

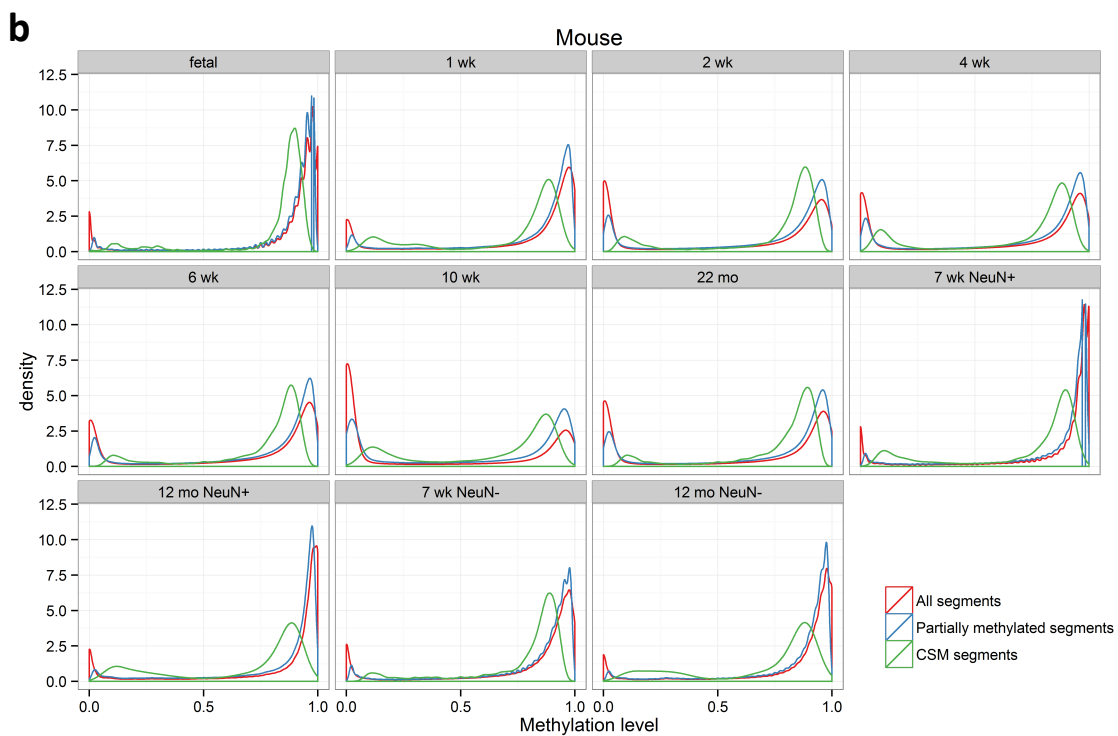
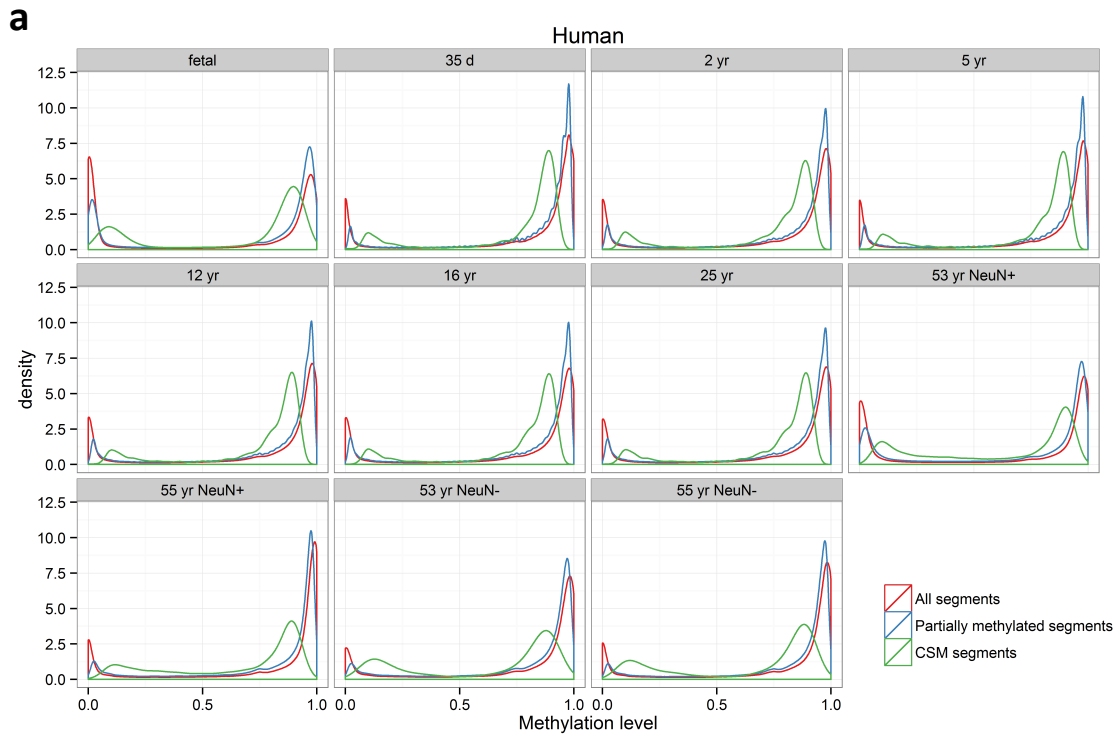
Supplementary figures and tables for:

Mammalian Brain Development is Accompanied by a Dramatic Increase in Bipolar DNA Methylation

Ming-an Sun^{1,9}, Zhixiong Sun^{1,2,9}, Xiaowei Wu³, Veena Rajaram^{4,5}, David Keimig¹, Jessica Lim^{1,2}, Hongxiao Zhu³, Hehuang Xie^{1,2, 6,7,8}

¹Epigenomics and Computational Biology Lab, Biocomplexity Institute of Virginia Tech; ²Department of Biological Sciences, ³Department of Statistics, ⁶School of Neuroscience, ⁷Interdisciplinary Ph.D. Program in Genomics, Bioinformatics, and Computational Biology, ⁸ Ph.D. Program in Translational Biology, Medicine, and Health, Virginia Tech, VA 24061; ⁴Department of Pathology, Children's Medical center, ⁵Department of Pathology, UT Southwestern Medical Center, Dallas, TX 75235; USA

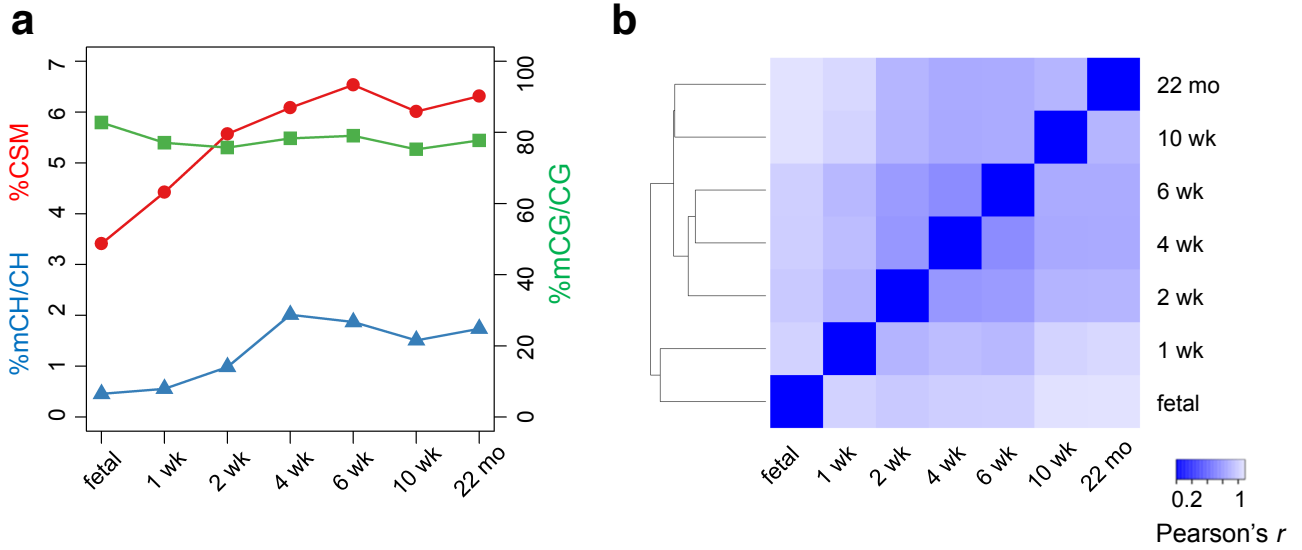
⁹These authors contributed equally to this work.



Supplementary Figure 1 (Xie)

The methylation level distribution for 4-CpG segments in human (a) and mouse (b)

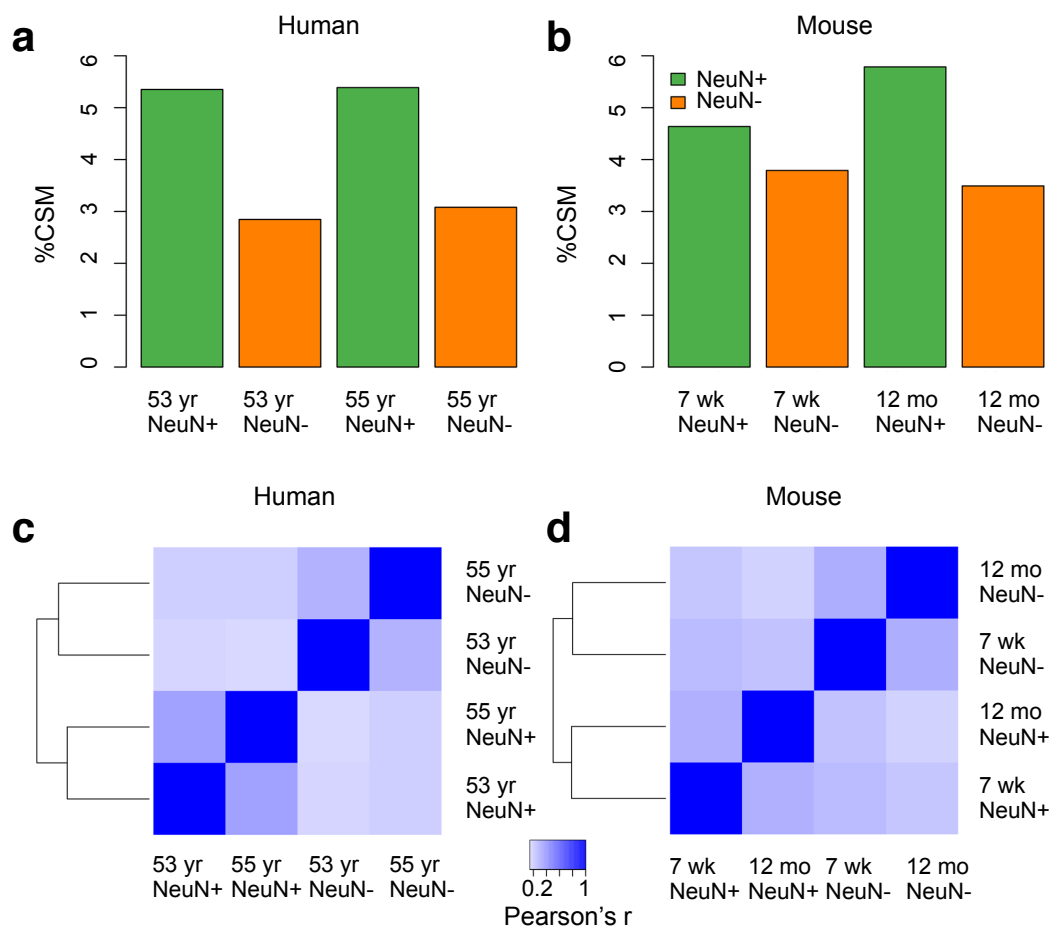
For each methylome, the distributions of methylation levels were shown for all 4-CpG segments, partially methylated 4-CpG segments and pCSM 4-CpG segments.



