

Supplementary Information

Supplementary Materials and Methods

Key resources table

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|----------------|--------------------------------|
| Antibodies | | |
| Mouse monoclonal to β -Actin | Chemicon | MAB1501; RRID: AB_2223041 |
| Mouse monoclonal to TUJ1 | Sigma | T8660; RRID: AB_477590 |
| Rat monoclonal to BrdU | Abcam | ab6326; RRID: AB_305426 |
| Mouse monoclonal to BrdU | Santa Cruz | sc-32323, RRID: AB_626766 |
| Rabbit monoclonal to Cleaved Caspase-3 | Cell Signaling | 9664; RRID: AB_2070042 |
| Rat monoclonal to CTIP2 | Abcam | ab18465; RRID: AB_2064130 |
| Rabbit polyclonal to GFP | Invitrogen | A11122; RRID: AB_221569 |
| Rabbit polyclonal to KDM2B | Millipore | 17-10264; RRID: AB_11205420 |
| Rabbit polyclonal to KDM2B | Millipore | 09-864; RRID: AB_10806072 |

| | | | |
|-----------------------------|------------------------------|---------------------------|-------|
| Rabbit polyclonal to KDM2B | Labaratory of JieKai Chen | N/A | |
| Rabbit polyclonal to PAX6 | Chemicon | AB2237; AB_1587367 | RRID: |
| Rabbit polyclonal to TBR2 | Abcam | ab23345; AB_778267 | RRID: |
| Rabbit polyclonal to SOX2 | Milipore | ab5603; AB_2286686 | RRID: |
| Rabbit polyclonal to NeuN | Abcam | ab177487 | |
| Rabbit polyclonal to GFAP | DAKO | Z0334 | |
| Mouse monoclonal to FLAG | Sigma-Aldrich | F1804; RRID: AB_262044 | |
| Mouse monoclonal to SATB2 | Abcam | ab51502; AB_882455 | RRID: |
| Goat polyclonal to UNC5D | R&D | AF1429; AB_2304199 | RRID: |
| Sheep anti-DIG AP | Roche | 11093274910; AB_514497 | RRID: |
| Mouse monoclonal to H3K4me3 | Active Motif | MABI0304; AB_514497 | RRID: |
| Mouse monoclonal to H3K27ac | Millipore | 17-683; AB_1977529 | RRID: |

| | | | |
|---------------------------------------|-------------|---------------------------|-------|
| Rabbit polyclonal to NEUROD2 | Abcam | ab104430; AB_10975628 | RRID: |
| Rabbit polyclonal to SATB1 | Abclonal | A5800 | |
| Mouse monoclonal to hnRNPAB | Santa Cruz | sc-32323 | |
| Rabbit polyclonal to GAPDH | Cwbio | CW0101M; AB_2665434 | RRID: |
| Rabbit polyclonal to β -TUBULIN | Proteintech | 10094-1-AP; AB_2210695 | RRID: |

Chemicals, Peptides, and Recombinant Proteins

| | | |
|--------------------------------|---------------|-----------|
| B27 | Thermo Fisher | 17504044 |
| N2 | Thermo Fisher | 17502048 |
| hEGF | Thermo Fisher | PHG0311 |
| hFGF2 | Thermo Fisher | PHG0261 |
| Papain | Worthington | LS003118 |
| DNase I | Sigma-Aldrich | DN-25 |
| DMEM-F12 medium | Thermo Fisher | 12634-010 |
| Protein G agarose | | |
| Streptavidin Agarose | Thermo Fisher | S951 |
| 5-Bromouridine 5'-triphosphate | Sigma-Aldrich | B7166 |
| Mung Bean Nuclease | Takara Bio | 2420A |
| Micrococcal Nuclease | NEB | M0247S |

| | | |
|--------------------------------|---------------|-------------|
| Vanadyl Ribonucleoside Complex | Sangon | B644221 |
| | Biotech | |
| Protease inhibitor | Bioutil | B14001 |
| PMSF | Sigma-Aldrich | P7626 |
| Proteinase K | Sigma | P4032 |
| NBT/BCIP | Roche | 11681451001 |
| DIG-NTP | Roche | 11277073910 |
| Biotin RNA labeling mix | Roche | 11685597912 |
| CDP-star | Roche | 11685627001 |
| paraformaldehyde | Sigma-Aldrich | P6148 |
| FastGreen | Sigma-Aldrich | F7252 |
| Lipofectamine 3000 | Invitrogen | L3000-150 |
| Trizol | Thermo Fisher | 15596026 |

Bacterial and Virus Strains

| | | |
|-----------------------------------|----------|------------|
| <i>E. coli</i> DH5 α | TransGen | CD201-01 |
| <i>E. coli</i> Stbl3 | TransGen | CD521-01 |
| Mus musculus BAC clone | BACPAC | RP23-21416 |
| Lentivirus vector, pLKO.1-zsGreen | | |

Critical Commercial Assays

| | | | | | |
|-------------------------------|--------|------|------|-------|-----------|
| Mouse | Neural | Stem | Cell | LONZA | VAPG-1004 |
| Nucleofector [®] Kit | | | | | |

HiScribe™ T7 High Yield RNA NEB E2040S

Synthesis Kit

Experimental Models: Cell Lines

Mouse Neuro-2a cell line The Cell Bank TCM29

of Chinese

Academy of

Sciences

Mouse NE-4C cell line The Cell Bank SCSP-1501

of Chinese

Academy of

Sciences

Human HEK293T cell line A gift from Dr.

