction (siDnmt1) at several KO-H3K4me1 sites compared to

control treated WT MEFs (Ctrl). TRC Dnmt1 shRNA retro-viral vector sets were used from Thermo Scientific (RMM4534-EG13433).

Fig.S3 GO term analysis for KO-H3K4me1 marked genes. Gene ontology analysis (Partek) was performed for genes that were enriched for KO-H3K4me1 peaks. The graph presents Biologic Process GO terms (P value>0.05) selected for developmental processes.

Fig.S4 Overlap of KO-H3K4me1 with enhancers present in brain tissues. UCSC genome browser representation of raw sequence data showing input (KO Input or WT Input), and H3K4me1 enrichment in KO MEFs (KO H3K4me1) compared to WT MEFs (WT H3K4me1) at selected neuronal genes. In the upper three panels the center position of cis-regulatory sites in murine cortex, adult murine cerebellum, and fetal murine brain is presented (4). In addition, the raw sequence reads of H3K4me1 and H3K27ac of cortex, cerebellum and fetal brain (3) are shown to visualize the precise overlap of reads at highest resolution.

Fig.S5 Characterization of Lsh KO iPS and WT iPS cells. (*A*) Experimental schema for reprogramming. The transcription factors Oct4 (Pou5f1), Sox2, c-myc and Klf4 were overexpressed in primary fibroblasts derived from *Lsh-/-* (KO) and WT littermates using a polycistronic vector. iPS cells were isolated for each genotype. (*B*) Microscopic image of WT iPS and KO iPS cells. (*C*) Real-time PCR analysis for detection of Nanog, Oct4, Sox2 mRNA in WT iPS

and KO iPS cells comparing three independently derived cell lines of each genotype. ES cells served as positive control and MEFs as negative controls. (*D*) Real-time PCR analysis for detection of Lsh mRNA expression in WT iPS and KO iPS cells, ES cells and MEFs. (*E*) Bisulfite sequencing to examine CG methylation at the Oct4 promoter region in WT iPS and KO iPS, ES cells and WT MEFs for comparison. (*F*) Three independently derived sets of WT iPS and KO iPS clones were examined for alkaline phosphatase expression. (*G*) Immunofluorescence analysis of iPS cells for pluripotency markers expression Oct4, Nanog, SSEA1, and Sox2.

Fig.S6 Characterization of neuronal marker gene expression in iPS cells. (*A*) Confocal immunofluorescence analysis for the indicated markers upon neuronal differentiation after RA treatment at day4. Neuronal marker Nestin and Tubb3 expression in three independently derived sets of KO and WT iPS clones. (*B*) The percentage of positive Tubb3 staining of the gated population is shown in the bar graph. (*C*) Teratoma were harvested four weeks after injection of three sets of independently derived KO iPS clones and WT iPS clones. The size was measured individually.

Fig.S7 Characterization of neuronal marker gene expression in teratoma. Teratoma were harvested after 4 weeks and teratoma with similar size were dissociated to single cells, examined for mRNA expression of neuronal progenitor cell markers (*A*), immunostained with the neuron specific marker Tubb3 and Nestin (*B*), and examined by FACS analysis (*C*). The bar

represents 20µm. mRNA levels were normalized to one WT iPS cells for each gene, standard deviation was shown from duplicate experiments for three independent sample sets.

Fig.S8 Transcription factor association at KO-H3K4me1 sites. (*A*) ChIPs analysis using Real-time PCR analysis to assess TF enrichment at neuronal lineage genes at KO-H3K4me1 sites in KO MEFs compared to WT MEFs. (*B*) Western analysis to assess TF expression in WT and KO MEFs. (mean and s.d., n=3). **, P-value<0.01 and *, P-value<0.05. (*C*) YY1 and Ets1 association at WT-H3K4me1 marked sites. ChIPs analysis using anti-YY1, Ets1 antibodies followed by Real-time PCR analysis to detect YY1 and Ets1 enrichment at different loci of WT-H3K4me1 marked sites in WT and KO MEFs. Anti-Rabbit IgG served as negative control. The percentages of each PCR product in the immunoprecipitated sample per input are shown (mean and s.d., n=2), and the differences between WT and KO were analyzed for significance by Student's t test. **, p<0.01 and *, p<0.05. (*D*) Western analysis examining protein expression level of H3K4m1 modifying enzymes in WT and KO MEFs.

Fig.S9 Mesoderm and Endoderm differentiation from different iPS clones. (*A*) KO iPS clones (KO1 and KO2) and WT iPS clones (WT1 and WT2) were differentiated into the mesoderm lineage. (*B*) mRNA expression analysis using q-PCR analysis of mesodermal marker genes before (control) and after differentiation (induced). (*C*) FACS analysis quantifying the expression of the mesoderm marker protein FLK1 before (red) and after differentiation (blue).(*D*) KO iPS clones (KO1 and KO2) and WT iPS clones (WT1 and WT2) were differentiated into the endoderm lineage. (*E*) mRNA expression analysis using q-PCR analysis of endodermal marker genes before

(control) and after differentiation (induced). (*F*) FACS analysis quantifying the expression of the endodermal marker protein CXCR4 before (red) and after differentiation (blue).

Fig.S10 Chondrogenic and osteogenic lineage differentiation for iPS clones. Two KO iPS clones and Two WT iPS clones were differentiated into the chondrogenic lineage and stained with Alcian blue (*A*) or into the osteogenic lineage and stained with Alizarin red (*B*). The wild type ES cell line CCE served as control. No significant differences based on staining properties were detected comparing WT and KO iPS cells.

Table S1 Mouse genome (MM9) is used for all analysis.

