www.sciencemag.org/cgi/content/full/329/5990/444/DC1

Supporting Online Material for

Dnmt3a-Dependent Nonpromoter DNA Methylation Facilitates Transcription of Neurogenic Genes

Hao Wu,* Volkan Coskun, Jifang Tao, Wei Xie, Weihong Ge, Kazuaki Yoshikawa, En Li, Yi Zhang, Yi Eve Sun*

*To whom correspondence should be addressed. E-mail: ysun@mednet.ucla.edu (Y.E.S.); haowu7@gmail.com (H.W.)

Published 23 July 2010, *Science* **329**, 444 (2010)
DOI: 10.1126/science.1190485

This PDF file includes:

Supporting Online Material

Materials and Methods

Mouse postnatal SEZ/SVZ NSC culture

Postnatal day 14-21 day mouse lateral ventricle walls were dissected and mechanically dissociated as previously described (*1*). Single cell suspensions were plated at low density (10,000~20,000 cells cm^{-2}) on cell culture dishes coated with poly-L-ornithine (PO)/fibronectin (FN), allowing individual cells to form spatially distinct colonies. Serial passage (less than 2 passages, or 2 weeks) allowed us to enrich undifferentiated NSCs to high purity (>95% $Nestin^{+}/Sox2^{+}$). The NSCs were maintained in serum-free B27 medium in presence of 10 ng/ml FGF-2/bFGF (Peprotech) and 10 ng/ml EGF (Invitrogen). The monolayer cultures were serially passaged every 4-7 days (up to 14 passages, 2 months) without losing their self-renewal ability and multipotency. To avoid the potential non-specific effect induced by culture adaption, only NSCs with limited *in vitro* culturing (less than 3 weeks) were used in the study. For tri-lineage differentiation, a cohort of undifferentiated NSCs, were re-plated on PO/FN substrate with mitogen for 2 days followed by withdrawal of mitogens for an additional 3 days. All animal experiments were done in accordance with the protocols approved by the Animal Research Committee of the University of California, Los Angeles.

Immunostaining

Cells on coverslip were fixed (4% paraformaldehyde (PFA) at room temperature), processed and stained (overnight at 4°C) for immunofluorescence analysis using following antibodies as previously described (*2*): mouse anti-Dnmt3a (1:400, Imgenex, C-terminus specific, for both Dnmt3a and Dnmt3a2), mouse anti-Dnmt3b (1:400, Imgenex, N-terminus specific, for all Dnmt3b isoforms), mouse anti-Map2 (1:1000, Sigma), mouse anti-Tubb3/TuJ1 (1:1000, Covance), mouse anti-CNPase/Cnp (1:400, Chemicon), mouse anti-Mbp (1:1000, Chemicon), rabbit anti-Nestin (1:1,000, a kind gift from Dr. Ron McKay at NINDS), goat anti-Sox2 (1:200, Santa Cruz), rabbit anti-Ki67 (1:1,000, Novocastra), anti-GFAP (guinea pig, 1:2,500, Advanced immunochemical; rabbit, 1:500, Chemicon), rabbit anti-H3K27me3 (1:1,000, Millipore), and rabbit anti-Ezh2 (1:1,000, Active Motif). Cy3-, Cy2- and Amca-conjugated donkey secondary antibodies were purchased from Jackson Immunoresearch. Hoechst 33342 dye was used to label the nuclei. Images were acquired on an Olympus fluorescent microscope.

Sagittal brain sections were prepared and processed for immunostaining as previously described (*3*) with minor modifications using following primary antibodies: rabbit anti-TuJ1 (1:1000, Covance), rabbit anti-Calbindin (1:100, Cell Signaling), sheep anti-TH (1:500, Pel-Freez), rabbit anti-Calretinin (1:500, Millipore), Guinea pig anti-Dcx (1:1,000, Millipore), guinea pig anti-Dlx2 (1:3,000) (*4*), anti-Sp8 (1:15,000, Millipore), rabbit anti-Ki67 (1:500, Novocastra), rabbit anti-GFAP (1:400, Millipore), mouse anti-Mbp (1:1,000, Millipore), rabbit anti-Dnmt3a (1:400, Santa Cruz), and rabbit anti-Dnmt3b (1:400, #157) (*5*). Briefly, wild-type and mutant mice (P21-24) were transcardially perfused by using 4% PFA. All brains were cryoprotected in 20% sucrose for 24 h, and cut in sagittal plane on a cryostat (Leica) at 10 μm. After several PBS washes, all slides were antigen-retrieved (95°C for 15min, Bio-Genex) and immersed in blocking solution (5% normal donkey serum and 0.2% Triton X-100 in PBS) for 1 h at room temperature. Primary antibody incubation was carried out overnight at 25°C with in dilution buffer (5% normal donkey serum and 0.5% Triton X-100 in PBS). The following day, slides were incubated with appropriate secondary antibodies as described above. The slides were examined by using Olympus fluorescent microscopes or a confocal system (Zeiss, LSM 510-META). Cell apoptosis in SVZ was examined using a Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) kit (Roche).

Immunoblotting

The immunoblotting was performed essentially as described with minor modifications (*6*). Whole cell lysate were prepared from cells or tissues using SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl, pH8.1) supplemented with protease inhibitor cocktails (Roche). The lysate was immediately sonicated using a microtip (Branson sonifier 450) to reduce viscosity. Protein concentration was measured by a BCA protein assay (Pierce). The cell lysates were stored at -80˚C before use. 10-30 μg whole cell lysate was fractionated with 10% SDS-PAGE. The following mouse monoclonal antibodies were used: anti-Dnmt3a (1:250, C-terminus, Imgenex), anti-Dnmt3b (1:250, Imgenex), anti-Gapdh

(1:2,500, Abcam), and anti-β-actin (1:2,000, Sigma). The following rabbit polyclonal antibodies were used: anti-Dnmt3a (1:1,000, N-terminus, Santa Cruz), anti-Dnmt1 (1: 5,000, Path52, a kind gift from Dr. T. Bestor at Collumbia University), anti-Suz12 (1:1,000, Abcam) and anti-Ezh2 (1:1,000, Active motif).

For immunoblotting of histone modifications, 10 µg whole cell lysate was fractionated with 4-20% NuPAGE gel (Invitrogen) or 15% SDS-PAGE. The following polyclonal antibodies were used to probe histone modifications: anti-H3K4me3 (Millipore/Upstate, 07-473, or Abcam, ab8580), anti-H3K27me3 (Millipore/Upstate, 07-449), anti-H3K9me2 (Millipore/Upstate, 07-441, or Abcam, ab1220), anti-H3K9me3 (Abcam, ab8898), anti-H3K36me3 (Abcam, ab9050), anti-panH3 (Millipore/Upstate, 07-690).

RNA isolation and quantitative real-time RT-PCR

Total RNA was isolated using Trizol (Invitrogen). After Turbo DNase (Ambion) treatment, 1 µg total RNA was used for cDNA synthesis using Superscript III (Invitrogen) following manufacturer's recommendations. Real-time PCR was performed in an iCycler using iQ SYBR Green Supermix (Bio-Rad). PCR efficiency and specificity of each primer pair was examined by standard curve of serially diluted cDNA and melting curve functionality respectively. Fold change was calculated based on $2^{-\Delta\Delta Ct}$ method after normalization to the transcript level of housekeeping gene *Gapdh*. Primer sequences used in the real-time RT-PCR are listed in the Table S2.

Gene expression microarrays

Two-color gene expression microarrays (Agilent) were used to directly compare relative changes in gene expression between WT and KO NSCs under proliferating or differentiation conditions. 500 ng of total RNA was labeled with Cy3- or Cy5-CTP using the Agilent Low RNA Input Fluorescent Linear Amplification Kit. After labeling and cRNA purification (Qiagen), cRNA was quantified using the NanoDrop spectrophotometer (Agilent). 825 ng Cy3- or Cy5-labelled cRNA were combined and hybridized to the Agilent 4x44K whole mouse genome microarray (G4122F, ~44,000 probes per array) for 17 h at 65 ˚C (10 r.p.m.). Replicate experiments were performed using biologically independent cultures with a dye-swap experimental design. Data was collected using Agilent microarray scanner and extracted using Feature Extraction 9.1 software (Agilent).

For identification of differentially expressed genes, we used NIA array analysis tool (http://lgsun.grc.nia.nih.gov/ANOVA). Of all the probes present on the microarray, signal intensity of redundant probes was averaged before analysis. Following parameters were used for analyzing statistically significant differential expression:

Threshold z-value to remove outliers: 10,000
Error Model: Max (Average, Bayesian)
Error variance averaging window: 200
Proportion of highest error variances to be removed: 0.01
Bayesian degrees of freedom: 5
False discovery rate (FDR) threshold: 0.05 (high stringency) or 0.2 (moderate stringency)

For clustering and heatmap display, Lowess-normalized signal intensity was log2 transformed and median-certered. Heatmaps were generated using Cluster3 and Java Treeview.

Bisulphite conversion followed by genomic sequencing or MALDI-TOF MS

We carried out bisulfite conversion of 1 µg Genomic DNA using EZ DNA methylation-Gold Kit (Zymo) according to manufacturer's instructions. Hot-start PCR was used to amplify the region of interest from the bisulfite converted genomic DNA. The PCR products were purified (Qiagen) and sub-cloned into the pCR4 TOPO vector (Invitrogen). 15-20 clones were randomly selected and sequenced using BigDye 3 terminator sequencing reagents (Applied Biosystems). Genomic sequences were retrieved from the UCSC genome database (mm8). All the bisulfite genomic sequencing primers were designed using the MethPrimer program. Primer sequences used in bisulfite genomic sequencing are listed in the Table S2.

Bisulphite conversion followed by Sequenom MassARRAY® MALDI-TOF Mass Spectrometry (MS) based quantitative DNA methylation analysis was performed in triplicate using standard protocol (Sequenom EpiTYPER DNA Methylation Analysis). The results were highly similar to those from bisulphite genomic sequencing analyses. Primers were designed using Sequenom's EpiDesigner tool (http://www.epidesigner.com). Primer sequences (without linker sequences) used in Sequenome analysis are listed in the Table S2. A T7-promoter tag is incorporated into each reverse PCR

amplification primer (5'-cagtaatacgactcactatagggaga-3'), and a 10-mer linker sequence (5'-aggaagagag-3') is added to each forward primer that balances the primer length.

Chromatin immunoprecipitation (ChIP) and DNA tiling microarrays (ChIP-chip)

To immunoprecipitate chromatin, 2-5x10^7 (for Dnmt3a and PRC2 subunits, i.e. Suz12 and Ezh2) or 1-2x10^7 (for histone modifications) cells were cross-linked with 1% formaldehyde for 10 minutes at room temperature followed by exposure to 0.125 M glycine. After two washes with cold PBS, cells were collected as pellets and stored at -80 ˚C before use. Nuclei were extracted and lysed sequentially with lysis buffer 1 (LB1, 50 mM Hepe2-KOH, pH7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP40, 0.25% Triton X-100), lysis buffer 2 (LB2, 10 mM Tris-HCl, pH8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA), and lysis buffer 3 (LB3, 10 mM Tris-HCl, pH8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine). Chromatin was sonicated using a microtip (Branson sonifier 450) until the DNA fragments were reduced to 200-1000 bp in length. 10 µg antibodies were pre-incubated with 100 µl Dynal protein-G beads (Invitrogen) for at least 6 hours. Immunoprecipitation was performed overnight at 4 ˚C with antibody conjugated protein-G beads. DNA/protein complexes were washed with RIPA buffer (50 mM Hepes-KOH, pH7.6, 500 mM LiCl, 1 mM EDTA, 1% NP-40, 0.7% Na-Deoxycholate) for 5 times and eluted from beads and reverse cross-linked at 65 ˚C overnight. The DNA was digested with RNase A and proteinase K sequentially, followed by phenol/chloroform extraction ethanol precipitation. Following antibodies were used in ChIP assays: anti-H3K4me3 (ab8580, Abcam or 07-473, Millipore), anti-H3K27me3 (07-449, Millipore), anti-panH3 (07-690, Millipore), anti-Dnmt3a (Santa Cruz), anti-Ezh2 (39103, Active Motif), and anti-Suz12 (a kind gift from Dr. Y. Zhang).

Real-time quantitative PCR was performed in an iCycler using iQ SYBR Green Supermix (Bio-Rad). For quantification of relative level of Dnmt3a, PRC2 and histone modification occupancy, we calculated the percentage of immunoprecipitated DNA over input DNA. Primer sequences used in ChIP-qPCR are listed in the Table S2.

For whole-genome DNA tiling microarray analyses of Dnmt3a (in WT, IP/WCE) and H3K27me3 (WT versus KO, IP/IP) occupancy, we amplified ChIP or input DNA samples using whole genome amplification kit (Sigma). Amplified ChIP or input samples were labeled (5' Cy5- or Cy3-random nonamers, TriLink Biotechnologies) using the standard protocol (NimbleGen Arrays User's Guide for ChIP-chip analysis). Hybridization of labeled samples to whole genome HD2 microarrays 4-array set (Roche/NimbleGen, ~2.1 million tiling probes per array, covering the entire non-repetitive portion of mouse genome) was carried out for 16-20 h at 42 ˚C using NimbleGen hybridization System 4. After stringent washes, microarrays were subsequently scanned using the Agilent scanner at 5 micron resolution. Data were extracted and analyzed using NimbleScan v2.5 (Roche/NimbleGen). Dnmt3a binding sites were identified by a sliding window-base statistical algorithm in NimbleScan v2.5 (FDR<1% or <5%). For identification of probes associated with significant increase in H3K27me3 levels in KO NSCs as compared to WT NSCs, a non-parametric one-sided Kolmogorov-Smirno (KS) test was used (KS score>2). Briefly, from the scaled log2-ratio data, a fixed-length window (750bp) is placed around each consecutive probe and the one-sided KS test is applied to determine whether the probes are drawn from a significantly more positive distribution of intensity log-ratios than those in the rest of the array. The resulting score for each probe is the -$\log_{10}$ P-value from the windowed KS test around that probe. Using NimbleScan v2.5, peak data files are generated from the P-value data files. NimbleScan software detects peaks by searching for at least 2 probes above a P-value minimum cutoff (-log10) of 2. Peaks within 500bp of each other are merged.

For extended promoter tiling microarray analyses of Dnmt3a, PRC2 (Suz12 and Ezh2) and histone modification (H3K4me3, H3K27me3 and histone H3) occupancy, we amplified input and ChIP DNA samples using two-step ligation-mediated PCR (LM-PCR) as previously described (*7*) or whole genome amplification kit (Sigma). For hybridization with mouse extended promoter tiling microarray 2-array set (Agilent, G4490A, ~244,000 probes per array, ~280 bp per probe, covering approximately -5.5kb to +2.5kb genomic regions relative to TSS of ~17,000 annotated Refseq genes), 2 µg of amplified input and IP DNA was labeled with Cy3- and Cy5-dUTP (Perkin Elmer) respectively using Bioprimer DNA labeling Systeim (Invitrogen). For direct comparison of chromatin immunoprecipitated from wild-type and *Dnmt3a*-deficient NSCs, WT and KO samples were labeled with Cy3- and Cy5-dUTP respectively. Hybridization was carried out for 40 h at 65 ˚C (10-15 r.p.m.) and subsequently scanned using an Agilent scanner. Data were extracted using Feature Extraction 9.1 (Agilent) and analyzed using ChIP analytics 1.3 (Agilent). For IP versus WCE (IP/WCE) comparison, we performed blank subtraction normalization,

inter-array median normalization and intra-array (dye-bias) median normalization by calculating a one-step Tukey biweight median before peak detection. For IP versus IP (IP/IP) comparison, we performed an intra-array Lowess (intensity-dependent) normalization. Two Dnmt3a ChIP-chip experiments using biological independent samples were performed and yielded highly reproducible results. We combined and processed the replicate array results to identify Dnmt3a high affinity targets. Probes associated with significant level of Dnmt3a signals were called on the basis of a built-in heuristic algorithm (P_{Xbar}<0.001) incorporating neighbor probes information using a Whitehead error model (Agilent ChIP analytics 1.3). Probes associated with significant level of histone modification signals were also called on the basis of the built-in heuristic algorithm (P_{Xbar}<0.002 for H3K27me3 or P_{Xbar}<0.2 for H3K4me3) (Agilent ChIP analytics 1.3). This statistical method generated highly similar lists of genes associated with given histone modification within promoter regions as compared to that of previously reported in NSCs using ChIP-seq method (*8*). For visualization, raw enrichment ratios of annotated probes (UCSC, mm7) were displayed as bar graph (log2 ratios). Moreover, IP/WCE comparison and IP/IP comparison of WT and KO samples generated highly similar results.

For calculating the averaged enrichment distribution of unprocessed probe signals (log2 Ratio), Dnmt3a binding sites or regions associated with significant changes in H3K27me3 levels data were binned to 500-bp intervals using a 250-bp sliding window within genomic regions (10-bk for NimbleGen whole-genome microarrays, 2.5kb for Agilent extended promoter microarrays) flanking TSSs of annotated Refseq genes. Heatmaps were generated and visualized using Cluster3 and Java Treeview, respectively.

