HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2018 November 25.

Published in final edited form as:

Nat Neurosci. 2018 June ; 21(6): 794–798. doi:10.1038/s41593-018-0155-8.

Genome Wide Distribution of Linker Histone H1.0 is Independent of MeCP2

Aya Ito-Ishida[#1,5,6,7], **Hari K. Yamalanchili**[#1,5,6], **Yingyao Shao**[1,5,6], **Steven A. Baker**[1,5,6], **Laura D. Heckman**[1,5,6], **Laura A. Lavery**[1,5,6], **Ji-yoen Kim**[1,5,6], **Laura M. Lombardi**[1,5,6], **Yaling Sun**[1,4,5,6], **Zhandong Liu**[1,5,6], and **Huda Y. Zoghbi**[1,2,3,4,5,6,*]

[1]Department of Molecular and Human Genetics

[2]Department of Neuroscience

[3]Department of Pediatrics

[4]Howard Hughes Medical Institute

[5]Jan and Dan Duncan Neurological Research Institute at Texas Children's Hospital, Houston, TX 77030, USA

[6]Baylor College of Medicine, Houston, Texas 77030, USA

[#] These authors contributed equally to this work.

Abstract

Previous studies suggested that MeCP2 competes with linker histone H1, but this hypothesis has never been tested *in vivo*. Here, we performed ChIP-Seq of Flag-tagged-H1.0 in mouse forebrain excitatory neurons. Unexpectedly, Flag-H1.0 and MeCP2 occupied similar genomic regions and the Flag-H1.0 binding was not changed upon MeCP2 depletion. Furthermore, mild overexpression of H1.0 did not alter MeCP2 binding, suggesting that the functional binding of MeCP2 and H1.0 are largely independent.

Rett syndrome (RTT) is a postnatal neurodevelopmental disorder caused by mutations in the gene encoding methyl-CpG binding protein 2 (MeCP2)[1,2]. Despite extensive research, the precise molecular function of MeCP2 remains unclear. MeCP2 contains a methyl-CpG binding domain and a transcription repression domain, and was initially thought to work as a transcriptional repressor[3,4]. However, transcriptomic analysis of mouse brain RNA revealed

[*]Correspondence: hzoghbi@bcm.edu.

[7]Present address: Department of Physiology, Keio University School of Medicine, Tokyo, Japan

Author Contributions

A. I-I., S.A.B. and H.Y.Z. conceived the project; A.I-I. performed experiments; A.I-I. and H.Y.Z. interpreted data, and wrote and edited the manuscript, respectively; Y.Shao performed MeCP2 ChIP-Seq; S.A.B and L.D.H. generated Fl-H1.0 mice; L.A.L and J.K. helped cell sorting; L.M.L interpreted data; Y.Sun genotyped the mice; H.K.Y. and Z.L. designed, conducted and interpreted the ChIP-Seq data analysis.

Accession Codes

GSE107533

Competing Interests

The authors declare no competing interests.

that MeCP2 depletion leads to both decreases and increases in gene expression[5], indicating that MeCP2 is not a typical repressor. Because MeCP2 is extremely abundant in neurons and is distributed genome-wide, tracking DNA methylation[6,7], an emerging possibility is that MeCP2 is an architectural chromatin component with specialized mechanistic properties in the brain.

How could MeCP2 regulate neuronal chromatin? One hypothesis posits that MeCP2 may function like the linker histone H1, which binds to linker DNA and compacts nucleosomes[3]. This hypothesis was supported by *in vitro* studies which showed that MeCP2 binds to the linker DNA[8] and displaces linker H1[3,9]. It was further supported by a study showing that H1 protein level in male *Mecp2*-knockout (*Mecp2*$^{-/y}$) mice is increased two-fold relative to wild-type mice[6]. Interestingly, HMGA proteins, which have AT-hook domains homologous to MeCP2[10], were also shown to compete with H1[11,12]. These observations suggest that MeCP2 depletion may lead to accumulation of H1 in genomic regions where MeCP2 normally binds, causing chromatin structural changes. This model has been widely discussed as a potential mechanism to describe MeCP2's function. However, it has never been tested *in vivo* due to lack of optimal antibodies for H1.