H3K4me1 and H3K27ac ChIP Primer Sequences

Name	Sequence(5'to 3')	H3K4me1 peaks position
Grik2	F: CCAAAGCAAAATCACCAACC R: TCTGACCTTCTCAGGTACCAGAC	chr10:48,911,935-48,912,873
	F: GCTCTGTTCCAAGACGACATC R: CCTTTCTTGTTGAAGCCTCTG	chr10:49,436,870-49,437,004 *[8]
	F: CCTTTTGTCTCCACTAGAATAGG R:TCAGCTACTCAAGGCAGGTTTG	chr10:49,188,952-49,189,609
Mib1	F: GGAGCGATAGAAGCATCAAAGC R: TTAGTGGTAGACAGCTGGCTGG	chr18:10,756,308-10,757,263
	F: TCTTTGTAGCTCTGGCTGTCC R: TAGCACGGCACCTAATGACC	chr18:10,734,896-10,735,973 *[9]
Nrcam	F: ACCTGATGGTTCTGTGTTTGATG R: CCAATGGAAGGGACTAGGCTA	chr12:45,471,308-45,473,346 *[10]
	F: CCTAGAGAATGAGGATTGTACCA R: GCCTTTTTGCTTCCTTCTCTG	chr12:45,557,942-45,559,905 *[11]
Negr1	F: AATACCCACTGCAGCCTGACT R: CAACCTTCTCCACCTCCCTC	chr3:156,564,826-156,567,781 *[12]
	F: CTTTTCTACAGCATCAATGGCAC R: TGGCATGTGGTACTTGTCCAC	chr3:156,269,101-156,271,059 *[13]
Nrxn1	F: CAGCCTGATTACCACCAGTTTC R: TGTCAGTTTTTGACACTGGATGG	chr17:91,235,787-91,236,788
	F: TGACATGGTTGCACTGCTGAG R: GGCAGTGAAAACTCGAATGTGA	chr17:91,156,795-91,158,285 *[4]
	F: ACCATTTCTTGGGCTTTGG R: GCTCTGATATTGGGCCTACTTC	chr17:91,380,477-91,380,989 *[16]
Nlgn1	F: GCTTTTTTTCTGCCTGTTGTG R:GAGCCAGCTTCAGAATGTAGTG	chr3:25,704,140-25,704,698 *[5]
	F: GGCATGTGTAAGCAAGAGTC R: GATATGGCTCTCAAAGCTGG	chr3:25,645,645-25,648,164 *[6]
Reln	F: TGAGCCTAAGAAGGTTAGCACG R: AATCTCCTTGCCTTCTTCCTG	chr5:21,729,464-21,731,808
	F: GTCAGAAGCAAAGCTAAGTGAC R: CCGAAACTTTCCGTGTTTCCT	chr5:21,834,299-21,835,333 *[7]
Rims1	F: GACAGCAAGAAAGCGAGCAT R:CCAACTGCTTGAGACCAGATAC	chr1:22,411,453-22,412,240
	F: GAGGGATTGCTGAGAAATTGAG R: GAAGCCACAACTGAAATCAAGG	chr1:22,619,415-22,624,96 *[14]
Slit2	F: TCCACCTTAGAAGCACCCAG R:GAGCAATGCAGTCAGCTTCTG	chr5:48,504,107-48,504,509
	F: CCTTAAGCTCCTGGACACTGC R:TGCTGGCACCTCAATCTCAC	chr5:48,567,966-48,568,749 *[15]
	F: AGATTCTGAGGTTGGAGGCAC R:GTGACTTTAGCTGGGTTTCTGTTC	chr5:48,653,094-48,653,847 *[1]

Shank2	F: CCTGGAAAGGATTTGGCAC R:GCAGGATGCTTTCAGGTTG	chr7:151,407,617-151,409,280
	F: GCACCCTCCTTACTCCCTACA R: TGTGGAATTTGGGGTTCTAGG	chr7:151,418,691-151,421,132 *[2]
	F: CCATTGCCTCCATTCTTCC R:GTGTTGGTGGCTTACTGGTTAG	chr7:151,447,540-151,449,273 *[3]

* presents the numbered primers location for Figure S2C and Figure 5A.

Gene Expression Primer Sequences

Name	Forward primers (5'to 3')	Reverse primers (5'to 3')
Lsh	GCAGATGAAATGGGTTTGGGA	GGGTTCCATGATACAGCAGAG
Gapdh	CCAGTATGACTCCACTCACGG	GAAGACACCAGTAGACTCCACG
Tubb3	TGGAGCGCATCAGCGTATACTACA	TGCAGGCAGTCACAATTATCACAC
Nlgn1	ACCCCTTGCGAATCACTGT	CTAGCGTATTTTGCAGGCTG
Nrxn1	CCAGTGACGATGAGGACATTG	ATAGAGGAGGATGAGGATGCAC
Shank2	GAGCTAAACTCCATTCTGCAGC	ACAGTGAAGGTGACCGTCGT
Mib1	GGTGACCTCAATGAAGAGCTG	CCAGCACACTGTCCGTTTAC
Slit2	CCTGGAATTGCACGCTGT	AATAGTGATATCCACGGGACCTT
Grik2	AGTTTTTGTGGCAGTGGGAG	ACGCTGGCACTTCAGAGAC
Negr1	GGCATATCTCCCCATCAGC	TCGTTTTCTGCACTGCATTC
Rims1	CAAGGCCTATTGGTGACATCC	CCCACCTTATCATACCACAGC
Reln	ACTGGAATTGAACCCCAACA	GCTCACAGAAGGTGACAGCA
Nrcam	GAAAAGGAGGATGCTCATGC	TTCACCCCCTCTCCATAGTC

Bisulfite Primer sequences

Name	Sequence (5'to 3')	Chromosome position
Oct4	F: ATGGGTTGAAATATTGGGTTTATTTA R: CCACCCTCTAACCTTAACCTCTAAC	ch17: 35642574-35642935
(1)	F: GGTTATAGTGGTTGTTTTGGTGGAGAG R: CTACCTCTAATTCCATAACATATATAAAC	chr3:25646617-25647257
(2)	F: GATTTTTATTTGAAAAGGATTAAGTG R: AAATAAACTTCTATTTTCACAACAACC	chr3:156566931-156567553
(3)	F: ATGTTAAAGGAATATGGAATTGTTTTATG R: CAACTCCCATCAATAAAAAACACAATC	chr5:21834299-21834827
(4)	F: AAGATTTTGAGGTTAATAAGAGAAATTT R: TATCTCACCAAATCCTCCCTAATAAAC	chr7:151448409-151448949

(5)	F: GTGATGTAGAAATTAATATTGTTTGTG R: ATATCAAAAACAATAACTAAATAACC	chr12:45558646-45559291
(6)	F: GTTGGGAGTTTGTGTGTTGGAATTGGA R: CTATAACAACAAACCACATCTACTAAC	chr12:45471607-45472267
(7)	F: GTTGTGTTATAATATGTAAGTGGTG R: CTTAAATACTCTCATCACAACTCC	chr3: 25704365-25704645
(8)	F: ATGTAGATGTTATGTTGTTTTAAATT R: AATCCTTTACATACACAACCCAAAA	chr3:156269810-156270349

Primers sequences and chromosome locations of potential enhancers used for enhancer reporter assay

Name	Sequence(5'to 3')	H3K4me1 peaks position
Mib1	F: GGAGGATCCGGAATGATGGCAACACTTATGC R: CAGTCGACGAGCAGAAGAAAGGGCACTAAAC	chr18:10,734,670-10,740,383
Negr1	F: CTTGGATCCCAAATGCCCAAATCTGTTCC R: CTCGTCGACAAAAATGACCTCCCCTCACAG	chr3:156,564,833-156,567,787
Nrxn1	F: CTTGGATCCGCCTGGTGAGTAGGTGTCTCTTC R: GGTGTCGACCAGTATGCCCTTGTGATGCAG	chr17:91,156,007-91,158,325
Nrcam	F: TTCGGATCCGGCAATTCTGTCTAGCATTG R: CTCGTCGACCAACATGGCTGTTATGAAGAC	chr12:45,471,230-45,476,393
Shank2	F: CAAGGATCCATGTCACACACCACAGGCAAG R: CAAGTCGACCATCTCTCCTAGTTCAGCAGAGC	chr7:151,418,700-151,421,123
Nlgn1	F: CTTGGATCCCACTCCTCTGACCTGTCAAGTG R: GACGTCGACCTAAGGAAATGGGGGTTGTG	chr3:25,702,653-25,707,030
Grik2	F: TTCGGATCCACATGAAGCTGTGGAGCAAAAC R: CTCGTCGACGTCTCAAACAACTTTTCAGCTCC	chr10:49,436,251-49,437,731