Methylated DNA immunoprecipitation (MeDIP) and DNA tiling microarrays (MeDIP-chip)

MeDIP asay was carried out as previously described with minor modifications (*9*). Briefly, genomic DNA was digested with proteinase K and RNase A sequentially, and purified by phenol/chloroform extraction. Purified genomic DNA was sonicated and heat-denatured (95 ˚C, 10 min). An aliquot of sonicated genomic DNA was saved as input. 5 µg fragmented genomic DNA was immunoprecipitated with 5 µl of monoclonal antibody against 5-methylcytidine (Eurogentec) at 4 ˚C overnight in a final volume of 500 µl of IP buffer (10 mM sodium phosphate (pH 7.0), 140 mM NaCl, 0.05% Triton X-100). We incubated the DNA-antibody mixture with 30 µl protein G Dynabeads (Invitrogen) for 2 h at 4 ˚C and washed it three times with 1 ml IP buffer. We then treated the beads with proteinase K for at least 3 h at 55 ˚C and purified the methylated DNA by phenol-chloroform extraction followed by ethanol precipitation. For real-time PCR analysis, 10 ng of input genomic DNA and 1/30 of the immunoprecipitated (IP) methylated DNA was used for each PCR reaction. MeDIP-qPCR was performed in an iCycler (Bio-Rad) using iQ SYBR Green Supermix (Bio-Rad). Reactions were done in duplicates and standard curves were calculated on serial dilutions of input genomic DNA. ChIP primers were used in MeDIP-qPCR to quantify the relative DNA methylation level within the regions occupied by Dnmt3a.

For microarray analyses, we amplified input and immunoprecipitated DNA using whole genome amplification kit (Sigma). Labeling, hybridization, array processing and data extraction were done following procedures described for ChIP-chip. Data analysis was performed using ChIP analytics 1.3 (Agilent, for extended promoter microarrays) or NimbleScan v2.5 (Roche/NimbleGen, for whole-genome microarrays). For identification of probes associated with significant decrease in DNA methylation levels in KO NSCs as compared to WT NSCs, a non-parametric one-sided Kolmogorov-Smirno (KS) test was used.

Lentiviral infection and acute ablation of Dnmt3a in *Dnmt3a*$^{flox/flox}$ NSCs

Control (UbiC-empty-IRES-EGFP) or Cre-expressing (UbiC-Cre-IRES-EGFP) FUIGW lentiviral vector was co-transfected with three other helper vectors (VSVG, RSV-REV and pMDL g/p RRE) into HEK-293T cells as previously described (http://www.sciencegateway.org/protocols/lentivirues/index.htm). Packaged lentiviral particles were harvested at 48h and 72h post transfection. Harvested lentiviral particles were then combined and concentrated using ultracentrifugation (Beckman SW28 rotor, 25,000 rpm for 90 min). The embryonic or postnatal *Dnmt3a*$^{flox/flox}$ NSCs were transduced with concentrated control or Cre-expressing lentiviruses in presence of 4 µg/ml polybrene (Sigma) in two consecutive days and analyzed 5 days after first infection. The typical infection efficiency of NSCs was 40-60% 3 days after first infection based on EGFP fluorescence. Depletion of Dnmt3a after infection was routinely monitored by immunostaining of EGFP, Cre and Dnmt3a.

Short-hairpin RNA (shRNA)-mediated suppression of PRC2 core subunits in *Dnmt3a*-null NSCs
A lentiviral vector encoding shRNA specific for Suz12 (NM_199196, Sigma MISSION^{TM} TRC shRNA set) was used for silencing Suz12 expression (TRCN0000123889, targeting 3' UTR, 5'-CCGGGCTGTCTTAGAGATGGAGAATCTCGAGATTCTCCATCTCCATCTCTAAGACAGCTTTTTG-3'). Two lentiviral vectors encoding shRNA specific for Eed (NM_021876, Sigma MISSION^{TM} TRC shRNA set) were used for inhibiting Eed expression (#1: TRCN0000095719, targeting 3' UTR, 5'-CCGGTCTTGCTAGTAAGGGCACATACTCGAGTATGTGCCCTTACTAGCAAGATTTTTG-3';
#2: TRCN0000095722, targeting CDS, 5'- CCGGGAAGCAACAGAGTAACCTTATCTCGAGATAAGG TTACTCTGTTGCTTCTTTTTG-3'). A control shRNA vector (SHC002, The MISSION^{TM} Non-Target shRNA Control Vector, Sigma), which contains a shRNA sequence that does not target human and mouse genes (>4 base pair mismatches to any known human or mouse gene), was used as a negative control for Suz12 and Eed knockdown shRNAs. Lentiviruses were packaged as described above and were used for infecting NSCs in presence of 4 μg/ml polybrene (Sigma). NSCs were infected for two consecutive days and subsequently selected in media containing 2 μg/ml puromycin for 7-10 days before analysis.

Re-expression of WT or catalytically inactive Dnmt3a in *Dnmt3a*-null NSCs
Dnmt3a-null NSCs were transfected with the control vector (CAG-empty-IRES-blasticidin) and expression vectors encoding catalytically active (CAG-Dnmt3a^{WT} -IRES-blasticidin) or catalytically inactive (CAG-Dnmt3a$^{P705V/C706D}$-IRES-blasticidin) *de novo* DNA methyltransferases using lipofectamine 2000 (Invitrogen) according to manufacturer's recommendations. As a control, wild-type NSCs were also transfected with the empty vector (CAG-empty-IRES-blasticidin). After 24 h, transfected NSCs were passaged and subsequently selected in 5 μg/ml blasticidin-containing medium for 8-10 days before analysis.

***In vitro* chromatin recruitment assay**
A 1.1kb genomic region corresponding to Dnmt3a bound regions within the *Dlx2* gene was PCR cloned, purified and biotinylated at one terminus (a NotI site) by Klenow filling reactions. An aliquot of DNA was in vitro methylated by the CpG DNA methylase, SssI (NEB) following manufacturer's instructions. Both methylated and unmethylated DNA was then assembled into chromatin arrays consisted of recombinant *Xenopus* histones using salt dialysis as previously described (*10*). After immobilization using strepavidin-conjugated magnetic beads (Invitrogen), chromatin arrays were then incubated with *in vitro* reconstituted PRC2 complexes (Ezh2, Eed, Suz12, Aebp2, and RbAp48) for 30 min with rotation at room temperature (*11*). After several washes with binding buffer, the beads and bound PRC2 complexes were re-suspended in SDS sampling buffer and fractionated by 10% or 15% SDS-PAGE. The bound fraction of PRC2 complexes was detected by western blotting. The following primers were used to clone the 1.1kb genomic fragment within the *Dlx2* gene: forward 5'- CCACATCTTCTTGAACTTGGATCG-3'; reverse 5'-GTAAGCATCGCAAGCATCTAGGG-3'.

Supporting Text

Gene expression analysis of differentiating WT and KO SEZ/SVZ NSC cultures

To further investigate the effect of Dnmt3a on cell fate specification during postnatal NSC differentiation, we performed expression profiling of differentiating NSCs and cross-referenced the datasets with a published database where cell-type specific gene expression was comprehensively examined in highly enriched postnatal neuronal and glial populations (*12*). Among 766 down-regulated genes in differentiating KO cells, the percentage of neuronal genes was significantly enriched in down-regulated genes (32.1% versus 11.7% expected by chance, $P=3.34\times10^{-48}$), but not in up-regulated genes (8.5%). GO analysis also confirmed that genes functionally related to neurogenesis were overly represented in genes down-regulated in differentiating KO cells (Table S1). In contrast, glial lineage genes showed an opposite trend (25.5% in down-regulated genes; 46.3% in up-regulated genes), and were only enriched in up-regulated genes ($P=3.58\times10^{-39}$), indicating that Dnmt3a may also be involved in regulating glial lineage differentiation. Indeed, several Dnmt3a targets critically involved in astroglial (e.g. *Gfap* and *S100b*) and oligodendroglial differentiation (e.g. *Nkx2-2, Pdgfra* and *Mbp*) were significantly up-regulated in differentiating KO cultures. Taken together, these findings support the hypothesis that Dnmt3a is involved in maintaining neurogenic potential in postnatal NSCs by directly promoting transcription of neurogenic targets while repressing differentiation genes of alternative glial lineages through different mechanisms.

Mechanisms of Dnmt3a-dependent gene activation

Although we cannot entirely exclude the possibility that *Dnmt3a*-deficiency may induce a transcriptional repressor, which in turn represses other genes including some Dnmt3a targets, this seems unlikely, because we did not find a known repressor in the list of up-regulated genes in KO NSCs. Futhermore, all the Dnmts may possess methyltransferase activity independent functions in developmental gene regulation. In our rescue experiments, catalytically inactive Dnmt3a mutant re-occupy most, if not all, Dnmt3a binding-sites, but fail to restore normal chromatin states (reduce aberrantly elevated PRC2/H3K27me3 levels at H3K4me3-high/CpG-rich Dnmt3a targets) and neurogenesis deficits in *Dnmt3a*-null postnatal NSCs. Thus, DNA methylation independent functions of Dnmts are unlikely to play a major role in the gene regulatory events described in our study.

Non-promoter DNA methylation and gene expression in other tissues

While we have focused our epigenomic analyses on postnatal SEZ/SVZ neural stem cells (NSCs), multiple lines of evidence suggest that the positive correlation between non-promoter DNA methylation and tissue-specific gene expression is not restricted to this system. First, partial inactivation of maintenance methyltransferase Dnmt1 in other lineage-committed stem/progenitor cells (blood or skin) via hypomorphic mutations or lentiviral shRNA inhibition results in similar number of genes that are up-regulated or down-regulated (*13, 14*). Second, previous efforts on identifying tissue-specific de novo DNA methylated regions in human adult tissues have also revealed a strong positive correlation between non-promoter methylation (i.e. intragenic DNA methylation) and tissue-specific transcription of a cohort of developmentally regulated genes, including genes (e.g. PAX6 and UNC13A) specifically expressed in brain tissues (*15*). Third, unbiased genome-wide analyses indicate that high levels of DNA methylation are found in gene bodies of transcriptionally active genes in both human fibroblasts and B-cells (*16, 17*). Finally, gene body DNA methylation is a widespread phenomenon on actively transcribed X-chromosome, whereas promoter DNA methylation, but not intragenic methylation is present on the inactive X-chromosome in female cells (*18*).

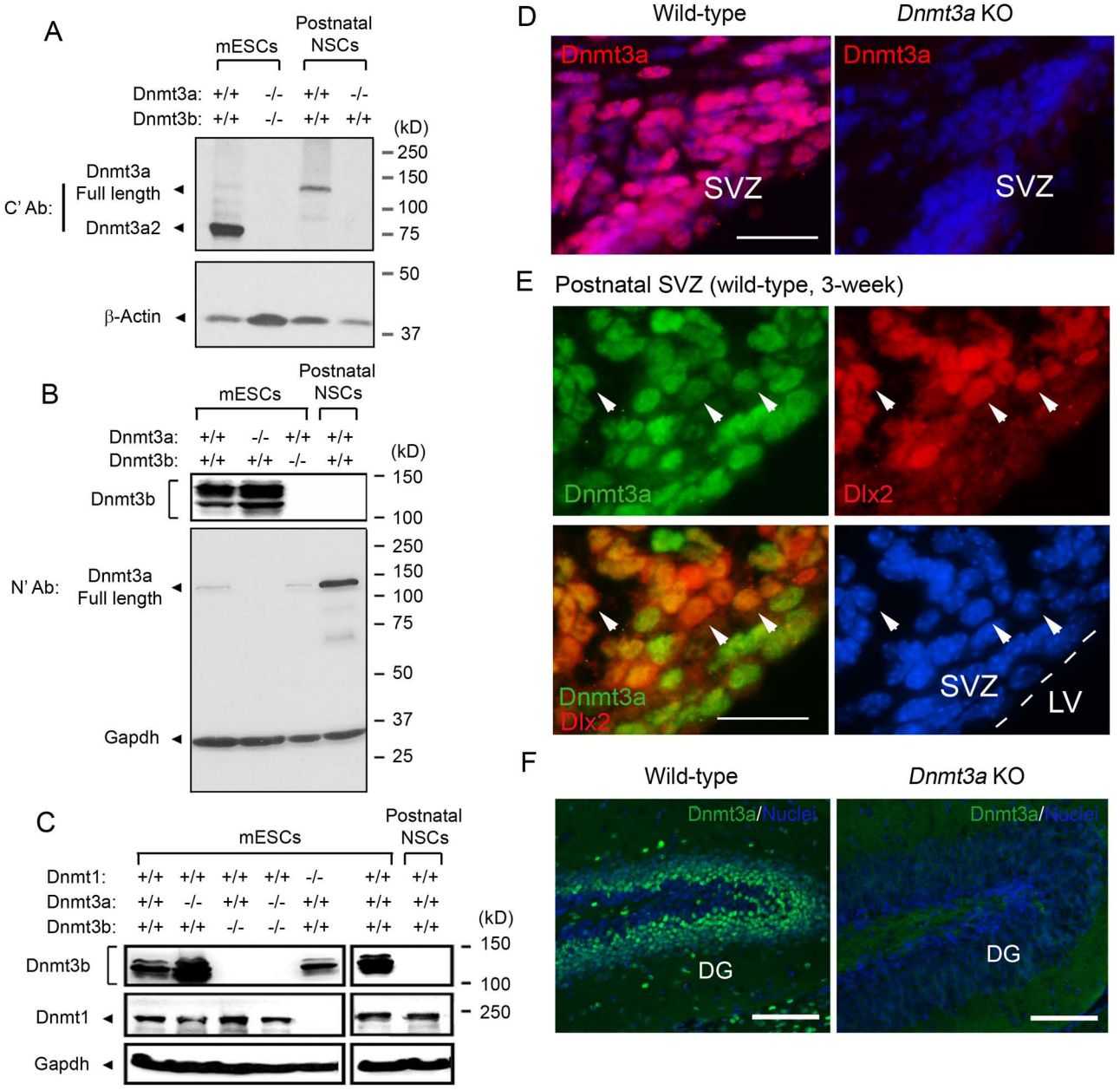

Fig. S1 Dnmt3a expression in embryonic stem cells (ESCs), postnatal neural stem cells (NSCs) and postnatal neurogenic zones.

(A-B) Immunoblotting using antibodies against Dnmt3a in ESCs and postnatal subventricular zone (SVZ) NSCs. The anti-Dnmt3a C-terminus monoclonal antibody (C' Ab in A) recognizes both full-length Dnmt3a (~130kD) and N' truncated Dnmt3a2 variant (~75kD). N' specific polyclonal antibodies (N' Ab in B) only recognize the full length Dnmt3a.

(C) Immunoblotting using antibodies against Dnmt3b and Dnmt1 in ESCs and NSCs with indicated genotypes.

(D) Immunohistochemistry for Dnmt3a (red) in subventricular zone (SVZ) sagittal brain sections of wild-type and *Dnmt3a*-null mice. Scale bars, 25 μm.

(E) Immunohistochemistry for Dnmt3a (green) and the neuronal progenitor marker Dlx2 (red) in sagittal brain sections of wild-type mice. Arrow heads indicate cells expressing both Dlx2 and Dnmt3a. Scale bars, 25 μm.

(F) Representative immunofluorescence images of Dnmt3a in WT and KO dentate gyrus (DG) at postnatal 3-weeks. Scale bar, 100 μm.

