Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain

Junjie U. Guo, Yijing Su, Joo Heon Shin, Jaehoon Shin, Hongda Li, Bin Xie, Chun Zhong, Shaohu Hu, Thuc Le, Guoping Fan, Heng Zhu, Qiang Chang, Yuan Gao, Guo-li Ming and Hongjun Song

Supplementary Information

Contents:

Supplementary Table 1. Primer sequences used in this study.

a. Quantitative PCR (FspEI digestion)	
Car10 (Forward)	GGAAATATGAGGGAGACATTAG
Car10 (Reverse)	CCACTTCCTTTCCCAAGTATC
Lrba (Forward)	GAGCTTTACAAACAATTTCTTATC
Lrba (Reverse)	GTTCTCTCGTGGTTTTCTTTTC
Ptprt (Forward)	CACAGAACAAGCTCAGAAGCAC
Ptprt (Reverse)	TCACCCATCTGATTTCAGTCC
Slmo1 (Forward)	GCAGGACTGTGCTCACTCAA
Slmo1 (Reverse)	GAGGCGTTCACCTTCTTAGC
Phf17 (Forward)	CTGTATCTCAGTGTATAGGCC
Phf17 (Reverse)	GAATTATAGAGAAAGGGGTAAC
Lgr6 (Forward)	AGGTGGGGACAGAGCAATG
Lgr6 (Reverse)	GTTGGGAAAGCTTCCTAGCG
b. Sanger bisulfite sequencing (mouse)	
mFzd3 (+) (Forward)	GTGGAGTAGGGATGAAGTTAGGAAA
mFzd3 (+) (Reverse)	CAACAATCATTATAAAATTACTAAAACCCA
mSox6 (+) (Forward)	AAGTTTGGTTAAGGGATTGAGGTAGG
mSox6 (+) (Reverse)	CTTTCACATCCTAAACATCCATAAAAATTC
mOpcml (+) (Forward)	GGGTTGGGGTTATAGGTTATTATGTTG
mOpcml (+) (Reverse)	ACCAAAAAACAATTCCCATTTCTTC
mC1ql1 (+) (Forward)	TTTTGGAATTAAAGAATTAGGTGTT
mC1ql1 (+) (Reverse)	TAAAACCAAAAAACCCACATAAAAA
mLrba (+) (Forward)	TTTTTGTTTTGGAATTATGTGGAAT
mLrba (+) (Reverse)	CAAAAAAAACTTAAAAACTACAACTTCCT
mFzd3 (−) (Forward)	TGAATTTTTATTTTTTATTTATTTAAAGAT
mFzd3 (−) (Reverse)	CAACAAATTCTACTATAAATTCTTCATATC
mSox6 (−) (Forward)	TGTATGGTTATGAAGTTTTTTATTTGTTGT
mSox6 (−) (Reverse)	AATATAATTTCCCCAATATCCAATATCCTA
mOpcml (−) (Forward)	AATTTTGGTTGTATTGTGGATGTTT
mOpcml (−) (Reverse)	TCAAACTCAAAATTATCTTATCTTCAACTC
mC1ql1 (−) (Forward)	AGGAAAGGAAAGGAAAGGAAAGAAG
mC1ql1 (−) (Reverse)	CCAAACCTCCAATACCCTCCTCTAA
mBdnf IV (+) (Forward)	GTGAATTTGTTAGGATTGGAAGTGAAAATA
mBdnf IV (+) (Reverse)	TTACCCACTACTCAAATCACACC
mFgf1B (+) (Forward)	TTTTGTTGATGAGTAAGGGTTAAGG
mFgf1B (+) (Reverse)	CAAACTAAAAAACTCTCTTCACTCCA
mHdac4 (+) (Forward)	TTAGGAAGTAGGAGTGGATGGGTTG
mHdac4 (+) (Reverse)	CTCACTAAAAAATTAATACTTTCTTCATCA
mMegf11 (+) (Forward)	ATGTGTAGATATGTGAGGGTATTAGGTATT
mMegf11 (+) (Reverse)	CTAAACAAACCAAAAAAAACAAACC
mBai3 (+) (Forward)	AATGAAAGTAATAGGAGTTAGGGGG
mBai3 (+) (Reverse)	ACCCAATTCATAATTAAAATAAAACATCTT
mSamd4 (+) (Forward)	TGAGTGTTATATTTATAAATAGTTGAAGGT
mSamd4 (+) (Reverse)	AAATATACCTCAAAAACAAACATAATT

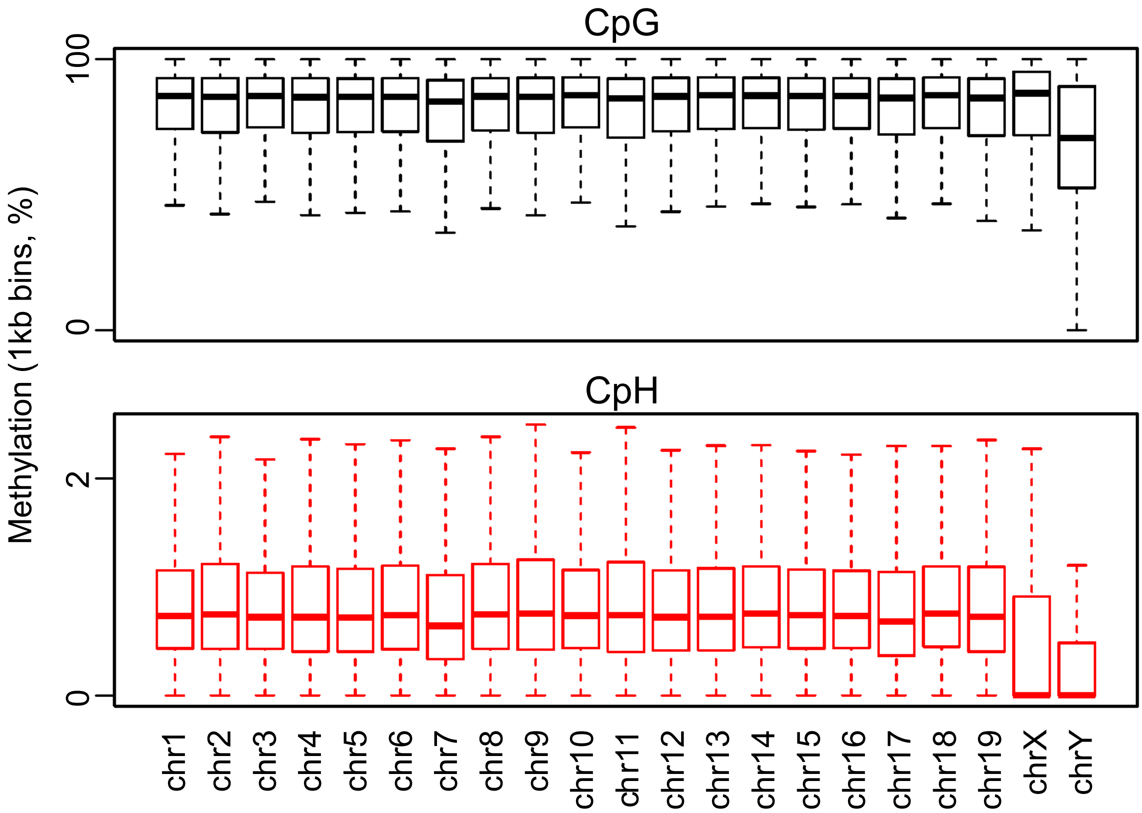
c. Sanger bisulfite sequencing (human)	
hC1QL1 (−) (Forward)	TGAGGTAGAAGAATGGTTGGAATTTG
hC1QL1 (−) (Reverse)	AACCTACCCAACAACCCCTAACATC
hOPCML (−) (Forward)	TTTAGTGGAGTTGTTTTGGTGTTGG
hOPCML (−) (Reverse)	CCCAAATATATATAATCCCTCTCTCCTTTA
hLRBA (−) (Forward)	TTATGATTGAAATGTTTGAAAATGTATTAG
hLRBA (−) (Reverse)	AACACTACAAAAACATAAATATCCTCATTT
hSOX6 (+) (Forward)	AAGTTTTGTTGGGTTTTGGGTAT
hSOX6 (+) (Reverse)	AACATATAAAAAAAACAAATTCACC
d. Quantitative PCR (gene expression)	
Dnmt1 (Forward)	AAGAATGGTGTTGTCTACCGAC
Dnmt1 (Reverse)	CATCCAGGTTGCTCCCCTTG
Dnmt3a (Forward)	TACATCAGCAAACGGAAACG
Dnmt3a (Reverse)	AGACTCTCCAGAGGCCTGGT
Actb (Forward)	TAGGCACCAGGGTGTGATGG
Actb (Reverse)	CATGGCTGGGGTGTTGAAGG
C1ql1(Forward)	ATCAGCACGGCCACCTATAC
C1ql1(Reverse)	CGCATCGTAGTTGTTGCCTA
Fzd3(Forward)	GAAGCAAAGCAGGGAGTGTC
Fzd3(Reverse)	ATGGCTGCCGTGAGGTAGTCT
Lrba(Forward)	GCAAGAAAACACACCAGCAGA
Lrba(Reverse)	GGAAGAAGCGACAGACAGACC
BDNF IV(Forward)	CAGGAGTACATATCGGCCACC
BDNF IV(Reverse)	TGGTCATCACTCTTCTCACCTG
BDNF IX(Forward)	GCAGCTGGAGTGGATCAGTAA
BDNF IX(Reverse)	TGGTCATCACTCTTCTCACCTG
Fgf1B (Forward)	GAGAGGCAGCTTCAGTCCAG
Fgf1B (Reverse)	TCACAAGACGGGAATGAAGTC