Primers sequences and chromosome locations of transcription factor binding motifs

Name	Sequence(5'to 3')	Chromosome postition
Negr1	F: GGTCCTATAGTGACCCAAATGCT R: AGCTGTGTCCACAGAGTGATTTG	chr3:156,567,708-156,567,943
Rims1	F: AGTGGGGTGGGGGGGTCTCC R: TTCCTGGGGTCCAGGCTGCAG	chr1:22,621,003-22,621,252
Nrxn1	F: GTACCTTGATCAATCACCCCTC R: GGGACAGTCGCTGATGATTTA	chr17:91,157,486-91,157,680
Nrcam	F: GAGAAACCCTGTCTCGAAAAAC R: GTAATCCAAAAACAAATCAAGCC	chr12:45,474,680-45,474,900

Mib1	F: GGGACTTCAGGATAGAATTTTCC R: GTAGACATCCACAACATGGTTGTT	chr18:10,736,394-10,736,572
Grik2	F: AGGCTCTGTTCCAAGACGAC R: TGACACTAGATGGGAATGAAAGAC	chr10:49,436,868-49,437,131
Nlgn1	F: GCTTTTTTTCTGCCTGTTGTG R: GAGCCAGCTTCAGAATGTAGTG	chr3:25,704,350-25,704,497
Slc25a21	F: CACTGGTCCCAAGGAACAAC R: CCTCTGCTGCCATGAAGAAA	chr12:58,209,900-58,210,100
C2	F: CAGACTGCTTTAACCATCCGA R: GTAATTTCACCCTCAGCCATG	chr17:35,018,520-35,018,720
Sh3bp4	F: CTTACGTGGTATGCGCCCTA R: GGAGGAACTGCTACGGGTTAG	chr1:90986650-90986850
Pou2f2	F: TGAAGCTGCCTAGCCCTGAT R: TTTTCTGGAGCTTTGGTCGAG	chr7:25,916,400-25,916,600

Fig S1

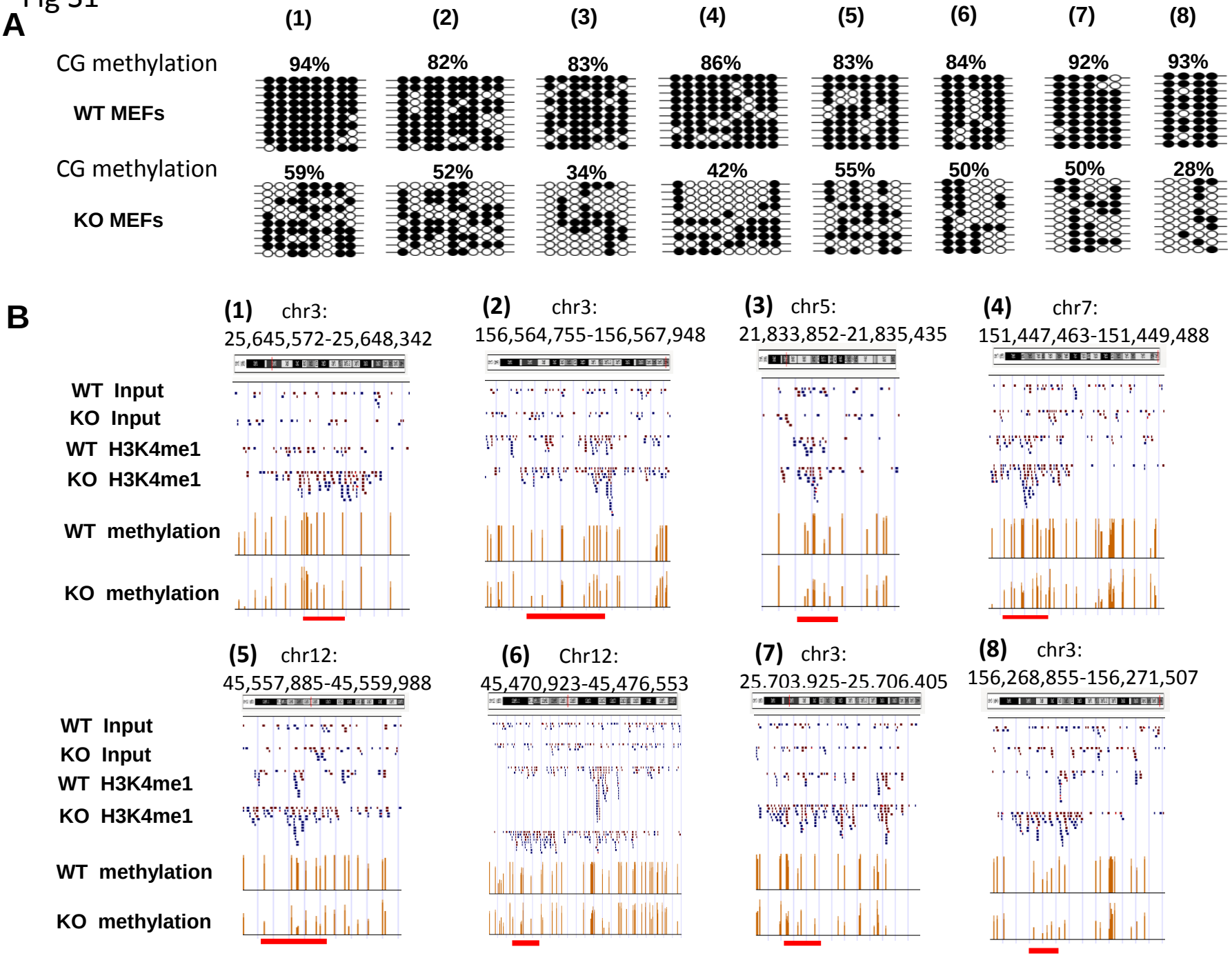
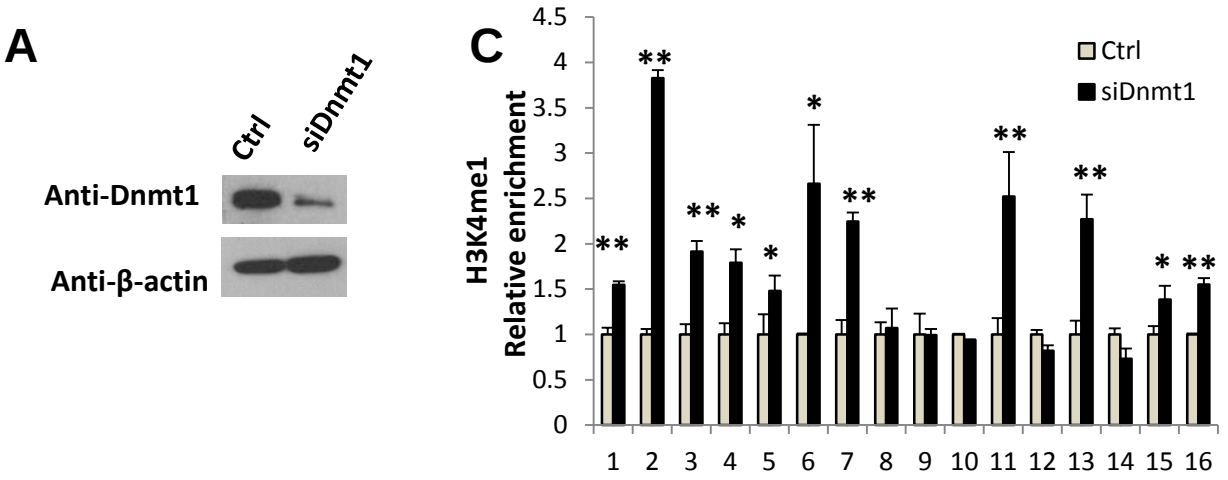