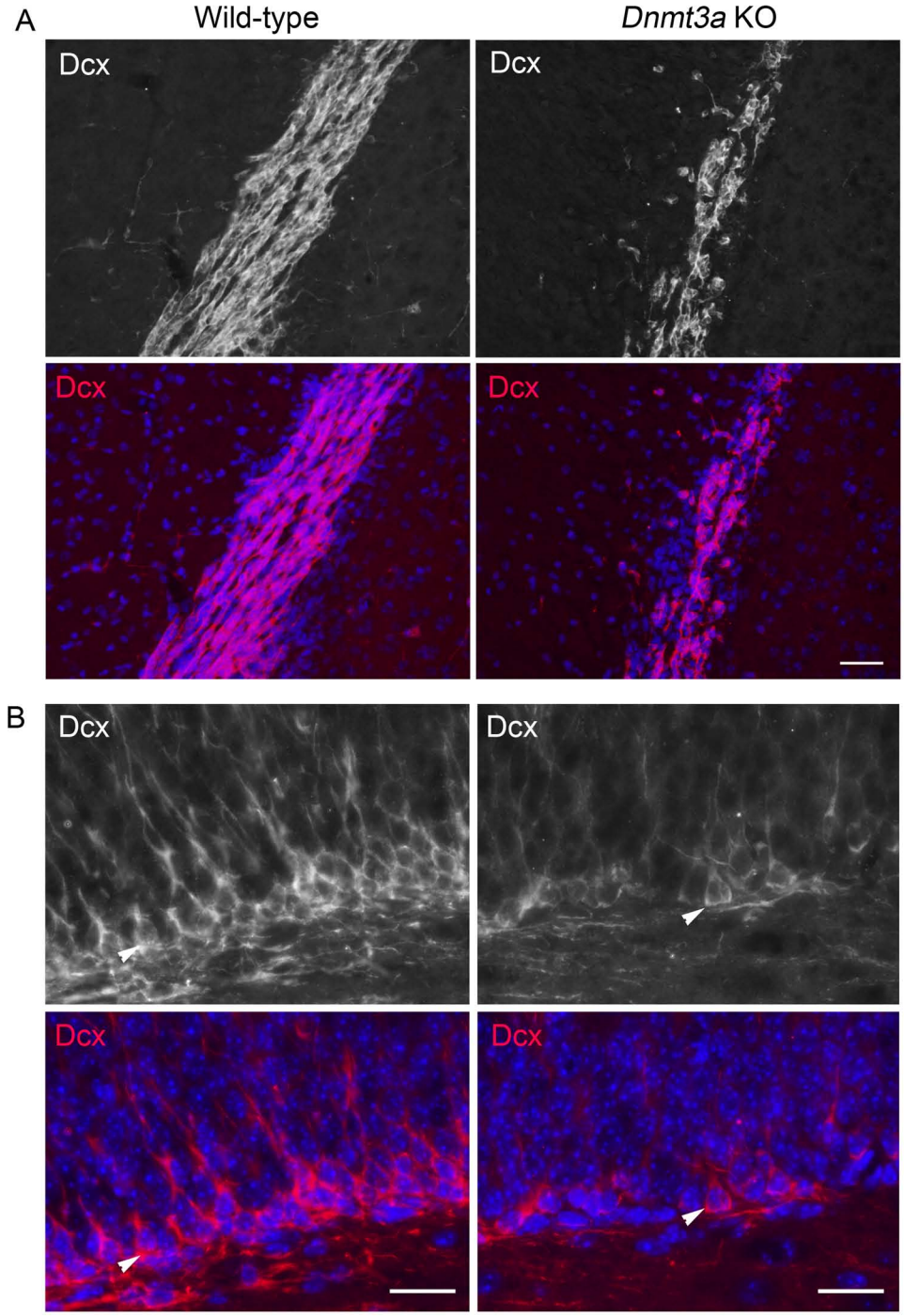

Fig. S2 *Dnmt3a*-deficiency impairs postnatal subventricular zone (SVZ) and hippocampal dentate gyrus (DG) neurogenesis.
(A) Immunohistochemistry for the neuroblast marker Dcx in SVZ sagittal brain sections of WT and KO mice at postnatal 3 weeks. Quantification of Dcx in SVZ is present in Fig. 1B. Scale bars, 50 μm.
(B) Representative immunostaining of the immature neuronal marker Dcx in WT and KO postnatal (3-weeks) DG (36 +/- 3.3 Dcx+ cells per section in WT versus 25 +/- 4.3 in KO, mean +/- s.d.; $P<0.05$). Arrow heads indicate Dcx+ cells. Scale bar, 50 μm.

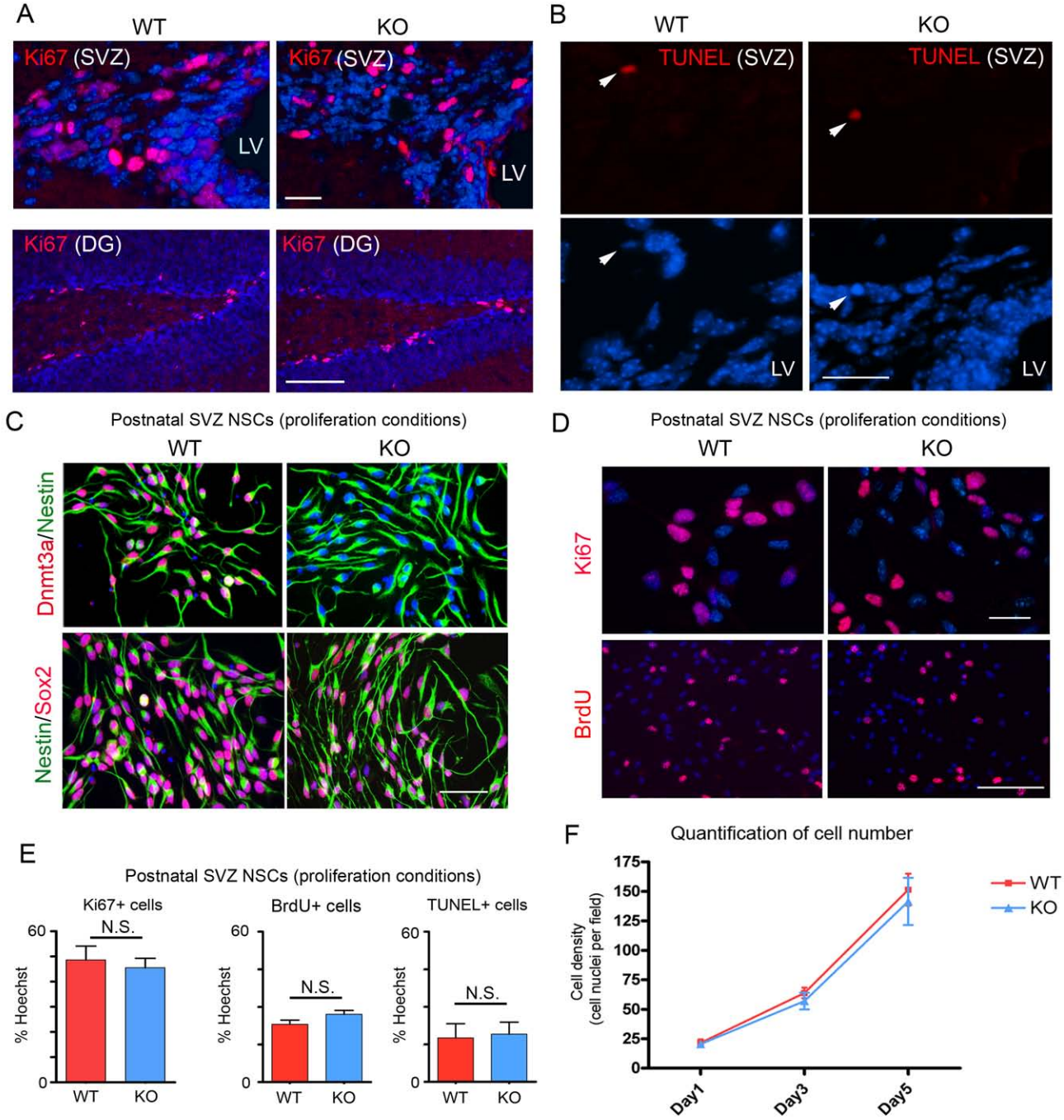

Fig. S3 Dnmt3a is dispensible for proliferation, suvival and maintenance of undifferentiated states of postnatal NSCs.

(A) No significant difference in number of Ki67+ proliferating cells was observed between WT and KO at both SVZ (82% of WT, 21% s.d.) and DG (96% of WT, 29% s.d.). Scale bars, 20 μm (SVZ) or 100 μm (DG).

(B) TUNEL assay in SVZ sagittal brain sections of WT and KO mice. Scale bars, 20 μm.

(C) Immunostainings of NSC markers Nestin and/or Sox2 in WT and KO NSCs showed that most of Dnmt3a-null NSCs (>90%) were still in undifferentiated states. Scale bars, 50 μm.

(D) Immunostaining of Ki67 and 5-Bromo-2-deoxyuridine (exposed to BrdU for 2h before analysis) in WT and KO SVZ NSC cultures at day 3 after passaging. Scale bar, 20 (upper) or 100 (lower) μm.

(E) Quantification of proliferation and cell death rate in WT and KO SVZ NSCs. Error bars, s.e.m. of triplicate experiments.

(F) Quantification of the cell growth in WT and KO SVZ NSC cultures at day 1, 3 and 5 after plating.

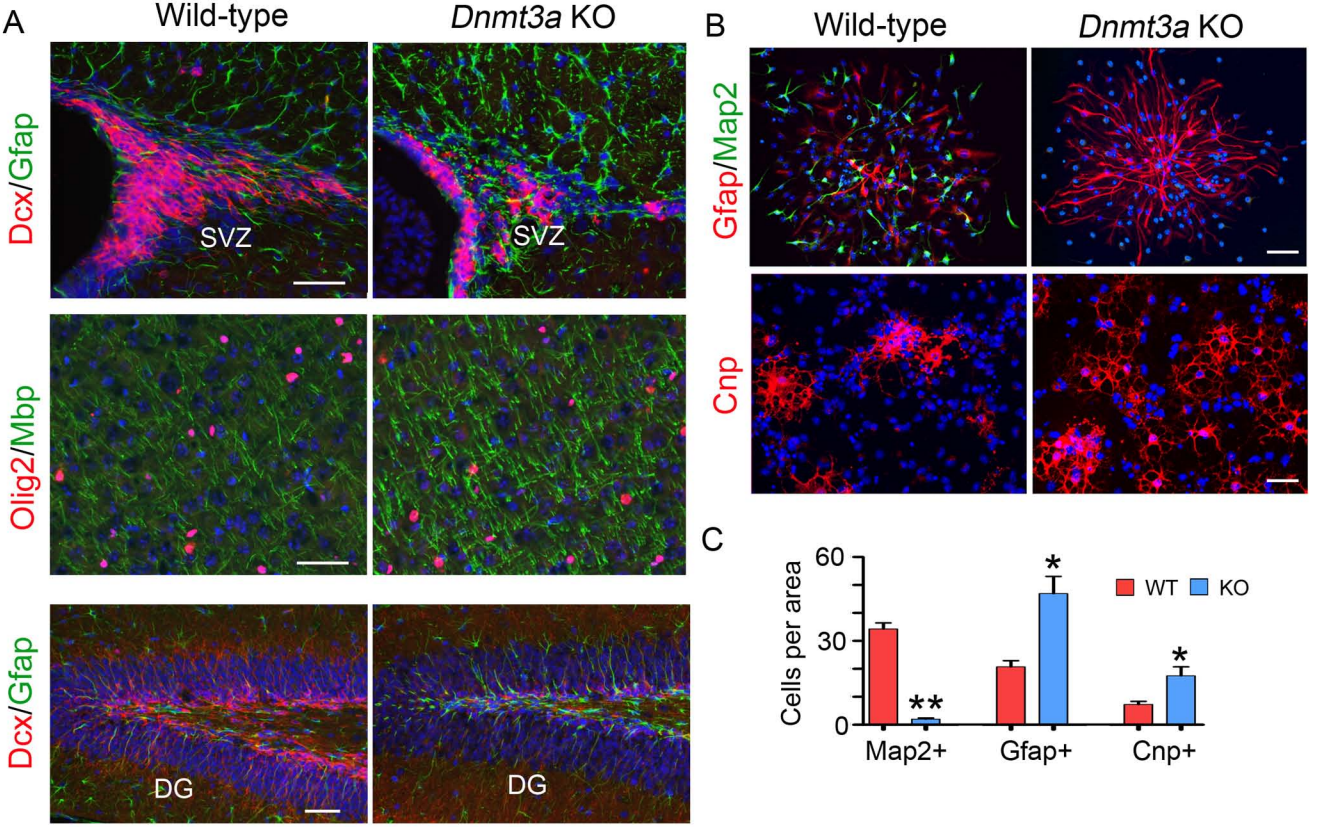

Fig. S4 Dnmt3a is dispensible for postnatal gliogenesis in vivo and in SVZ NSC cultures.
(A) Immunohistochemistry for the neuroblast marker Dcx, the astroglial marker Gfap and oligodendroglial markers Olig2 as well as Mbp in sagittal brain sections of WT and KO mice at P24. Shown for Olig2 and Mbp immunostaining are cortical regions immediately above corpus callosum (CC). Scale bars, 50 μm.
(B) Immunocytochemistry for the neuronal marker Map2, the astroglial marker Gfap and the oligodendroglial marker CNPase/Cnp in differentiating WT and KO SVZ NSC cultures. Scale bars, 50μm.
(C) Quantification of the number of Map2+ neurons, Gfap+ astrocytes and Cnp+ oligodendrocytes after 3 days of differentiation. Error bars, s.e.m. (*, P<0.05; **, P<0.01; n=4).

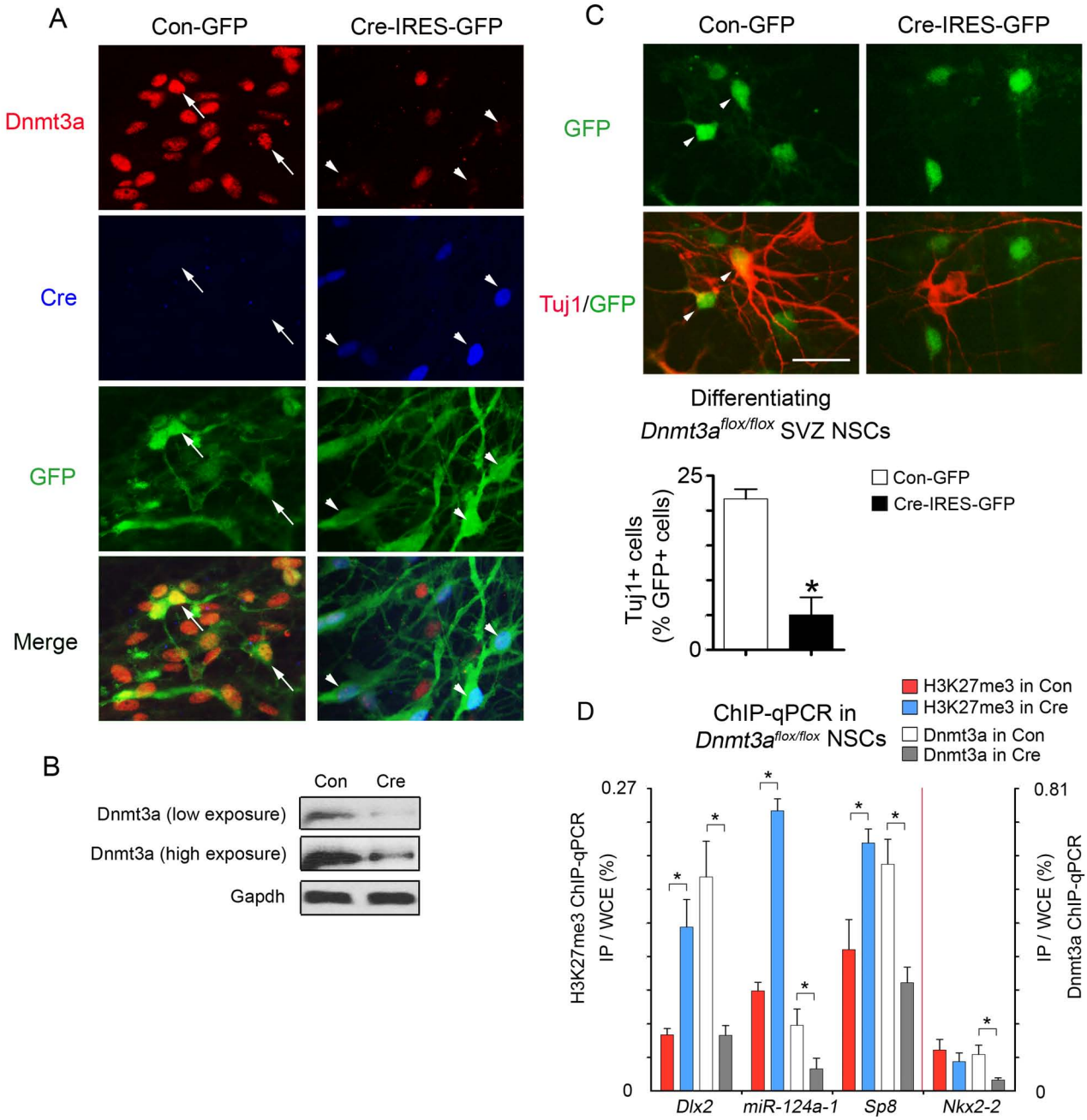

Fig. S5 Cre recombinase mediated acute ablation of Dnmt3a in *Dnmt3a*$^{flox/flox}$ NSCs impairs neurogenesis and causes aberrant increase in H3K27me3 on neurogenic targets.

(A-B) Immunostaining (A) and western blotting (B) of Dnmt3a expression levels in postnatal *Dnmt3a*$^{flox/flox}$ NSCs infected with control (empty-IRES-GFP) and Cre (Cre-IRES-GFP) lentivirues for 5 days before analysis. Arrow (left) and arrowhead (right) in (A) marked GFP-positive cells in con- and Cre-lentivirus infected *Dnmt3a*$^{flox/flox}$ NSCs. Gapdh was used in (B) as a loading control.

(C) Tuj1+GFP+ double-positive cells were significantly fewer in *Dnmt3a*$^{flox/flox}$ NSCs infected with Cre lentivirues than those infected with control. Error bars, s.e.m. (*, as compared to control infected cells, $P<0.005$, n=3). Scale bar, 20 μm.

(D) Dnmt3a occupancy and H3K27me3 levels of 4 representative Dnmt3a target genes were determined by ChIP-qPCR in control and Cre lentivirus infected postnatal NSCs (*, as compared to WT, $P<0.01$, n=3). Error bars represent s.e.m. determined from three independent experiments.

