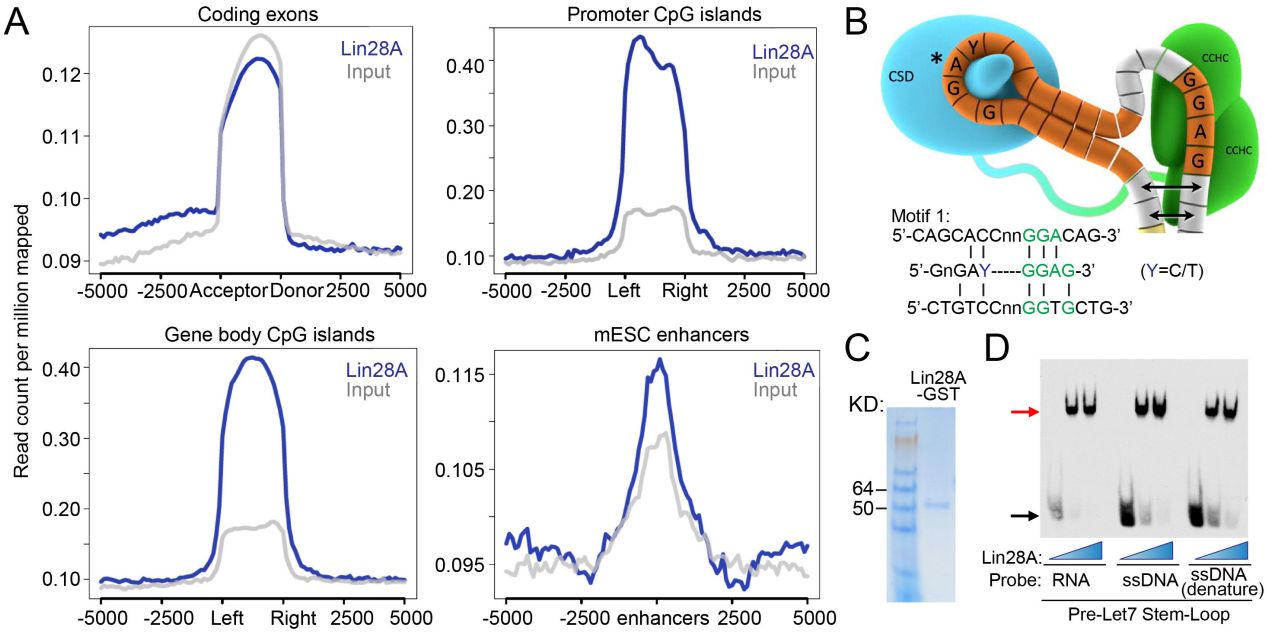

E

Endogenous Lin28A vs FLAG-Lin28A

Samples	# of peaks or # of bins	# of regions overlap with FLAG-Lin28	Percentage of overlapped bins to total selected bins
FLAG-Lin28 ChIP	7874 peaks		
Endogenous Lin28 ChIP	3111 bins	2682 bins	2682/3111=86.2%
Endogenous H3K4me1 ChIP	27778 bins	3414 bins	3414/27778=12.3%
Input	16698 bins	2008 bins	2008/16698=12.0%
Random selected bins from endogenous Lin28 ChIP-1	3111 bins	89 bins	89/3111=2.8%
Random selected bins from endogenous Lin28 ChIP-2	3111 bins	110 bins	110/3111=3.5%

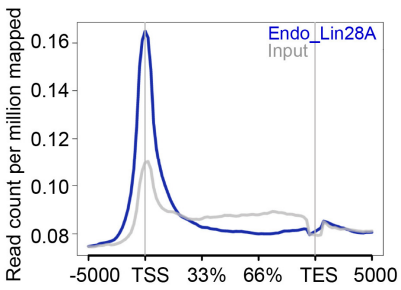

Figure S1. Enrichment of Lin28A binding in different genomic features, direct binding of Lin28A to Pre-let-7 RNA and DNA analogs in vitro, and comparison between Flag-Lin28A ChIP-seq and endogenous Lin28A ChIP-seq datasets, related to Figure 1.

(**A**) Normalized Lin28A ChIP-seq read densities on various genomic features including coding exons, promoter-associated CpG islands, gene body-associated CpG islands and defined mESC enhancers (Shen et al., 2012). 5 kb upstream and downstream of each feature were included. Regions between Acceptor and Donor define boundaries of coding exons. Regions between Left and Right define boundaries of corresponding genomic features.

(**B**) A structural model of Lin28A-DNA interaction proposed based on Lin28A-let7 microRNA (upper panel; adopted from published Lin28A structure (Nam et al., 2011)). Lin28A binds two distinct regions of pre-let7, the 5' GNGA(C/T) in stem-loop conformation and GGAG in single-strand conformation. Sequence alignment between let7 microRNA and DNA motifs identified in this study (lower panel).