A
(1)
(2)
(3)
(4)
(5)
(6)
(7)
(8)
CG methylation
WT MEFs
94%
82%
83%
86%
83%
84%
92%
93%
CG methylation
KO MEFs
59%
52%
34%
42%
55%
50%
50%
28%
B
(1) chr3: 25,645,572-25,648,342
(2) chr3: 156,564,755-156,567,948
(3) chr5: 21,833,852-21,835,435
(4) chr7: 151,447,463-151,449,488
WT Input
KO Input
WT H3K4me1
KO H3K4me1
WT methylation
KO methylation
(5) chr12: 45,557,885-45,559,988
(6) Chr12: 45,470,923-45,476,553
(7) chr3: 25,703,925-25,706,405
(8) chr3: 156,268,855-156,271,507
WT Input
KO Input
WT H3K4me1
KO H3K4me1
WT methylation
KO methylation

Fig S2


A
Ctrl
siDnmt1
Anti-Dnmt1
Anti-β-actin
C
H3K4me1
Relative enrichment
Ctrl
siDnmt1
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

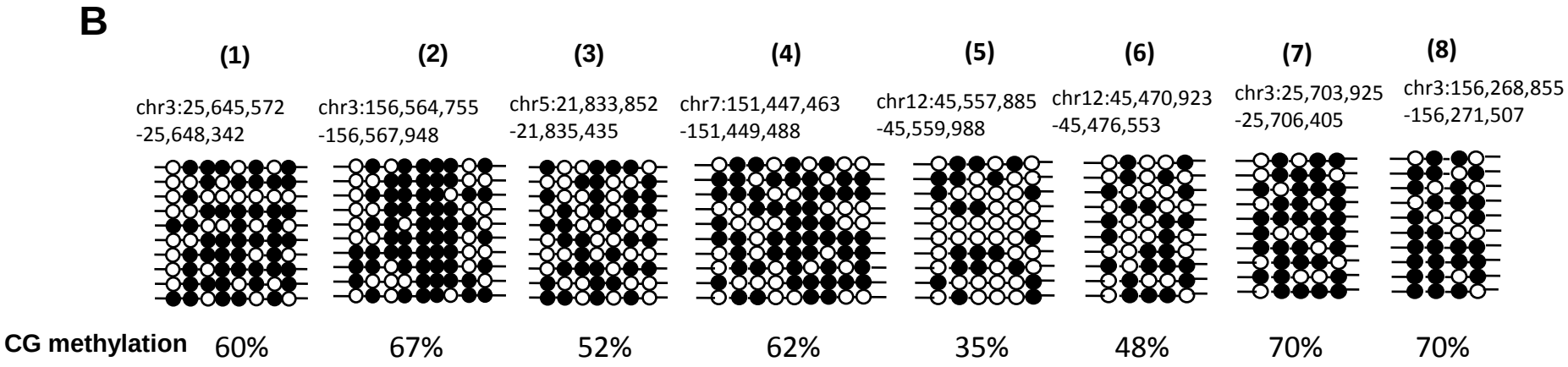

B
(1)
chr3:25,645,572
-25,648,342
(2)
chr3:156,564,755
-156,567,948
(3)
chr5:21,833,852
-21,835,435
(4)
chr7:151,447,463
-151,449,488
(5)
chr12:45,557,885
-45,559,988
(6)
chr12:45,470,923
-45,476,553
(7)
chr3:25,703,925
-25,706,405
(8)
chr3:156,268,855
-156,271,507
CG methylation
60%
67%
52%
62%
35%
48%
70%
70%