e. Primers for Dnmt3a ChIP-qPCR

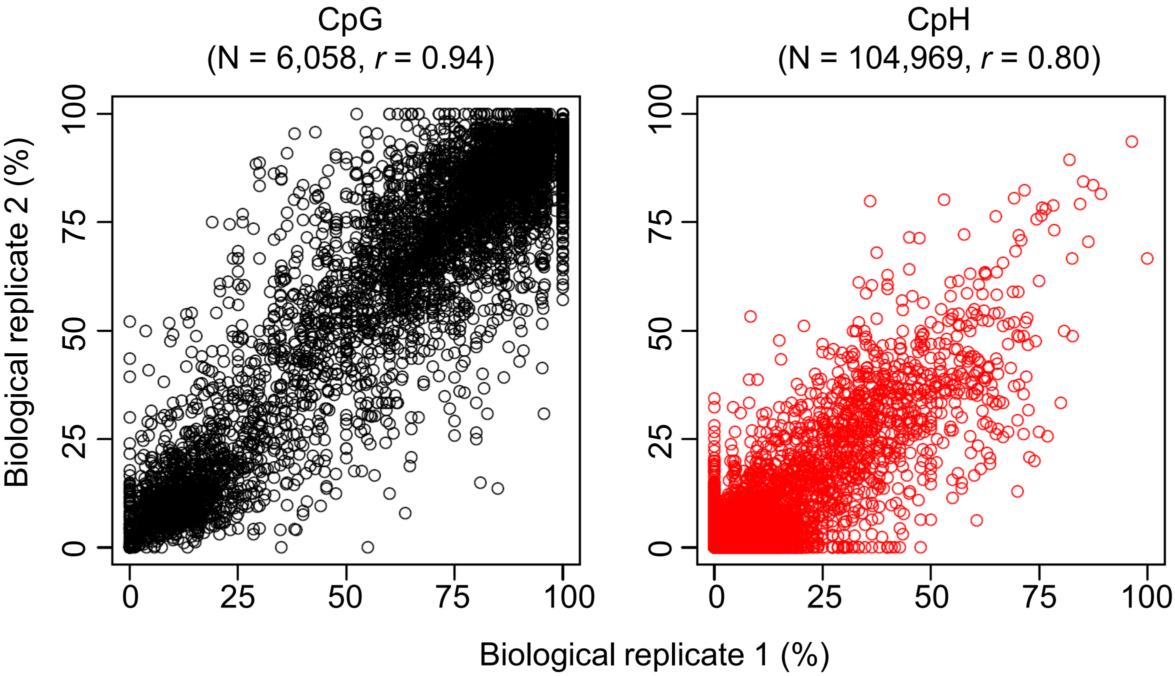
Car10 (forward)	TGAAACCCATTATTTTTGTATGC
Car10 (reverse)	GTCCAGTCGCTATTTGTCCA
Fzd3 (forward)	TGCTGCTTTATCAGCCTTTACTT
Fzd3 (reverse)	TTCCAAGGCCTGATGTCTCT
Lrba (forward)	GCTGAAAGCTGCACCACTCT
Lrba (reverse)	GTCCTTCTAGGAGGGGCAAA
Opcml (forward)	TCCACAGGTTGAGCATTTCA
Opcml (reverse)	TTTGAATCTTCCACCAGCTTC
Sox6 (forward)	TGGTCTCATTTTCGTCATTTCA
Sox6 (reverse)	GCTTCCAATTGCTTATGCTTTC
Actb (forward)	CATGGTGTCCGTTCTGAGTG
Actb (reverse)	CAGCTTCTTTGCAGCTCCTT
Mapk (forward)	GAGACTCCGCCCTCTCTACC
Mapk (reverse)	AATTAACCGCCGGTAGAACC
Tubb2a (forward)	GGATGGGACTACCTCATCCA
Tubb2a (reverse)	CTCCACCCCTTCTACAACCA
Ubb (forward)	CCACCTCAAGCAGGAAACAT
Ubb (reverse)	TTCTCGGGCAGTTTAACGTC
Ubc (forward)	CGAAAGCGACAGGCTAAAAC
Ubc (reverse)	CACACAAAGCCCCTCAATCT

Supplementary Table 2. Protein-DNA interaction data sources.