(C) Coomassie Blue staining of purified recombinant Lin28A-GST protein used in EMSA.

(**D**) Binding of Lin28A to RNA and single-strand DNA forms of biotin-labeled pre-let7 stem-loop sequences. Equal amounts (0.1 pmol) of biotin-labeled RNA or DNA oligo were added to each reaction, together with increasing amounts (0, 0.5, 1.0 pmol) of Lin28A protein. Red arrow: bound probe; black arrow: unbound probe.

(**E**) FLAG-Lin28A ChIP-seq peaks were called by MACS. Endogenous Lin28A, H3K4me1 (negative control) and input ChIP-seq mapped reads were calculated in 10 kb bins, normalized to their sequencing coverage and ranked in descending order. Bins with a normalization value > 10 were selected to cross-compare with FLAG-Lin28A peaks. 86.2% of endogenous Lin28A ChIP-seq

selected bins overlapped with FLAG-Lin28A peaks whereas ~12% of H3K4me1 and input bins were correlated with FLAG-Lin28A. Two randomly selected pools of bins from endogenous Lin28A ChIP-seq showed 2-3% correlation with FLAG-Lin28A.
(**F**) Normalized endogenous Lin28A mapped read densities on UCSC mm9 Refseq gene bodies. Gene bodies were normalized to 0-100% as relative position. 5 kb upstream and downstream are also shown.

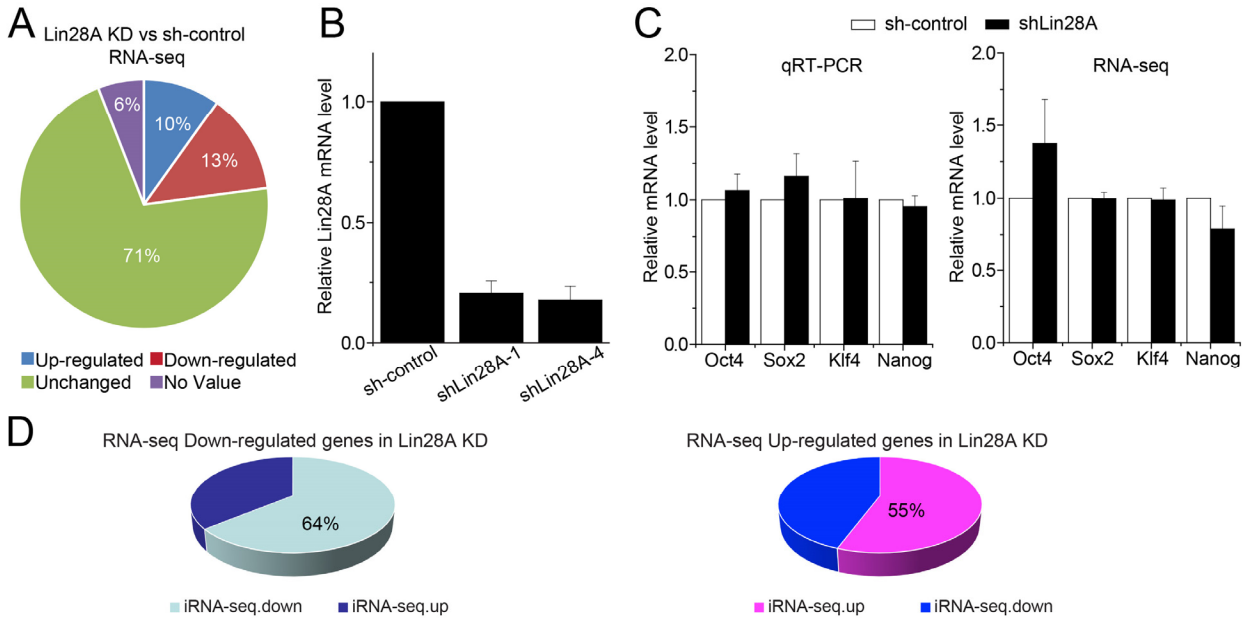

Figure S2. Effect of Lin28A knockdown on gene expression in mESCs, related to Figure 2.
(**A**) Percentage of genes that are up-regulated (10%), down-regulated (13%), unchanged (71%) or no value (6%) in RNA-seq of Lin28A KD cells compared to scrambled sh-control.
(**B**-**C**) Knockdown efficiency of Lin28A (**B**) and expression levels of pluripotency genes (**C**) in stable mESC lines. mRNA levels were normalized to control mESCs transduced with scrambled shRNA controls. Values represent mean ± SD. Expression levels of pluripotency genes were measured by both reverse transcription followed by quantitative PCR, and RNA-sequencing. Values represent mean ± SD (n = 3 for quantitative PCR; n = 2 for RNA-seq).
(D) Comparison between traditional RNA-seq and iRNA-seq analyses of Lin28A KD and shRNA control mESCs. RNA-seq data was further analyzed by iRNA-seq, which performs genome-wide assessment of acute transcriptional activity based on analysis of intron coverage from total RNA-seq data (Madsen et al., 2015).

