APPENDIX

Table of contents

Appendix Figure S1

Appendix Figure S2

Appendix Figure S3

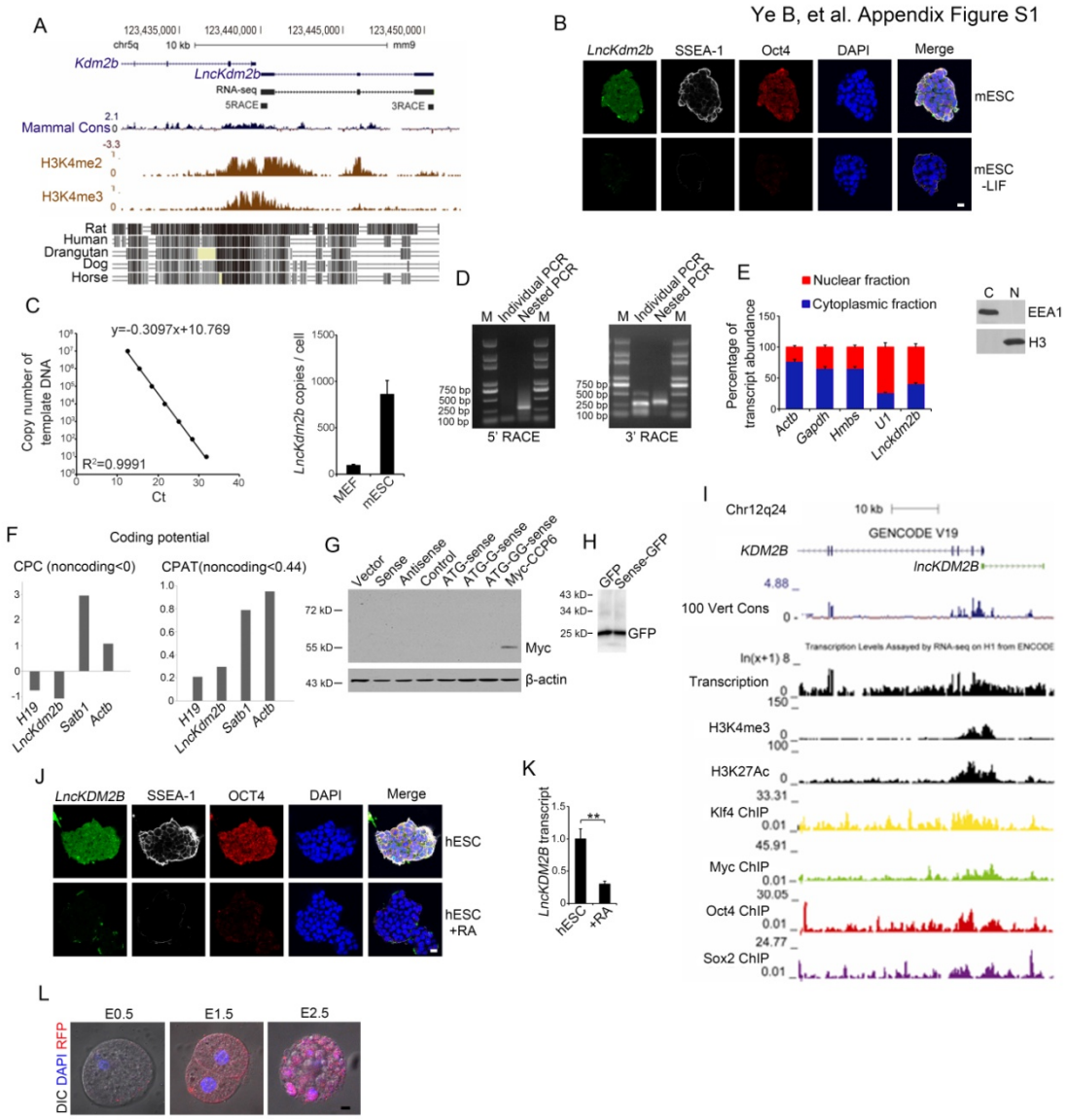
Appendix Figure S4

Appendix Figure S5

Appendix Table S1

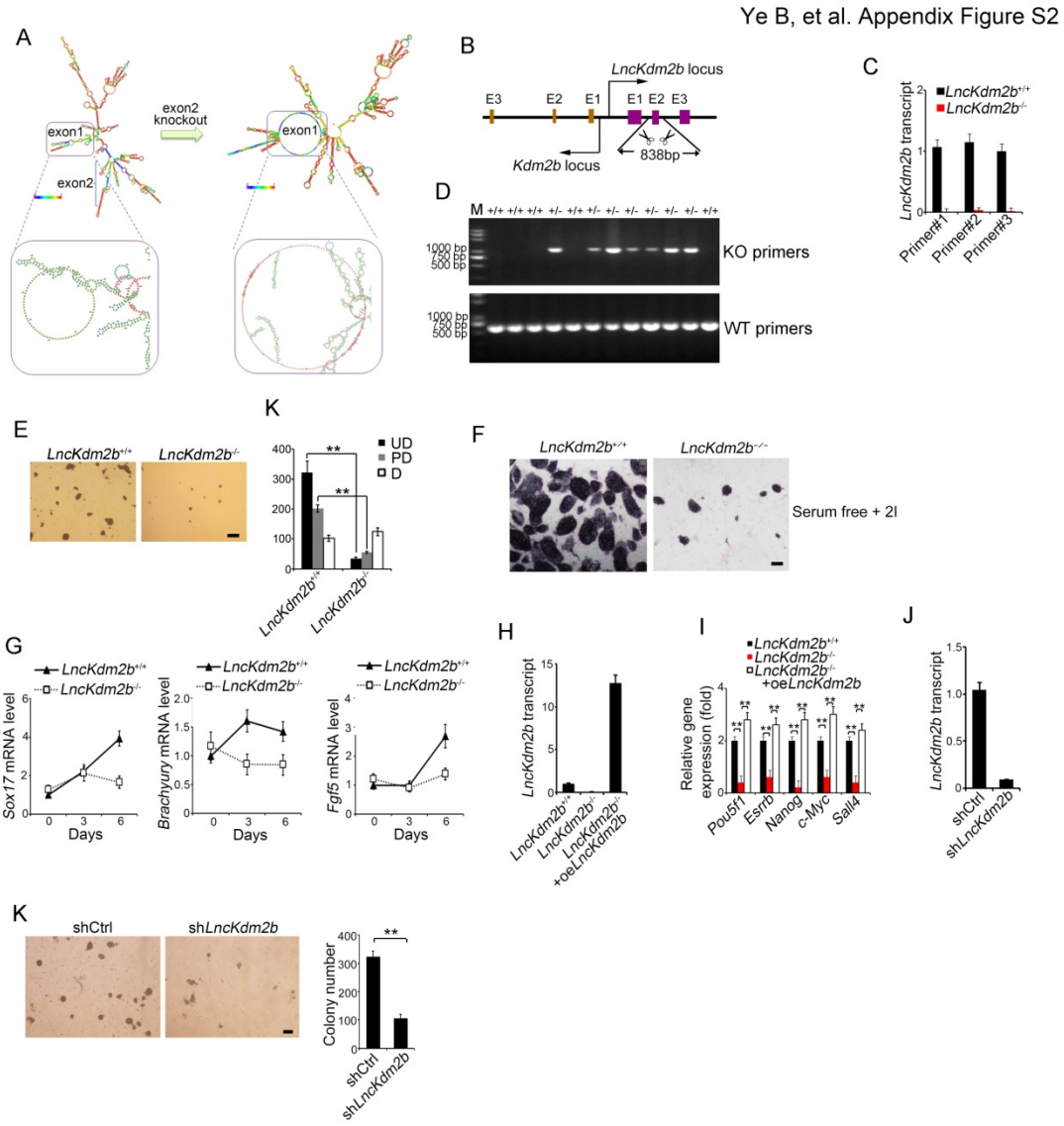
Appendix Table S2

Appendix Figures



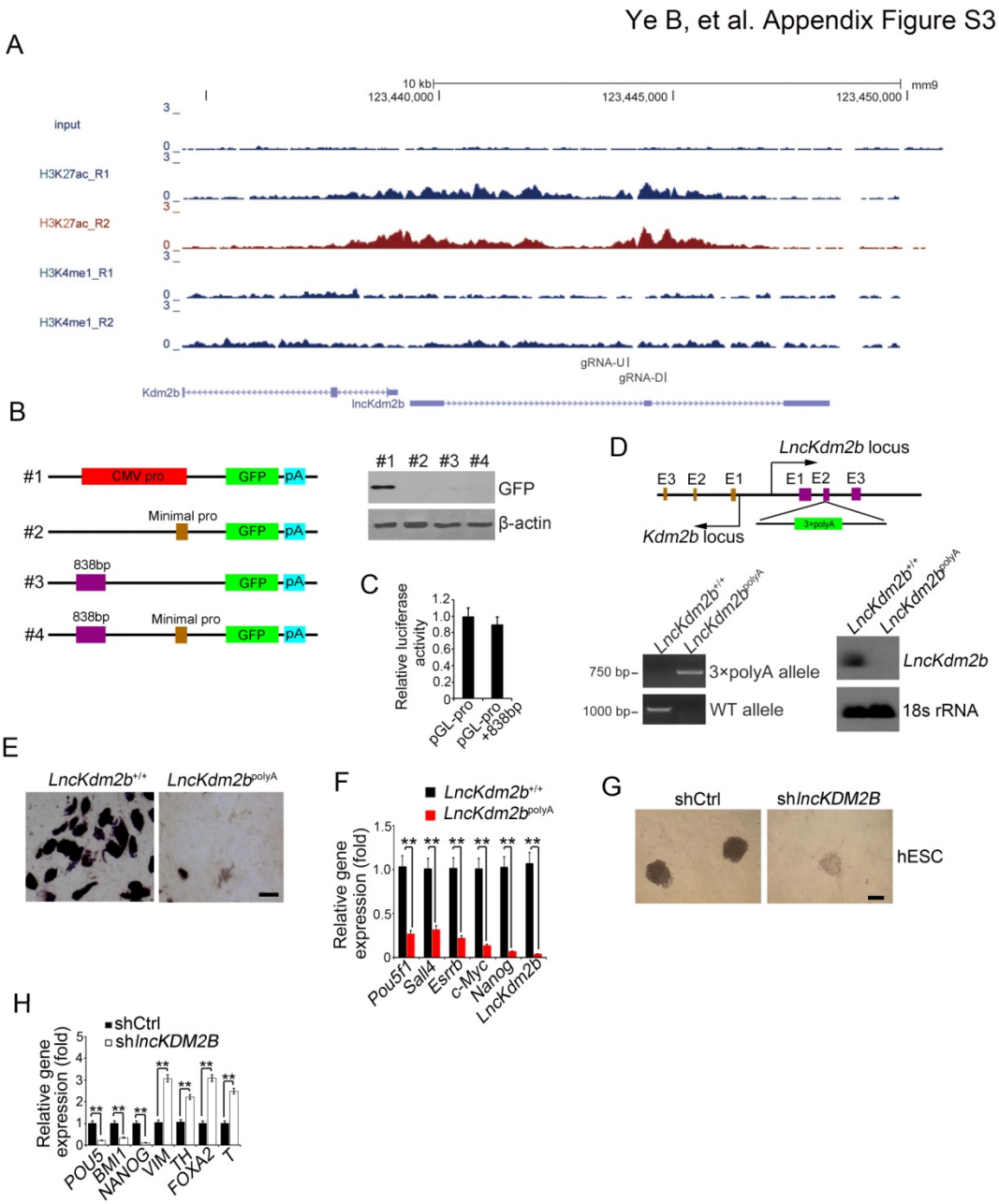
Appendix Figure S1. Characterization of *lncKdm2b*. (A) Location of *lncKdm2b* in mouse genome. *LncKdm2b* locates on mouse chromosome 5 adjacent to gene *Kdm2b*, comprising 3 exons. *LncKdm2b* was highly conserved in rat, human, orangutan, dog and horse. (B) *LncKdm2b* was visualized in mESCs and differentiated ESC cells by RNA-FISH assays followed with immunofluorescence staining. Green: *lncKdm2b* probe; Gray: SSEA-1; Red: Oct4; nuclei were counterstained by DAPI. Scale bar, 10 μm. Sequences of probes were listed in Appendix Table S1. (C) Quantitative estimation of lncKdm2b copy numbers per cell. PCR products of lncKdm2b were quantified in Nanodrop. A 1:10 fold serial dilution was performed as templates to generate standard curve in real-time qPCR. ESCs were counted and lysed to extract RNA. cDNA was obtained by reverse transcription. Ct values were counted to calculate original template