Hongbing Shu

LncKdm2b polyA Knock-in mouse This paper N/A

ES cells

LncKdm2b polyA Knock-in and This paper N/A

Kdm2b indels Mouse ES Cells

Experimental Models: Organisms/Strains

Mouse: CD-1 Hunan SJA

Laboratory

Animal Co

| | | |
|--|------------------|-----|
| Mouse: C57BL/6 | Hunan | SJA |
| | Laboratory | |
| | Animal Co | |
| Mouse: <i>Ai14</i> reporter | (Madisen et al., | |
| | 2010) | |
| Mouse: <i>Kdm2b</i> ^{CreERT2/+} | This paper | N/A |

Recombinant DNA

| | | |
|--------------|---------|-------|
| pGEM-Teasy | Promega | A1360 |
| pCAGGS | | |
| pCAG-mir30 | | |
| pMD2.G | Addgene | 12259 |
| psPAX2 | Addgene | 12260 |
| pGL3-basic | Promega | E1751 |
| phRL-TK | Promega | E6921 |
| pCALNL-DsRed | Addgene | 13769 |

Software and Algorithms

| | | |
|----------------------------------|----------------------|---|
| DAVID functional annotation tool | (Huang et al., 2009) | https://david.ncifcrf.gov/ |
| UCSC Genome Browser | (Kent et al., 2002) | http://genome.ucsc.edu/ |

| | | |
|--------------------|---------------------|---|
| CPAT | (Wang et al., 2013) | http://lilab.research.bcm.edu/cpat/index.php/ |
| PhyloCSF | (Lin et al., 2011) | http://compbio.mit.edu/PhyloCSF/ |
| RNAfold web server | | http://rna.tbi.univie.ac.at/ |
| Prism | GraphPad | Ver 6 |

Supplementary tables

Table S1. Divergent lncRNAs identified in this study.

Table S2. Significantly-enriched proteins in *LncKdm2b*-precipitating extracts compared to the antisense-*LncKdm2b*.

Table S3. Statistical analyses of electroporated cortices.

Table S4. Sequences for all primers used in this study.

Supplementary figure titles and legends

Figure S1. *LncKdm2b* and *Kdm2b* are transiently expressed in the developing mouse embryonic cortex.

(A) Gene ontology (GO) analysis of coding genes associated with divergent lncRNAs.

The top GO terms (>1.5-fold and $P < 1 \times 10^{-6}$) are shown.

(B) Protein-coding scores of *LncKdm2b* using CPAT and PhyloCSF programs.

(C) Thirteen putative ORFs of *LncKdm2b* were cloned into the pFLAG-N3 vector with

3 × Flag tag sequence fused to their 3' for HEK293T cell transfection. After 48 hours, immunoblotting was performed to detect Flag-tagged proteins. PRRX1B tagged with 3 × Flag tag served as a positive control. Data are representative of three independent experiments.

(D) Fragments per kilobase per million mapped reads (FPKM) values for *LncKdm2b*, and *Kdm2b* in mouse ESCs, mouse NPCs and mouse cortices at indicated developmental stages.

(E-L) RT-qPCR analyses of indicated markers of E10.5, E12.5, E14.5 and adult (6 weeks) mouse dorsal forebrains. The y-axis represents relative expression normalized to *Gapdh*.

(M) Representative immunoblotting of mouse dorsal forebrains using antibodies against KDM2B and β -TUBULIN.

(N) Northern blots of *LncKdm2b* and *Gapdh* using poly(A) RNAs extracted from mouse dorsal forebrains.

(O) Schematic illustration of *in situ* hybridization probes for mouse *LncKdm2b* and *Kdm2b*.

(P) *In situ* hybridization (ISH) of *LncKdm2b* on coronal sections of E12.5 mouse forebrain (the CD-1 strain). Scale bars, 100 μ m. The higher-magnification image of the boxed area shows immunofluorescent staining for TBR2 (green) and TUJ1 on ISH section of *LncKdm2b* (red). Scale bars, 50 μ m.

(Q) Southern blot analysis of genomic DNA from wild-type (WT) or *Kdm2b*^{CreERT2/+}

knock-in mice.

(R-S) Immunofluorescent staining for EGFP (green), TBR2 (red), TUJ1 (blue), UNC5D (red), and DAPI (blue) on cortical sections of E14.5 *Kdm2b*^{CreERT2/+} mice. Boxed areas are enlarged at the bottom-right corners. Scale bars, 50 μ m.

In (E-L), quantification data are shown as mean + SD (n = 3 unless otherwise indicated).

Ctx, cortex; LV, lateral ventricle; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone.

Figure S2. *Kdm2b*-expressing cortical cells are fated to be cortical projection neurons.

(A-D) Immunofluorescent staining for SATB2 and CTIP2 on coronal cortical sections of E16.5 (A) and P0 (C) *Kdm2b*^{CreERT2/+}; *Ai14* embryos. Tamoxifen (TAM, 100 mg/kg) was injected at E12.5 or E14.5 respectively. Boxed areas are enlarged at the bottom-right corners. Quantification of SATB2 or CTIP2 expression in tdTomato+ recombined cells (B and D). A total of 3372 cells from 2 embryos were analyzed in (B), and 2282 cells from 2 animals in (D).

(E) Immunofluorescent staining for SOX2 and TUJ1 on head sections of E10.5 *Kdm2b*^{CreERT2/+}; *Ai14* embryo. Boxed areas are enlarged at the bottom-right corners.

(F) Immunofluorescent staining for PAX6 (top) and TBR2 (bottom) on head sections of E13.5 *Kdm2b*^{CreERT2/+}; *Ai14* embryo. Tamoxifen (TAM, 100 mg/kg) was injected at E12.5. Boxed areas are enlarged at the bottom-right corners.

(G-H) Immunofluorescent staining for SATB2 and CTIP2 on coronal cortical sections of P7 *Kdm2b*^{CreERT2/+};*Ai14* mouse brain. Boxed areas are enlarged at the bottom-right corners. Quantification of SATB2 or CTIP2 expression in tdTomato+ recombined cells (H). A total of 373 cells were analyzed.

In (B), (D), and (H), quantification data are shown as mean + SEM.