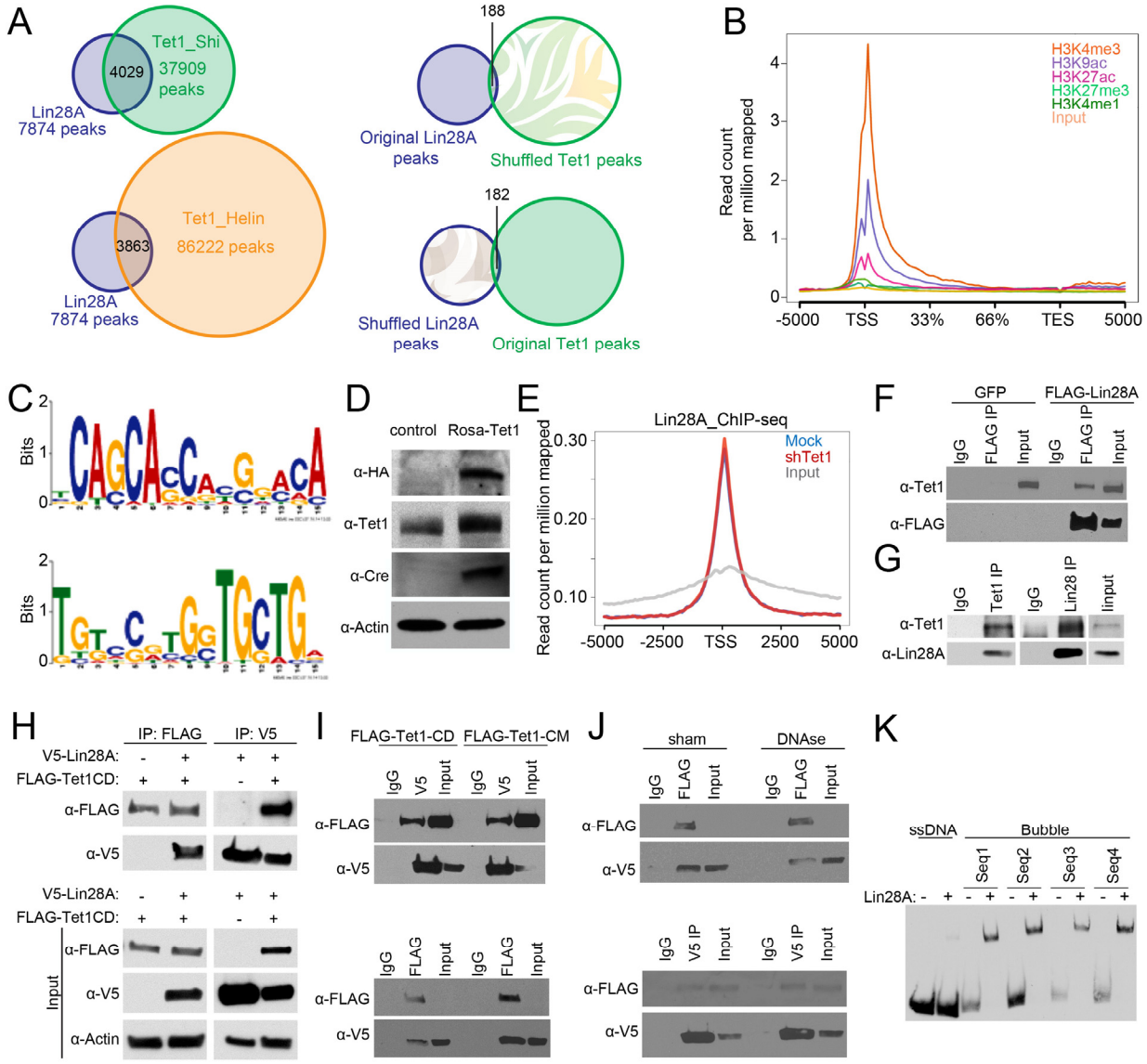

Figure S3. Features shared by Lin28A and Tet1 and their physical association, related to Figure 3.

(**A**) Substantial overlap of Lin28A-binding sites with Tet1 ChIP-seq peaks from two additional publications (Williams et al., 2011; Xu et al., 2011). Also shown is result from Shuffleband Command for the datset shown in **Figure 3A**. Venn diagrams were constructed proportional to the number of peaks.