Supplementary Figure 2 (Xie)

Dynamic of CSM during mouse brain development

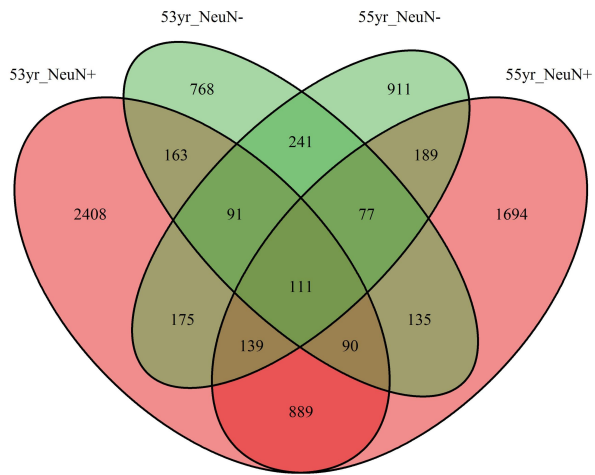
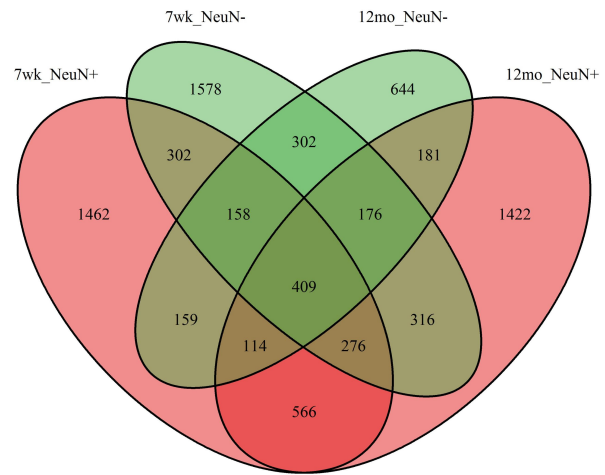
(a) Changes of mC levels in CG and CH context and the percentage of pCSM segments during mouse brain development. **(b)** Hierarchical clustering based on Pearson's correlations of pCSM statuses predicted for all 4-CpG segments in different mouse brain methylomes.



Supplementary Figure 3 (Xie)