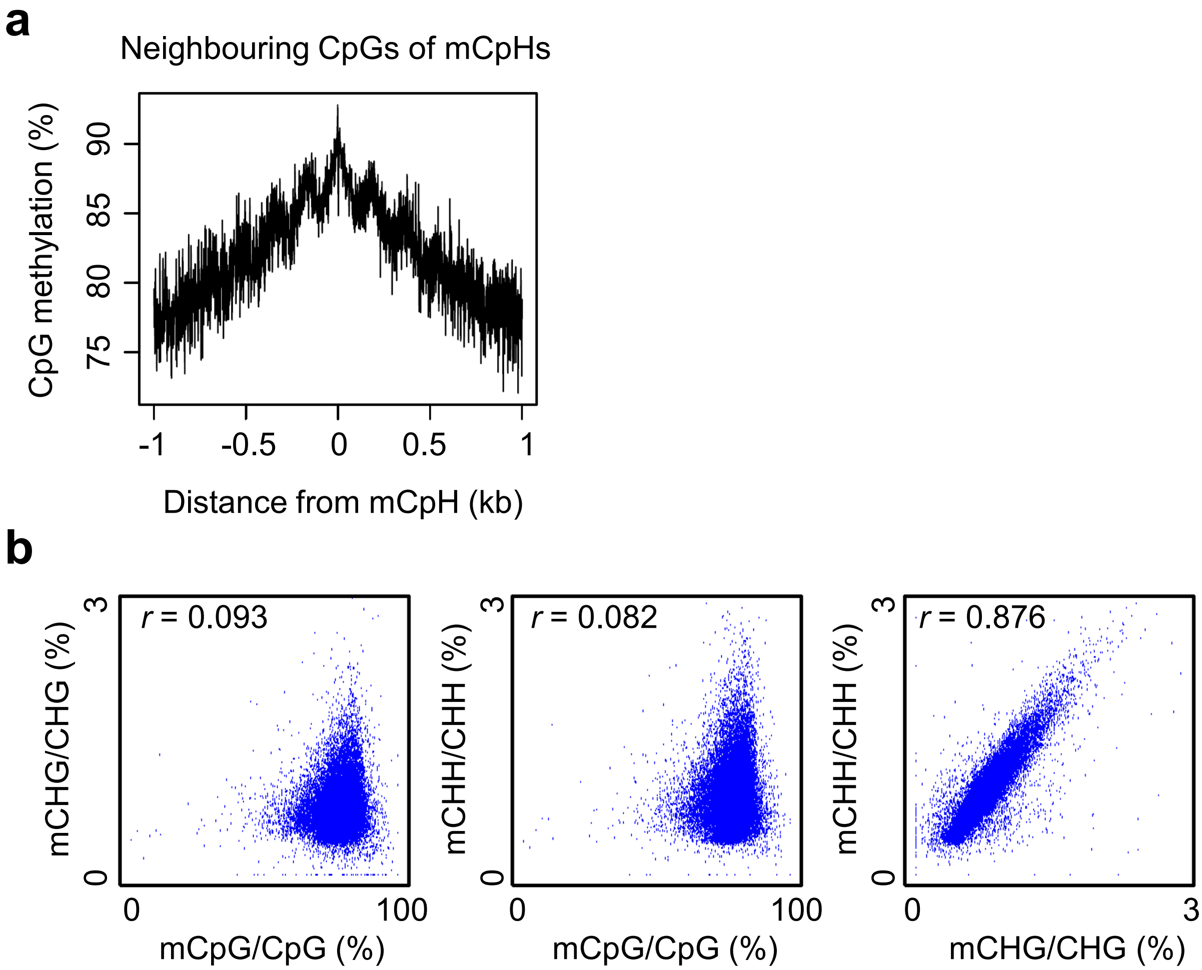
Protein	Accession	Reference
CBP	GSE21161	Kim et al., *Nature* (2010)
CREB	GSE21161	Kim et al., *Nature* (2010)
NPAS4	GSE21161	Kim et al., *Nature* (2010)
POL2	GSE21161	Kim et al., *Nature* (2010)
SRF	GSE21161	Kim et al., *Nature* (2010)
CTCF	wgEncodeEM001690	ENCODE Project Consortium et al., *Nature* (2012)
MeCP2	GSE19786	Skene et al., *Mol Cell* (2010)
TCF3	GSE11724	Marson et al., *Cell* (2008)
NANOG	GSE11724	Marson et al., *Cell* (2008)
SOX2	GSE11724	Marson et al., *Cell* (2008)
OCT4	GSE11724	Marson et al., *Cell* (2008)



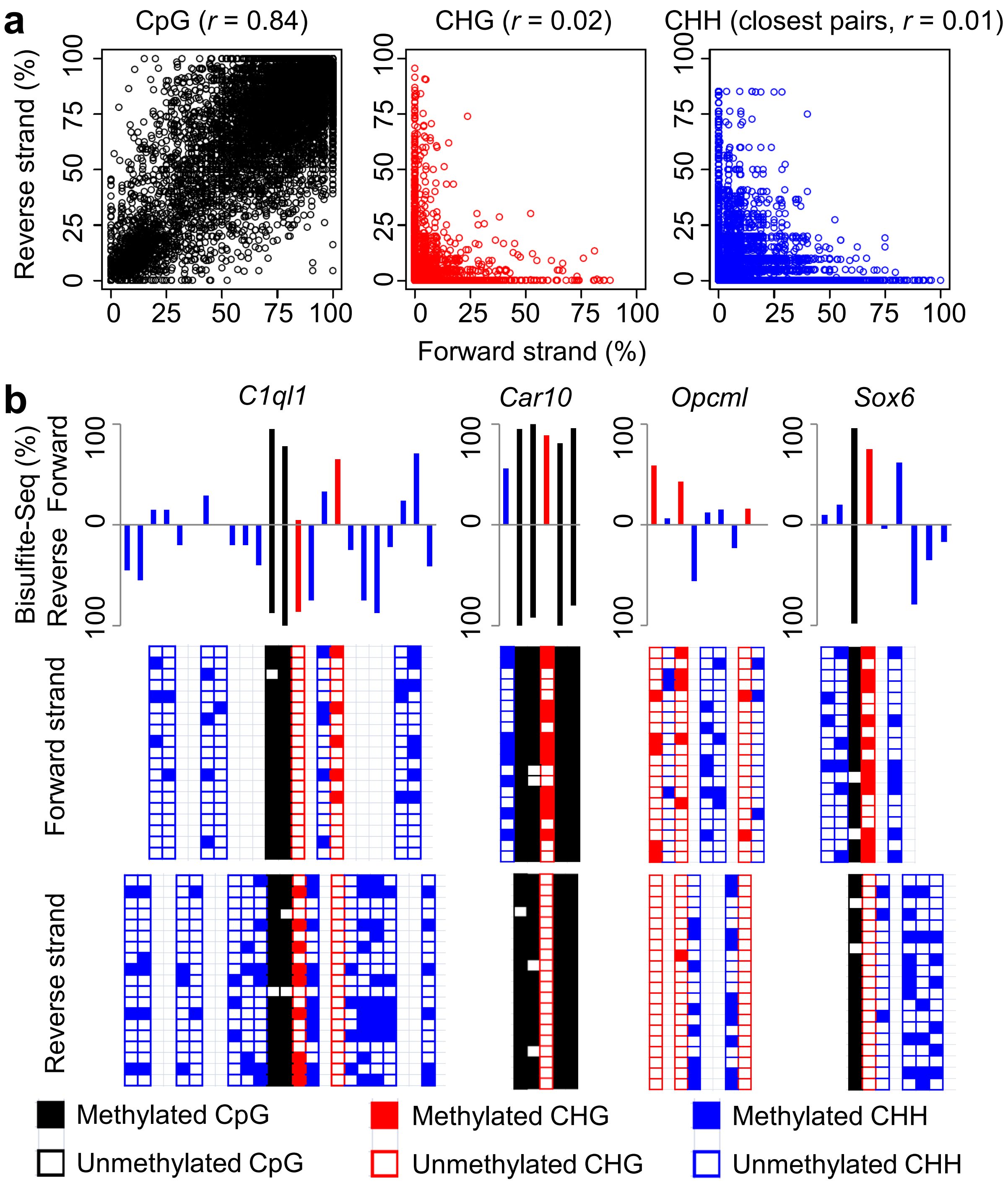
Supplementary Figure 1. Global levels of CpG and CpH methylation across mouse chromosomes. Interquartile boxplots of 1 Kb-bin-averaged methylation levels of each mouse chromosome are shown. Note that chrY shows lower levels of CpG methylation, whereas both sex chromosomes exhibit lower levels of CpH methylation.