(**B**) Enrichment of distinct histone markers on Lin28A-Tet1 co-bound genes. Gene bodies were normalized to 0-100% as relative position. 5 kb upstream and downstream are also shown.

(**C**) Tet1 ChIP-seq motifs were predicted using MEME suite from the top 1000 peaks based on their fold enrichment using a published dataset (Wu et al., 2011).

(**D**) mESCs with flox-stop-flox-HA-Tet1 knocked into the endogenous Rosa locus expressing Cre were analyzed for protein levels by immunoblotting.

(**E**) Tet1 KD did not affect Lin28A binding on target genomic loci. mESCs stably expressing FLAG-Lin28A were transfected with siRNA targeting Tet1 for 2 days, then evaluated for Lin28A-chromatin

binding by ChIP-seq. Mapped ChIP-seq reads within 5 kb upstream and downstream of transcription start sites (TSS) are shown.
(**F-J**) Lin28A was in a complex with Tet1 in mESCs and 293T cells, independent of catalytic activity of Tet1 or the presence of DNA. Binding of Lin28A with Tet1 was detected in mESCs stably expressing FLAG-tagged Lin28A (**F**), parental E14 mESCs (**G**), and 293T cells transfected with both V5-Lin28A and FLAG-Tet1CD (**H**). mESCs and 293T cells were lysed and immunoprecipitated by agarose coupled with antibodies specific to FLAG, V5, Lin28A or Tet1, then precipitated complexes were analyzed by immunoblotting. To test whether catalytic activity of Tet1 is required for binding to Lin28A (**I**), V5-Lin28A was co-transfected with FLAG-Tet1 catalytic domain with (CD) or without (CM) catalytic activity. To test whether the presence of DNA is required for Lin28A to bind Tet1 (**J**), DNase I (RNase-free) was added to 293T cell lysate, then complexes immunoprecipitated by agarose coupled with FLAG or V5 antibodies were analyzed by immunoblotting.
(**K**) Affinity of Lin28A to bubble DNAs with different insertion sequences (CACAA, TTTTT, GGGAA, CCCTT). Seq 1 is the one used in **Figure 3E**.