Global pCSM profiles in NeuN+ and NeuN- cells

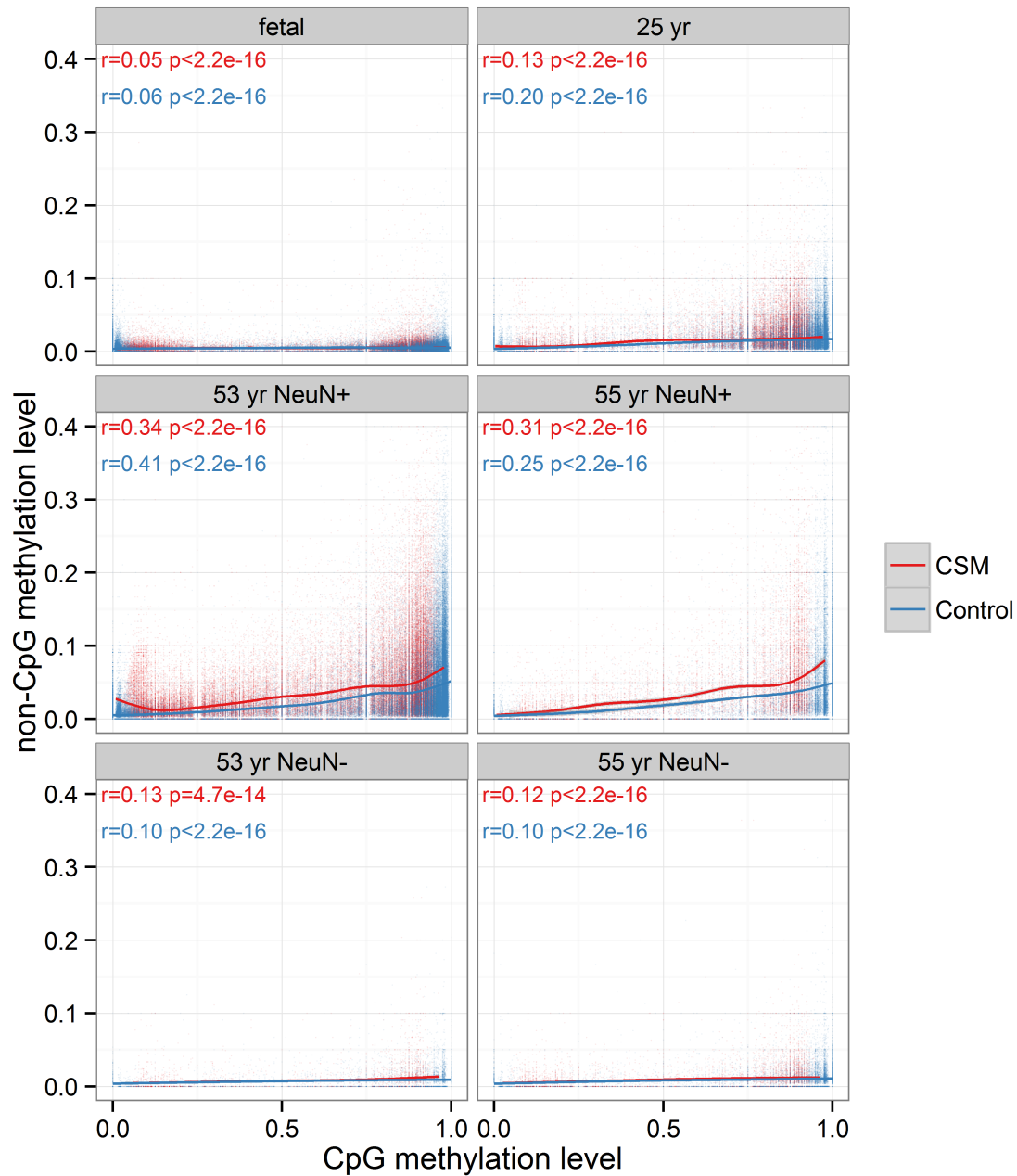
(a,b) The percentage of pCSM 4-CpG segment in the methylomes of NeuN+ and NeuN- cells for human and mouse, respectively. (c,d) Hierarchical clustering based on Pearson's correlations of pCSM statuses predicted for all 4-CpG segments in the methylomes of NeuN+ and NeuN- cells for human and mouse, respectively.

a**b**

Supplementary Figure 4 (Xie)

Venn diagrams show the overlapping of pCSM segments between the NeuN+ and NeuN- cells

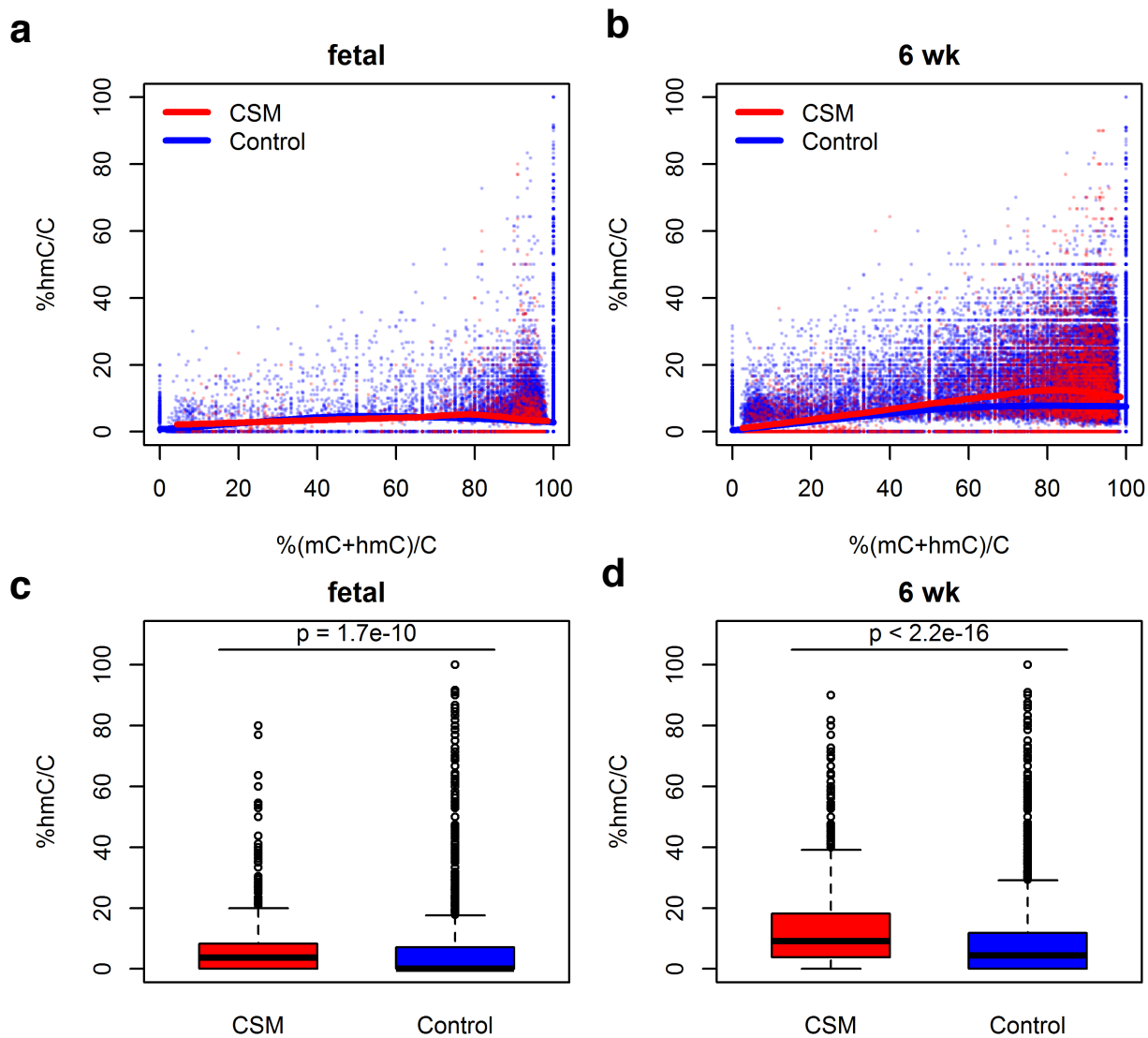
The figures show the overlapping of pCSM segments between the purified NeuN+ and NeuN- samples for human (a) and mouse (b), respectively.