concentrations. Copy numbers per cell were calculated as means ± S.D. (D) *LncKdm2b* was cloned via 5'-RACE and 3'-RACE in mouse ESCs. (E) Fractionation of ESCs followed by qPCR. U1 RNA served as a positive control for nuclear gene expression. Actb RNA served as a positive control for cytoplasmic gene expression. EEA1, endosome antigen 1. H3, Histone H3. N: nuclear fraction. C: cytoplasmic fraction. Primers were listed in Appendix Table S1. (F) *LncKdm2b* displayed no coding potentiality by CPC/CPAT analysis. *H19* transcript served as a non-coding gene control. SATB homeobox 1 (*Satb1*) and *Actb* served as conding gene controls. (G) *LncKdm2b* displayed no coding potentiality by eukaryotic cell expression assay. *LncKdm2b* transcripts were cloned into pcDNA4-Myc-His plasmid and transfected into 293T cells for 48 h. Expression of Myc-tagged protein was analyzed by immunoblotting. Cytosolic carboxypeptidase 6 (CCP6) served as a coding gene control. (H) *LncKdm2b* transcript was cloned into pEGFP-C3 plasmid and transfected into 293T cells for 48 h. Expression of GFP-fused protein was analyzed by immunoblotting. (I) *LncKDM2B* locates on human chromosome 12 adjacent to gene *KDM2B*, encompassing 2 exons. The *lncKDM2B* locus exhibited enrichment of activation marks (H3K4me3, H3K27Ac) and core pluripotent transcriptional factors (OCT4/SOX2/MYC/KLF4) in local chromatin environments. (J) *LncKDM2B* was visualized in hESCs and differentiated cells by RNA-FISH assays. Green: *lncKDM2B* probe; Gray: SSEA-4; Red: OCT4; nuclei were counterstained by DAPI. Scale bar, 10 μm. Sequences of probes were listed in Appendix Table S1. (K) *LncKDM2B* transcript was analyzed in indicated cells by real time qPCR. Relative gene expression folds were normalized to endogenous β-actin and shown as means ± S.D. $p<0.01$. hESCs H1 were treated with RA (1 μM) to induce differentiation for 3 days. RA: retinoic acid. (L) The specificity of RFP fluorescence was confirmed by anti-RFP antibody staining in indicated embryo stages. Scale bar, 10 μm. All data represent five independent experiments.