LV, lateral ventricle; MB, midbrain. Scale bars, 50 μ m.

Figure S3. *LncKdm2b* maintains *Kdm2b* transcription in *cis*.

(A) Percentages of marker-expressing neuronal cells in adherent cultures derived from E12.5 cortices.

(B) RT-qPCR analysis of *LncKdm2b* and *Kdm2b* RNA levels in adherent cultures derived from E12.5 cortices. The cultures were treated with indicated ASOs. The y-axis represents relative expression normalized to *Gapdh*.

(C) RT-qPCR analysis of *Zfp292* mRNA levels in Neuro-2a cells treated with indicated *LncKdm2b* ASOs for three days. The y-axis represents relative expression normalized to *Gapdh*.

(D) Genotyping of NE-4C clones with *LncKdm2b*'s exon2 knocked out.

(E) RT-qPCR analysis of *Zfp292* mRNA levels in NE-4C clones with *LncKdm2b*'s exon2 knocked out. The y-axis represents relative expression normalized to *Gapdh*.

(F) RT-qPCR analysis for *LncKdm2b*, *Gapdh*, *Actb*, *Xist*, and *Neat2* from cytoplasmic and nuclear RNA fractions of primary E14.5 cortical neural progenitor cells (NPCs)

cultured *in vitro* for four days.

(G) Fluorescent *in situ* hybridization of *LncKdm2b* on cortical NPCs treated with or without RNase A. The nuclei were counter-stained with DAPI.

(H-I) RT-qPCR analysis of *LncKdm2b* and *Kdm2b* mRNA levels in Neuro-2a and NE-4C cells transfected with empty or *LncKdm2b*-expressing vectors for 48 hours. The y-axis represents relative expression normalized to *Gapdh* (n = 5).

(J) RT-qPCR analysis of *LncKdm2b* and *Kdm2b* mRNA levels in parental and *LncKdm2b*-KO NE-4C cells transfected with empty or *LncKdm2b*-expressing vectors for 48 hours. The y-axis represents relative expression normalized to *Gapdh*.

(K) RT-qPCR analysis of *LncKdm2b* and *Kdm2b* RNA levels in Neuro-2a cells treated for two days with Scramble ASOs or ASOs targeting *Kdm2b*.

In (A) and (H-I), quantification data are shown as mean + SEM. In (B-C), (E-F), and (J-K) quantification data are shown as mean + SD (n = 3 unless otherwise indicated).

In (B) and (H-I), statistical significance was determined using 2-tailed Student's t test.

In (C), (E), and (J-K), statistical significance was determined using 1-way ANOVA followed by the Turkey's post hoc test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, "NS" indicates no significance.

Figure S4. *LncKdm2b* maintains *Kdm2b* transcription in *cis*.

(A) Relative crosslinking frequency between the T5 and *Kdm2b*'s TSS measured by 3C-qPCR in Neuro-2a cells treated for two days with Scramble ASOs or ASOs

targeting *LncKdm2b*. The y-axis shows fold enrichment normalized to the scramble control.

(B) Luciferase activities in experiments where indicated vectors were transfected into HEK293T cells for 24 hours. 'Forward' and 'Reverse' indicate directions same as or opposite of *Kdm2b*'s transcription orientation.

(C) Genotyping of NE-4C cells with the T5 region knocked out.

(D) RT-qPCR analysis of *Zfp292* mRNA levels in NE-4C clones with the T5 region knocked out. The y-axis represents relative expression normalized to *Gapdh*.

(E) Genotyping of cortical cells subjected to Cas9-mediated knockout of the T5 region.

In (A-B) and (D), quantification data are shown as mean + SD (n = 3 unless otherwise indicated). In (A-B) and (D), statistical significance was determined using 1-way ANOVA followed by the Turkey's post hoc test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, "NS" indicates no significance.

Figure S5. Characterization of *LncKdm2b*-associated proteins.

(A) The illustration describing the Gal4- λ N/BoxB RNA tethering system.

(B-D) Relative luciferase activities in experiments where Neuro-2a cells were transfected with plasmids expressing BoxB-tagged *LacZ*, full-length *LncKdm2b*, 5' *LncKdm2b*, or 3' *LncKdm2b* along with Gal4- λ N and 5xUAS-TK-Luciferase-expressing plasmids for 24 hours.

(E) RT-qPCR analysis of *Kdm2b* mRNA levels in Neuro-2a cells treated for three days

with scramble siRNA (siNC) or siRNA targeting indicated molecules. The y-axis represents relative expression normalized to *Gapdh*.

(F) Fluorescent *in situ* hybridization of *LncKdm2b* on cortical NPCs followed by co-staining of hnRNPAB and DAPI.

(G) RNA secondary structure prediction by *RNAfold* showed two putative stem-loop structures.

(H) CHIP-qPCR analysis of indicated primer sets enriched by anti-hnRNPAB antibodies in Neuro-2a cells. The y-axis shows fold enrichment normalized to the IgG control.

In (B–E) and (H), quantification data are shown as mean + SD (n = 3 unless otherwise indicated). In (B–E), statistical significance was determined using 1-way ANOVA followed by the Turkey's post hoc test. In (H), statistical significance was determined using 2-tailed Student's t test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, "NS" indicates no significance.

Figure S6. *Kdm2b* promotes cortical neurogenesis.

(A) Representative immunoblots of HEK293T cells transfected with empty vector or KDM2B-expressing vector for two days using antibodies against KDM2B and ACTIN.

(B) Immunoblotting of E15.5 embryonic cortices with indicated genotypes using antibodies against KDM2B and β -TUBULIN.

(C-D) Quantification of relative location (C) and percentiles (D) of NEUROD2+ transduced cells in scramble or KDM2B shRNA electroporated sections. Three embryos each.