Fig S3

GO term analysis for KO-H3K4me1 marked genes

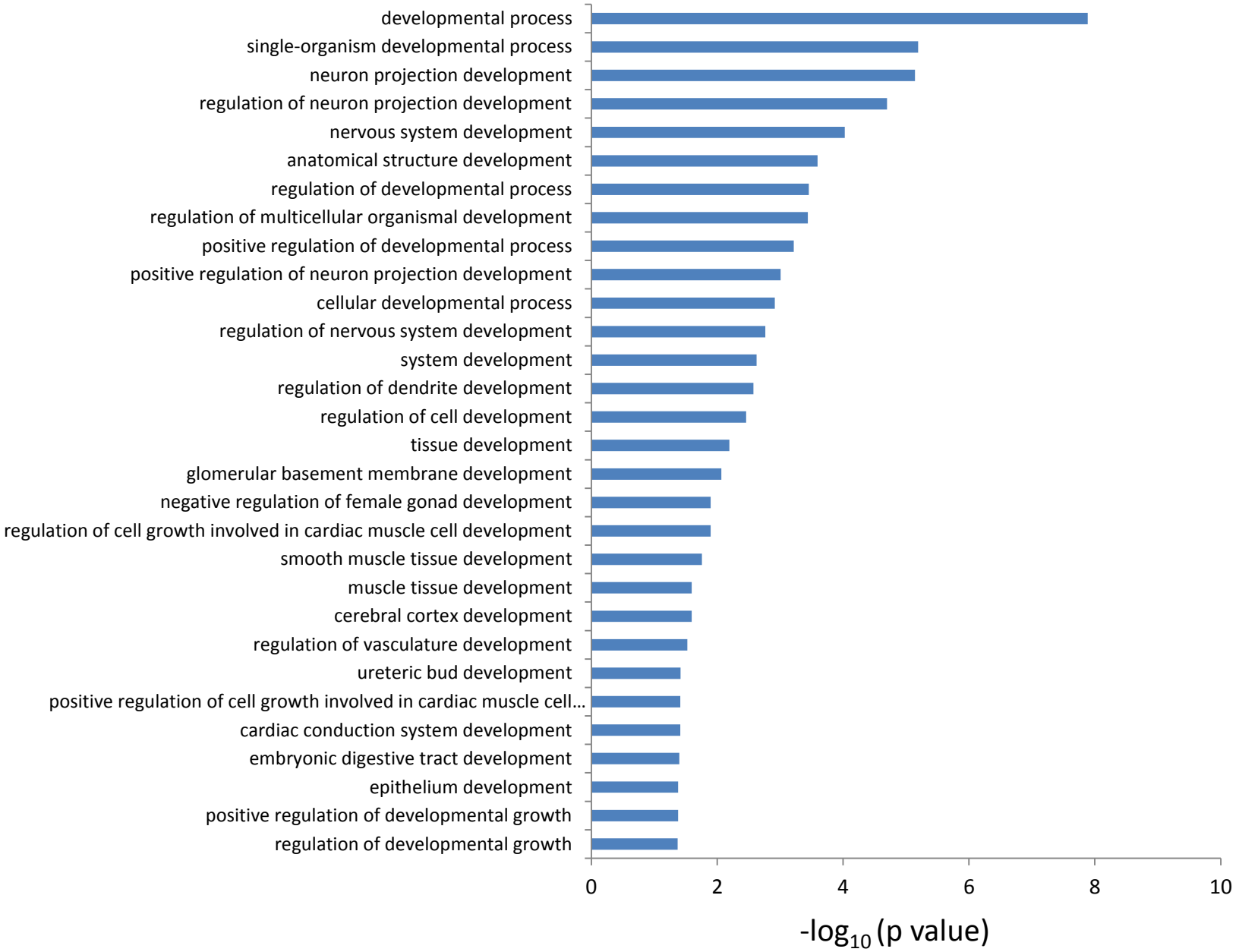


Fig S4

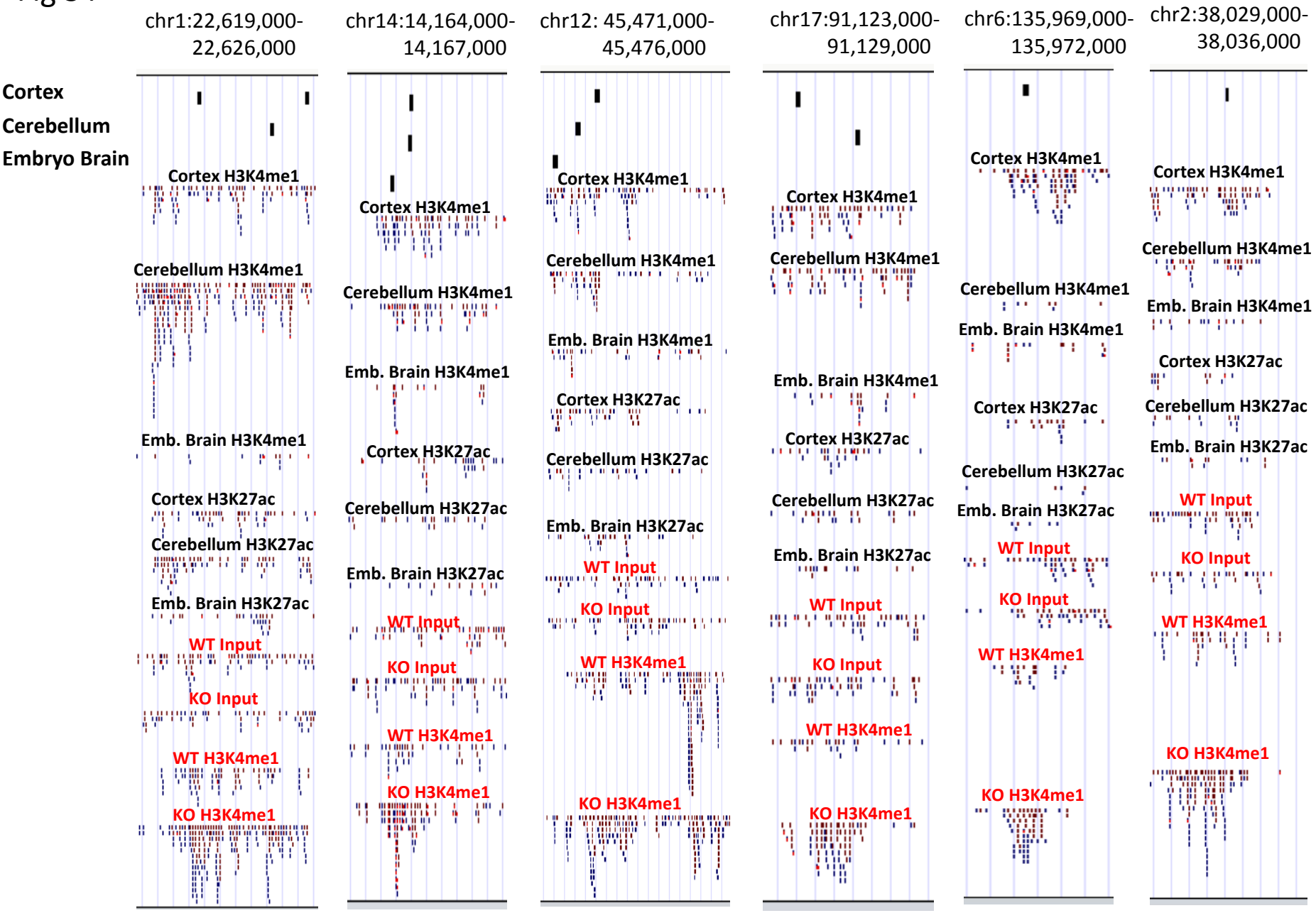
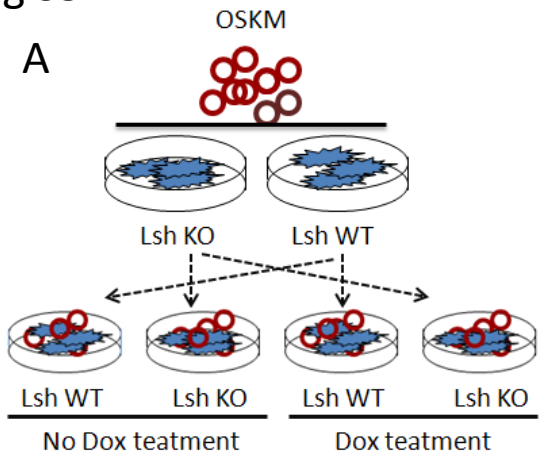
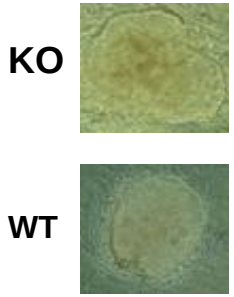
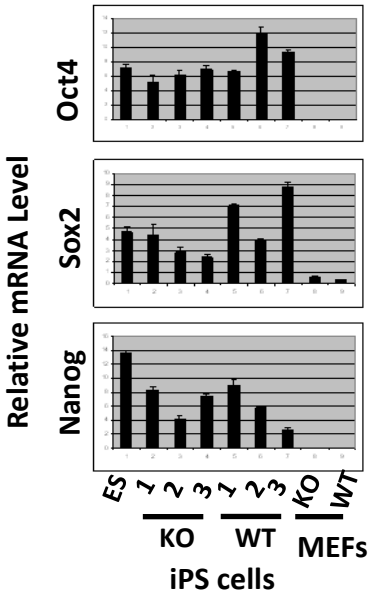
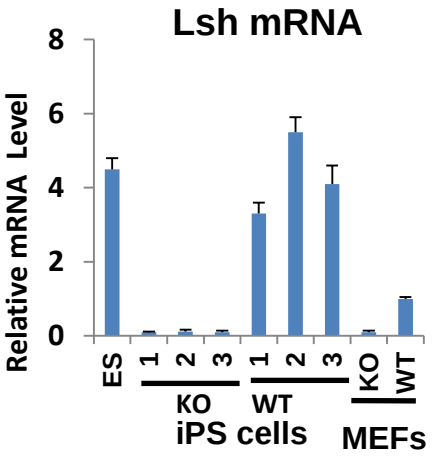

chr1:22,619,000-22,626,000
chr14:14,164,000-14,167,000
chr12: 45,471,000-45,476,000
chr17:91,123,000-91,129,000
chr6:135,969,000-135,972,000
chr2:38,029,000-38,036,000
Cortex
Cerebellum
Embryo Brain
Cortex H3K4me1
Cerebellum H3K4me1
Emb. Brain H3K4me1
Cortex H3K27ac
Cerebellum H3K27ac
Emb. Brain H3K27ac
WT Input
KO Input
WT H3K4me1
KO H3K4me1