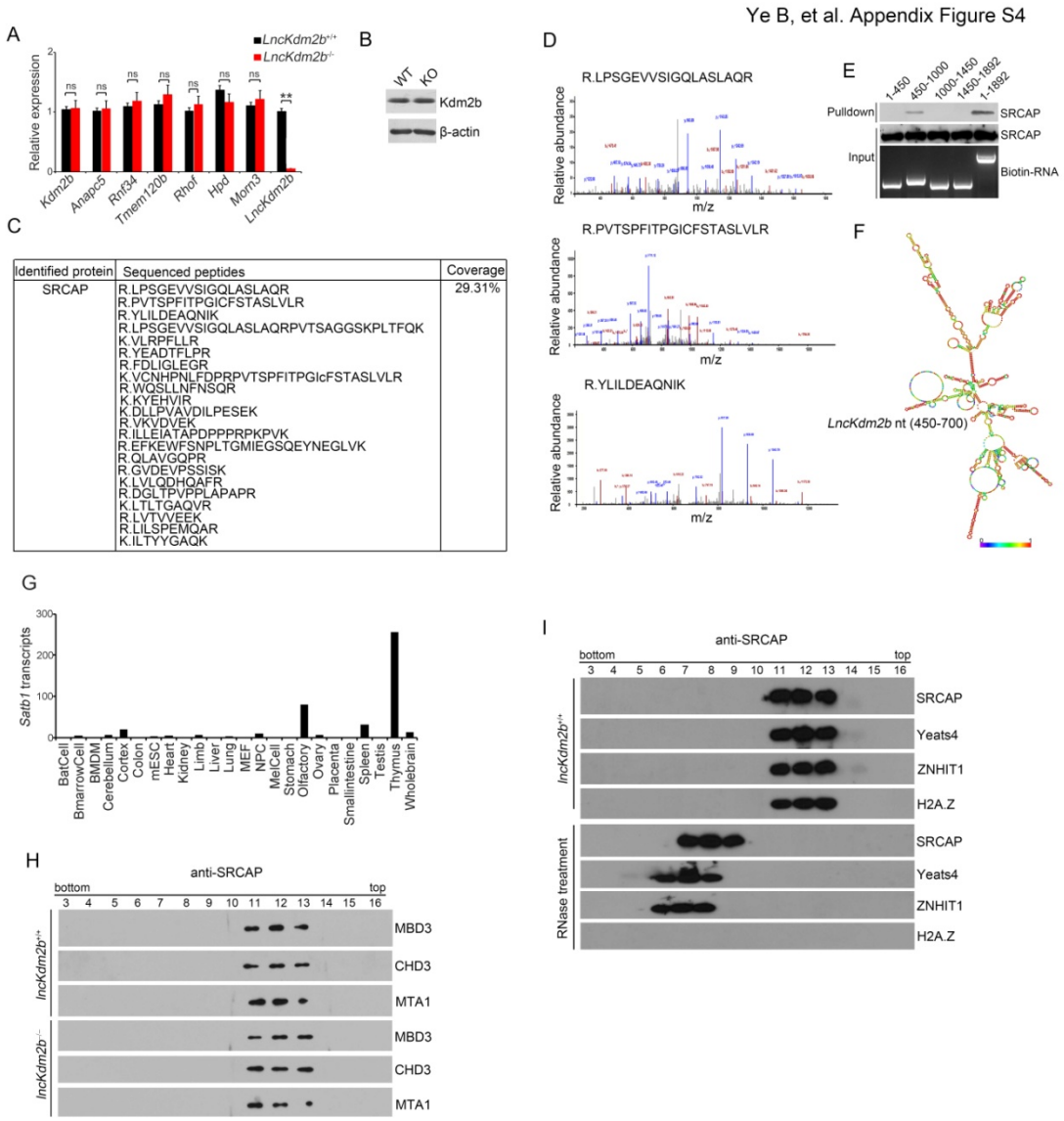
Appendix Figure S2. Generation strategy of *lncKdm2b* KO mice and rescue of *lncKdm2b* in *lncKdm2b* KO ESCs. (A) The deficiency of exon2 causes a drastic change of lncKdm2b structure predicted by RNA folding analysis. (B) The schematic diagram of *lncKdm2b* KO mouse construction. (C) *LncKdm2b* was examined in *lncKdm2b* KO cells by real time PCR. Primers were listed in Appendix Table S1. (D) *LncKdm2b*-deficient pups were genotyped after heterozygotes crossing. (E) *LncKdm2b* knockout ESCs were generated by CRISPR/Cas9 system. ESC pluripotency was analyzed by AP staining. Colony numbers for undifferentiated (UD), partially differentiated (PD) or differentiated (D) clones were calculated as means ± S.D. **, $p<0.01$. Scale bar, 50 μm. (F) Mouse ESCs were cultured in serum-free 2i condition for 5 days, followed by AP staining. Scale bar, 50 μm. (G) mESCs were induced to EB differentiation by LIF withdrawal for 5 days, followed by analysis of indicated mRNAs by real time qPCR. Primers were listed in Appendix Table S1. (H) *LncKdm2b* KO ESC cells were infected with *lncKdm2b* lentivirus for 5 days. *LncKdm2b* transcripts were analyzed by real time qPCR. Relative gene expression fold changes were counted as means ± S.D. **, $p<0.01$. Primers were listed in Appendix Table

S1. (I) *LncKdm2b* KO ESC cells were infected with *lncKdm2b* lentivirus for 5 days. mRNA levels of the indicated genes were analyzed by real time qPCR. Relative gene expression fold changes were counted as means ± S.D. **, $p<0.01$. Primers were listed in Appendix Table S1. (J) *LncKdm2b* was silenced by shRNA-containing lentivirus infection in mESCs. *LncKdm2b* transcripts were analyzed by real time qPCR as in (J). (K) ESC self-renewal was analyzed by AP staining. All data are representative of five independent experiments.