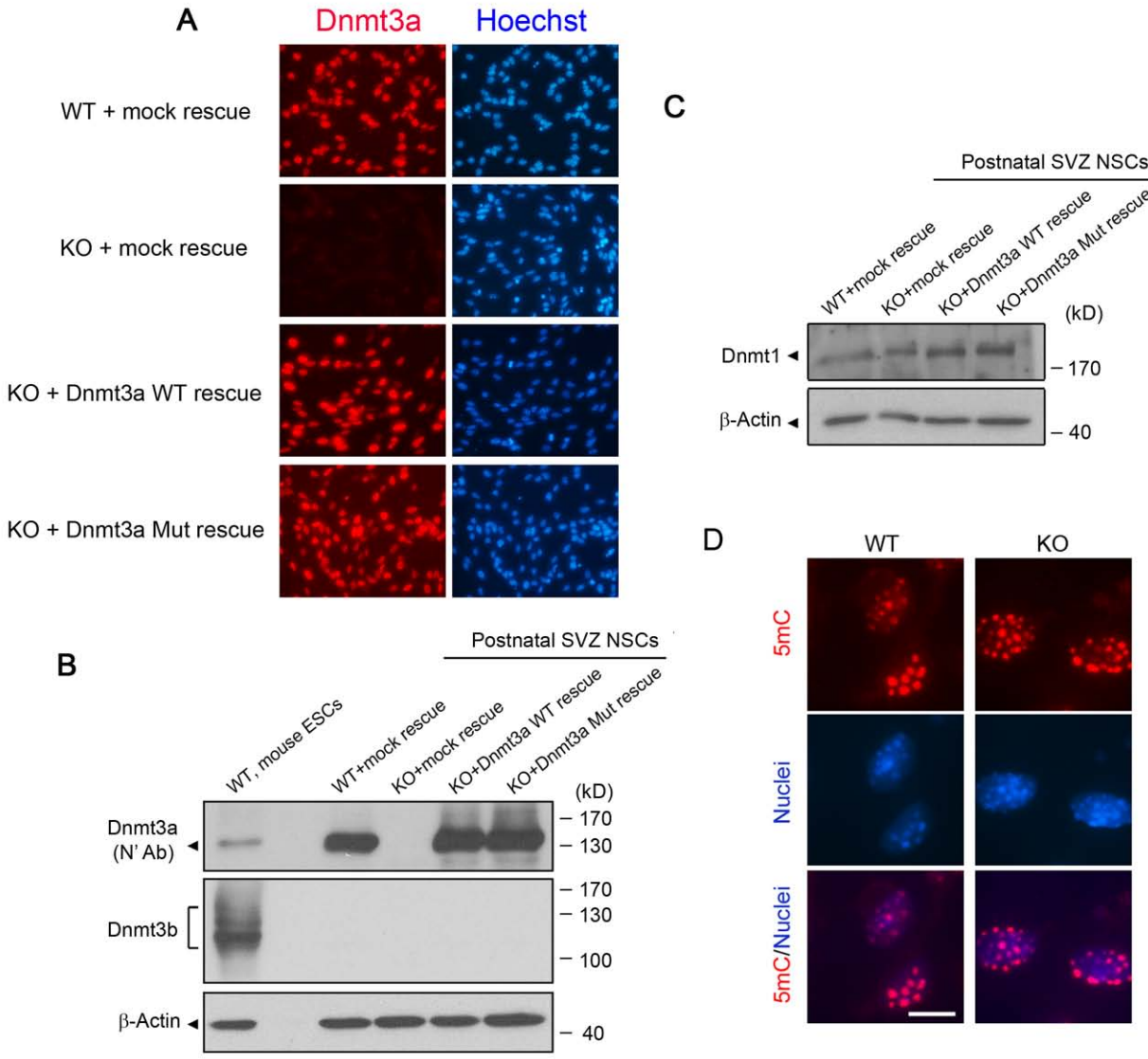

Fig. S6 *Dnmt3a*-deficiency does not alter expression of other DNA methyltransferases (Dnmt3b and Dmnt1) and global DNA methylation patterns in postnatal NSCs.

(A) Immunostaining of Dnmt3a in WT, KO and rescued NSCs.
(B) Immunoblotting analysis of de novo DNA methyltransferases, Dnmt3a and Dnmt3b, in WT, KO and rescued SVZ NSCs. Whole cell lysate from mouse ESCs were used as a positive control for Dnmt3b proteins, which are undetectable in postnatal SVZ NSCs.
(C) Comparable expression levels of Dnmt1 are detected in WT, KO and rescued SVZ NSCs.
(D) Global DNA methylation pattern is not altered in KO NSCs as measured by immunostaining of 5-methylcytosine (5mC) in WT and KO postnatal SVZ NSCs. Nuclear punctate loci that are highlighted by Hoechst 33342 likely represent regions enriched with hypermethylated repetitive DNA sequences. Scale bar, 10 μm.

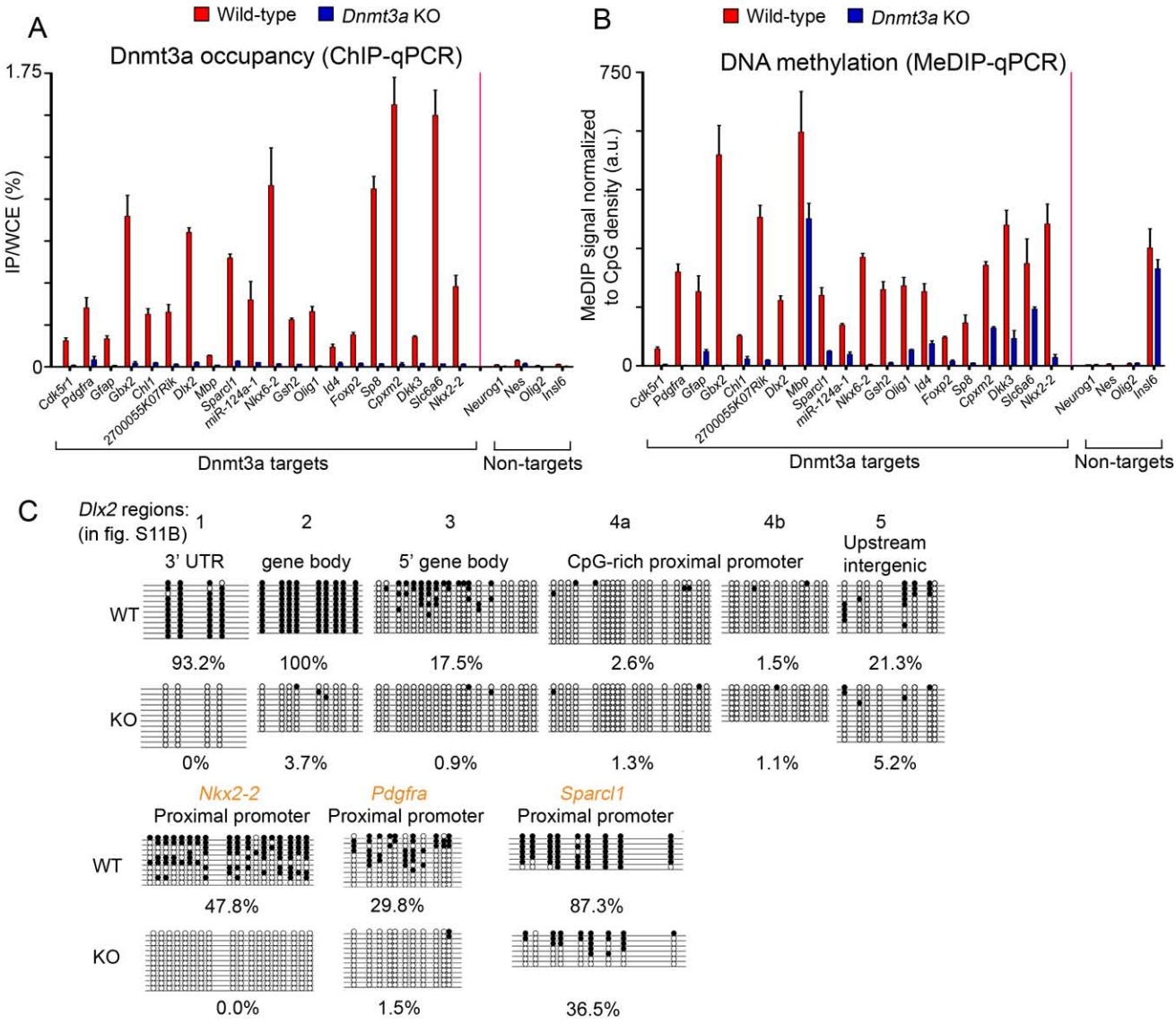

Fig. S7 Locus-specific validation of Dnmt3a ChIP-chip and MeDIP-chip.
(A-B) Quantitative PCR (qPCR) validation of Dnmt3a occupancy (A) and DNA methylation (B) for Dmnt3a bound (n=20) and unbound (n=4) genes in WT and KO NSCs. The MeDIP-qPCR signals (IP/Input, %) are normalized to CpG-density (CpG-per-bp) of PCR amplicon and flanking genomic regions (200bp, up- and down-stream) and normalized DNA methylation levels are shown as arbitrary units (a.u.).
(C) Bisulphite sequencing analyses of selected genomic regions in wild-type (WT) and *Dnmt3a*-null (KO) NSCs. Methylated (filled circles) and unmethylated (open circles) CpG dinucleotides are shown for multiple independently sequenced alleles (horizontal lines). Methylation level is quantified as the percentage of methylated over total CpGs. Results were pooled from two independent NSC cultures.

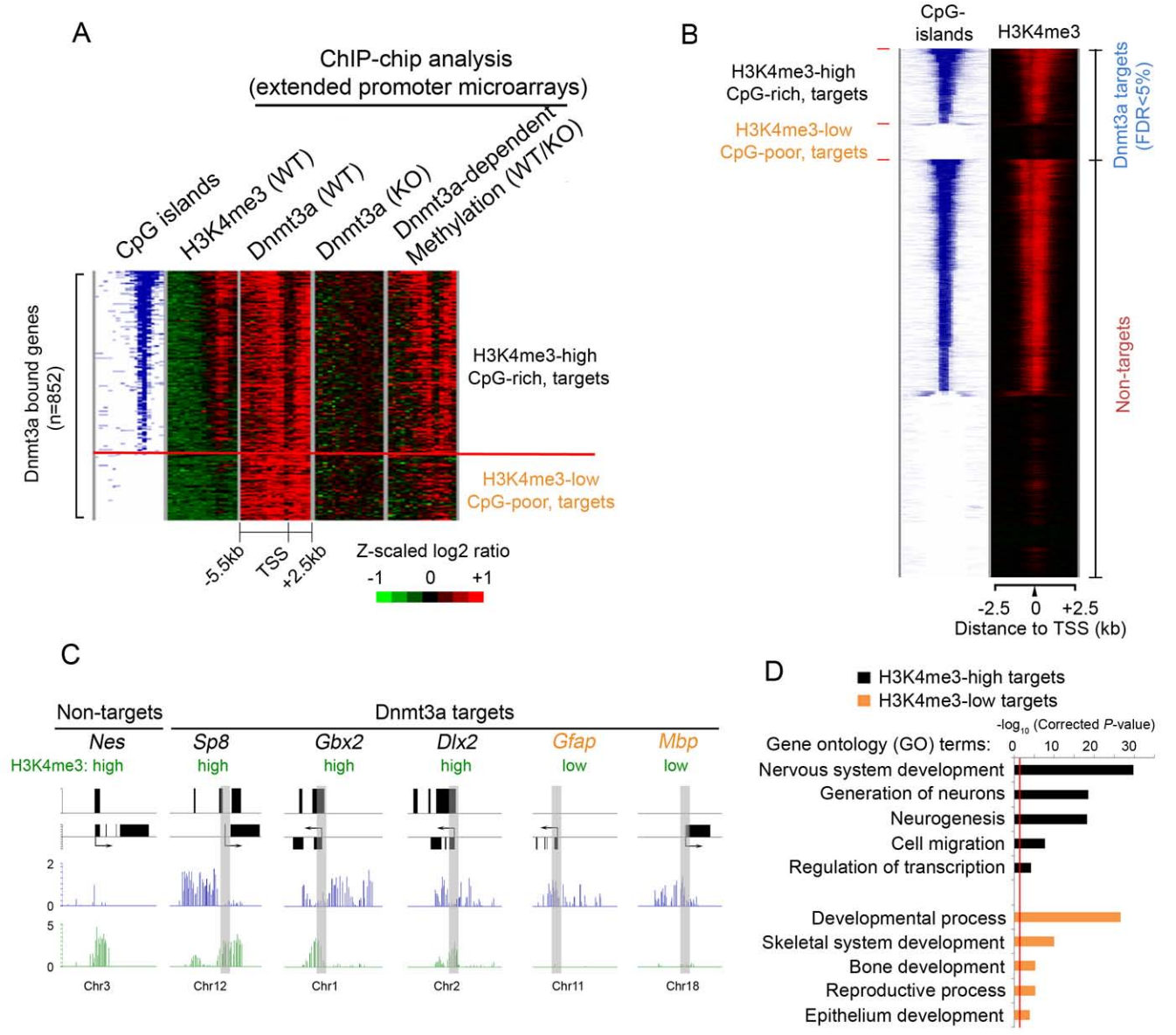

Fig. S8 Genomic analysis of Dnmt3a occupancy and H3K4me3 within extended promoters in SVZ NSCs. (A) ChIP-chip analysis using extended promoter microarrays (-5.5kb to +2.5kb) of Dnmt3a (IP/WCE in WT and KO), H3K4me3 (IP/WCE in WT) and Dnmt3a-dependent DNA methylation (IP/IP, WT/KO). Note that 87.4% of Dnmt3a targets identified by extended promoter arrays are also found in the Dnmt3a target list generated from whole-genome experiment, and the specific Dnmt3a binding sites in WT NSCs is absent in *Dnmt3a*-null NSCs. Heatmap representation are shown for Dnmt3a targets identified by extended promoter DNA microarray, and the relative enrichment for normalized ChIP-chip signals are represented by color intensity. Dnmt3a targets were ranked by the CpG-density of CpG-islands within -2.5kb to +2.5kb promoter regions. (B) The transcriptionally permissive histone mark H3K4me3 is highly enriched on CpG-rich gene promoters. (C) Relative spatial distribution of CpG-islands, Dnmt3a and H3K4me3 within representative Dnmt3a targets with CpG-rich and CpG-poor promoters. Regions shaded in gray indicate 1-kb core promoter regions (-0.5kb to +0.5kb). A non-target, the NSC marker gene Nes/Nestin, is shown as a control. (D) Gene ontology (GO) analysis of CpG-rich/H3K4me3high (black) and CpG-poor/H3K4me3low (orange) Dnmt3a targets.

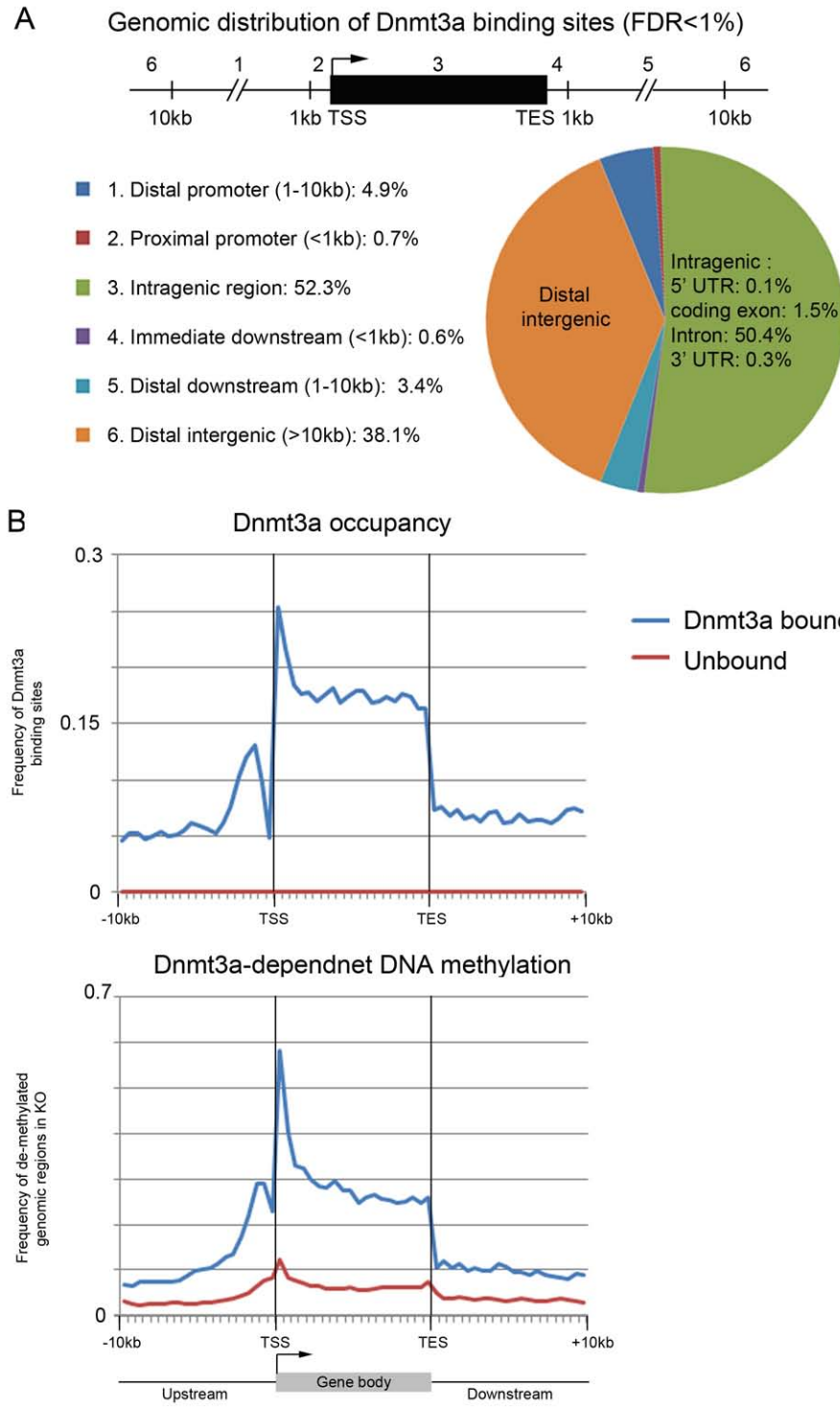

Fig. S9 Genomic analysis of Dnmt3a occupancy and Dnmt3a-dependent DNA methylation in SVZ NSCs. (A) Genomic distribution of Dnmt3a bound sites (FDR<1%) relative to UCSC RefSeq genes (mm8). (B) Averaged peak densities (the percentage of statistically significant peaks) of Dnmt3a (FDR<1%, mapped to 3,271 Refseq genes, mm8) and Dnmt3a-dependent DNA methylation (KS>2) are binned into 500bp windows across genomic regions associated with UCSC Refseq genes (mm8). TSS, transcription start site; TES, transcription end site. The gray bar represents the full-length gene body of all annotated RefSeq genes (n=16,946).