Supplementary Figure 5 (Xie)

Correlation between CpG and non-CpG methylation levels within pCSM and control regions

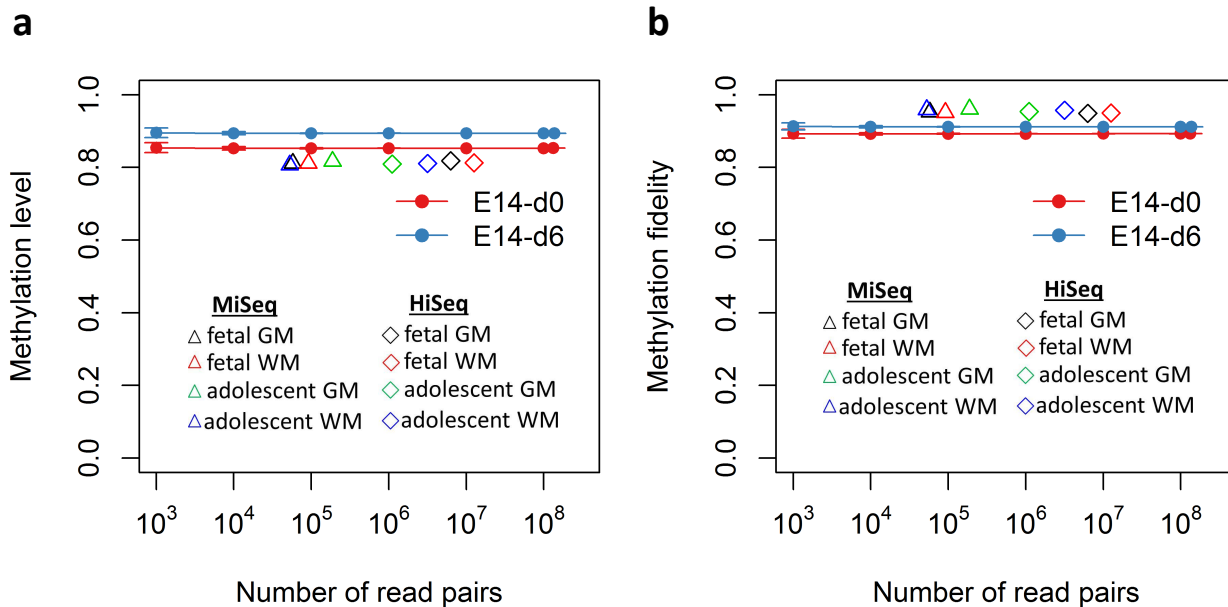
The x-axis and y-axis show the mC levels in CpG and non-CpG context for each 4-CpG segment, respectively. The curves show the smoothed spline. Pearson's r and p-values are shown.



Supplementary Figure 6 (Xie)