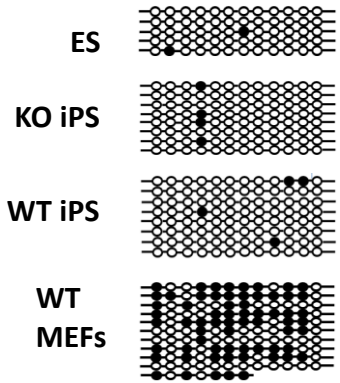
Fig S5

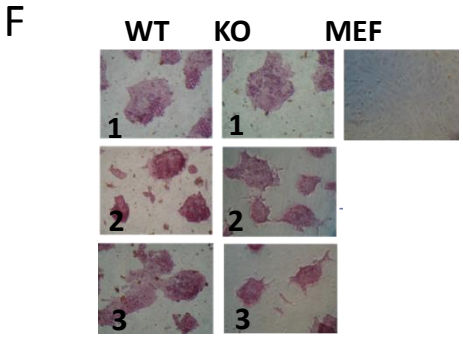

A
OSKM
Lsh KO
Lsh WT
Lsh WT
Lsh KO
Lsh WT
Lsh KO
No Dox teatment
Dox teatment


B
KO
WT


C
Oct4
Sox2
Nanog
Relative mRNA Level
ES
1
2
3
1
2
3
KO
WT
KO
WT
MEFs
iPS cells


D
Lsh mRNA
Relative mRNA Level
8
6
4
2
0
ES
1
2
3
1
2
3
KO
WT
KO
WT
iPS cells
MEFs


E
ES
KO iPS
WT iPS
WT
MEFs


F
WT
KO
MEF
1
1
2
2
3
3

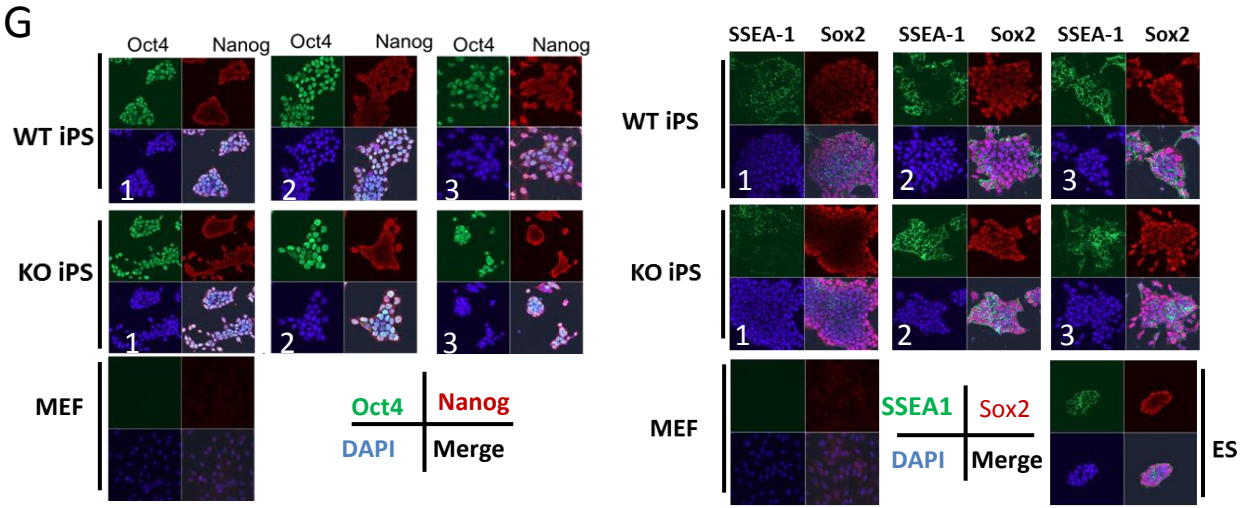
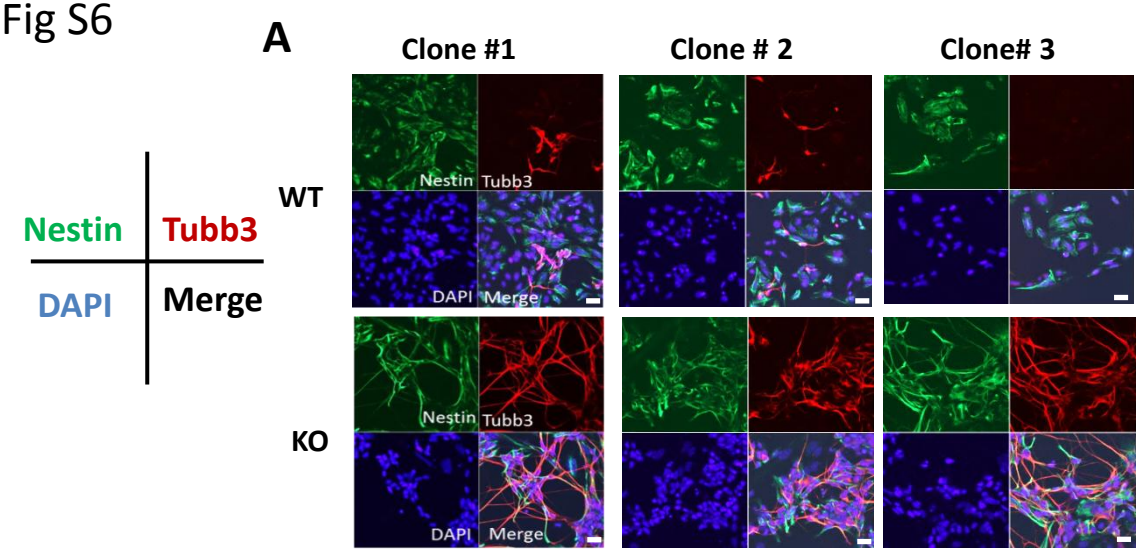
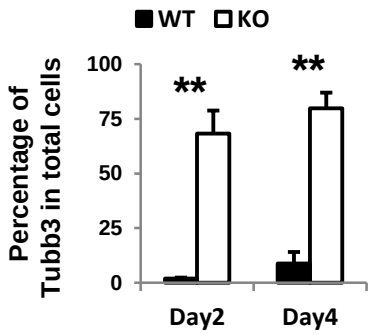