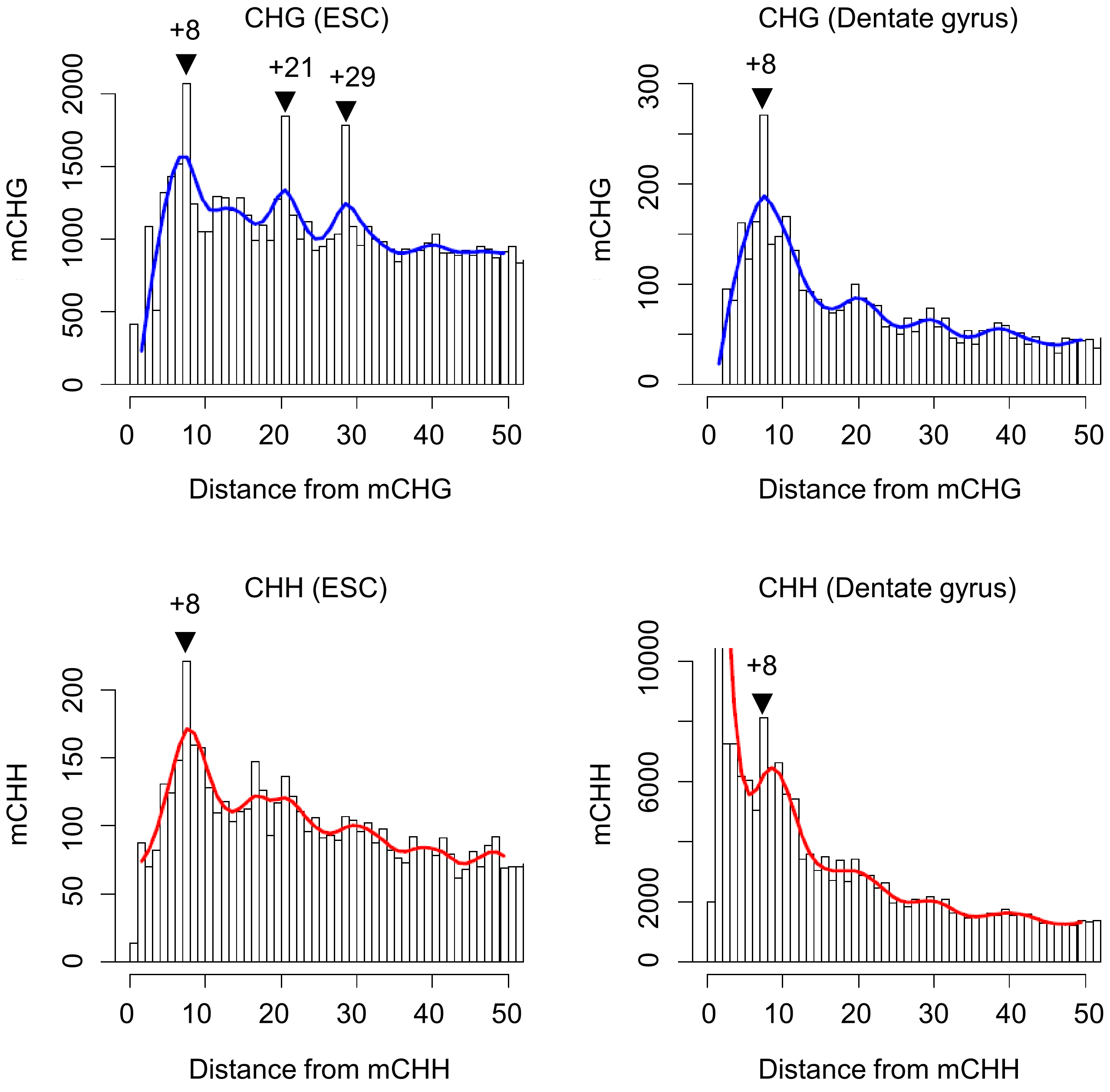
Supplementary Figure 2. Inter-sample correlation of CpG and CpH methylation. Methylation levels of all individual CpG (left) and CpH (right) loci were compared between two biological replicates. Both classes of methylation exhibit high correlation between individuals. Pearson's correlation coefficients (*r*) are shown.



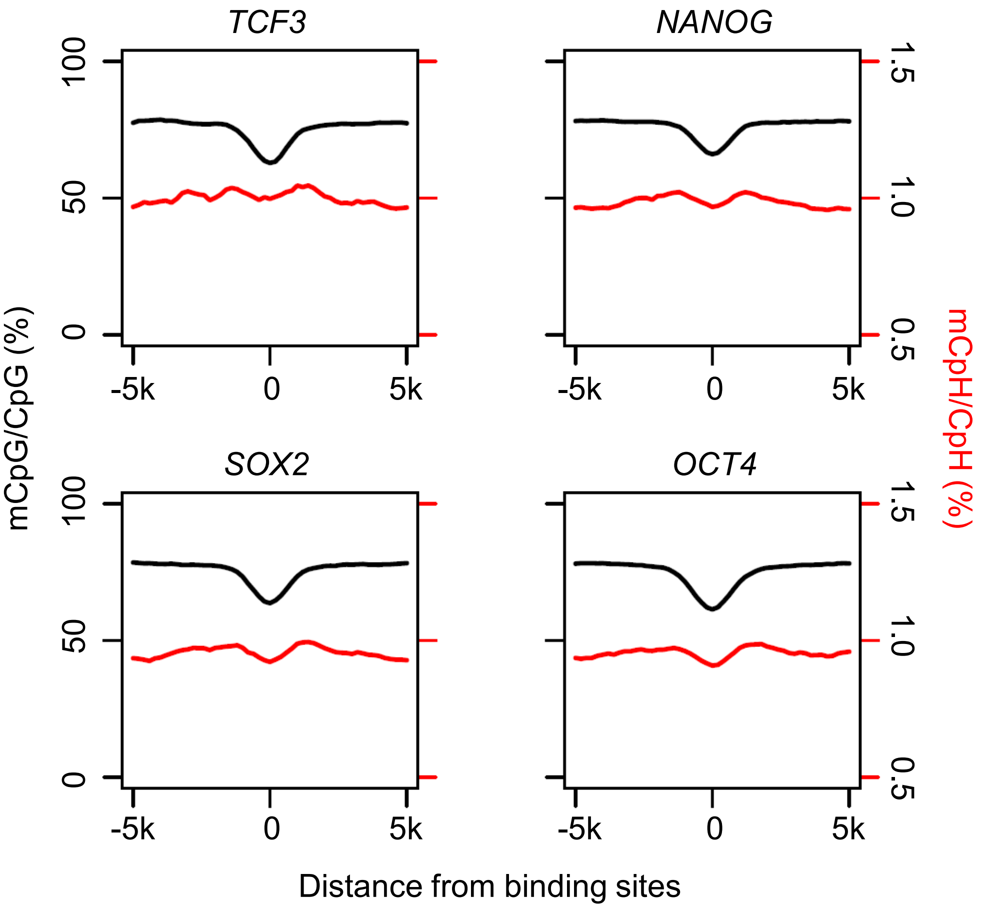
Supplementary Figure 3. CpH methylation is spatially associated with CpG methylation. (**a**) Methylation levels of neighboring CpGs of all mCpHs with ≥ 20% methylation. Note the anti-correlation between CpG methylation and its distance from mCpH. (**b**) Scatter plots of three classes of cytosine methylation. Pearson's correlations (*r*) calculated using non-overlapping 100 Kb bins are shown.