(E-F) E13.5 mouse cortices were electroporated with indicated combination of vectors, with transduced cells labeled with EGFP. Embryos were sacrificed at E16.5 for immunofluorescent analysis. Representative VZ/SVZ images of sections immunostained with PAX6 (E) and quantification of PAX6+ transduced cells (F) were shown. Three embryos in control and shKDM2B, five embryos in shKDM2B plus KDM2B. Scale bars, 50 μ m.

(G) Quantification of Cleaved Caspase3+ transduced cells in scramble or KDM2B shRNA electroporated sections. Three embryos each.

In (C-D) and (F-G), quantification data are shown as mean + SEM. In (C-D) and (G), statistical significance was determined using 2-tailed Student's t test. In (F), statistical significance was determined using 1-way ANOVA followed by the Turkey's post hoc test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, "NS" indicates no significance.

VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate.

Figure S7. *LncKdm2b* maintains mouse cortical neurogenesis through KDM2B.

(A) E13.5 mouse cortices were electroporated with Scramble ASO or ASOs targeting *LncKdm2b*, with transduced cells labeled with EGFP. RT-qPCR analysis of *LncKdm2b*

and *Kdm2b* mRNA levels in EGFP+ and EGFP- cells from E15.5 electroporated cortices. The y-axis represents relative expression normalized to *Gapdh*.

(B) Quantification of Cleaved Caspase3+ transduced cells in Scramble or *LncKdm2b* ASO electroporated cortices. Three embryos each.

(C-E) E13.5 mouse cortices were electroporated with indicated siRNAs, with transduced cells labeled with EGFP. Embryos were sacrificed at E16.5 for PAX6 immunofluorescent stainings (C). The relative location of EGFP+ cells (D) and percentiles of PAX6+ transduced cells (E) were quantified. Three embryos in control (siNC), five embryos in sihnRNPA2B1. Scale bars, 50 μ m.

(F-G) E13.5 mouse cortices were electroporated with empty or *LncKdm2b*-expressing vector, along with mCherry-expressing vector to label transduced cells. Embryos were sacrificed at E15.5 followed by DAPI staining of coronal sections (F). The locations of mCherry+ cells were quantified (G). Seven embryos each. Scale bars, 50 μ m.

In (A), quantification data are shown as mean + SD. In (B), (D-E), and (G), quantification data are shown as mean + SEM. In (A), statistical significance was determined using 2-way ANOVA followed by the Bonferroni's post hoc test. In (B), (D-E), and (G), statistical significance was determined using 2-tailed Student's t test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, "NS" indicates no significance.

VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate.

Figure S8. *LncKdm2b* regulates cortical neuronal differentiation and migration.

(A) A schematic illustration of the piggyBac-CRISPR/Cas9 toolkit for *LncKdm2b* knockout by *in utero* electroporation. Two sgRNAs were used for deletion of the *LncKdm2b*'s second exon.

(B) Quantification of neuron to glia ratio in brains at P10. Four brains each.

(C) Quantification of the distribution of SATB2+EGFP+ cells in cortices at P10. Four brains each.

(D) Quantification of Cleaved Caspase3+ transduced cells. Three brains each.

In (B-D), quantification data are shown as mean + SEM. In (B-D), statistical significance was determined using 2-tailed Student's t test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, "NS" indicates no significance.

REFERENCES

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44-57

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002) The human genome browser at UCSC. *Genome Res* 12:996-1006

Lin, M.F., Jungreis, I., and Kellis, M. (2011) PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* 27:i275-282

Madisen, L., Zwingman, T.A., Sunkin, S.M., Oh, S.W., Zariwala, H.A., Gu, H., Ng, L.L., Palmiter, R.D., Hawrylycz, M.J., Jones, A.R., *et al.* (2010) A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. Nat Neurosci 13:133-140

Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.P., and Li, W. (2013) CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res 41:e74