5hmC occurrence in pCSM and control regions

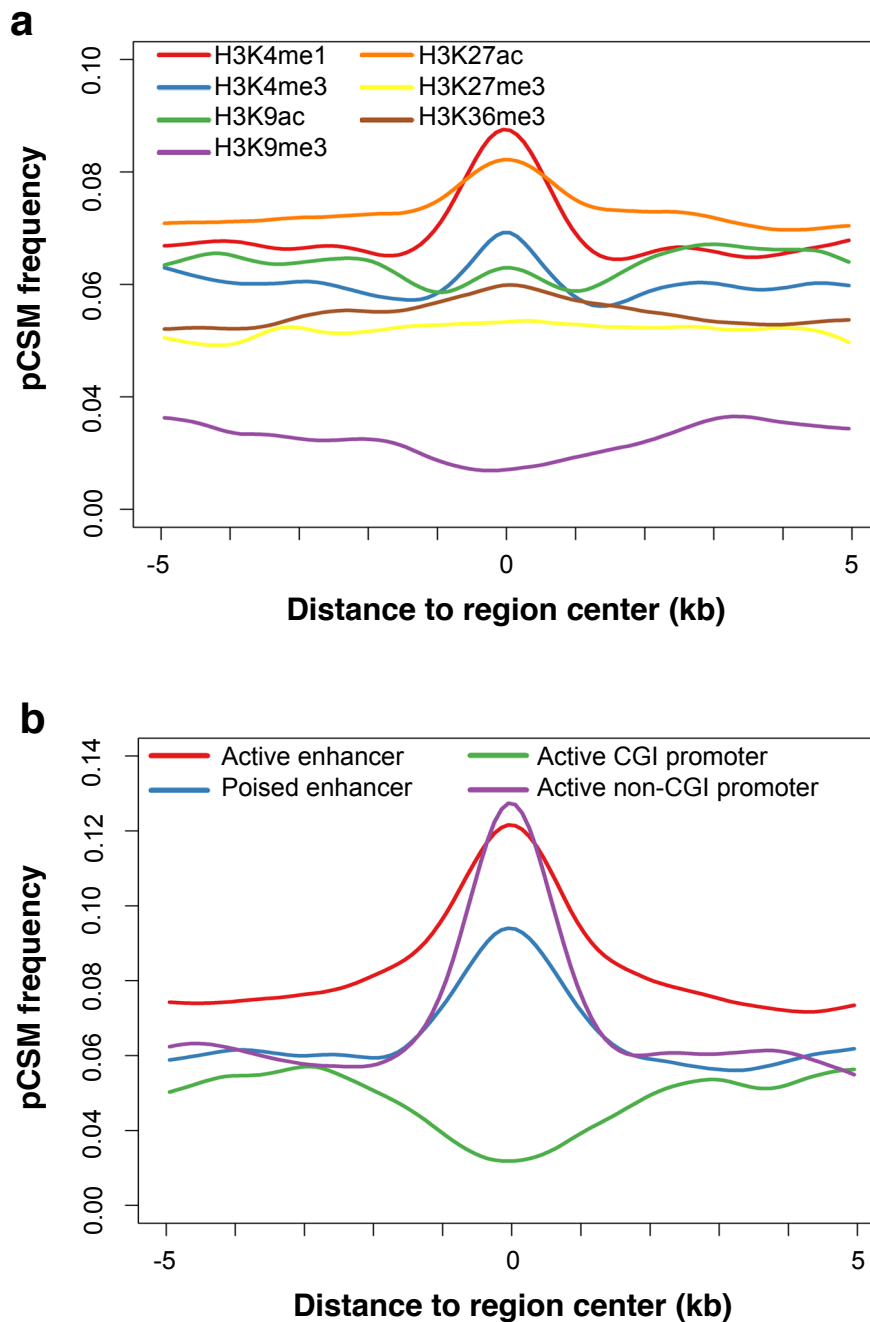
(a,b) Scatter plots show the relationship between 5mC and 5hmC in CpG context within pCSM or control regions of fetal **(a)** and 6 wk **(b)** mouse brain methylomes, respectively. The lines are the smoothed curve computed using loess method. **(c,d)** Boxplots show the distributions of 5hmC level within pCSM or control regions of fetal **(c)** and 6 wk **(d)** mouse brain methylomes, respectively.



Supplementary Figure 7 (Xie)

Methylation level and fidelity estimated with hairpin bisulfite sequencing data at different sequencing depths

Various numbers of read pairs were selected randomly from the two previously published data sets generated for mouse embryonic stem cell E14 with genome-scale hairpin bisulfite sequencing technique²⁰. The error bars show the standard deviation estimated from 100 times of simulations. The filled circles represent methylation level (**a**) or fidelity (**b**) determined with hairpin bisulfite sequencing data for mouse ESCs (E14-d0: mouse ESCs; E14-d6: mouse ESCs at day 6 with LIF withdrawn). In each line, the first six circles represent simulation results and the last circle shows the result using all reads generated. The open triangles and diamonds represent hairpin bisulfite sequencing data generated for human brain tissues in this study using MiSeq and HiSeq, respectively. “fetal GM” represents human grey matter from a fetal frontal cortex. “fetal WM” represents human white matter from a fetal frontal cortex. “Adolescent GM” represents human grey matter from an adolescent frontal cortex. “Adolescent WM” represents human white matter from an adolescent frontal cortex.



Supplementary Figure 8 (Xie)

Aggregate plots of pCSM segments surrounding different histone modifications (a) and regulatory elements (b) in adult human brain

The regulatory elements were annotated based on the combinations of histone modifications and the presence of CGI. Active promoters are with H3K4me3 mark, Active enhancers are with H3K27ac but without H3K4me3 mark, and poised enhancers are with H3K4me1 but without H3K27ac or H3K4me3.

Supplemental Table 1 Statistics for pCSM predicted from mammalian brain methylomes