G
Oct4
Nanog
Oct4
Nanog
Oct4
Nanog
WT iPS
KO iPS
MEF
1
2
3
Oct4
Nanog
DAPI
Merge
SSEA-1
Sox2
SSEA-1
Sox2
SSEA-1
Sox2
WT iPS
KO iPS
MEF
1
2
3
SSEA1
Sox2
DAPI
Merge
ES

Fig S6



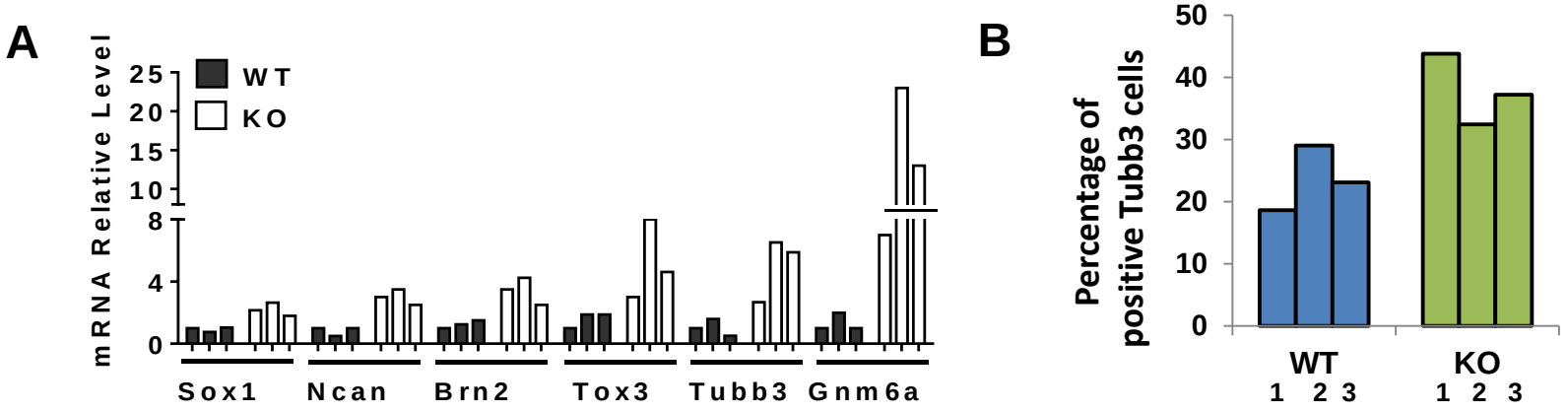


C

Table of teratoma sizes (Length x Width mm)

	1	2	3	4	5	6
WT1	9x6	10x7	19x17	9x6	9x7	12x7
WT2	7x6	8x5	11x10	n/a	n/a	n/a
WT3	5x5	4x4	1x1	n/a	n/a	n/a
KO1	8x5	8x7	10x8	2x2	5x4	n/a
KO2	9x7	12x7	19x6	9x7	n/a	n/a
KO3	7x7	11x8	4x4	2x2	n/a	n/a

Fig S7


A
mRNA Relative Level
25
20
15
10
8
4
0
WT
KO
Sox1
Ncan
Brn2
Tox3
Tubb3
Gnm6a
B
Percentage of positive Tubb3 cells
50
40
30
20
10
0
WT
1 2 3
KO
1 2 3