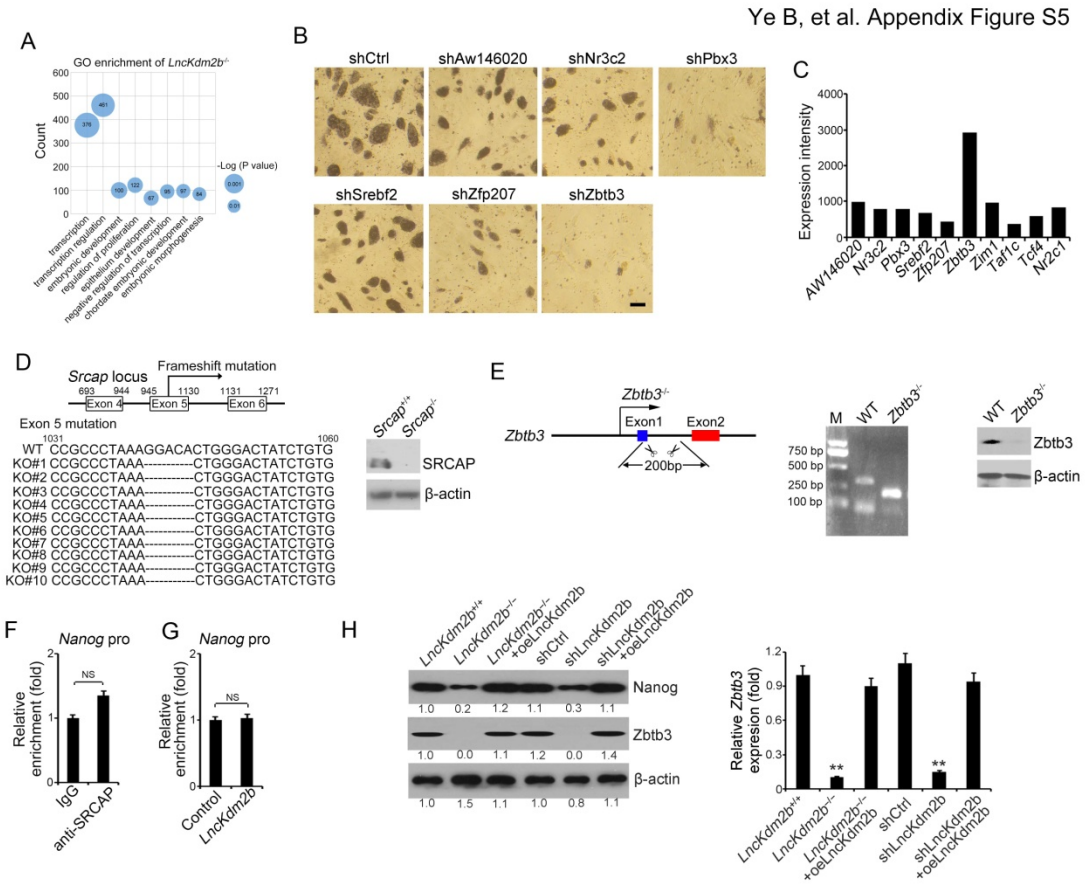
Appendix Figure S3. Confirmation of *LncKdm2b* KO strategy. (A) Chromatin modifications associated with active enhancers were mapped in the 800bp exon 2 deletion region. A published dataset GSE98063 was used to analyze modifications of H3K27ac (marking transcriptional active region) and H3K4me1 (marking active enhancer) on the 800bp region. (B) Enhancer capacity of 838bp exon 2 deletion region was measured by enhancer assay. Indicated fragment were cloned into upstream of GFP. Expression of GFP was analyzed by immunoblotting. (C) Enhancer capacity of 838bp exon 2 deletion region was validated by an enhancer reporter assay. The 838bp exon 2 region was cloned into a pGL3-promoter vector and transfected into 293T cell. Cells were lysed for 24 h after transfection, followed by luciferase activity assay. (D) polyA insertion ESC clones were generated by CRISPR/Cas9. gRNA sequences were listed in Appendix Table S2. Positive clones were confirmed by DNA sequencing and Northern blotting. (E)

polyA insertion in mouse ESC D3 cells abolished ESC self-renewal. Indicated ESCs were cultured in ESGRO-2i medium for 5 days followed by AP staining. (F) polyA insertion in mouse ESC D3 cells inhibited expression of pluripotency factors. Indicated ESCs were cultured in ESGRO-2i medium for 5 days followed by real-time PCR. Primers were listed in Appendix Table S1. Relative gene expression fold changes were counted as means ± S.D. **, $p<0.01$. (G, H) Human ESC H1 cells were infected with lentivirus expressing sh*lncKDM2B* and cultured in human ESC media for 3 weeks (G). Scale bar, 100 μm. mRNAs levels of the indicated genes were analyzed by real time qPCR (H). Relative gene expression fold changes were counted as means ± S.D. **, $p<0.01$. Primers were listed in Appendix Table S1. All data are representative of five independent experiments.