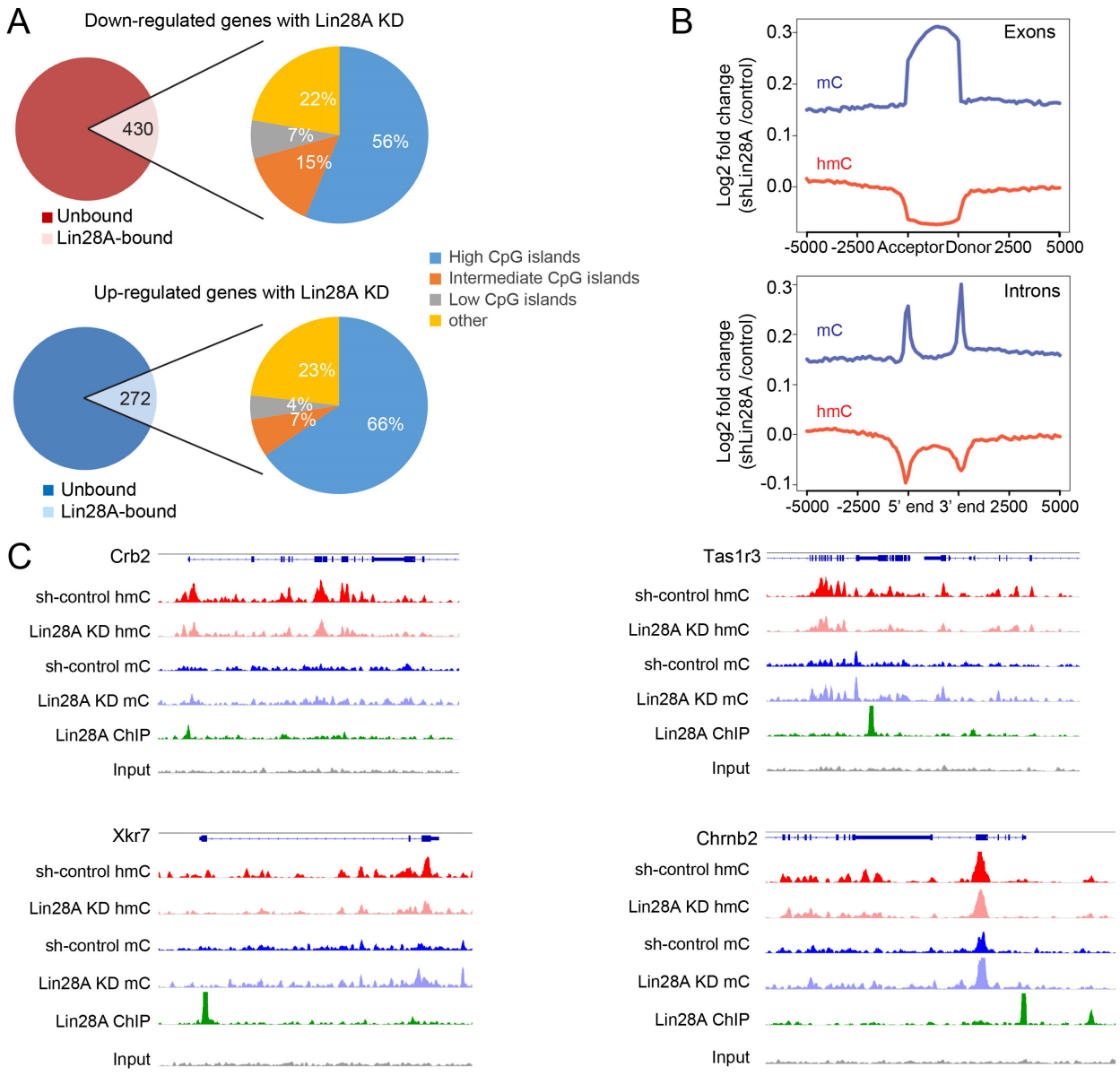

Figure S4. CpG islands abundance in Lin28A-regulated genes, exon/intron methylation features, and four additional ChIP-seq examples of genes that displayed reduced expression upon Lin28A KD in mESCs, related to Figure 4.

(**A**) Lin28A-bound genes, either down-regulated or up-regulated by Lin28A KD in mESCs were analyzed for abundance of CpG islands (Ku et al., 2008).

(**B**) 5hmC and 5mC dynamics in the exons and introns with or without Lin28A KD in mESCs.

(**C**) Locus-specific 5hmC and 5mC abundance in mESCs with scrambled sh-control or Lin28A shRNA are shown. FLAG-Lin28A ChIP-seq peaks are also shown side by side for comparison. These genes were bound by Lin28A and showed inverse changes of 5hmC and 5mC, mostly on their gene bodies.

Table S1. Summary Table for Lin28A ChIP-seq of FLAG-Lin28A mESCs and RNA-seq of Lin28A KD mESCs, related to Figure 1. (See Excel file)

Table S2. Motif predction of ChIP-seq of FLAG-Lin28A mESCs with three algorithms, related to Figure 1.

Three motif prediction algorithms were utilized to predict Lin28A DNA binding motifs. "**GGACAG**" was consistently predicted as the top hit. The significance index, e-value or p-value is indicated.

LSAT motif	Sig. score
	58.32
	43.07
	26.06
	20.11
	16.68
MEME motif	**E-value**
	6.3e-156
	9.0e-035
HOMER motif	**p-value**
	1e-1423
	1e-882
	1e-88
	1e-82
	1e-43
	1e-35

SUPPLEMENTARY EXPERIMENTAL PROCEDURES

Cell Cultures

The E14 mESC line was cultured on Mouse Embryonic Fibroblast (MEF) feeder cells as previously described (Ma et al., 2008). The medium contained DMEM supplemented with 15% FBS, GlutaMAX™-I (2 mM), sodium Pyruvate (1 mM), MEM nonessential amino acid (0.1 mM), Penicillin/Streptomycin (50 ug/ml each), β-mercaptoethanol (100 uM) and ESGRO-LIF (Millipore, ESG1106; 1000 U/ml). 293T cells were cultured in DMEM supplemented with 10% FBS and Penicillin/Streptomycin. mESCs were transfected by electroporation using Mouse ES Cell Nucleofector® Kit from Lonza (VPH-1001). 293T cells were transfected by calcium phosphate.

Lentiviral Transduction

Lentiviral pLKO.1 vectors encoding shRNAs targeting mouse Lin28A (Open Biosystems) were co-transfected with pMDL gag/pol RRE, pRSV-Rev and pCMV VSV-G into 293T cells using the calcium phosphate method. Transfected cells were incubated at 37°C in DMEM supplemented with 10% newborn calf serum for 24 hours. The medium was replaced with DMEM supplemented with 2% newborn calf serum and incubated for an additional 24 hours at 37°C. Viral supernatants were collected and filtered through a 0.45-μm syringe filter, supplemented with 10% FBS and 10 μg/mL polybrene. mESCs were infected with the viral supernatant and selected by 3 μg/ml puromycin for 3 days.

RNA Extraction, Quantitative-RT-PCR and Genomic DNA Extraction

Total RNA was extracted from mESCs by RNeasy (Qiagen), and 1 μg total RNA was primed with Oligo(dT)20 primers (Invitrogen) for cDNA synthesis. The cDNA was added to a q-RT–PCR mixture that contained 1× SYBR Green PCR master mix (Applied Biosystems) and 150 nM gene-specific primers. Assays were performed in triplicate on a 7900 HT sequence detection system (Applied Biosystems). The PCR includes 2 min incubations at 50°C and 10 min at 95°C, followed by 40 cycles, each consisting of 15 sec at 95°C and 1 min at 60°C. The expression level of each gene was normalized by the amount of glyceraldehyde-3-phosphate dehydrogenase (GAPDH).