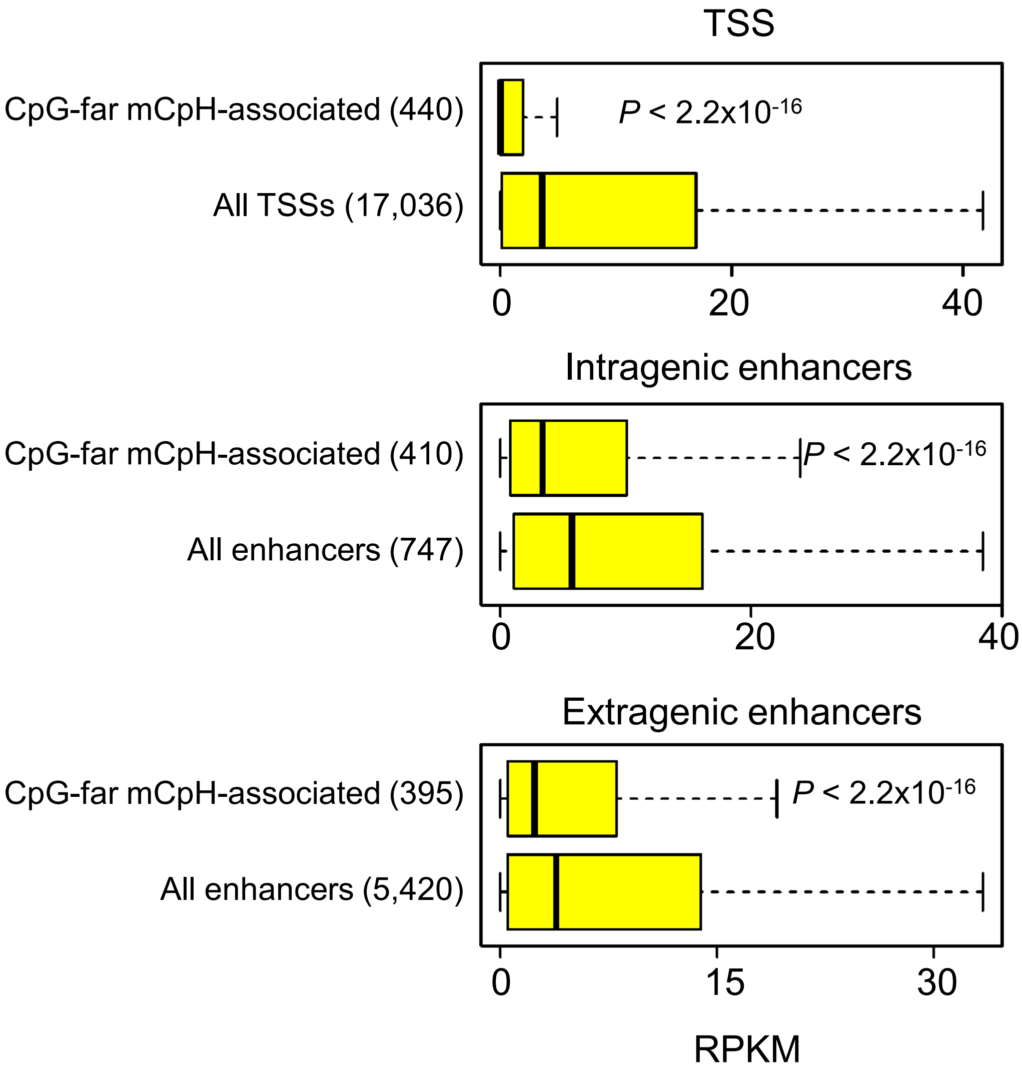
Supplementary Figure 4. Mouse neuronal CpG and CpH methylation on two opposite DNA strands. (**a**) Correlation between individual cytosines of three different classes. All top strand CpG and CHG loci were compared to the cytosines in the bottom strand within the motifs, whereas CHH loci were compared to the closest CHH in the opposite strand. Pearson's correlation coefficients are shown. (**b**) Bisulfite-Seq (top) and Sanger bisulfite sequencing (bottom) results of CpH methylated regions from both strands. Note that CpH methylation is present on both strands in most regions.



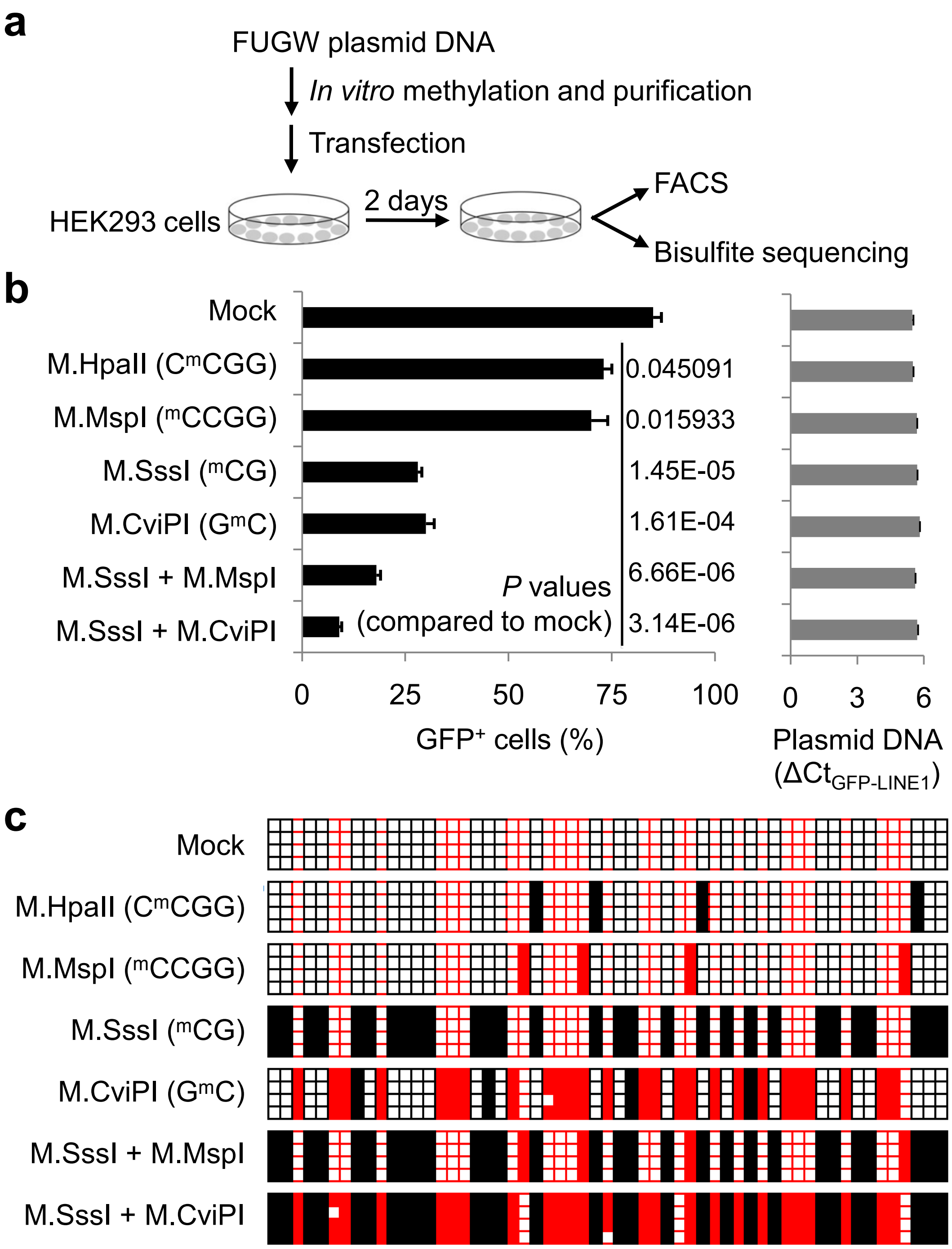
Supplementary Figure 5. Distance analysis of CHG and CHH methylation. Numbers of mCHG (top) and mCHH (bottom) are plotted against the base-pair distance from any mCHG/mCHH in H1 ESCs (left) and in the adult dentate gyrus (right). Note that only CHG methylation in ESCs exhibits the distinct 8, 21, 29 bp spacing pattern. Cubic spline smoothing curves are shown.