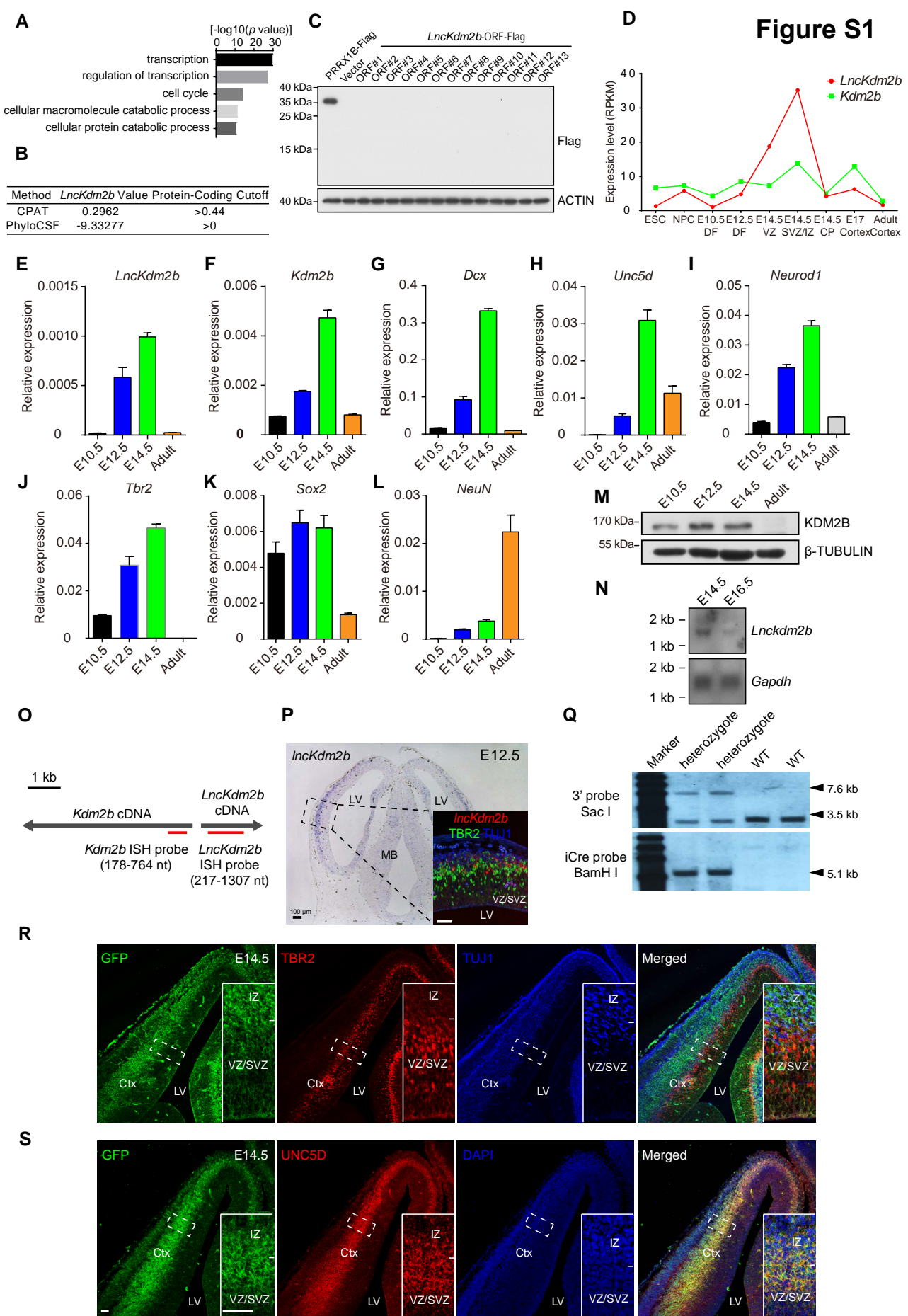
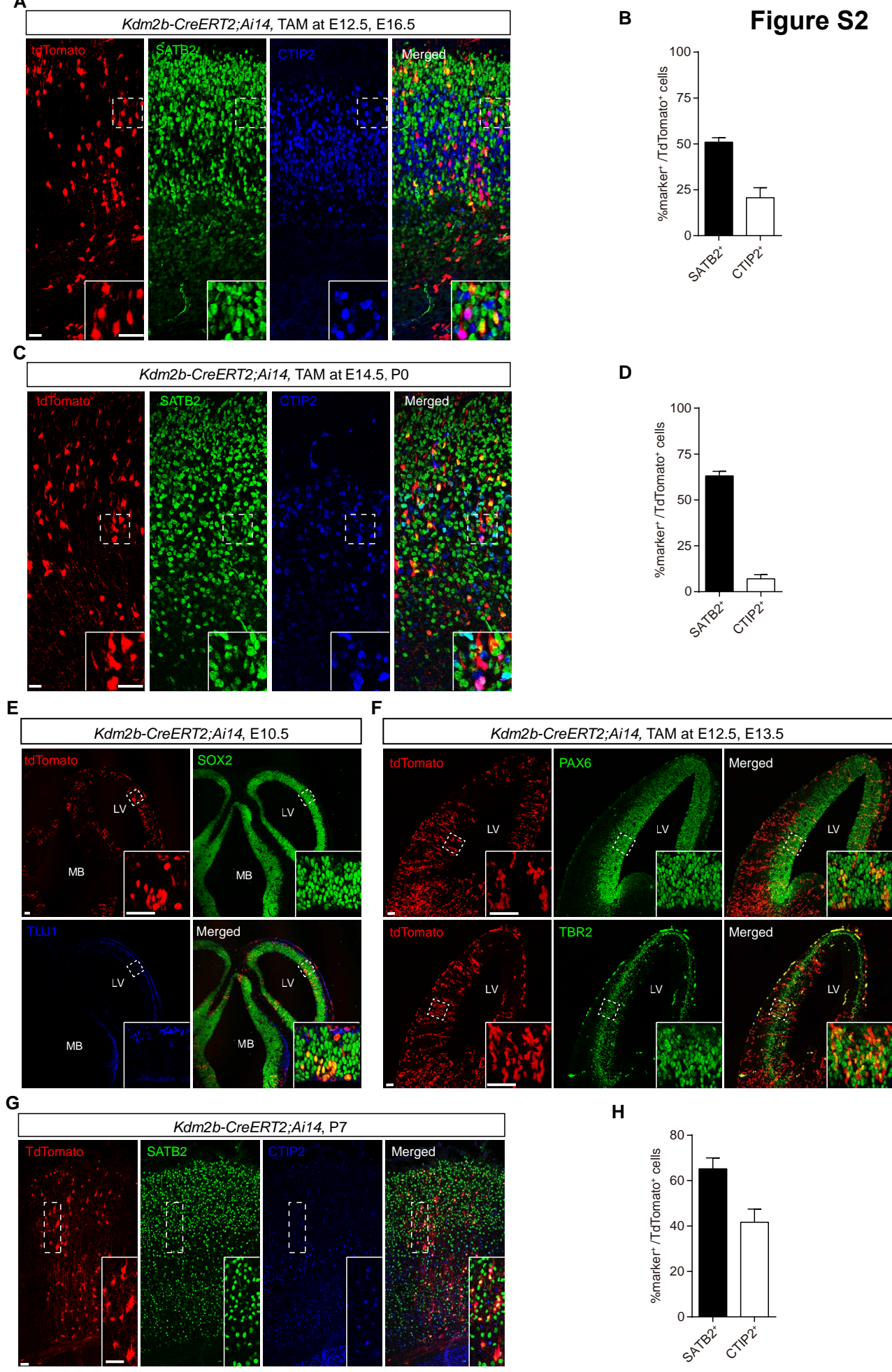