Total genomic DNA of mESCs was extracted by Qiagen DNeasy Blood and Tissue Kit as previously described (Guo et al., 2011). Extracted genomic DNA was then subject to HpaII digestion, 5hmC-capture or MeDIP.

Chromatin Immunoprecipitation (ChIP)

Five million mESCs were treated with 1% formaldehyde for 10 min at room temperature with gentle shaking, then 0.125 M final concentration of Glycine was added to stop the fixation, and cells were incubated for an additional 5 min at room temperature before lysed on ice in a NP-40 lysis buffer (10 mM Hepes/pH7.9, 0.5% NP-40, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT and protease inhibitor cocktail). After centrifugation at 4000 rpm for 5 min, the nuclear pellets were lysed by sonication on ice in a nuclear lysis buffer (20 mM Hepes/pH7.9, 25% glycerol, 0.5% NP-40, 0.5% Triton X-100, 0.42 M NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA and protease inhibitor cocktail), then centrifuged at 13,000 rpm for 10 min in a cold room. The resulting nuclear lysate was diluted with 2 volumes of dilution buffer (0.01% SDS, 1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl, 16.7 mM Tris-HCl/pH8.0 and protease inhibitor cocktail). Immunoprecipitation was performed with an antibody specific to FLAG (M2, Sigma), HA (Sigma), Lin28A (Abcam) or Tet1 (Millipore) for 6 hours or overnight at 4ºC, with normal mouse or Rabbit IgG (Millipore) as a control. After antibody incubation, 20 μl salmon sperm DNA/protein G agarose (Upstate) were added and incubated for another 1 hour. Precipitates were sequentially washed with TSE I (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl/pH 8.1), TSE II (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl/pH 8.1), TSE III (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl/pH 8.1), and washed twice with TE buffer before eluted with 1% SDS and 0.1 M $NaHCO_3$. The elution was incubated at 65 ºC for at least 6 hours to reverse the formaldehyde

cross-linking, then DNA fragments were purified by using PCR Purification Kit (Qiagen). ChIPed DNA was either subjected for library preparation or qPCR validation. For ChIP-q-PCR, purified DNA was added to a q-RT–PCR mixture that contained 1× SYBR Green PCR master mix and 150 nM gene-specific primers. Assays were performed in triplicate on a 7900 HT sequence detection system.

Electrophoretic Mobility Shift Assay (EMSA)

Each binding reaction was carried out with 100 fmol of biotinylated dsDNA probe and 1 pmol of purified protein in 20 μl of binding buffer containing 25 mM HEPES at pH 8.0 with 50 mM KGlu, 0.1% Triton X-100, 2 mM $MgAC_2$, 3 mM DTT, 50 ng/μl PolydI.dC and 5% glycerol as previously described (Guo et al., 2014). Human Lin28A protein was purified as GST-His_6 fusion proteins from yeast using a high-throughput protein purification protocol as described previously (Zhu et al., 2001). Reactions were carried out for 30 min at room temperature, followed by overnight incubation at 4°C. Reaction mixtures were loaded onto 6% DNA retardation gels (Invitrogen) and separated at 100 V until the dye front migrated two-thirds of the way to the bottom of the gel. Nucleic acids were transferred to nylon membranes (Pierce, USA) and visualized with the LightShift EMSA Kit (Pierce, USA) according to the manufacturer's recommendations. Control DNA is a 60 bp biotin end-labeled EBNA duplex in the Pierce kit. Sequences of 5'-biotinylated DNA probes are: Motif1 ssDNA (sense), 5'-CTTCTGGCGCTGTCCATGGTGCTGAACCCAG-3', antisense 5'-CTGGGTTCAGCACCATGGACAGCGCCAGAAG-3'; Motif2 ssDNA (sense), 5'- GAA ACTGTCCGTGGTGCTGAAACACAGCAGCA-3', antisense 5'-TGCTGCTGTGTTTCAGCACCACGGACAGTTTC-3'. dsDNA was made by annealing sense strand and antisense strand ssDNA. Mismatch duplex of motif1 was constructed by annealing sense strand with mutant antisense strand (5'-CTGGGTTCAGCACCATTTTTTGCGCCAGAAG-3'); bubble duplex of motif1 was made by annealing mutant sense strand (5'-CTTCTGGCGCTGTCCATCACAAGGTGCTGAACCCAG-3') with mutant antisense strand (5'-CTGGGTTCAGCACCATTTTTTGCGCCAGAAG-3').