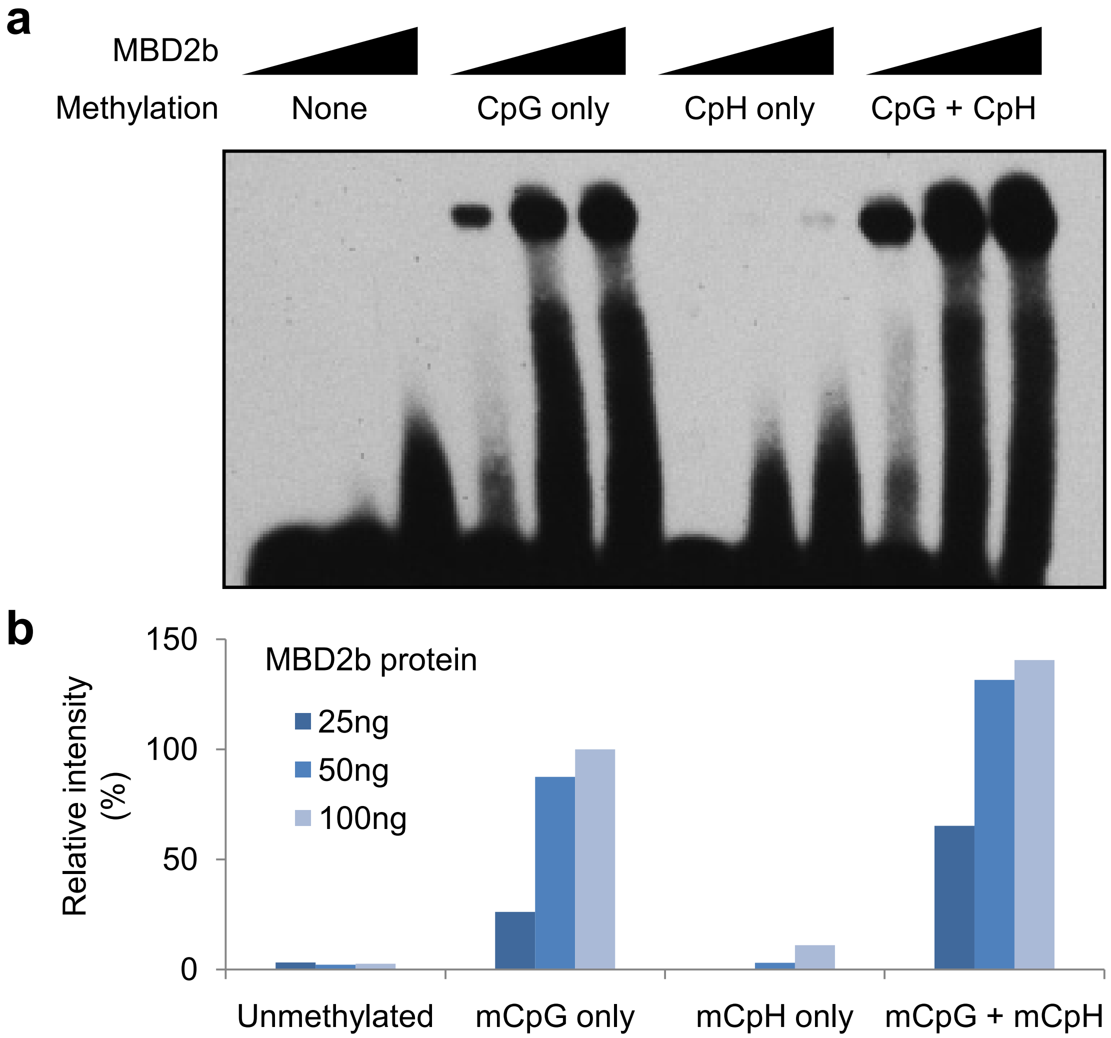
Supplementary Figure 6. Neuronal CpG and CpH methylation around ESC-specific transcription factor binding sites. CpG and CpH methylation levels were averaged across each of the four sets of ESC-specific transcription factor binding sites (Marson et al., *Cell* 2008). Note that CpH hypomethylation is much less pronounced than that around neuronal transcription factor binding sites (**Fig. 4a**). Modest hypomethylation is still observed because many of these binding sites map closely to TSSs, which are intrinsically hypomethylated.



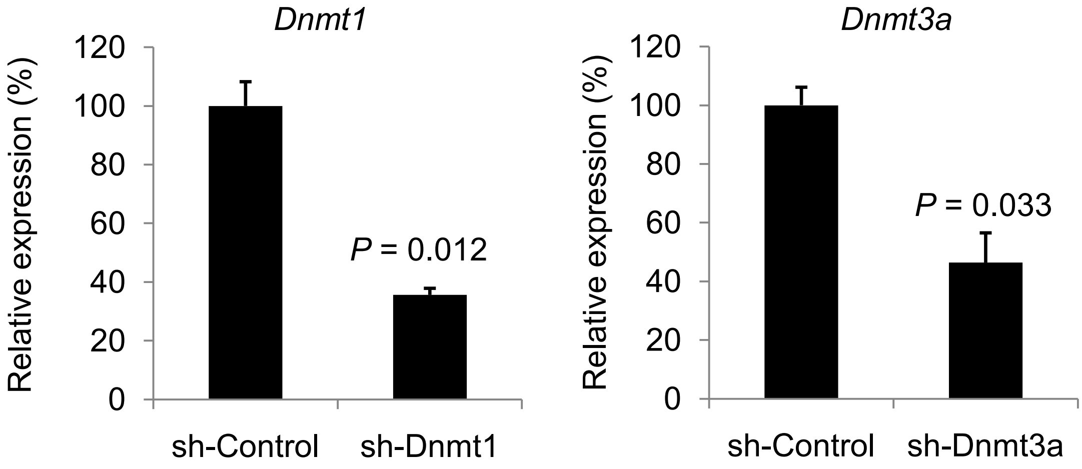
Supplementary Figure 7. Anti-correlation between CpG-far CpH methylation in regulatory regions and gene expression. 440, 410 and 395 of 70,364 mCpHs that do not have any CpG in their 500 bp flanking sequence are mapped within 2 kb from TSS, intragenic and extragenic enhancers (Kim et al., *Nature* 2010), respectively. These mCpH-containing regulatory regions exhibit lower nearest gene expression levels (*P* values are indicated; Mann-Whitney U-tests).