Appendix Figure S4. *LncKdm2b* interacts with SRCAP. (A) *LncKdm2b* deletion did not alter the expression of *Kdm2b* and other neighboring genes by RT-qPCR. Relative gene expression fold changes were calculated as means ± S.D. **, $p<0.01$. ns, no significance. Primers were listed in Appendix Table S1. (B) *LncKdm2b* deletion did not alter the expression of Kdm2b by immunoblotting.(C, D) Identification of SRCAP as a *lncKdm2b* binding protein. SRCAP protein sequences were identified by LTQ Orbitrap XL. (E) Mapping analysis of SRCAP-binding domains of *lncKdm2b*. Full-length and truncated fragments of *lncKdm2b* were in vitro transcribed to biotin-labeled RNA, followed by RNA pulldown and immunoblotting. Fragment 450-1000 was required for the interaction with SRCAP. (F) The region (450-700nt) of *lncKdm2b* harbored a stable stem-loop structure by RNA folding analysis. (G) Satb1 is not expression in mESCs. Satb1 expression was detected by qPCR. Primers were listed in Appendix Table S1. (H) ESCs treated for ChIP as in Fig. 4I. Anti-MBD3 antibody was incubated with treated lysates followed by size fractionation with sucrose gradient ultracentrifugation. Eluent gradients were examined by Western blot. Elution pattern of NuRD complex was not affected by lncKdm2b deletion. (I)

ESCs were lysed with/without RNase treatment and then treated with 1% formaldehyde for crosslinking as in Fig. 4I. Anti-SRCAP antibody was incubated with treated lysates for ChIP assays, followed by size fractionation with sucrose gradient ultracentrifugation. Eluent gradients were examined by Western blot. Complex assembly was impaired upon RNase treatment. Data are representative of five independent experiments.



Appendix Figure S5. *LncKdm2b* activates Zbtb3 transcription to initiates *Nanog* expression. (A) GO analysis of differentially expressed genes in *lncKdm2b*$^{-/-}$ versus WT control ESCs. (B) Mouse ESC D3 cells were infected with scrambled shRNA (shCtrl) and indicated shRNAs and cultured in mouse ESC media for 5 days, followed by AP staining. Scale bar, 50 μm. (C) *Zbtb3* was most highly expressed in mouse ESCs among the top 10 upregulated transcription factors by transcriptome microarray analysis. (D) Schematic diagram of *Srcap* KO construction by CRISPR/Cas9 technology in mESCs. Frameshift of *Srcap* was confirmed by DNA sequencing. SRCAP deletion was validated by immunoblotting. (E) Schematic diagram of *Zbtb3* KO construction by CRISPR/Cas9 technology in mESCs. Deletion of Zbtb3 was confirmed by genotyping and immunoblotting. (F) ESCs were lysed for ChIP assays with anti-SRCAP antibody. SRCAP protein enriched to *Nanog* promoter was examined by real time qPCR. Signals were normalized to input DNA. Enrichment fold changes were counted as means ± S.D. NS, no significance. (G) ESCs were lysed for ChIRP assays with biotin-labeled *lncKdm2b*(450-700) probe. *LncKdm2b* enriched to *Nanog* promoter was examined by real time qPCR. Signals were normalized to input DNA. Enrichment fold changes were counted as means ± S.D. NS, no significance. (H) LncKdm2b restoration in LncKdm2b KO or knockdown cells restore Zbtb3 and Nanog expression to its WT levels. Expression of Zbtb3 and Nanog were detected by immunoblotting. Fold changes of relative expression of indicated proteins compared with that of *lncKdm2b*$^{+/+}$ cells were caculated by band signal intensities using ImageJ (left panel). *Zbtb3* mRNA expression levels were examined by real time qPCR and relative

gene expression fold changes were counted as means ± S.D. **, $p<0.01$ (right panel). Data are representative of five independent experiments.