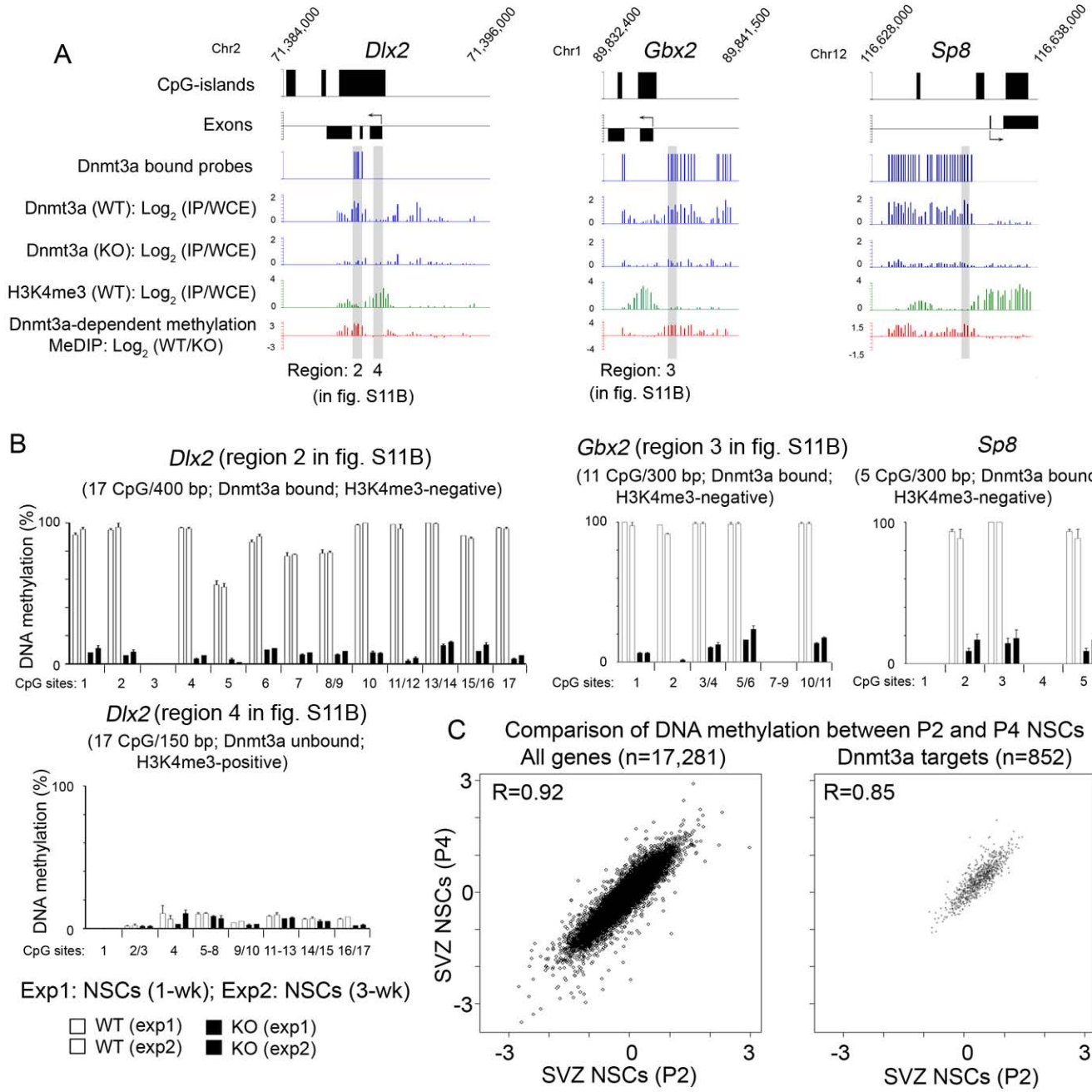

Fig. S10 Limited in vitro culturing does not alter DNA methylation levels on Dnmt3a targets.
(A) CpG-islands, annotated exons (UCSC Refseq, mm7), Dnmt3a bound probes, and raw enrichment (log2 ratios) of Dnmt3a, H3K4me3, and DNA methylation in NSCs at passage (P) 4 are shown for three Dnmt3a targets. The Dnmt3a bound regions analyzed by MALDI-TOF MS (shown in B) are shaded in gray.
(B) Using bisulphite conversion followed by MALDI-TOF MS (Sequenom), DNA methylation levels were determined for three Dnmt3a bound regions (shown in A) with different CpG densities in WT and KO NSCs with short-term (exp1: 1-week/P2) or extended (exp2: 3-week/P4) in vitro culturing. The DNA methylation level of individual CpG site was quantified by averaging two replicate measurements of the same sample. Closely clustered CpG sites are reported as a single value. CpG sites not analyzed are shown as missing in histograms.
(C) MeDIP-chip analyses of genomic DNA extracted from NSCs at P2 (1-week) or P4 (3-week). The promoter DNA methylation level (MeDIP signal) of each gene was estimated by averaging the log2 ratio (IP versus total input) of all the probes assigned to the same gene. Scatterplots of promoter DNA methylation between P2 and P4 NSCs are shown for all genes (n=17,281) and for Dnmt3a bound genes (n=852).

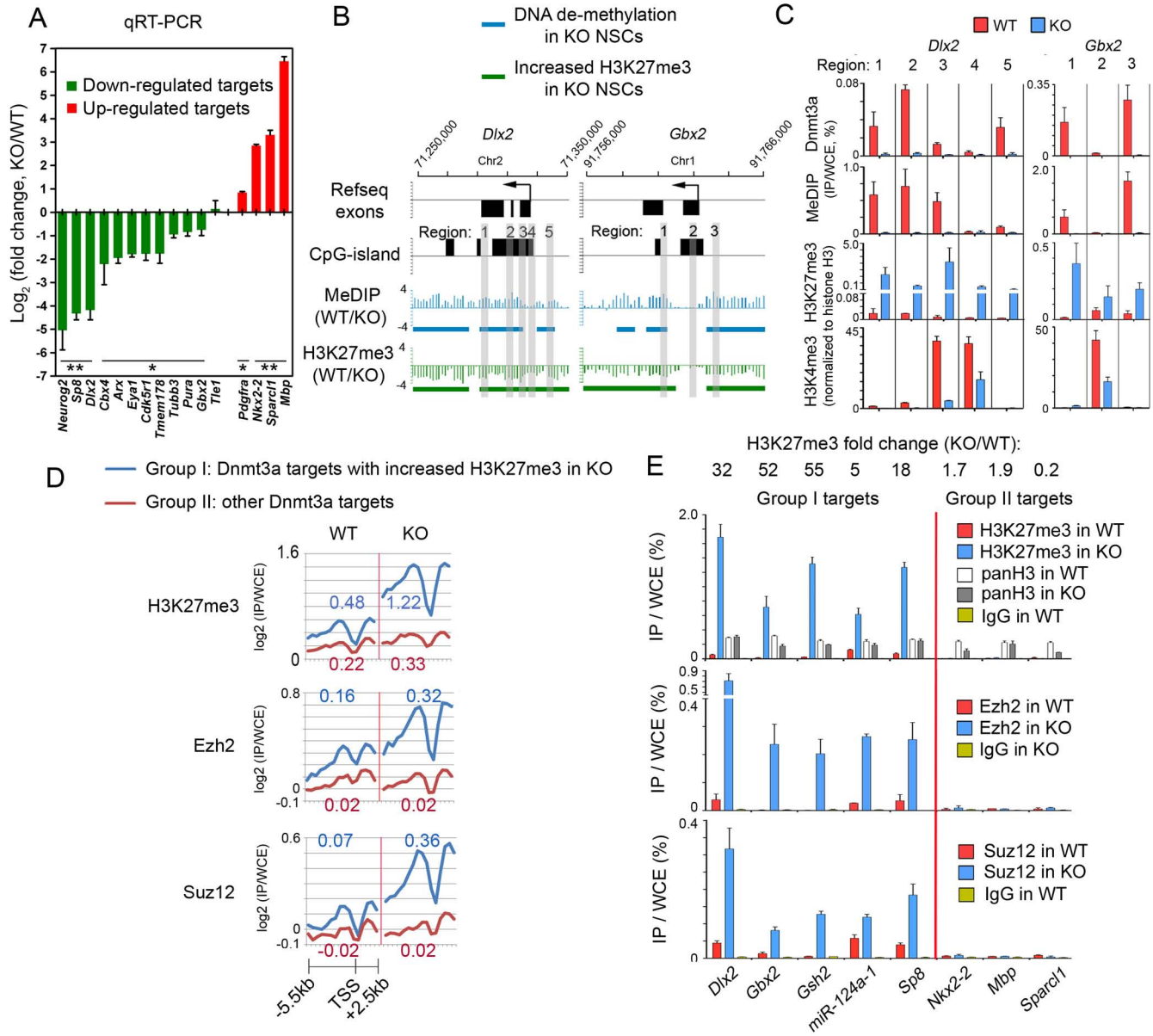

Fig. S11 Dnmt3a antagonizes PRC2/H3K27me3 on Dnmt3a targets down-regulated in KO NSCs.
(A) Quantitative PCR (qPCR) analysis of mRNA levels of 16 Dnmt3a direct targets (identified by expression microarrays as differentially expressed between WT and KO NSCs) in WT and KO NSCs under proliferating conditions. All but one target (Tle1) are confirmed as differentially expressed targets. Error bars represent s.e.m. determined from four independent cultures (*, $P<0.05$; **, $P<0.001$; paired two-tailed Student's t-test).
(B) Relative spatial distribution of relative changes in DNA methylation and H3K27me3 within two representative down-regulated Dnmt3a targets in KO NSCs. Regions examined by locus-specific ChIP/MeDIP-qPCR (C) are shaded in gray.
(C) QPCR analysis of Dnmt3a occupancy, DNA methylation, H3K27me3 and H3K4me3 levels within selected loci of Dnmt3a targets in WT and KO NSCs. Error bars, s.e.m of two experiments.
(D) Distributions of H3K27me3, Ezh2, and Suz12 are shown for these two groups of Dnmt3a targets in WT and KO NSCs. H3K27me3, Ezh2 and Suz12 level increases markedly on a cohort of Dnmt3a targets (Group I: 35.3% of Dnmt3a targets identified by extended promoter microarrays) in *Dnmt3a*-null NSCs. Mean values of log2R (within -2.5 to +2.5kb) in WT and KO NSCs are also shown for two groups of Dnmt3a targets.Note that the Group I targets are also enriched in the down-regulated Dnmt3a targets in KO NSCs.
(E) ChIP-qPCR analysis of H3K27me3, total histone H3 (panH3), Ezh2 and Suz12 occupancy in WT and KO NSCs for representative Group I or II Dnmt3a targets. Mock ChIP using non-specific rabbit IgG is shown as negative controls. Error bars, s.e.m. of three independent experiments.

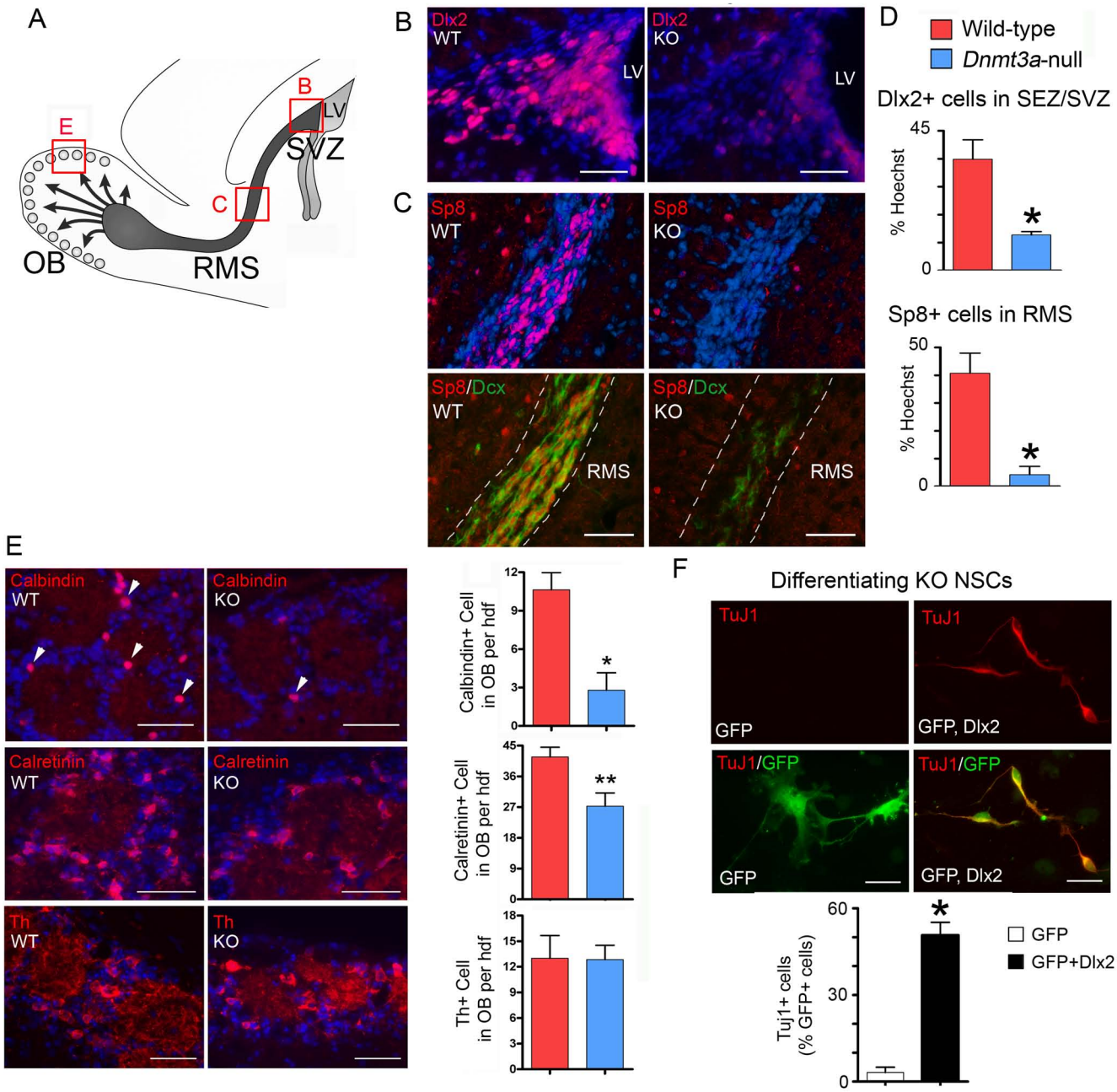

Fig. S12 *Dnmt3a*-deletion impairs expression of neurogenic Dnmt3a targets (Dlx2 and Sp8) and generation of Calbindin+ and Calretinin+ periglomerular interneurons in OB in vivo.

(A) Schematics of olfactory bulb (OB) neurogenesis from postnatal neural stem cells in the SVZ. Red boxes indicate regions shown in (B), (C) and (E).

(B-C) Immunohistochemistry for the Dnmt3a targets Dlx2 (B) and Sp8 (C) in postnatal 3-week sagittal brain sections of WT and KO. Note that Sp8 was double-labeled with Dcx. Scale bars, 50 μm.

(D). Quantification of Dlx2+ cells in SEZ and Sp8+ cells in RMS of WT and KO mice at postnatal 3-week. Error bars, s.e.m. (*, $P<0.05$; $n\geq3$ animals, paired two-tailed Student's *t*-test).