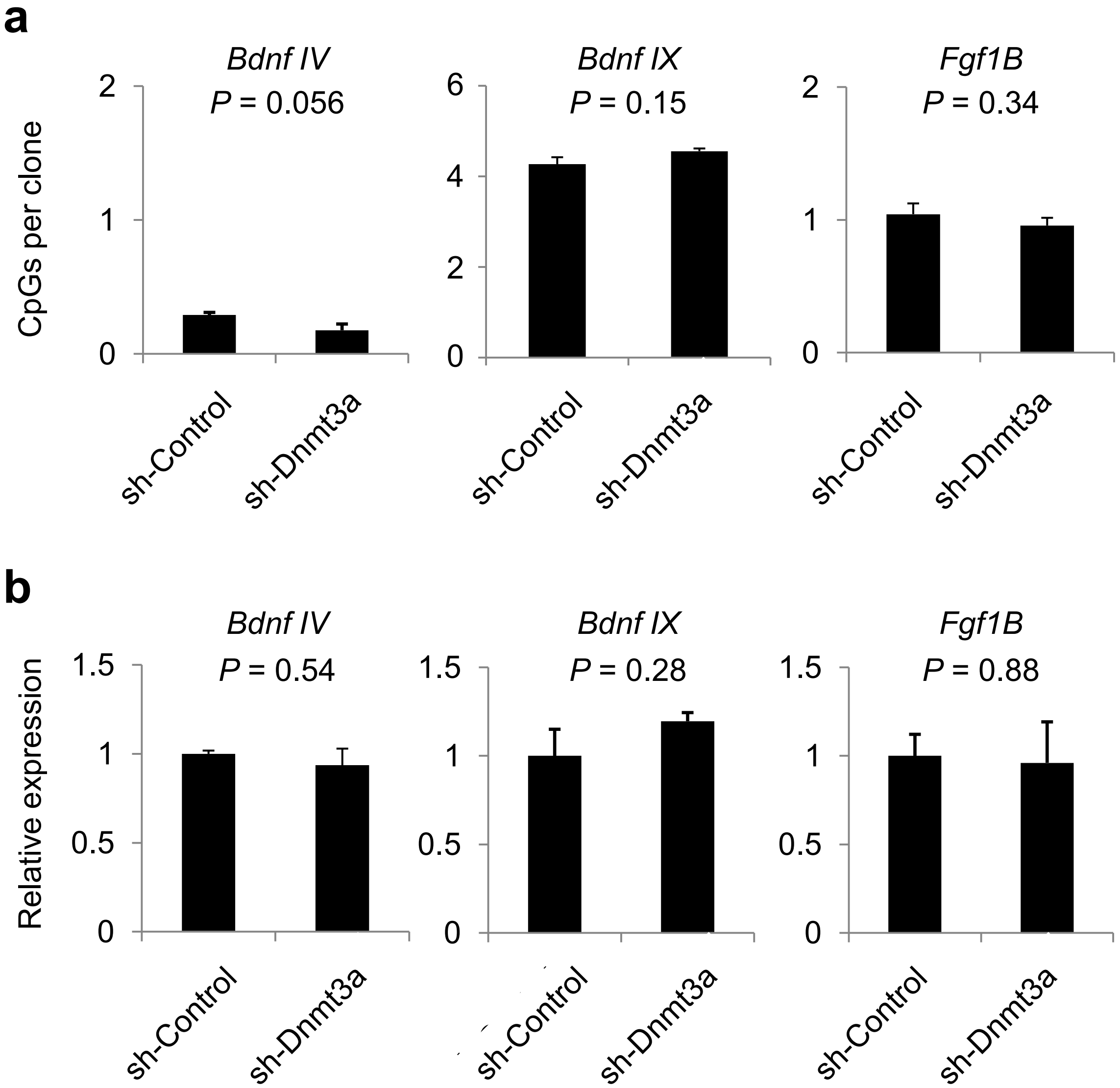
Supplementary Figure 8. *In vitro* methylated reporter assay in HEK293 cells. (**a**) A schematic illustration of experimental design. (**b**) FACS results (left) of GFP$^+$ cells for each *in vitro* methylated reporter. *In vitro* methylation does not alter transfection efficiencies as measured by qPCR (right). Values represent mean ± s.e.m. (P values are indicated, ANOVA). (**c**) Methylation patterns were determined by bisulfite sequencing of plasmids recovered 2 days after transfection.



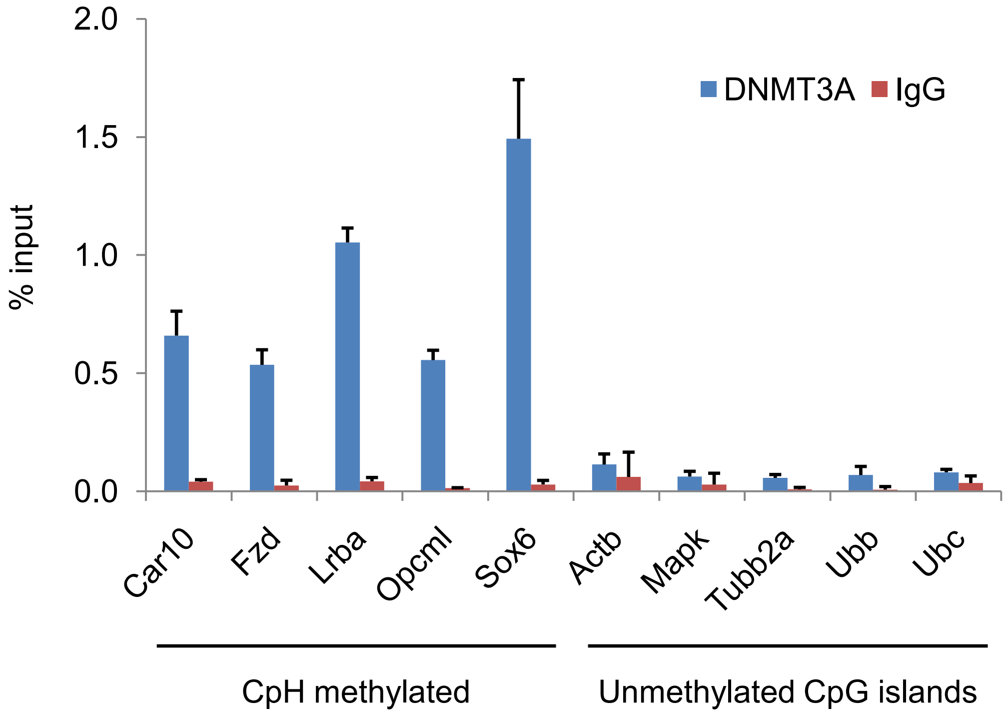
Supplementary Figure 9. Effects of CpH methylation on MBD2b-DNA interaction. (**a**) An EMSA experiment using different amounts of recombinant MBD2b proteins and the same set of synthetically methylated oligos as in **Fig. 6b.** (**b**) Quantification of MBD2b-bound oligos. Note that MBD2b exhibits higher selectivity towards mCpGs than MeCP2 does (**Fig. 6b**).