Figure S2

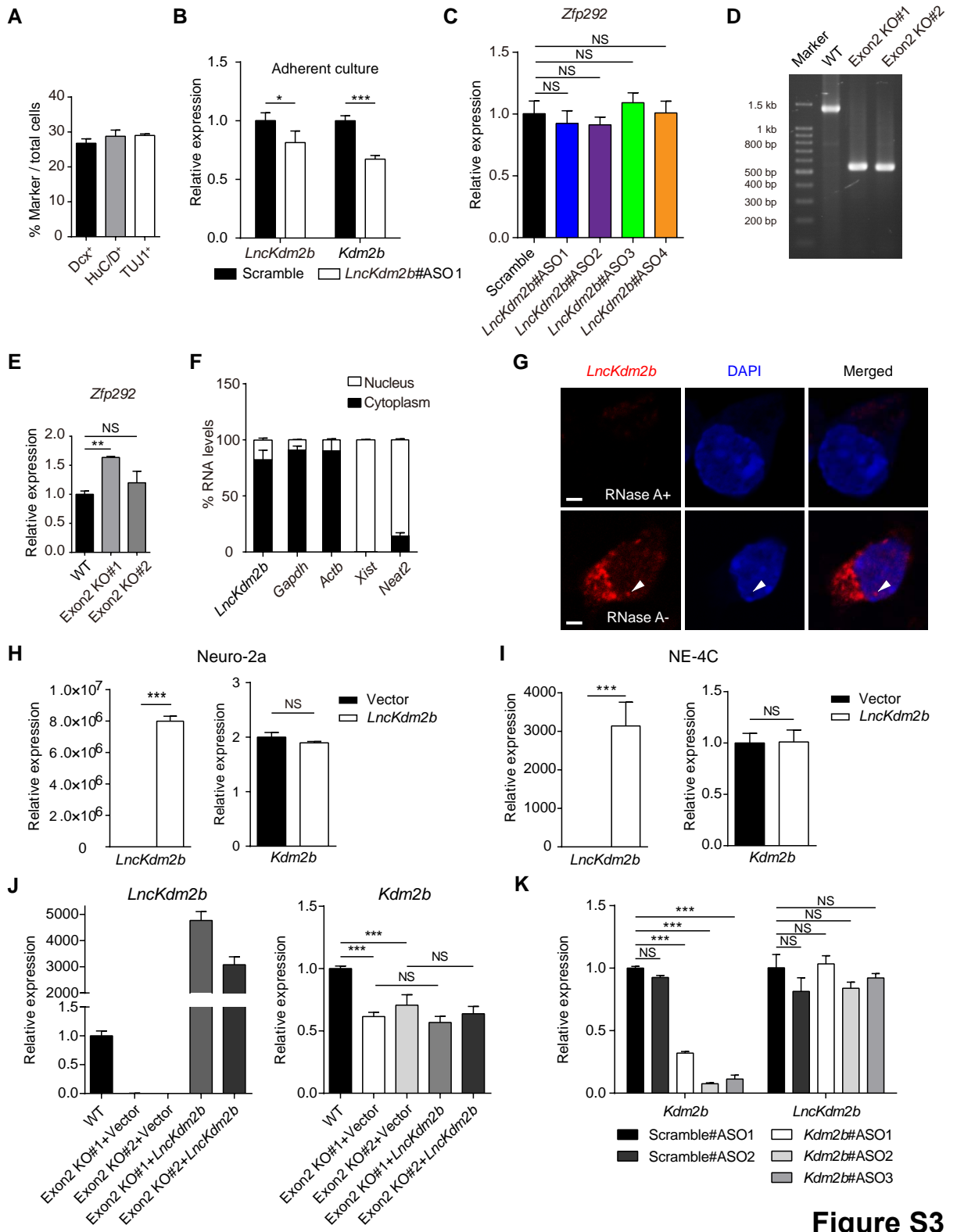


Figure S3

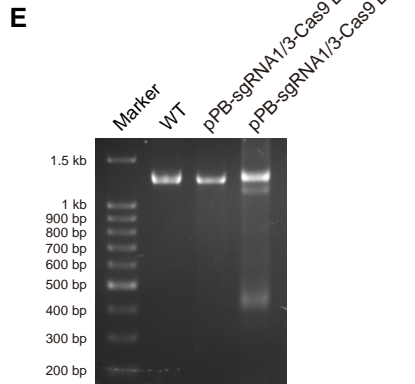
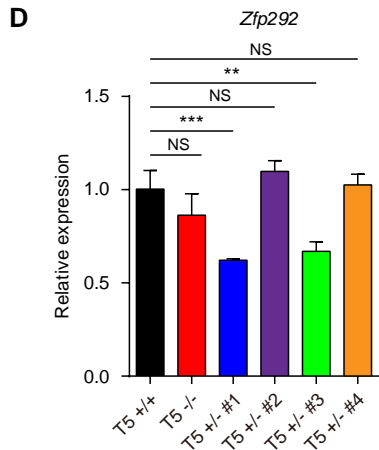
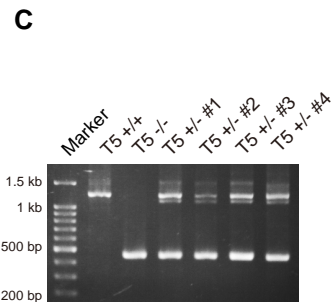
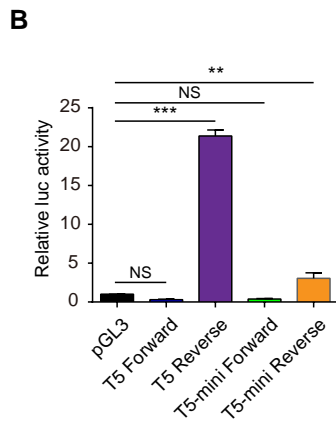
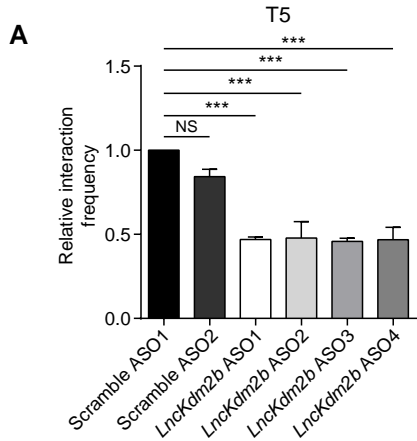