(E) Sample immunofluorescence images of Calbindin+ (paritally Dlx2-dependent), Calretinin+ (paritally Sp8-dependent) and Th+ periglomerular neurons in sagittal sections of P24 OB are shown for WT and KO mice. Scale bars, 50 μm. Quantification of Calbindin+, Calretinin+ and Th+ cells in glomerular layers of WT and KO OB. Error bars, s.e.m. (*, $P=0.0105$; **, $P<0.005$, $n=7$ sections from three pairs of mice).

(F) Quantification of TuJ1+/GFP+ cells after transfection of GFP alone or both pCAG-Dlx2 and GFP expression plasmids into Dnmt3a-null SEZ NSCs. Transfected KO NSCs were differentiated for 3 days before analysis. Error bars, s.e.m. of triplicate experiments (*, $P<0.0001$).

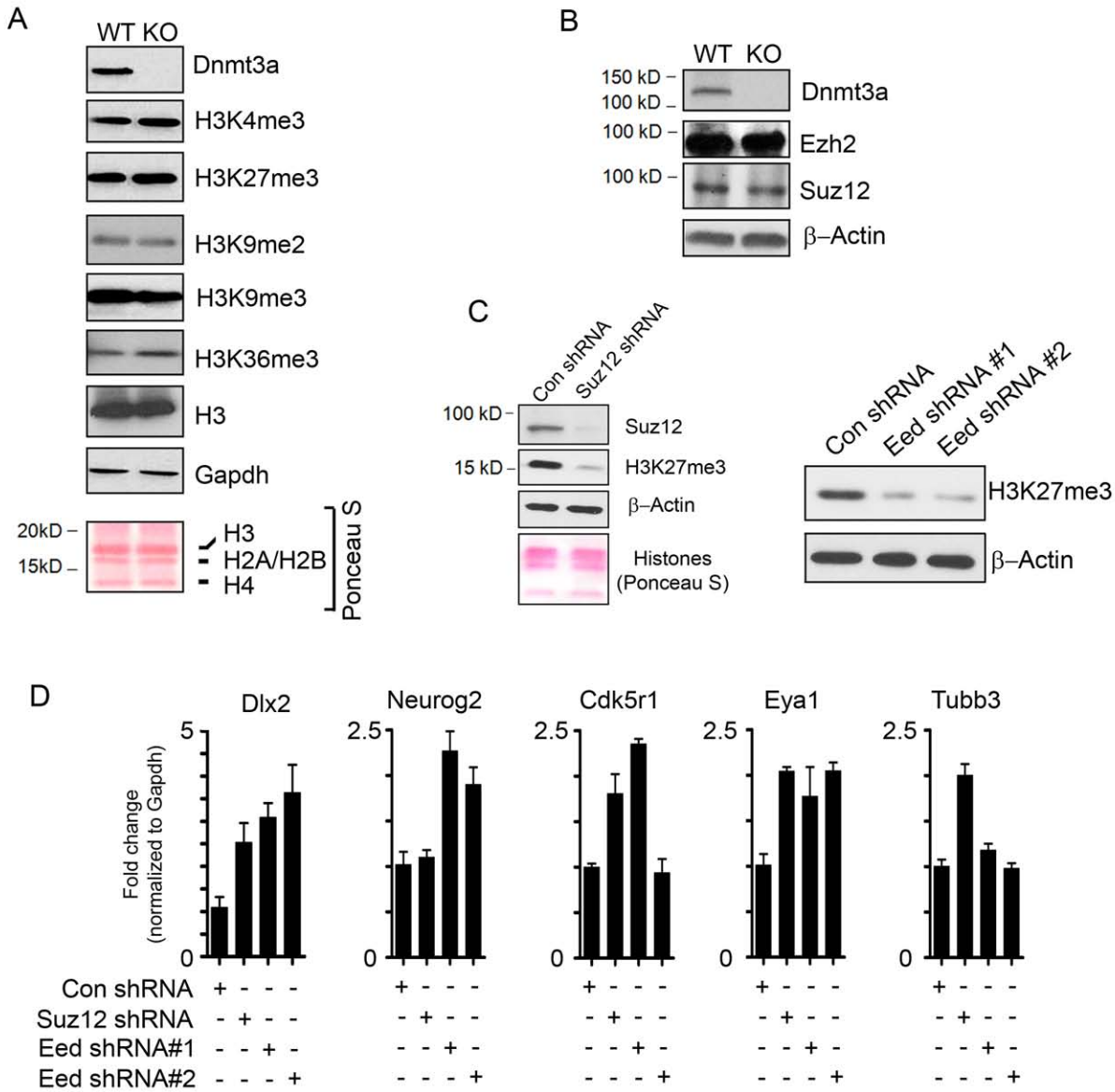

Fig. S13 Inhibiting PRC2 in Dnmt3a-null NSCs partially rescue expression of neurogenic Dnmt3a targets.
(A) Global levels of histone lysine methylation and PRC2 core subunits are not altered in *Dnmt3a*-null postnatal NSCs. Immunoblotting of site-specific histone lysine methylation in WT and KO NSCs. Immunoblotting of total histone H3, house-keeping gene Gapdh and Ponceau S staining of total histones were used as loading controls.
(B) Immunoblotting of core PRC2 subunits (Ezh2 and Suz12) in WT and KO NSCs.
(C) Immunoblotting of PRC2/H3K27me3 in KO NSCs infected with lentiviruses expressing control shRNAs or shRNAs specific for the core PRC2 subunit Suz12 or Eed.
(D) Quantitative RT-PCR analyses of five neurogenic Dnmt3a target in control and PRC2-shRNA infected postnatal *Dnmt3a*-null (KO) NSCs under proliferation conditions.

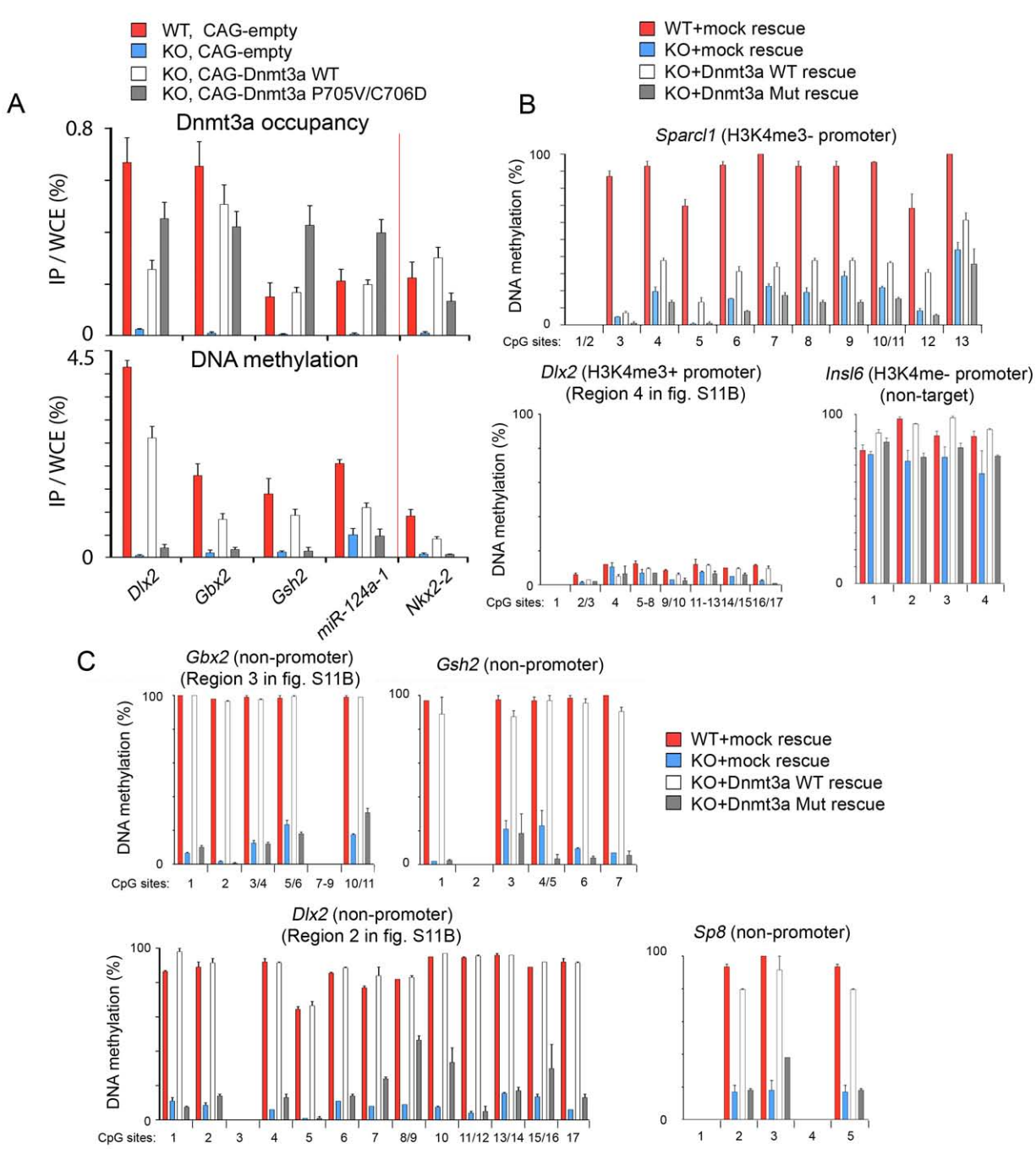

Fig. S14 Charaterization of *Dnmt3a*-null NSCs that were rescued with wild-type or catalytically inactive mutant Dnmt3a.

(A) Quantitative PCR analysis of Dnmt3a and DNA methylation levels in WT, KO, and rescued KO NSCs (transient re-expression of wild-type or catalytically inactive mutant (P705V/C706D) full-length Dnmt3a) are shown for six Dnmt3a targets. Note that *Nkx2-2* is a Dnmt3a target that is not associated with increase in H3K27me3 in KO NSCs.

(B-C) Quantitative DNA methylation analysis by bisulphite Mass Spectrometry (Sequenom MassArray assays) of five Dnmt3a bound regions and two unbound regions (*Dlx2*, region 4: a H3K4me3-high promoter region, and *Insl6*: a sperm-specific gene) in rescued NSCs.

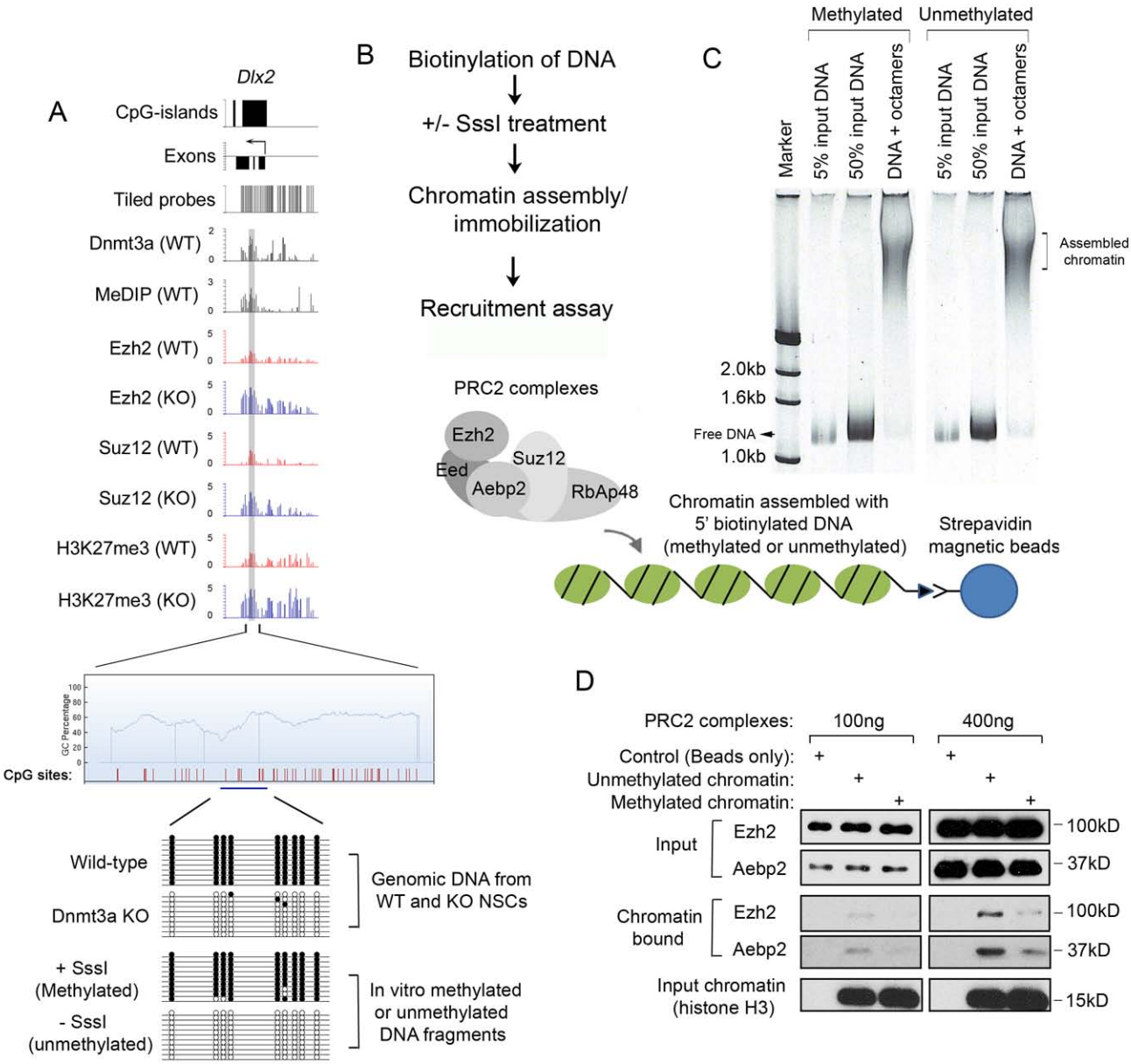

Fig. S15 Recruitment of PRC2 to chromatin is partially inhibited by DNA methylation *in vitro*.
(A) Based on ChIP-chip and MeDIP-chip results, a 1.1 kb genomic region containing Dnmt3a binding sites within the *Dlx2* gene locus was cloned, purified and incubated with or without CpG DNA methylase SssI *in vitro*. CpG sites within the 1.1 kb region were shown as red vertical bars. SssI treatment was confirmed by bisulphite sequencing of a sub-region within the 1.1kb region (blue horizontal bar). As a reference, bisulphite sequencing results of the same region in wild-type and Dnmt3a KO NSCs were shown.
(B) Diagram of experimental design of *in vitro* PRC2 repression complex recruitment assay.
(C) Both methylated and unmethylated DNA (5' biotinylated) can be efficiently assembled into chromatin arrays (shifted smear) using recombinant *Xenopus* octamers.
(D) Western blotting of bound fraction of PRC2 subunits on methylated or unmethylated chromatin arrays. Equal ammount of PRC2 complexes (100ng or 400ng) was incubated with only beads (without chromatin), and beads with immobilized unmethylated or methylated chromatin (100ng). After 30 min incubation, the beads were stringently washed and resuspended in the SDS sampling buffer for SDS-PAGE. The chromatin bound fraction of PRC2 complexes was detected by immunoblotting of Aebp2 (Flag-tagged) and Ezh2. The relative binding was quantifed by densitometry (in Fig. 4C).

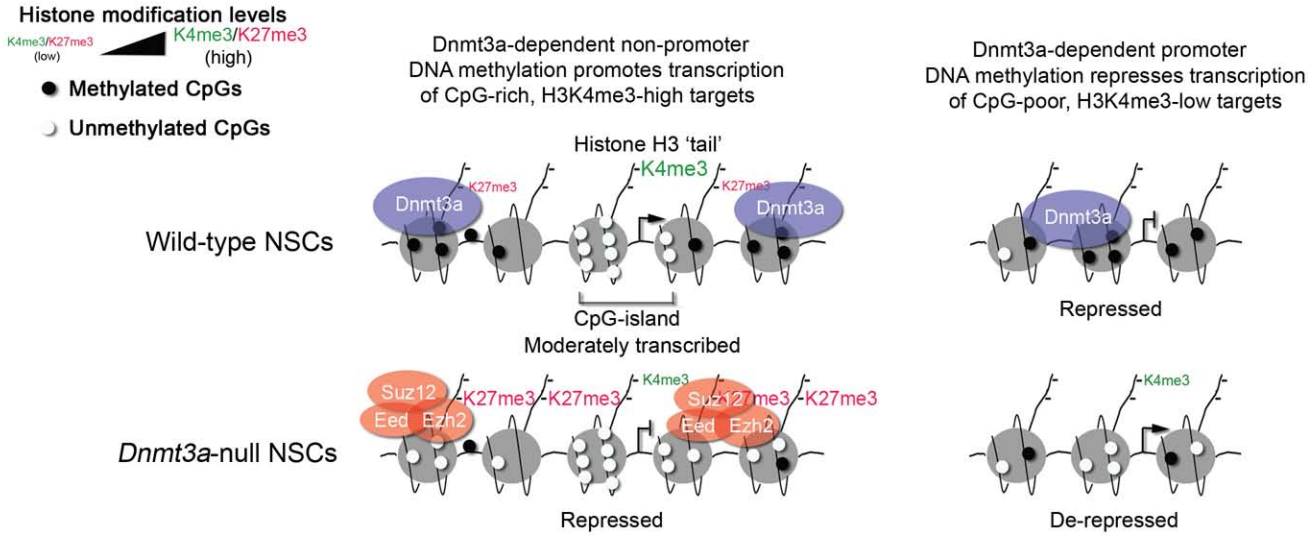

Fig. S16 A proposed model of Dnmt3a-dependent non-proximal promoter DNA methylation in antagonizing Polycomb repression and promoting transcription in postnatal neural stem cells.