| Sample | Species | #mapped reads | #CG segment (>=10X, noXY, noASM) | #predicted CSM segments | %pCSM frequency (unnormalized) | #pCSM segments after merging | #CPGs covered in pCSM segments | Total length of pCSM segments (bp) |
|---------------|---------|---------------|----------------------------------|-------------------------|--------------------------------|------------------------------|--------------------------------|------------------------------------|
| Hs fetal | human | 718,715,354 | 3,575,021 | 41,369 | 1.20% | 26,960 | 124,217 | 1,260,254 |
| Hs 35 do | human | 754,934,101 | 1,014,430 | 25,030 | 2.50% | 18,375 | 80,778 | 761,814 |
| Hs 2 yr | human | 769,797,789 | 1,255,212 | 47,202 | 3.80% | 32,822 | 147,195 | 1,413,735 |
| Hs 5 yr | human | 749,639,389 | 1,105,550 | 40,181 | 3.60% | 28,236 | 126,072 | 1,198,411 |
| Hs 12 yr | human | 757,884,985 | 1,133,642 | 46,538 | 4.10% | 32,216 | 144,736 | 1,377,415 |
| Hs 16 yr | human | 746,971,503 | 1,289,687 | 57,503 | 4.50% | 39,051 | 176,787 | 1,687,696 |
| Hs 25 yr | human | 725,663,402 | 1,133,157 | 50,024 | 4.40% | 34,366 | 154,856 | 1,455,879 |
| Hs 53 yrNeuN+ | human | 841,935,270 | 2,210,113 | 96,033 | 4.30% | 61,121 | 284,305 | 2,891,375 |
| Hs 53 yrNeuN- | human | 476,665,175 | 217,362 | 5,636 | 2.60% | 4,276 | 18,583 | 159,715 |
| Hs 55 yrNeuN+ | human | 558,556,927 | 471,011 | 21,566 | 4.60% | 15,516 | 68,659 | 619,658 |
| Hs 55 yrNeuN- | human | 539,415,159 | 362,534 | 9,925 | 2.70% | 7,369 | 32,229 | 282,263 |
| Mm fetal | mouse | 553,383,395 | 162,533 | 5,218 | 3.20% | 3,821 | 16,771 | 140,672 |
| Mm 1 wk | mouse | 385,489,360 | 235,332 | 8,455 | 3.60% | 5,441 | 24,994 | 187,504 |
| Mm 2 wk | mouse | 618,425,417 | 1,439,393 | 55,568 | 3.90% | 39,091 | 174,828 | 1,784,255 |
| Mm 4 wk | mouse | 672,968,248 | 1,197,542 | 57,001 | 4.80% | 40,094 | 179,238 | 1,847,952 |
| Mm 6 wk | mouse | 667,130,394 | 805,885 | 40,088 | 5.00% | 29,524 | 129,877 | 1,310,969 |
| Mm 10 wk | mouse | 448,962,696 | 797,566 | 23,794 | 3.00% | 17,580 | 77,231 | 714,279 |
| Mm 22 mo | mouse | 682,097,075 | 944,270 | 51,795 | 5.50% | 33,086 | 152,841 | 1,249,124 |
| Mm 7 wkNeuN+ | mouse | 633,803,213 | 357,027 | 15,569 | 4.40% | 9,864 | 45,567 | 383,268 |
| Mm 7 wkNeuN- | mouse | 676,970,294 | 587,176 | 20,815 | 3.50% | 13,900 | 63,013 | 552,001 |
| Mm 12 moNeuN+ | mouse | 436,577,319 | 194,533 | 9,765 | 5.00% | 6,951 | 30,825 | 253,737 |
| Mm 12 moNeuN- | mouse | 332,614,011 | 102,534 | 3,343 | 3.30% | 2,166 | 9,927 | 70,782 |

Supplementary Table 2 Pearson's correlations of pCSM statuses predicted for all 4-CpG segments in the methylomes of different samples.

a. human developmental stages (related to Figure 1E).

| | | | | | | | |
|-------|-------|------|------|------|------|------|------|
| | fetal | 35d | 2yr | 5yr | 12yr | 16yr | 25yr |
| fetal | | 0.20 | 0.16 | 0.15 | 0.15 | 0.15 | 0.15 |
| 35d | 0.20 | | 0.23 | 0.21 | 0.21 | 0.20 | 0.20 |
| 2y | 0.16 | 0.23 | | 0.31 | 0.30 | 0.31 | 0.30 |
| 5y | 0.15 | 0.21 | 0.31 | | 0.31 | 0.31 | 0.30 |
| 12y | 0.15 | 0.21 | 0.30 | 0.31 | | 0.32 | 0.31 |
| 16y | 0.15 | 0.20 | 0.31 | 0.31 | 0.32 | | 0.33 |
| 25y | 0.15 | 0.20 | 0.30 | 0.30 | 0.31 | 0.33 | |

b. mouse developmental stages (related to Supplemental Figure 2B).

| | | | | | | | |
|-------|-------|------|------|------|------|------|------|
| | fetal | 1wk | 2wk | 4wk | 6wk | 10wk | 22mo |
| fetal | | 0.18 | 0.21 | 0.19 | 0.19 | 0.12 | 0.11 |
| 1wk | 0.18 | | 0.29 | 0.26 | 0.28 | 0.17 | 0.15 |
| 2wk | 0.21 | 0.29 | | 0.41 | 0.39 | 0.29 | 0.29 |
| 4wk | 0.19 | 0.26 | 0.41 | | 0.45 | 0.34 | 0.33 |
| 6wk | 0.19 | 0.28 | 0.39 | 0.45 | | 0.33 | 0.33 |
| 10wk | 0.12 | 0.17 | 0.29 | 0.34 | 0.33 | | 0.29 |
| 22mo | 0.11 | 0.15 | 0.29 | 0.33 | 0.33 | 0.29 | |

c. human cell types (related to Supplemental Figure 3C)..

| | | | | |
|-----------|------------|------------|------------|------------|
| | 53yr NeUN+ | 55yr NeUN+ | 53yr NeUN- | 55yr NeUN- |
| 53y NeUN+ | | 0.37 | 0.16 | 0.19 |
| 55y NeUN+ | 0.37 | | 0.15 | 0.19 |
| 53y NeUN- | 0.16 | 0.15 | | 0.30 |
| 55y NeUN- | 0.19 | 0.19 | 0.30 | |

d. mouse cell types (related to Supplemental Figure 3D).