***K_D* Measurement by Octet® QK^e system**

Octet® QK^e systems (FortéBIO) utilize the recently developed technique of Bio-Layer Interferometry (BLI), which enables real-time, label-free analysis of affinity, kinetics and concentration from biomolecular interactions in 96- or 384-well microplates. Streptavidin (SA) Biosensors were pre-wetted in the buffer for 10 minutes before use and system temperature was set at 25°C. First, we scanned 120s baseline in the buffer for each SA biosensor by OCTET QK system and loaded 500 nM dsDNA or bubble DNA for 500 s in the 96 well-plate. One of the biosensors was set as reference to remove the atmosphere drift. Then we scanned another 120 s baseline in the buffer and dipped the biosensors in Lin28A protein solutions and control buffer/reference, before we scanned the association curves for 1200 s. After 1200 s of association in the Lin28A protein solution, we moved the biosensors to buffer wells to monitor the dissociation curves for 2400 s. We used the data analysis software of the OCTET QK system to extract association and dissociation curves for Lin28 protein and fit the curves by 1:1 model to obtain the KD for each condition.

Co-mmunoprecipitation

mESCs and 293T cells were lysed with NP-40 lysis buffer (50 mM Tris-HCl/pH 8.0, 150 mM NaCl, 1% NP-40, 1 mM DTT, and protease inhibitor cocktail) and pre-cleared with protein-G agarose for 1 hour at 4^{o}C. Immunoprecipitation was carried out with antibodies specific to FLAG (M2, Sigma), V5 (Invitrogen), Lin28A (Abcam) or Tet1 (Millipore) for overnight at 4^{o}C, with normal mouse or Rabbit IgG (Millipore) as a control. Twenty microliters of protein G agarose (Upstate) were then added and incubated for another 1 hour. Precipitates were washed with NP-40 lysis buffer for 4 times followed by PBS wash for twice. SDS-loading buffer was added to precipitates and boiled at 95^{o}C for 1 min to release the protein complex.

5hmC-specific Enrichment and Methylated DNA Immunoprecipitation (MeDIP)
Five μg of genomic DNA was sonicated to 100-500 bp, and then mixed with 100 μl reaction buffer (50 mM HEPES at pH 8.0, 25 mM $MgCl_2$, 250 μM UDP-6-N_3-Glu and 2.25 μM wild-type β-glucosyltransferase (β-GT)). Reactions were incubated at 37°C for 1 hour, and DNA substrates were then purified by Qiagen DNA purification kit. 150 μM dibenzocyclooctyne modified biotin was mixed with β-GT-modified DNA. The labelling reaction was performed at 37°C for 2 hours. The biotin-labeled DNA was enriched by Strepavidin-coupled Dynabeads (Dynabeads MyOne™ Streptavidin T1, Life Technologies) and purified by Qiagen DNA purification kit. MeDIP experiments were performed according to the manufacturer's protocol (Active Motif). The enriched methylated DNA was purified by Qiagen DNA purification kit. 5hmC- and 5mC-containing DNA was subjected to library preparation.

Library Preparation and High-throughput Sequencing
Enriched DNA from ChIP, 5hmC-capture and MeDIP was subjected to library construction using the NEBNext ChIP-Seq Library Prep Reagent Set for Illumina according the manufacturer's protocol. In brief, 25 ng of input genomic DNA or experimental enriched DNA were utilized for each library construction. 150-300 bp DNA fragments were selected by AMPure XP Beads (Beckman Coulter) after the adapter ligation. An Agilent 2100 BioAnalyzer was used to quantify the amplified DNA, qPCR was applied to accurately quantify the library concentration. 20 pM diluted libraries were used for sequencing. 50-cycle single-end sequencings were performed using Illumina HISeq 2000. Image processing and sequence extraction were done using the standard Illumina Pipeline.

RNA-seq libraries were generated from duplicated samples per condition using the Illumina TruSeq RNA Sample Preparation Kit v2 following manufacturer's protocol. The RNA-seq libraries were sequenced as 50-cycle pair-end runs using Illumina HISeq 2000.

Bioinformatics
Bioinformatics analyses for ChIP-seq, 5hmC-seq and MeDIP-seq were described previously (Szulwach et al., 2011; Yao et al., 2014). Briefly, FASTQ sequence files were aligned to mm9 reference genome using Bowtie (Langmead et al., 2009). Peaks were identified by Model-based Analysis of ChIP-Seq (MACS) software (Zhang et al., 2008). Unique ChIP-seq, 5hmC-seq and MeDIP-seq mapped reads were plotted to various genomic regions by ngsplot (Shen et al., 2014). Annotation and motif analysis were performed by Hypergeometric Optimization of Motif EnRichment (HOMER) v4.7 (http://homer.salk.edu/homer/) (Heinz et al., 2010) or MEME suite 4.10.1 (http://meme-suite.org/tools/meme) (Bailey et al., 2009). RNA-seq reads were aligned using tophat v2.0.8 (Trapnell et al., 2009) and differential RPKM expression values were extracted using cuffdiff v2.2.1 (Trapnell et al., 2009). Acute transcription was further evaluated using iRNA-seq package (Madsen et al., 2015).