Appendix Table S1

Primer oligonucleotides and probes used in this study

Genes	Primer Forward	Primer Reverse
m*lncKdm2b*#1	5'- AGGAAATACAGCCAGTCC -3'	5'- ATTCCCTCCTGTGTATCG -3'
m*lncKdm2b*#2	5'- AACTCTTTAACAACTGTCCAGC -3'	5'- ATCCTCATTCGTCATCTCG -3'
m*lncKdm2b*#3	5'- TTAGCTGGAAGCCCACACC -3'	5'- TACTGAATGCCATCTGAAAAGG -3'
m*Pou5f1*	5'- TCAGTGATGCTGTTGATCAGG-3'	5'- GCTATCTACTGTGTGTCCCAGTC-3'
m*Nanog*	5'- AAACCAGTGGTTGAAGACTAGCAA-3'	5'- GGTGCTGAGCCCTTCTGAATC-3'
m*Sox2*	5'- CCGTTTTCGTGGTCTTGTTT-3'	5'- TCAACCTGCATGGACATTTT-3'
m*cMyc*	5'- ATGCCCCTCAACGTGAACTTC-3'	5'- CGCAACATAGGATGGAGAGCA-3'
m*Klf4*	5'- GTGCCCCGACTAACCGTTG-3'	5'- GTCGTTGAACTCCTCGGTCT-3'
m*Lin28a*	5'- GAAGAACATGCAGAAGCGAAGA-3'	5'- CCGCAGTTGTAGCACCTGTCT-3'
m*Sox18*	5'-CCTGTCACCAACGTCTCGC-3'	5'- GCAACTCGTCGGCAGTTTG-3'
m*Sall4*	5'- CCCTGGGAACTGCGATGAAG-3'	5'- TCAGAGAGACTAAAGAACTCGGC-3'
m*Zfp281*	5'- CCTTCCCCCAGAGTATGGTTA-3'	5'- GAGGGTCTGGCAGGTACTC-3'
m*Srcap*	5'-TCAGCTCCCAATCCTACAGAC-3'	5'- CCTCCTGCCCCACTAGAAT-3'
m*Zbtb3*	5'- TCCTACGTTGGGAGTCCTTAG-3'	5'- TCCTTCTTGAAAGTCATGTCCG-3'
m*Actb*	5'- TGACGGGGTCACCCACACTGTGCCCATCTA-3'	5'- CTAGAAGCATTTGCGGTGGACGATGGAGGG-3'
m*Gapdh*	5'- GCAGTTAAGTTCAGGAGCTTCAGG-3'	5'- GAAGCACGGTTGTATGTGCAAGTG-3'
m*Hmbs*	5'- ATGAGGGTGATTCGAGTGGG-3'	5'- TTGTCTCCCGTGGTGGACATA-3'
m*U1*	5'- CTTACCCACGATTCTCCATCTGT-3'	5'- CCGTTGTCTTGTCAATCAGGC-3'
m*Sox17*	5'- GATGCGGGATACGCCAGTG-3'	5'- CCACCACCTCGCCTTTCAC-3'
m*Bruchyury*	5'- CTCTAATGTCCTCCCTTGTTGCC-3'	5'- TGCAGATTGTCTTTGGCTACTTTG-3'
m*Fgf5*	5'- TGTGTCTCAGGGGATTGTAGG-3'	5'- GCAGTTAAGTTCAGGAGCTTCAGG-3'
m*Anapc5*	5'- TTGACCGCCTGATTCTCACTG-3'	5'- AGCTGTTTTCTTGGAATCTCTCC -3'
m*Rnf34*	5'- AAGGCGGGTGCTACTTCTATG-3'	5'- CACAGCAGACATGCTTCTTTCTA-3'
m*Kdm2b*	5'- GATGCTGAGCGGTATCATCCG-3'	5'- GAGACAGCGATCCATGAGCAG-3'
m*Tmem120b*	5'- AGGCACCTGAAGGACCTGAA-3'	5'- GGCCTCCATGTCAAAGAAGAC-3'
m*Rhof*	5'- AAGATAGTGATCGTAGGTGACGG-3'	5'- GTACTTCTCAAACACCGACGG-3'
mHpd	5'- CGCCTGTGTCACATCGCTT-3'	5'- CTCATTCACTAGAAAGACAGCGT-3'
m*Morn3*	5'- ATGGCCGTTTCTTCCACCTG-3'	5'- CCACCTTAGGTATGGGGAACT-3'
m*AW146020*	5'- CTGTGACCATCAGGAAGGGAT-3'	5'- AATCAGGAAACTTTTCTCGCCA-3'
m*Nr3c2*	5'- CTCCGGGACCGAACAGAGT-3'	5'- ACAACAACCCTTTGGTAGCAG-3'
m*Pbx3*	5'- CGAGGCGCAAGCAAAGAAAC-3'	5'- TGCCAAAAGCATATTGTCCAGT-3'
m*Srebf2*	5'- GCAGCAACGGGACCATTCT-3'	5'- CCCCATGACTAAGTCCTTCAACT-3'
m*Zfp207*	5'- CAGCAACAACTACAGAACCCC-3'	5'- TCCTCATCTGGATGGATCAACT-3'
m*Hdac8*	5'- ACTATTGCCGGAGATCCAATGT-3'	5'- CCTCCTAAAATCAGAGTTGCCAG-3'
m*Zim1*	5'- GGAGAACTACGAGAACCTGATCT-3'	5'- TGTCTTAGAATTGTCTGGCTTCC-3'
m*Taf1c*	5'- CGGTCCACTTGGCATGACT-3'	5'- GGCAAAGGACAAGGTCGGA-3'
m*Tcf4*	5'- CGAAAAGTTCCTCCGGGTTTG-3'	5'- CGTAGCCGGGCTGATTCAT-3'
h*POU5*	5'- CTTGCTGCAGAAGTGGGTGGAGGAA-3'	5'- CTGCAGTGTGGGTTTCGGGCA-3'
h*BMI1*	5'- CCACCTGATGTGTGTGCTTTG-3'	5'- TTCAGTAGTGGTCTGGTCTTGT-3'