| | | | | |
|------------|-----------|------------|-----------|------------|
| | 7wk NeUN+ | 12mo NeUN+ | 7wk NeUN- | 12mo NeUN- |
| 7wk NeUN+ | | 0.31 | 0.27 | 0.22 |
| 12mo NeUN+ | 0.31 | | 0.24 | 0.18 |
| 7wk NeUN- | 0.27 | 0.24 | | 0.32 |
| 12mo NeUN- | 0.22 | 0.18 | 0.32 | |

Supplemental Table 3a Summary of CSM predicted for human and mouse purified cells

| Human (54,276 common 10X 4CG segment) | | Mouse (36,692 common 10X 4CG segment) | |
|---------------------------------------|---------------|---------------------------------------|---------------|
| Sample | #pCSM segment | Sample | #pCSM segment |
| 53yr NeuN+ | 4066 | 7wk NeuN+ | 3446 |
| 55yr NeuN+ | 3324 | 12mo NeuN+ | 3460 |
| 53yr NeuN- | 1676 | 7wk NeuN- | 3517 |
| 55yr NeuN- | 1934 | 12mo NeuN- | 2143 |

Supplemental Table 3b Number of shared pCSM segments between different samples

| Human | | | | |
|-----------|-----------|-----------|-----------|-----------|
| | 53y NeuN+ | 55y NeuN+ | 53y NeuN- | 55y NeuN- |
| 53y NeuN+ | | 1229 | 455 | 516 |
| 55y NeuN+ | 37% | | 413 | 516 |
| 53y NeuN- | 27% | 25% | | 520 |
| 55y NeuN- | 27% | 27% | 31% | |

| Mouse | | | | |
|------------|-----------|------------|-----------|------------|
| | 7wk NeuN+ | 12mo NeuN+ | 7wk NeuN- | 12mo NeuN- |
| 7wk NeuN+ | | 1365 | 1145 | 840 |
| 12mo NeuN+ | 40% | | 1177 | 880 |
| 7wk NeuN- | 33% | 34% | | 1045 |
| 12mo NeuN- | 39% | 41% | 49% | |

Supplemental Table 4 Statistics of the hairpin-bisulfite sequencing data

| Sample | Age | Tissue/cell type | Uniquely mapped read pairs (M) | Bisulfite conversion rate (%) | Genome covered (%) | CpG covered (%) | Genome sequencing depth for genome covered (X) | Average methylation level (%) | Average methylation fidelity (%) |
|----------|--------------------|-----------------------|--------------------------------|-------------------------------|--------------------|-----------------|--|-------------------------------|----------------------------------|
| fetal CG | 22 weeks gestation | Cortical grey matter | 6.4 | 99.8 | 8.2 | 9.7 | 10 | 81.8 | 94.9 |
| fetal CW | 22 weeks gestation | Cortical white matter | 12.7 | 99.8 | 15 | 17.7 | 10.9 | 81.3 | 95 |
| 17Y CG | 17 years | Cortical grey matter | 1.3 | 99.8 | 2.6 | 3.1 | 6.5 | 81.1 | 95.6 |
| 17Y CW | 17 years | Cortical white matter | 3.2 | 99.7 | 4.7 | 5.2 | 8.9 | 81.1 | 95.8 |

Supplemental Table 6 Summary of "omics" data used in this study

| GEO accession | Organism | Tissue/cell type | Age | Data type | Brief description | Reference |
|--------------------|----------|------------------------------------|-----------------------------------|-----------------------|---|---|
| GSE67482 | Human | brain cortex | fetal (22wk of gestation) and 17y | hairpin bisulfite-Seq | Hairpin bisulfite-Seq data for human brain | This study |
| GSE47966 | Human | brain frontal cortex, neuron, glia | fetal, adolescent and adult | methyl/C-Seq | Methylomes for several developmental stages and two cell types (neuron and glia) of human and mouse brain | Lister et al, Science 2013 ¹ |
| | Mouse | | | | | |
| GSE29184 | Mouse | brain cortex | 8wk | ChIP-Seq | Chromatin maps for H3K4me1, H3K4me3 and H3K27ac | Shen et al, Nature 2014 ² |
| GSE16368 | Human | brain | 75y/81y | ChIP-Seq | Chromatin maps for H3K4me1, H3K4me3, H3K9ac, H3K9me3, H3K27ac, H3K27me3 and H3K36me3 | Bernstein et al, Nat Biotechnol 2010 ³ |
| NHGRI GWAS catalog | Human | Various | Various | GWAS | NHGRI GWAS catalog | Welter et al, NAR 2014 ⁴ |

Reference:

1. Lister R, Mukamel EA, Nery JR, Urich M et al. Global epigenomic reconfiguration during mammalian brain development. *Science* 2013 Aug 9;341(6146):1237905. PMID: 23828890
2. Shen Y, Yue F, McCleary DF, Ye Z et al. A map of the cis-regulatory sequences in the mouse genome. *Nature* 2012 Aug 2;488(7409):116-20. PMID: 22763441
3. Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B et al. The NIH Roadmap Epigenomics Mapping Consortium. *Nat Biotechnol* 2010 Oct;28(10):1045-8. PMID: 20944595
4. Welter D, MacArthur J, Morales J, Burdett T, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, 2014, Vol. 42: D1001-D1006.