In multipotent postnatal NSCs, antagonism between Dnmt3a-mediated DNA methylation and PRC2-mediated repressive histone modification H3K27me3 primarily takes place on Dnmt3a targets with CpG-island containing gene promoters, many of which encode regulators of neurogenesis.
The font size of 'H3K4me3' (green) or 'H3K27me3' (red) represents relative levels of indicated histone modifications. Black circles represent methylated CpG dinucleotides, whereas white circles represent unmethylated CpG dinucleotides.

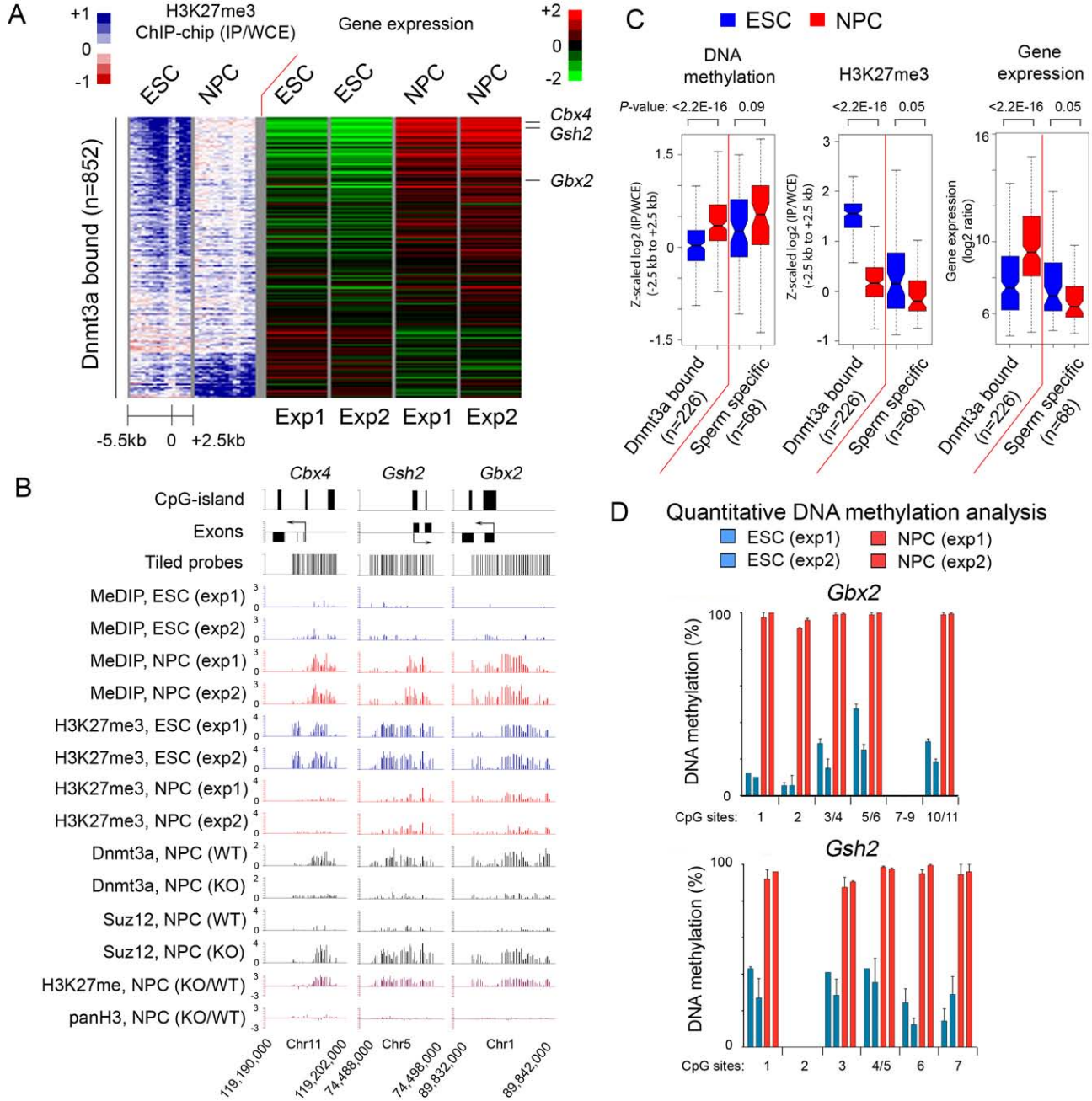

Fig. S17 Reduction in H3K27me3 levels on many genes (PcG targets in ESCs, Dnmt3a targets in NPCs) during the conversion of ESCs to NPCs is accompanied by *de novo* DNA methylation and a concomitant increase in basal transcriptional levels.

(A) Composite heatmaps of H3K27me3 and gene expression profiling in ESCs and NPCs are shown for Dnmt3a targets in NPCs (n=852). Genes were ranked by the difference (Z-scaled log2 ratios) in H3K27me3 between ESCs and NPCs.

(B) Raw enrichment of DNA methylation and H3K27me3 in ESCs and NPCs with indicated genotypes are shown for three Dnmt3a targets in NPCs.

(C) Reduction in H3K27me3 on many Dnmt3a targets (in NPCs) during ESC differentiation is accompanied by *de novo* DNA methylation and a concomitant increase in basal transcriptional levels. Boxplots of the normalized levels of DNA methylation, H3K27me3, and gene expression in ESCs and NPCs are shown for Dnmt3a targets associated with a significant loss of H3K27me3 from ESCs to NPCs (n=226). A group of sperm-specific genes is also shown as positive controls for DNA methylaiton (n=68).

(D) Sequenom MassArray DNA methylation analyses of two Dnmt3a targets during ESC to NPC conversion.

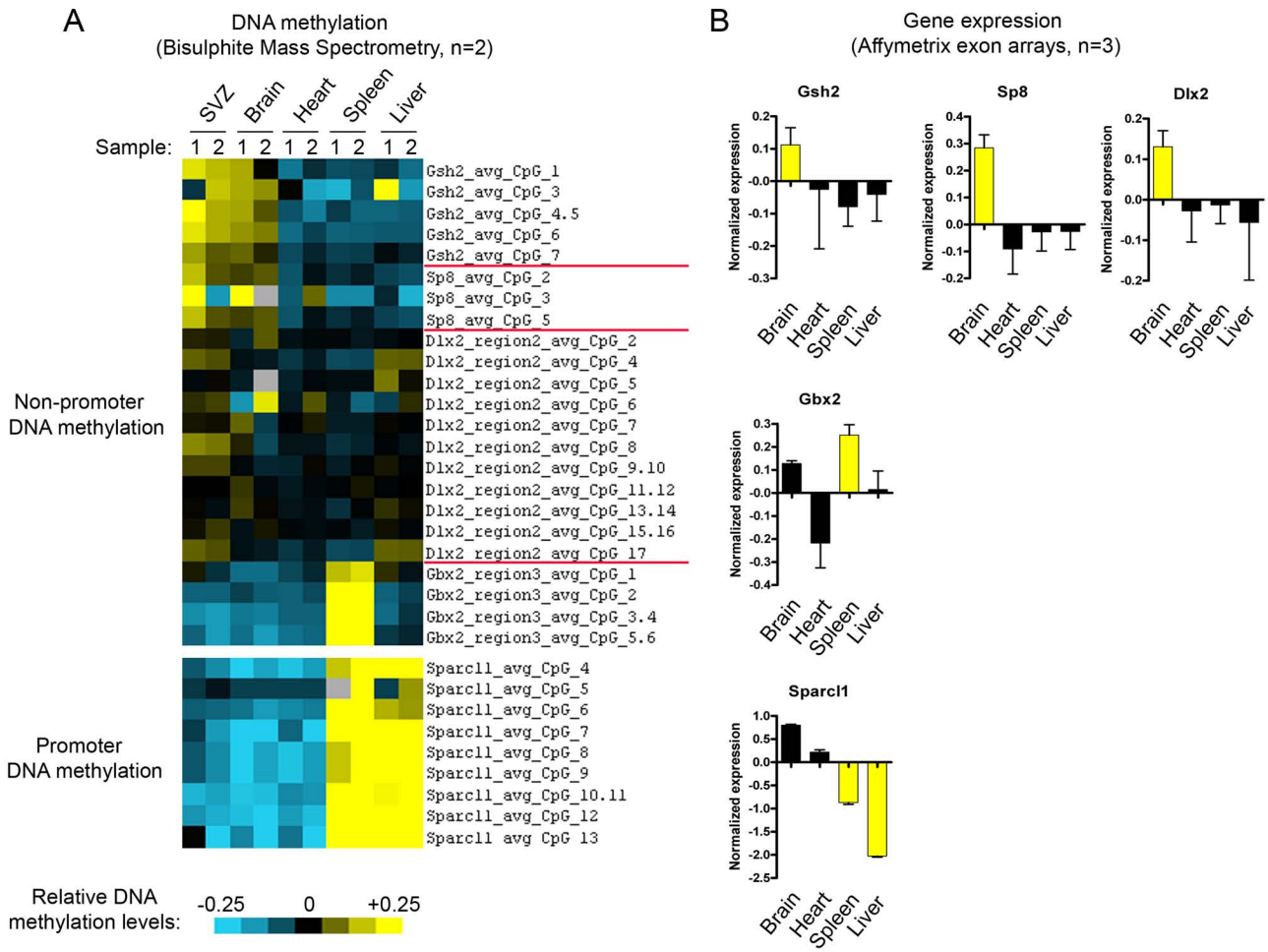

Fig. S18 Non-promoter DNA methylation positively correlates with tissue-specific transcription in vivo.

(A) Heatmap representation of quantitative DNA methylation analysis by Sequenom MassArray analyses of five Dnmt3a bound regions (amplicons in fig. S14B-C; 4 non-promoter: *Dlx2, Gsh2, Sp8* and *Gbx2*; one promoter: *Sparcl1*) in five postnatal tissues (SVZ at postnatal day (P) 18; brain, heart, spleen and liver at P45). The relative methylation levels was represented by color intensity (mean-centered values). Yellow color represents higher methylation levels, whereas blue represents lower methylation levels. Missing values are in gray. Each row represents a CpG unit (single or closely clustered CpG sites) measured by Sequenom bisulphite Mass Spectrometry assays. Each column represents a tissue sample. Two biological independent samples (one male and one female wild-type littermates) were measured in duplicate for each tissue. While promoter DNA methylation at Sparcl1 is negatively correlated with gene expression, non-promoter DNA methylation at other genes is generally associated with higher levels tissue-specific transcription.
(B) Relative gene expression data (in triplicates) in adult mouse brain, heart, spleen and liver are retrieved from a published database via UCSC Table browser (A. Pohl, *et al.*, *Bioinformatics* **25**, 2442 (2009)). The gene expression was measured by Affymetrix Mouse Exon 1.0 ST arrays. The quantile normalized data were further converted to median-centered log-ratios. The expression of each target was estimated by averaging the expression values of all anti-sense exon probes covered by the gene body. The tissues associated with relatively high level of DNA methylation are highlighted in yellow.

Table S1 Gene ontology (GO) analysis of differentially expressed genes between differentiating WT and KO SEZ cells (DAVID http://david.abcc.ncifcrf.gov)

Note: P-value were corrected for multiple testing and the cutoff for significantly enriched GO term was P-value<0.01. The total number of genes included in GO analysis is 14,977.

Down-regulated genes in differentiating KO SEZ cells

GO Term	% Total	Enrichment	Corrected P-value
GO:0007399~nervous system development	8.64%	2.58	1.34E-15
GO:0007275~multicellular organismal development	18.19%	1.61	1.85E-10
GO:0030182~neuron differentiation	4.65%	3.08	1.80E-10
GO:0048666~neuron development	3.99%	3.34	4.46E-10
GO:0032502~developmental process	22.26%	1.48	1.16E-09
GO:0022008~neurogenesis	4.98%	2.69	4.38E-09
GO:0048699~generation of neurons	4.73%	2.76	4.63E-09
GO:0048731~system development	14.12%	1.62	4.26E-08
GO:0030154~cell differentiation	14.70%	1.60	4.35E-08
GO:0048869~cellular developmental process	14.70%	1.60	4.35E-08
GO:0031175~neurite development	3.41%	3.22	3.69E-08
GO:0048856~anatomical structure development	15.95%	1.56	3.90E-08
GO:0065007~biological regulation	29.73%	1.30	5.67E-07
GO:0030030~cell projection organization and biogenesis	3.90%	2.65	7.68E-07
GO:0048858~cell projection morphogenesis	3.90%	2.65	7.68E-07
GO:0032990~cell part morphogenesis	3.90%	2.65	7.68E-07
GO:0050789~regulation of biological process	27.41%	1.31	1.10E-06
GO:0007409~axonogenesis	2.74%	3.24	1.67E-06
GO:0007156~homophilic cell adhesion	2.24%	3.73	2.45E-06
GO:0048667~neuron morphogenesis during differentiation	2.74%	3.05	7.26E-06
GO:0048812~neurite morphogenesis	2.74%	3.05	7.26E-06
GO:0016337~cell-cell adhesion	3.41%	2.61	1.04E-05
GO:0007268~synaptic transmission	2.91%	2.77	2.49E-05
GO:0019226~transmission of nerve impulse	3.24%	2.57	3.21E-05
GO:0000904~cellular morphogenesis during differentiation	2.82%	2.79	3.20E-05
GO:0050794~regulation of cellular process	24.42%	1.29	6.35E-05
GO:0007267~cell-cell signaling	4.07%	2.20	7.08E-05
GO:0048519~negative regulation of biological process	8.31%	1.62	2.33E-04
GO:0048523~negative regulation of cellular process	7.89%	1.65	2.28E-04
GO:0019222~regulation of metabolic process	17.77%	1.34	3.37E-04
GO:0048468~cell development	9.39%	1.53	6.65E-04
GO:0031323~regulation of cellular metabolic process	17.03%	1.33	8.31E-04
GO:0032989~cellular structure morphogenesis	4.90%	1.86	9.27E-04
GO:0000902~cell morphogenesis	4.90%	1.86	9.27E-04
GO:0009653~anatomical structure morphogenesis	9.05%	1.53	9.32E-04
GO:0010468~regulation of gene expression	16.45%	1.34	9.93E-04
GO:0016126~sterol biosynthetic process	0.91%	5.93	0.001085264
GO:0022610~biological adhesion	5.81%	1.72	0.001249485
GO:0007155~cell adhesion	5.81%	1.72	0.001249485
GO:0003001~generation of a signal involved in cell-cell signaling	1.58%	3.18	0.003020177
GO:0045449~regulation of transcription	15.28%	1.33	0.003401544

GO:0009892~negative regulation of metabolic process	3.82%	1.93	0.003472134
GO:0006355~regulation of transcription, DNA-dependent	14.62%	1.33	0.003775592
GO:0048489~synaptic vesicle transport	0.83%	5.76	0.003713862
GO:0032501~multicellular organismal process	25.08%	1.22	0.003954148
GO:0007269~neurotransmitter secretion	1.16%	3.97	0.00397036
GO:0006351~transcription, DNA-dependent	14.70%	1.33	0.004286769
GO:0032774~RNA biosynthetic process	14.70%	1.32	0.004838642
GO:0001505~regulation of neurotransmitter levels	1.50%	3.10	0.005750093
GO:0007411~axon guidance	1.41%	3.23	0.005848214