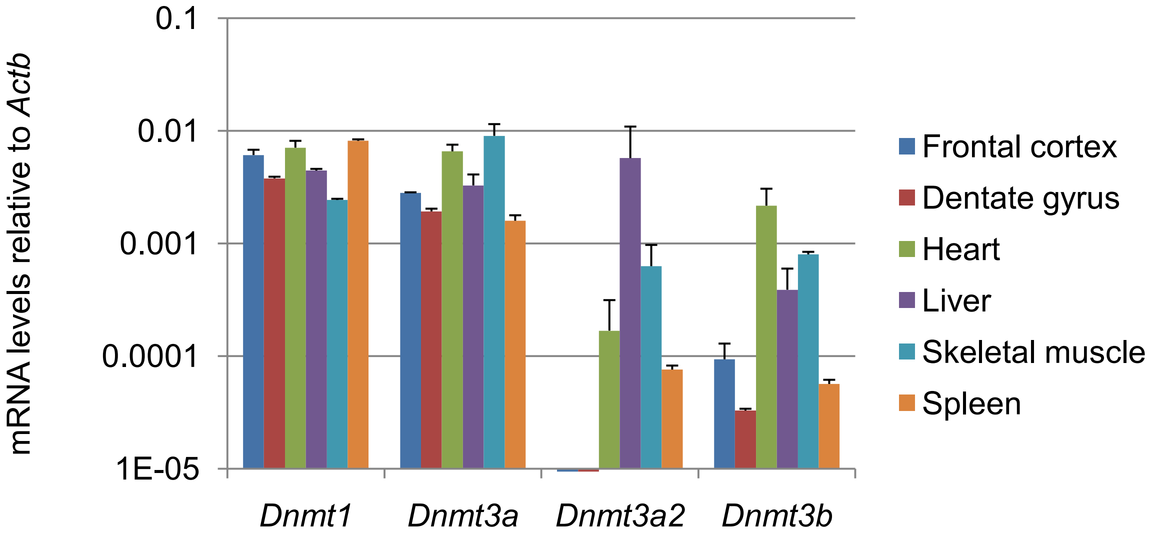
Supplementary Figure 10. Knock-down efficiency of shRNAs. AAVs expressing shRNAs targeting *Dnmt1* (left) or *Dnmt3a* (right) or control AAVs expressing a scrambled shRNA were injected into the adult mouse dentate gyrus. Knock-down efficiency was determined by qPCR using *Dnmt1*- or *Dnmt3a*-specific primers. Values represent mean ± s.e.m. (n = 3; *P* values are indicated; Student's t-test).



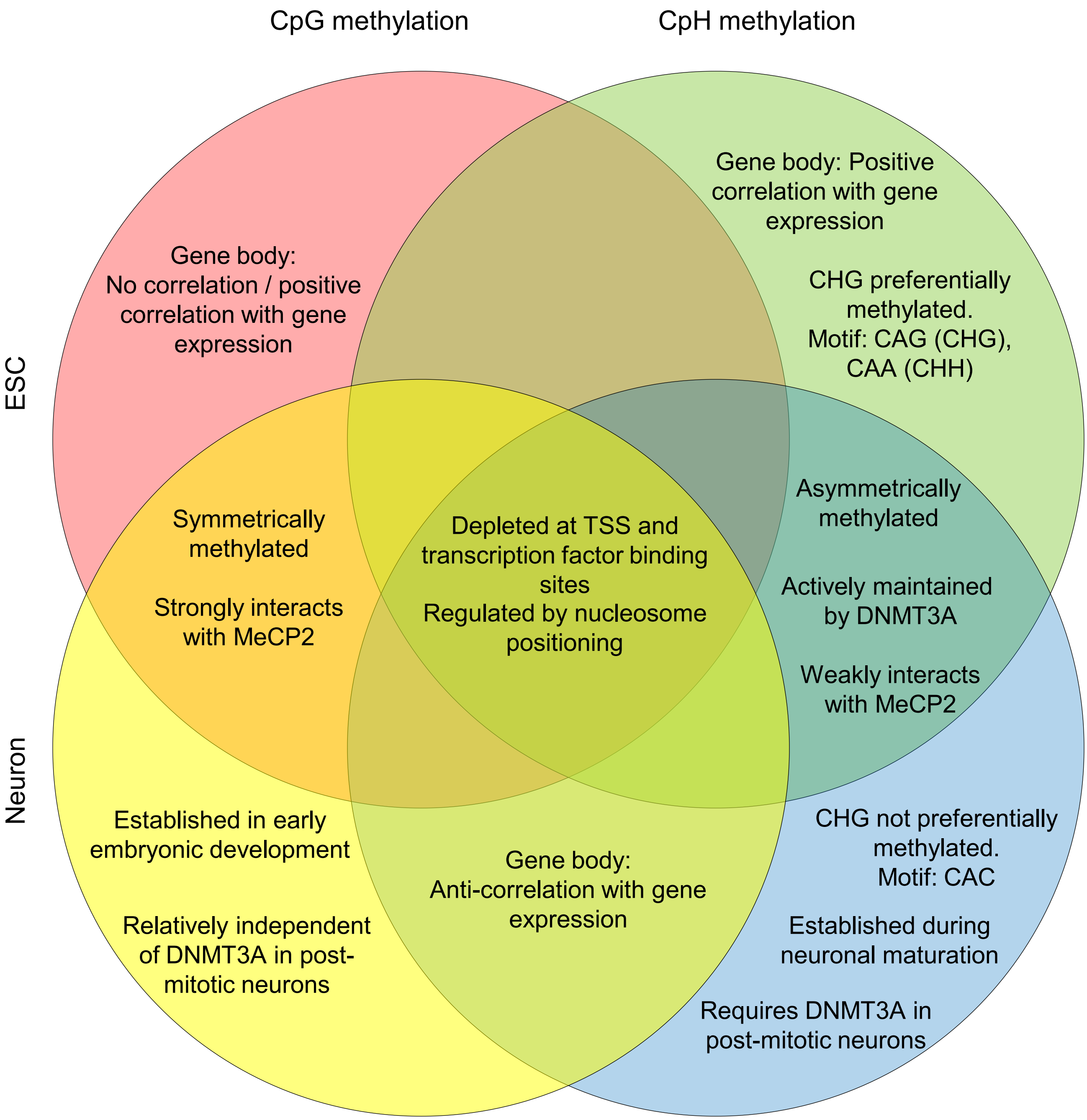
Supplementary Figure 11. Lack of effects of DNMT3A knock-down on unmethylated and CpG-methylated CpH-unmethylated regions. (**a**) Methylation levels of *Bdnf IV* (unmethylated, left), *Bdnf IX* (CpG-methylated/CpH-unmethylated, middle) and *Fgf1B* (CpG-methylated/CpH-unmethylated, right) promoters after DNMT3A knock-down. (**b**) Expression levels of *Bdnf IV*, *Bdnf IX* and *Fgf1B* transcripts after DNMT3A knock-down (*P* values are indicated; Student's t-tests).



Supplementary Figure 12. DNMT3A binds to CpH-methylated regions in adult dentate gyrus *in vivo*. Occupancy of DNMT3A in CpH-methylated regions, measured by the fraction of input neuronal chromatin immuno-precipitated by a DNMT3A antibody, was compared to unmethylated CpG island regions. A rabbit IgG antibody was used to control for unspecific binding.



Supplementary Figure 13. *Dnmt* gene expression in adult mouse tissues. *Dnmt* gene expression was measured by RT-qPCR in multiple adult mouse tissues. Note that the adult brain expresses a similar level of *Dnmt1* and *Dnmt3a*, a lower level of *Dnmt3b* compared to other tissues, and no *Dnmt3a2*. Values represent mean ± s.e.m. (n = 3).



Supplementary Figure 14. Summary of similarities and differences between CpG and CpH methylation in ESCs *in vitro* and neurons *in vivo*.