Figure S4

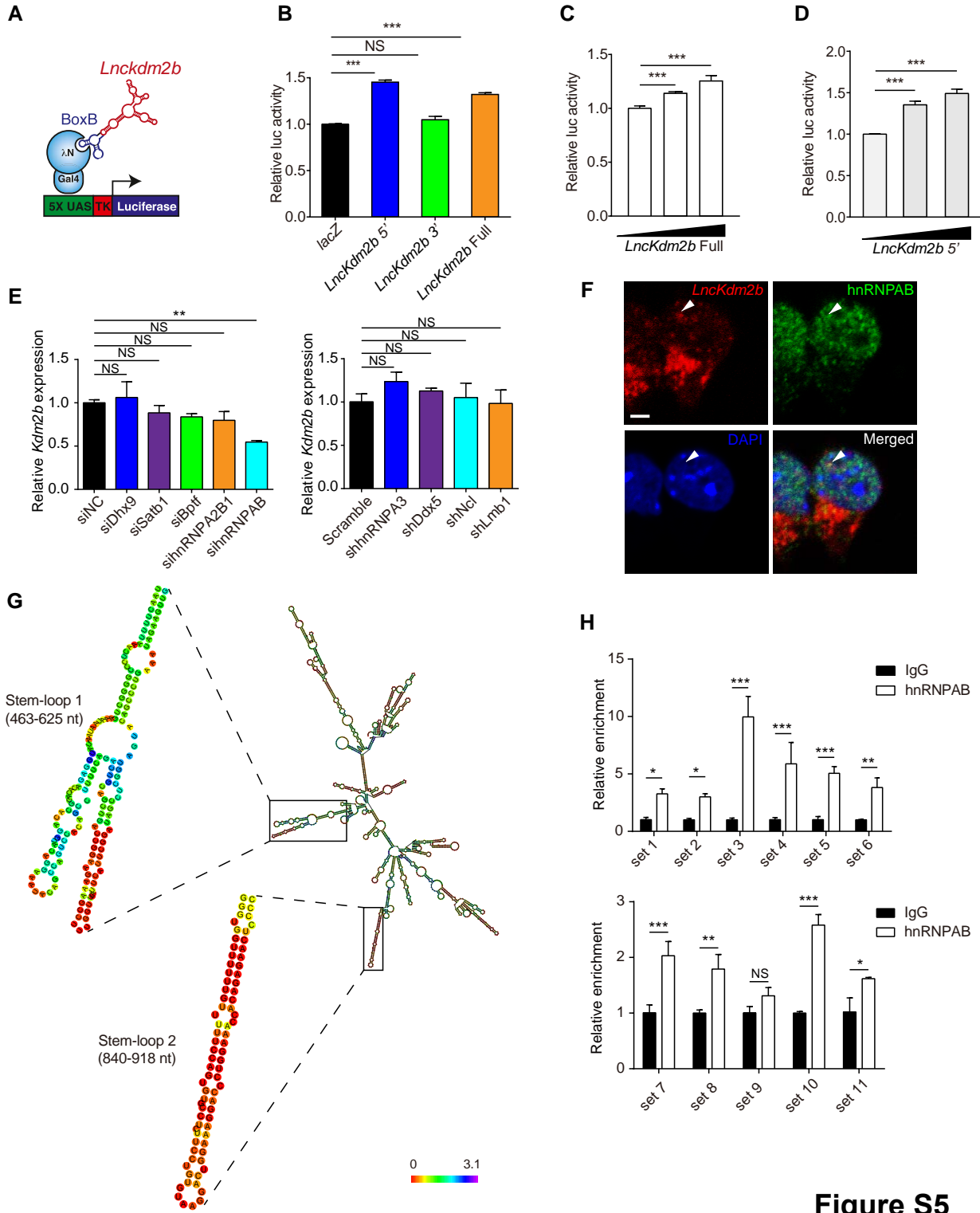
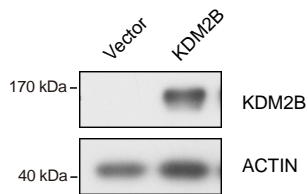
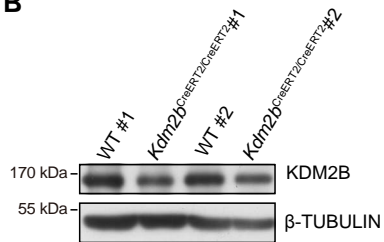
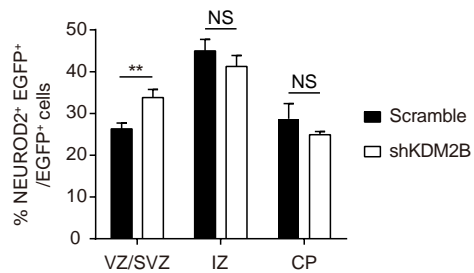
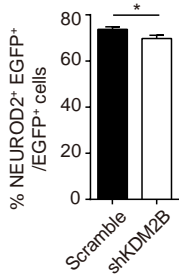
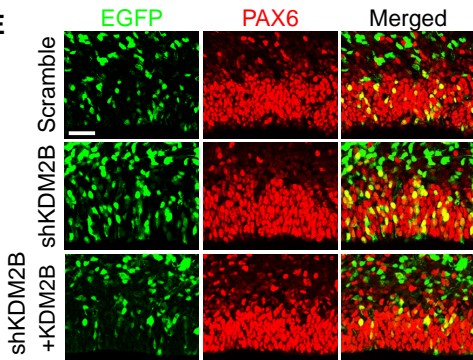
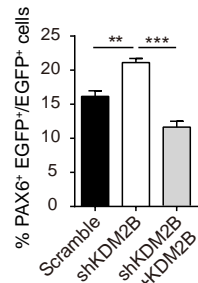
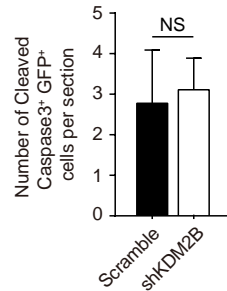