h*NANOG*	5'- CAAAGGCAAACAACCCACTT-3'	5'- TCTGCTGGAGGCTGAGGTAT-3'
h*VIM*	5'- GGGACCTCTACGAGGAGGAG-3'	5'- CGCATTGTCAACATCCTGTC-3'
h*TH*	5'- TCATCACCTGGTCACCAAGTT-3'	5'- GGTCGCCGTGCCTGTACT-3'
h*FOXA2*	5'- GACAAGTGAGAGAGCAAGTG-3'	5'- ACAGTAGTGGAAACCGGAG-3'
h*T*	5'- TAAGGTGGATCTTCAGGTAGC-3'	5'- CATCTCATTGGTGAGCTCCCT-3'
m*Zbtb3 pro*#1	5'- GGTGTGCTGCAGTTTCTTGCTG-3'	5'- TCCACTGAGTCCTACTTATAC-3'
m*Zbtb3 pro*#2	5'- AGTCTGCTTATCTCATCCCTGA-3'	5'- TAGCAGTGGAACGACCTAGCA-3'
m*Zbtb3 pro*#3	5'- ATCAGGACTCAGATAATAGA-3'	5'- TTCCCCCACCCTGTACCTTC-3'
m*Zbtb3 pro*#4	5'- ACTCCATTCGTGGAACTATTTC-3'	5'- AATCCCAACACTTGGGAAATAG-3'
m*Zbtb3 pro*#5	5'- GATGCTCTGTCTTCTACCTC-3'	5'- TGCAGACCCAGCTAAAGTTCT-3'
m*Zbtb3 pro*#6	5'- TTGCACATGACCCCAGGATCT-3'	5'- CCACCTATAGAATTCTGGGTAT-3'
m*Zbtb3 pro#7*	5'- ACACTCCTGTAAGTAACCCCTC-3'	5'- CTTCTGTAAGGTCCCAATGAT-3'
m*Zbtb3 pro*#8	5'- ATCTTCATAGAGTTAGAGGTA-3'	5'- CCAAGCTCTGCCGAAACAGTG-3'
m*Zbtb3 pro*#9	5'- AGCCAATCCTGAGGTCTCTGA-3'	5'- CTTCTCCAAATGGGCAATCC-3'
m*Nanog pro*	5'- ATCCACCTGCCTCTGCCGCCTAA-3'	5'- GCATTGGTGTTTTGCCTGCATGG-3'
m*Sox2 pro*	5'- TTTTCGTTTTTAGGGTAAGGTACTGGGAAG-3'	5'- CCACGTGAATAATCCTATATGCATCACAAT-3'
m*cMyc pro*	5'- GAACAGGAAGCTGGGGAAAT-3'	5'- TGCAAGGAGGCTTTTCCTAA-3'
m*Klf4 pro*	5'- CCACGTGCGCCGAGTTTGTTT-3'	5'- GCTCTTTCGGCCGGGGAACTG-3'
m*Pou5f1 pro*	5'- GGAACTGGGTGTGGGGAGGTTGTA-3'	5'- AGCAGATTAAGGAAGGGCTAGGACGAGAG-3'

Genes	probe
m*lncKdm2b* Northern	m*lncKdm2b* 81-358 nt
m*lncKdm2b* REMSA	m*lncKdm2b* 450-700 nt
m*lncKdm2b*RNA-FISH#1	5'-TCGCCAGCATAGGTAATGAG-3
m*lncKdm2b*RNA-FISH#2	5'-ACTGGCTGTATTTCCTGTAC-3'
m*lncKdm2b*RNA-FISH#3	5'-CAGAGGAGGTTTTAGGCATC-3'
m*lncKdm2b*RNA-FISH#4	5'-AAAGAGGAACTACACCCCGT-3'
h*lncKDM2B*RNA-FISH#1	5'- GGCCGACTCAAATTCAAAG-3'
h*lncKDM2B*RNA-FISH#2	5'- CACTCAAAGATGTGGACAC-3'
h*lncKDM2B*RNA-FISH#3	5'- GGAAGACGTTGCCTTTCATTC-3'
h*lncKDM2B*RNA-FISH#4	5'-GCTCAGGAGCGAAAGACG-3'
m*lncKdm2b*CHIRP#1	5'- TCGCCAGCATAGGTAATGAG-3'
m*lncKdm2b*CHIRP#2	5'- TATGGCAGGTTCGCAAATGG-3'
m*lncKdm2b*CHIRP#3	5'- CTATTTCAGGGCTGTACTAC-3'
m*lncKdm2b*CHIRP#4	5'- GTTTCCTATAAAGGCTGCTC-3'
m*lncKdm2b*CHIRP#5	5'- CCATATTATTATTTCAGCCC-3'
m*nanog pro* EMSA	5'-CATATGGCCAGGCTGGTTAGAGGGATGCACTTGGTGGCCATTAGCCACCCTGGCCTGGT-3'

M, mouse; H, human.

Appendix Table S2

sgRNA or shRNA sequences for gene editing by CRISPR/Cas9 technology

Genes	gRNA-Up	gRNA-Down
m*lncKdm2b*$^{-/-}$	5'- GGTCAAAACTTGTATAAGATTGG-3'	5'- GACCGGATCTCTCCTATCAAAGG-3'
lncKdm2b-RFP	5'- CTTAGTCTCTGAGTTAGATGTGG-3'	
lncKdm2b-polyA	5'- TTGCAGGCTTCTGTGCCCTGTGG-3'	
Srcap$^{-/-}$	5'- CCATGAAGAGCAGCGACAAAAGG -3'	
Zbtb3$^{-/-}$	5'- CCTGCGGGAGCAACGGTCCC-3'	5'- TTTCTACAAGGAGCGGGAAT-3'
shLncKdm2b(m)	5'-GGTAATCACTTAGCTACAT	
shLncKDM2B (h)	5'- GGAAGACGTTGCCTTTCATTC	
shAw146020	5'- CCTCAAACACTCCGGCAAA	
shNr3c2	5'- CCGAGTTATGAGAACCCTT	
shPbx3	5'- GCGTCTTGTGTGAGATCAAAG	
shSrebf2	5'- GCTGGTAAATGGTGTGATT	
shZfp207	5'- GCGTCAACCACTAGTACAA	
shZbtb3	5'- CCTGCAAGACTTGTGGAAA	