Up-regulated genes in differentiating KO SEZ cells

GO Term	% Total	Enrichment	Corrected P-value
GO:0007049~cell cycle	7.99%	2.17	7.53E-15
GO:0048856~anatomical structure development	16.70%	1.63	8.25E-15
GO:0048731~system development	14.55%	1.67	4.10E-14
GO:0032502~developmental process	22.06%	1.47	1.44E-13
GO:0022402~cell cycle process	6.62%	2.23	6.92E-13
GO:0007275~multicellular organismal development	17.29%	1.53	3.65E-12
GO:0022403~cell cycle phase	4.17%	2.74	8.15E-12
GO:0000279~M phase	3.70%	2.90	1.85E-11
GO:0048513~organ development	11.69%	1.66	1.67E-10
GO:0051301~cell division	3.34%	2.73	3.88E-09
GO:0000278~mitotic cell cycle	3.46%	2.67	4.14E-09
GO:0007067~mitosis	2.86%	3.00	4.18E-09
GO:0000087~M phase of mitotic cell cycle	2.86%	2.98	4.70E-09
GO:0009653~anatomical structure morphogenesis	9.78%	1.65	2.21E-08
GO:0006950~response to stress	7.10%	1.70	2.47E-06
GO:0000074~regulation of progression through cell cycle	3.88%	2.10	4.95E-06
GO:0051726~regulation of cell cycle	3.88%	2.08	6.41E-06
GO:0048869~cellular developmental process	13.00%	1.42	9.82E-06
GO:0030154~cell differentiation	13.00%	1.42	9.82E-06
GO:0009887~organ morphogenesis	4.83%	1.84	2.50E-05
GO:0065007~biological regulation	27.79%	1.21	3.98E-05
GO:0048514~blood vessel morphogenesis	2.39%	2.47	4.06E-05
GO:0048468~cell development	9.06%	1.48	1.02E-04
GO:0048522~positive regulation of cellular process	7.04%	1.58	1.03E-04
GO:0001568~blood vessel development	2.56%	2.29	1.04E-04
GO:0048518~positive regulation of biological process	7.81%	1.53	1.06E-04
GO:0001525~angiogenesis	1.97%	2.60	1.39E-04
GO:0001944~vasculature development	2.56%	2.26	1.40E-04
GO:0050789~regulation of biological process	25.34%	1.21	1.82E-04
GO:0048646~anatomical structure formation	2.33%	2.28	3.58E-04
GO:0022008~neurogenesis	3.52%	1.90	3.96E-04
GO:0008283~cell proliferation	4.53%	1.69	9.30E-04
GO:0009966~regulation of signal transduction	3.94%	1.76	0.001108842
GO:0007399~nervous system development	5.37%	1.60	0.001129143
GO:0006629~lipid metabolic process	5.25%	1.61	0.001168061
GO:0048699~generation of neurons	3.22%	1.88	0.001301783
GO:0032787~monocarboxylic acid metabolic process	2.33%	2.05	0.003762921
GO:0007010~cytoskeleton organization and biogenesis	4.35%	1.63	0.004565965

GO:0007059~chromosome segregation	0.95%	3.31	0.007267371
GO:0001501~skeletal development	1.97%	2.12	0.008465342
GO:0048519~negative regulation of biological process	7.22%	1.41	0.008892936
GO:0009611~response to wounding	2.86%	1.81	0.008990888
GO:0050793~regulation of developmental process	2.74%	1.84	0.009077248
GO:0006468~protein amino acid phosphorylation	5.01%	1.53	0.009810011

Table S2 Primer sequences

ChIP/MeDIP-qPCR primers

2700055K07Rik_ChIP_F	CTGTCCAGAGACCTTTCCTAACCC
2700055K07Rik_ChIP_R	TACCCAAGACTCAGCAATATCCGC
Cdk5r1_ChIP_F	GTGTCAGCGAAGAAGAAGAACTCC
Cdk5r1_ChIP_R	TCATTGTTGAGATGCGCGATGTTG
Chl1_ChIP_F	ACCAGACAGATCGGATGTCAAGAG
Chl1_ChIP_R	TTTCCTTCCACCTCCCTCAAAGAG
Cpxm2_ChIP_F	TGCACCAGTATTCTATGGCTCCTC
Cpxm2_ChIP_R	TCTCACCACAAATGAGTCCCTCAG
Dlx2 (region 1)_ChIP_F	AAACATAGGGACTGCTGAGGTC
Dlx2 (region 1)_ChIP_R	GCAGAGTTTGTAAATAAGGGTGTC
Dlx2 (region 2)_ChIP_F	ATGTGACCAAGGTTCTGGGTAGTG
Dlx2 (region 2)_ChIP_R	CATGGGAGGAAAGAAAATGACCCG
Dlx2 (region 3)_ChIP_F	CGGAGTAGGAGACATTGTTGAG
Dlx2 (region 3)_ChIP_R	TACTACACCAACCAGCAGCAC
Dlx2 (region 4)_ChIP_F	GAGCAGCTTTACGATTGTCTG
Dlx2 (region 4)_ChIP_R	ATTCCTTCTGATTGGCTGTTG
Dlx2 (region 5)_ChIP_F	GACAGTGTAGGAGGTTGTTACAGG
Dlx2 (region 5)_ChIP_R	GTTATGTCTTTAGGAAGGGATGC
Dkk3_ChIP_F	GGAAATCAACCTATTCCAGCAGGC
Dkk3_ChIP_R	CAAGTGGGTTGGGTGTTTACCATC
Foxp2_ChIP_F	TCTGGAGGCACTAAAACAGGAGAC
Foxp2_ChIP_R	AACTTGCTGGAGACAAAACAGCAG
Gbx2 (region 1)_ChIP_F	CCAGCAGATTGTCTTTCTACCAC
Gbx2 (region 1)_ChIP_R	TCTACTGTCTCCTCCTCTCCTTTG
Gbx2 (region 2)_ChIP_F	TCTATGCTGAAGGCGGTACTACTC
Gbx2 (region 2)_ChIP_R	TCGTCTCTCCCTGGCTTCTC
Gbx2 (region 3)_ChIP_F	GGGGTTAGGAGATTGGAGGGATTG
Gbx2 (region 3)_ChIP_R	CAAGGAATCTGAAGAACAGGCAGG
Gfap_ChIP_F	AGATTTAGTCCAACCCGTTCCTC
Gfap_ChIP_R	GGAGTCATTCGAGACAAGGAGAAG
Gsh2_ChIP_F	CTGGAGTGCTTGGTTTCTAAGCTG
Gsh2_ChIP_R	CACCACATTCATATCTGCCCTTGC
Id4_ChIP_F	TCAGTGAAAATGAAGGCAGCAGTC
Id4_ChIP_R	AAACAACCAAGCTCTTCAGATGGC
Insl6_ChIP_F	TTATTTGCATGCCTAATTTTGAGTG
Insl6_ChIP_R	CCTTTGTAGGAAAATGGATACAAGTG
Mbp_ChIP_F	GGAAATGCTTTGGCTTGATTGCTG
Mbp_ChIP_R	CTAACCTGGATTGAGCTTGCTTGG
miR-124a-1_ChIP_F	TATGAAAGTGGTCAGTGCTCCCAG
miR-124a-1_ChIP_R	ATTGAGAATCTTAGGAGGGGTGGG
Nes_ChIP_F	TATGAATACCCTCGCTTCAGCTCG

Nes_ChIP_R	CGGACGTGGAGCACTAGAGAAG
Neurog1_ChIP_F	GCCGTACTTAAGGGGTCCTG
Neurog1_ChIP_R	AGGGGCTCAGGGAGTGAAT
Nkx2-2_ChIP_F	GGACTTCTTTGAGGATCAGGGGAG
Nkx2-2_ChIP_R	AAGGAAATGGACAGGGTCTAGGTG
Nkx6-2_ChIP_F	GAAATCCTTGAAGTTGGGTTGGGG
Nkx6-2_ChIP_R	TTGCCAGAAAGCAATAAGCAGGAG
Olig1_ChIP_F	AGCAAGTGTGTTCATAGCAAACGG
Olig1_ChIP_R	CAGAACAGCAGATGTGGGTCCTC
Olig2_ChIP_F	GGGTTTCATTGAGCGGAATTAG
Olig2_ChIP_R	CTCGGTCTGTAATAAGCATCCAC
Pdgfra_ChIP_F	TAAGTGGCTCCGAAGGGATAAAGG
Pdgfra_ChIP_R	TTTTGAAAGGCTGGATTGTGAGGC
Slc6a6_ChIP_F	TAGGACAATGGATTCAGGCTGGTC
Slc6a6_ChIP_R	TGCTTGCCCACTTCCTTTTCTTTC
Sp8_ChIP_F	GCTTGCTCAAACCATAGATGCTCC
Sp8_ChIP_R	TTTTCTCAGGTAACCCCACTGTGC
Sparcl1_ChIP_F	GTGTAAGCATTTGGCTCTTCATGC
Sparcl1_ChIP_R	TGATCCTTCTGTCCTTTCACCCTG

RT-qPCR primers

Arx_RT_F	ATCACCCTGCGCTTGATTCGGCCTG
Arx_RT_R	AAGAGCCTGCCAAATGCTGGGCTG
Cbx4_RT_F	AAAGCCGAAGTGCCCTGCAAACGC
Cbx4_RT_R	TCTGGCTTGGTGGTGAGCTGTAGCG
Cdk5r1_RT_F	CCCTTCCTGGTAGAGAGCTG
Cdk5r1_RT_R	GTGAAATAGTGTGGGTCGGC
Dlx2_RT_F	CAAAAGCAGCTACGACCTGG
Dlx2_RT_R	CGGACTTTCTTTGGCTTCCC
Eya1_RT_F	AACAACGTGGGAGGTCTGCTTGGC
Eya1_RT_R	TTGCCAATGCTGGGATGAGCTGCG
Gapdh_RT_F	CTGAGTTCGTGGAGTCTACTGG
Gapdh_RT_R	GTCATATTTCTCGTGGTTCACACC
Gbx2_RT_F	CAAACTCAGCGAGGTGCAAG
Gbx2_RT_R	AATCTTGGGGTTCCGAGAGG
Gfap_RT_F	CAGGAGTACCACGATCTACTCAAC
Gfap_RT_R	GGAGAAAGTCTGTACAGGAATGGT
Mbp_RT_F	ACAGAAGAGACCCTCACAGC
Mbp_RT_R	CTGTCACCGCTAAAGAAGCG
Neurog2_RT_F	TCAAGAAGACCCGCAGGCTCAAGGC
Neurog2_RT_R	GATGTAATTGTGGGCGAAGCGCAGC
Nkx2-2_RT_F	GACACCAACGATGAAGACGG
Nkx2-2_RT_R	TGCTGTCGTAGAAAGGGCTC

Pura_RT_F	TGGGCCAAGTTCGGACACACCTTC
Pura_RT_R	TTGCTGCTGCTGTTGCTGCTGGTG
Sp8_RT_F	ATAGCTCCCATCAGCCAGTG
Sp8_RT_R	AACCAGGACTCATACGGGTG
Sparcl1_RT_F	TGCAAAACCGATCCACAAGG
Sparcl1_RT_R	CTGGCGTAGGTTTGGTTGTC
Teme178_RT_F	TCACGGCCATCTTCACCGACCACTG
Teme178_RT_R	ACAGCGGCATCAGGCGGTTCTTCTG
Tle1_RT_F	AATTCCAGGGCCACACAGACGGAGC
Tle1_RT_R	TTCCATGCCCACAGCAAGCCACTC
Tubb3_RT_F	ATACTACAATGAGGCCTCCTCTCAC
Tubb3_RT_R	GCACCACTCTGACCAAAGATAAAG

Bisulphite sequencing primers

Dlx2 (region 1)_BS-Seq_F	TTTTGTATTTTAAGAATAAAGATTTTAGGATAAG
Dlx2 (region 1)_BS-Seq_R	ATAAAAATATCTATACAAATTTCCCCACAT
Dlx2 (region 2)_BS-Seq_F	TAAGTATTGGGTTTAGAAGTTTATTTTGTT
Dlx2 (region 2)_BS-Seq_R	ACCATCTACTCCAATTTCCAACTAAC
Dlx2 (region 3)_BS-Seq_F	GGATTGTTTTTTATTTTATGTTTATT
Dlx2 (region 3)_BS-Seq_R	ACAACTACTACACCAACCAACAACA
Dlx2 (region 4a)_BS-Seq_F	TGTTGTTGGTTGGTGTAGTAGTTGT
Dlx2 (region 4a)_BS-Seq_R	TAATTACATTAATTACTAAAAAAAACCC
Dlx2 (region 4b)_BS-Seq_F	TTTAGTAATAGTTAATTAGAAGGAATGAGT
Dlx2 (region 4b)_BS-Seq_R	CCCACTTTCTACAATACTTATACAC
Dlx2 (region 5)_BS-Seq_F	TTTTTTTTAAGTGATGAGTTTTTTTT
Dlx2 (region 5)_BS-Seq_R	TTTTAAAATATCAACTTATTTTCCATTAAAC
Gbx2 (region 1)_BS-Seq_F	TTTTTTTTATTTGGGTATTTTATTTTTATTTT
Gbx2 (region 1)_BS-Seq_R	AAAAAAATATTCTCTTAAACCTCAAATCC
Gbx2 (region 2)_BS-Seq_F	GAGTGGTTATGAGCGTGGAGGTGAG
Gbx2 (region 2)_BS-Seq_R	GCTAAAAAATAATACCGCCTTCAACATAAAC
Gbx2 (region 3)_BS-Seq_F	TTGAGTTTTTTAAGATATTTTTTTTGTTTT
Gbx2 (region 3)_BS-Seq_R	ACTCCTTAACCTAATACTATCAACAATTCC
Gfap_BS-Seq_F	TGTAGGTAAGTAATTTATGGATTAA
Gfap_BS-Seq_R	AAAAAAAATCCACCCTAACAAAAA
Nkx2-2_BS-Seq_F	GGGTTATTGTTTTAGTAATGGAGGTAGTTATAGT
Nkx2-2_BS-Seq_R	ATTTTCAATCAAAAACATCTTAAACCTTCC
Sparcl1_BS-Seq_F	ATTGATTGGGTTGGGTAAGAGTAGGTTAT
Sparcl1_BS-Seq_R	AAACTTCACACTCTTATTTATTCCTAAAAC
Pdgfra_BS-Seq_F	GTTTGGTTTTTTGGTAGTGGGTTA
Pdgfra_BS-Seq_R	TCCACACACACACTTCTCCTAATA

Sequenom BS-MS analysis primers

Dlx2 (region 2)_BS-MS_F	

Dlx2 (region 2)_BS-MS_R	AATACCCACTTTCTACAATACTTATACAC
Dlx2 (region 4)_BS-MS_F	TTTTAAGTATGTGATTAAGGTTTTGGG
Dlx2 (region 4)_BS-MS_R	CCATCTACTCCAATTTCCAACTAAC
Gbx2 (region 3)_BS-MS_F	GATTGTTGGATGGAGTTGAGTTTT
Gbx2 (region 3)_BS-MS_R	CAATACCCTCCTTTACTCCTTAACC
Gfap_BS-MS_F	ATTAGTTAGTTTGGTGGGTTTTTTG
Gfap_BS-MS_R	ATCAAAAACAAATTTAATCCAACCC
Gsh2_BS-MS_F	ATAAGGATATTTGAGAATTGATGTAGTTTG
Gsh2_BS-MS_R	ACCATTTAAACTAAAATAACCAATACATTT
Insl6_BS-MS_F	AGGTTATGAAATTATAAAAGGGATTT
Insl6_BS-MS_R	CCCTATACCAACAAAACAACACATA
Nkx2-2_BS-MS_F	ATAGAAAGGAGGGGGTAAAGAATTT
Nkx2-2_BS-MS_R	CCTATTCCTCTCCTAAATTTCCAAC
Sp8_BS-MS_F	TATTGGTATTGGTTTTTTTGTGGTT
Sp8_BS-MS_R	TCTCAAATAACCCCACTATACCTTTC
Sparcl1_BS-MS_F	AATTTGGAGGTTGGTATTATAGGGT
Sparcl1_BS-MS_R	AAAAAACAACAACTTCATTTTCCAA
Pdgfra_BS-MS_F	GTGTTTTGGGTTTTTTGGTTTGT
Pdgfra_BS-MS_R	CCACTTATCTAATCCTACCTTTATTTTC

References for Supporting Online Material

1. V. Coskun *et al.*, *Proc Natl Acad Sci U S A* **105**, 1026 (Jan 22, 2008).
2. W. Ge *et al.*, *J Neurosci Res* **69**, 848 (Sep 15, 2002).
3. S. J. Cokus *et al.*, *Nature* **452**, 215 (Mar 13, 2008).
4. M. Kurita, T. Kuwajima, I. Nishimura, K. Yoshikawa, *J Neurosci* **26**, 12003 (Nov 15, 2006).
5. T. Chen, Y. Ueda, S. Xie, E. Li, *J Biol Chem* **277**, 38746 (Oct 11, 2002).
6. F. He *et al.*, *Nat Neurosci* **8**, 616 (May, 2005).
7. T. H. Kim *et al.*, *Nature* **436**, 876 (Aug 11, 2005).
8. T. S. Mikkelsen *et al.*, *Nature* **448**, 553 (Aug 2, 2007).
9. M. Weber *et al.*, *Nat Genet* **37**, 853 (Aug, 2005).
10. A. Smallwood, P. O. Esteve, S. Pradhan, M. Carey, *Genes Dev* **21**, 1169 (May 15, 2007).
11. R. Cao, Y. Zhang, *Mol Cell* **15**, 57 (Jul 2, 2004).
12. J. D. Cahoy *et al.*, *J Neurosci* **28**, 264 (Jan 2, 2008).
13. G. L. Sen, J. A. Reuter, D. E. Webster, L. Zhu, P. A. Khavari, *Nature* **463**, 563 (Jan 28).
14. A. M. Broske *et al.*, *Nat Genet* **41**, 1207 (Nov, 2009).
15. R. Straussman *et al.*, *Nat Struct Mol Biol* **16**, 564 (May, 2009).
16. R. Lister *et al.*, *Nature* **462**, 315 (Nov 19, 2009).
17. T. A. Rauch, X. Wu, X. Zhong, A. D. Riggs, G. P. Pfeifer, *Proc Natl Acad Sci U S A* **106**, 671 (Jan 20, 2009).
18. M. M. Suzuki, A. Bird, *Nat Rev Genet* **9**, 465 (Jun, 2008).