Figure S5

A**B****C****D****E****F****G****Figure S6**

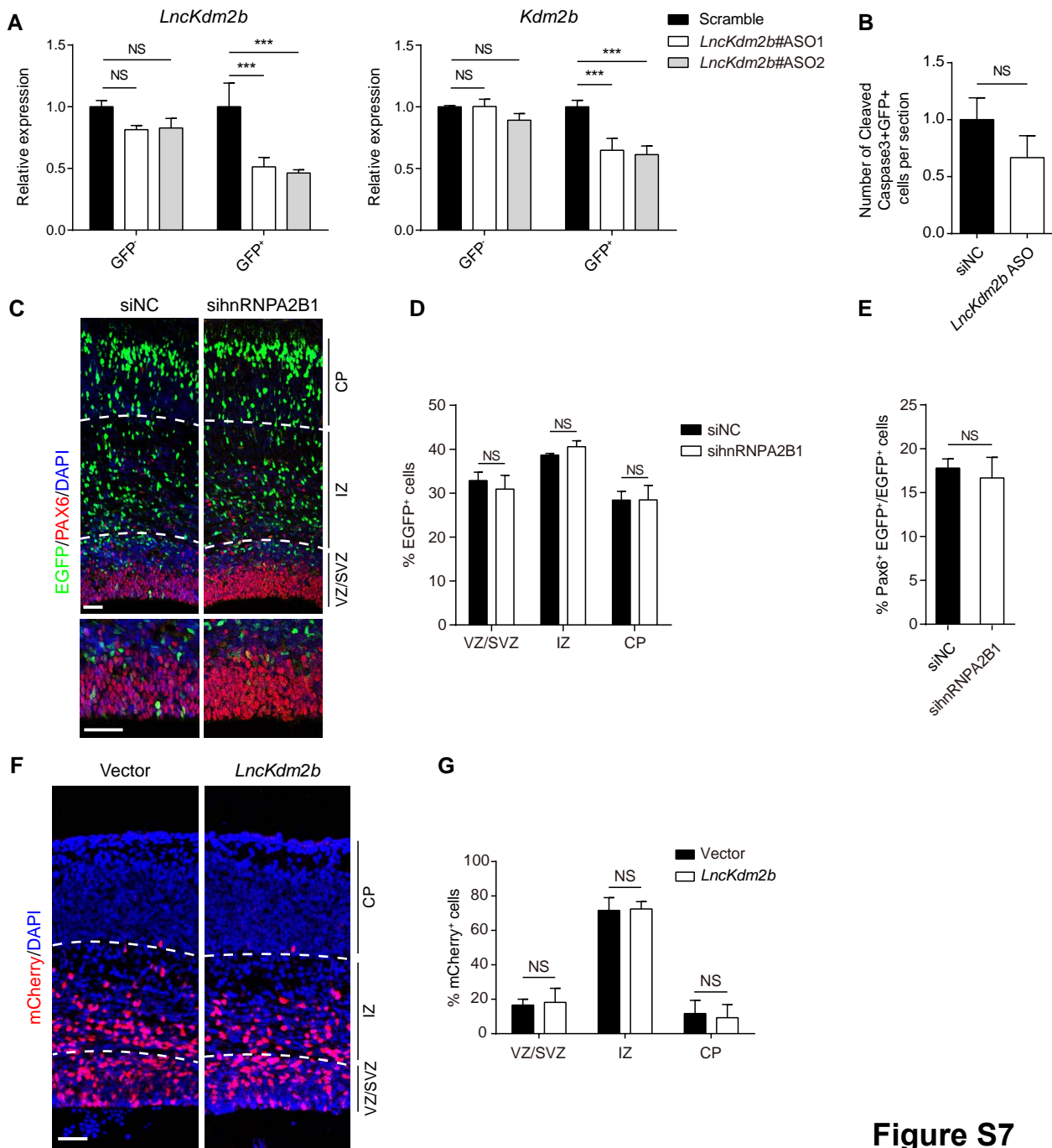
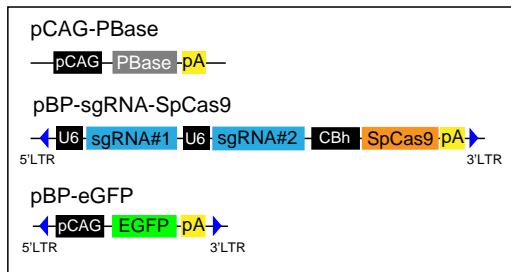
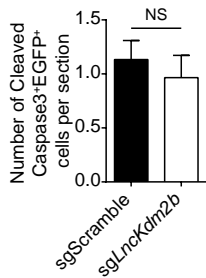
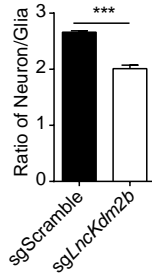
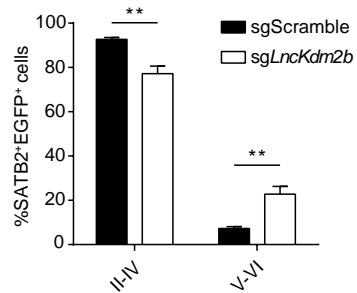


Figure S7

A**D****B****C****Figure S8**