Quantification of 5mC/5hmC by Glucosylation and HpaII/MspI Digestion
Two μg of gDNA were treated with Epimark 5hmC and 5mC Analysis Kit (New England Biolabs) according to the manufacturer's protocol. For quantitative PCR, 1 μl final product of treated DNA was added to a Q-PCR mixture that contained 1× SYBR Green PCR master mix and 150 nM gene-specific primers.

Statistic Analyses
T-tests were performed in Figure 2E, 3B, 3E and 4E, and significance was indicated by an asterisk. ShuffleBed command from bedtools was used to generate random regions for Figure S3A, and randomization statistics were performed using R. Three motif prediction algorithms were utilized to predict Lin28A DNA binding motifs. For LSAT, "significance index" instead of P-value was used because it is more accurate, considering the false discovery rate. With a sig > 0, we expect 1 FP per analysis; with a sig > 1, we expect 1FP per 10 analyses and with a sig > 5, we expect 1 FP

every 105 analyses. MEME usually finds the most statistically significant (low E-value) motifs first. The E-value of a motif is based on its log likelihood ratio, width, sites, the background letter frequencies (given in the command line summary), and the size of the training set. The E-value is an estimate of the expected number of motifs with the given log likelihood ratio (or higher), and with the same width and site count, that one would find in a similarly sized set of random sequences. Homer uses p-values to define significance, and the lowest p-value indicates the most significant motif prediction.

SUPPLEMNTARY REFERENCES

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME SUITE: tools for motif discovery and searching. Nucleic acids research *37*, W202-208.

Guo, J.U., Su, Y., Shin, J.H., Shin, J., Li, H., Xie, B., Zhong, C., Hu, S., Le, T., Fan, G.*, et al.* (2014). Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat Neurosci *17*, 215-222.

Guo, J.U., Su, Y., Zhong, C., Ming, G.L., and Song, H. (2011). Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. Cell *145*, 423-434.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell *38*, 576-589.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S.*, et al.* (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS genetics *4*, e1000242.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biology *10*, R25.

Ma, D.K., Chiang, C.H., Ponnusamy, K., Ming, G.L., and Song, H. (2008). G9a and Jhdm2a regulate embryonic stem cell fusion-induced reprogramming of adult neural stem cells. Stem cells *26*, 2131-2141.

Madsen, J.G., Schmidt, S.F., Larsen, B.D., Loft, A., Nielsen, R., and Mandrup, S. (2015). iRNA-seq: computational method for genome-wide assessment of acute transcriptional regulation from total RNA-seq data. Nucleic Acids Res *43*, e40.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular basis for interaction of let-7 microRNAs with Lin28. Cell *147*, 1080-1091.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC genomics *15*, 284.

Shen, Y., Yue, F., McCleary, D.F., Ye, Z., Edsall, L., Kuan, S., Wagner, U., Dixon, J., Lee, L., Lobanenkov, V.V.*, et al.* (2012). A map of the cis-regulatory sequences in the mouse genome. Nature *488*, 116-120.

Szulwach, K.E., Li, X., Li, Y., Song, C.X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I.*, et al.* (2011). 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. Nat Neurosci *14*, 1607-1616.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.

Williams, K., Christensen, J., Pedersen, M.T., Johansen, J.V., Cloos, P.A., Rappsilber, J., and Helin, K. (2011). TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature *473*, 343-348.

Wu, H., D'Alessio, A.C., Ito, S., Xia, K., Wang, Z., Cui, K., Zhao, K., Sun, Y.E., and Zhang, Y. (2011). Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. Nature *473*, 389-393.

Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A.J., Zheng, L., Zhang, H., Huang, S.*, et al.* (2011). Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. Mol Cell *42*, 451-464.

Yao, B., Lin, L., Street, R.C., Zalewski, Z.A., Galloway, J.N., Wu, H., Nelson, D.L., and Jin, P. (2014). Genome-wide alteration of 5-hydroxymethylcytosine in a mouse model of fragile X-associated tremor/ataxia syndrome. Human molecular genetics *23*, 1095-1107.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W.*, et al.* (2008). Model-based analysis of ChIP-Seq (MACS). Genome biology *9*, R137.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T.*, et al.* (2001). Global analysis of protein activities using proteome chips. Science *293*, 2101-2105.