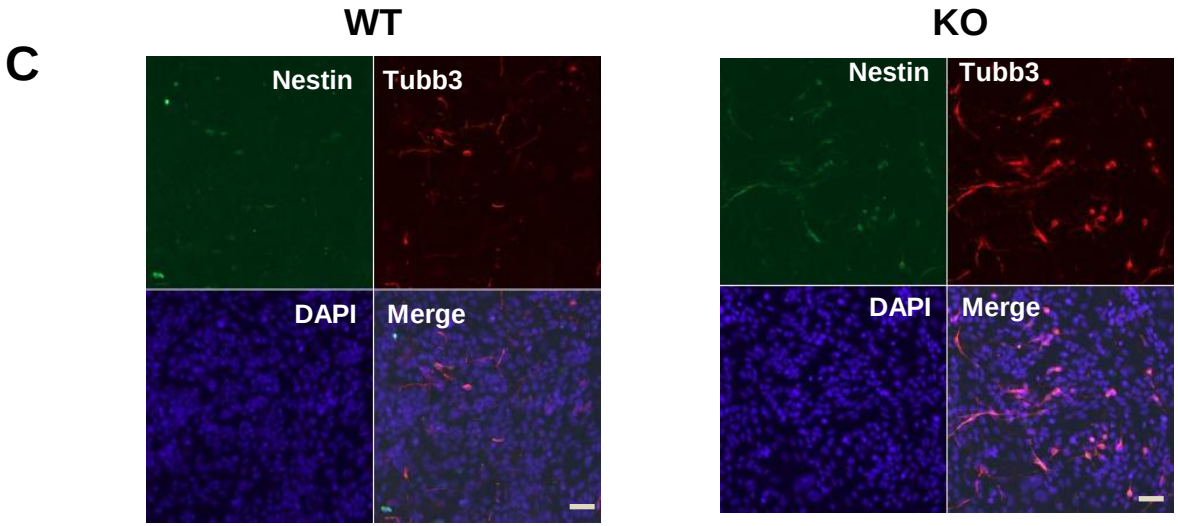

C
WT
KO
Nestin
Tubb3
DAPI
Merge
Nestin
Tubb3
DAPI
Merge

Fig S8

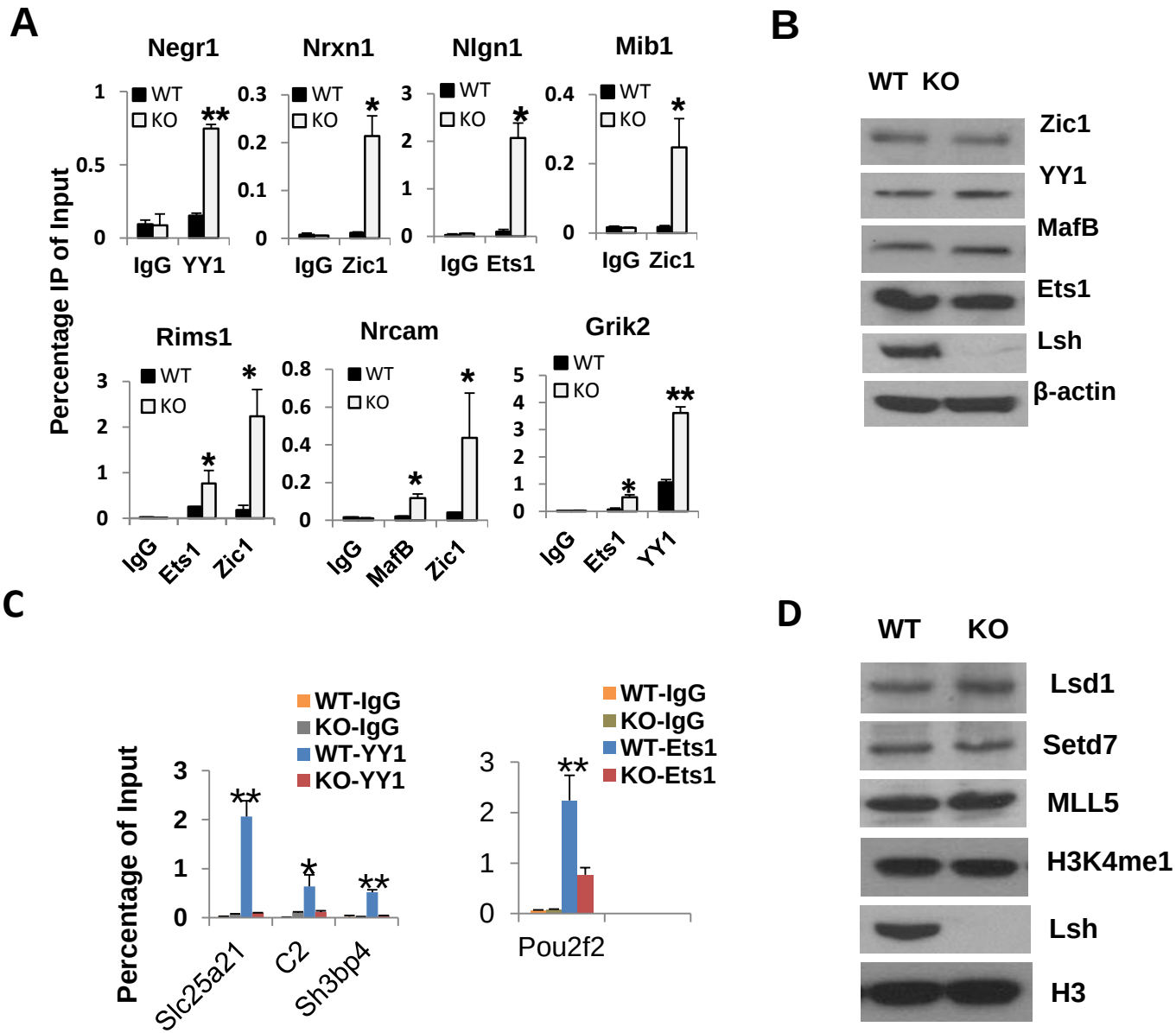

A
Negr1
Nrxn1
Nlgn1
Mib1
WT
KO
Percentage IP of Input
IgG YY1
IgG Zic1
IgG Ets1
IgG Zic1
Rims1
Nrcam
Grik2
IgG Ets1 Zic1
IgG MafB Zic1
IgG Ets1 YY1
B
WT KO
Zic1
YY1
MafB
Ets1
Lsh
β-actin
C
WT-IgG
KO-IgG
WT-YY1
KO-YY1
Percentage of Input
Slc25a21
C2
Sh3bp4
WT-IgG
KO-IgG
WT-Ets1
KO-Ets1
Pou2f2
D
WT KO
Lsd1
Setd7
MLL5
H3K4me1
Lsh
H3

Fig S9

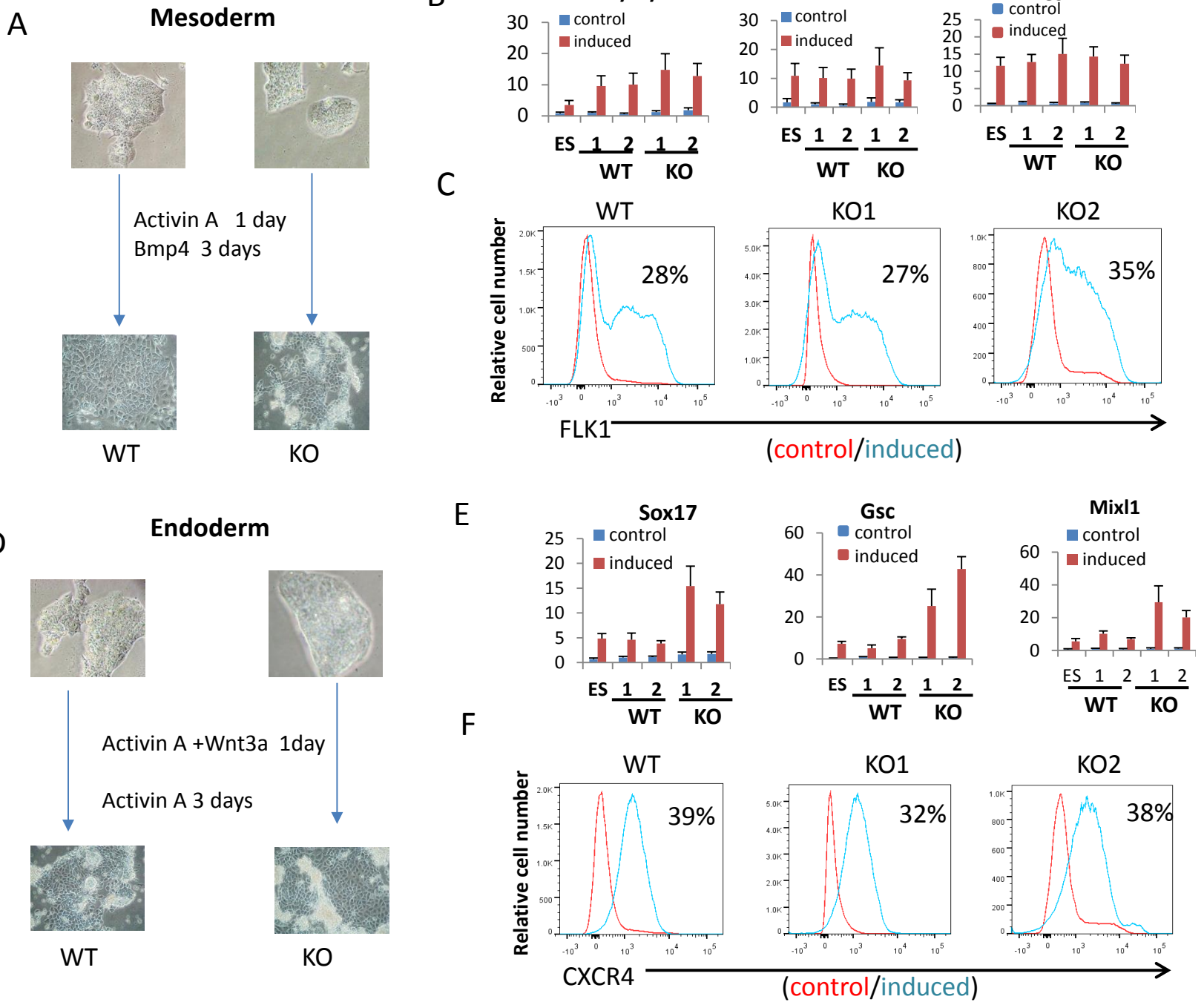

A
Mesoderm
Activin A 1 day
Bmp4 3 days
WT
KO
B
Brachyury
Meox1
Des
control
induced
ES 1 2 1 2
WT KO
C
WT
KO1
KO2
Relative cell number
28%
27%
35%
FLK1
(control/induced)
D
Endoderm
Activin A +Wnt3a 1day
Activin A 3 days
WT
KO
E
Sox17
Gsc
Mixl1
F
39%
32%
38%
CXCR4
(control/induced)

Fig S10