Linker H1 has multiple subtypes, among which H1.0, H1.c, and H1.x are expressed in the brain. To obtain insight into which H1 might be involved in RTT, we examined H1 levels in neuronal nuclei from the cortex of *Mecp2*$^{-/y}$, MeCP2-overexpressing (*Mecp2*$^{Tg3/y}$), and wild-type mice. The result showed a mild trend toward increase in H1.0 in *Mecp2*$^{-/y}$ mice (Supplementary Fig. 1f, g and Supplementary Fig. 5), while H1.c and H1.x levels remained unchanged. Neither H1 subtype was affected in *Mecp2*$^{Tg3/y}$ mice (Supplementary Fig. 1h, i, and Supplementary Fig. 5). Given that H1.0 was shown to compete with MeCP2 for DNA binding *in vitro*[9], we chose to study H1.0 *in vivo*.

To study the effects on H1.0, we generated mice expressing Flag-tagged H1.0 (Fl-H1.0) in forebrain excitatory cells, given the contributions of this cell type to RTT pathogenesis[13]. Fl-H1.0 mice, which carried Fl-H1.0 transgene under the tetracycline responsive element, were crossed with the Camk2α-promoter-driven tetracycline transactivator (Cam-tTA) line to induce Fl-H1.0 expression in Camk2α-positive cells (Fig. 1a-f). Total protein level of H1.0 was increased to 1.5-fold in the cortex of Cam-tTA:Fl-H1.0 mice (Fig. 1e, f), while the levels of MeCP2, other H1s, and H3 modifications associated with linker H1 remained unchanged (Fig. 1e, f and Supplementary Fig. 5)[14]. Consistent with previous reports[15], ChIP-qPCR showed that Fl-H1.0 enrichment was high at the heterochromatic major satellite repeats and low at the transcriptional start site (TSS) of *Gapdh* (Fig. 1g). To address whether Fl-H1.0 expression affected MeCP2 binding to DNA, we measured MeCP2 enrichment by ChIP-qPCR at sites identified as rich for MeCP2 binding[7,16], and found no change in Cam-tTA:Fl-H1.0 mice (Fig. 1h) [7,16]. To further examine the effect of H1.0 overexpression on the genomic distribution of MeCP2, we performed ChIP-Seq using frontal cortex of wild-type and Cam-tTA:Fl-H1.0 mice (3 months old). We used three mice per genotype for sequencing and obtained sufficiently high reading depth to cover the broad distribution of MeCP2 (Supplementary Table 1). Consistent with the previous studies[7,17,18], we found that MeCP2 enrichment positively correlated with the levels of both CG- and CH-methylation (where H = A, T or C)[19] while it was lower around CG-islands (Supplementary Fig. 2). These results

were similarly observed in both wild-type and Cam-tTA:Fl-H1.0 mice (Supplementary Fig. 2b-e). Further analysis revealed little change in MeCP2 distribution upon Fl-H1.0 expression: Principal component analysis revealed no clear separation between wild-type and Cam-tTA:Fl-H1.0 mice (Supplementary Fig. 2f); Spearman correlation coefficients of MeCP2 enrichment between wild-type and Cam-tTA:Fl-H1.0 mice was 0.93 at the binning size of 10 kb (Fig. 1i). We also performed window-based comparison, where average MeCP2 enrichment from each mouse was computed at every 200-bp window throughout the genome and was tested for significant differences. The result showed that MeCP2 enrichment was comparable in 99.96% of the genome at a *p*-value threshold of 0.001 (uncorrected). Together, these results suggested that Fl-H1.0 was properly expressed with no measured effect on MeCP2.

Next, we crossed Fl-H1.0 and Cam-tTA mice with *Mecp*$^{+/-}$ mice to obtain Cam-tTA:Fl-H1.0:*Mecp2*$^{+/y}$ (*fh-Mecp2*$^{+/y}$) and Cam-tTA:Fl-H1.0:*Mecp2*$^{-/y}$ (*fh-Mecp2*$^{-/y}$) mice, which expressed Fl-H1.0 in the presence and absence of MeCP2, respectively. Overall health of *Mecp2*$^{+/y}$ and *Mecp2*$^{-/y}$ mice was not affected by Fl-H1.0 expression (Supplementary Fig. 3a, b). The level of Fl-H1.0 was similar in *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice (Supplementary Fig. 3c-e and Supplementary Fig. 5). Together with our results on the endogenous H1.0 (Supplementary Fig.1f, g), it is likely that H1.0 level is not robustly affected by MeCP2. However, because binding patterns might change even in the face of unchanged protein levels, we reasoned that clarifying genome-wide distribution of Fl-H1.0 and its relations to MeCP2 is needed. We therefore performed Fl-H1.0 ChIP-Seq using the frontal cortex of *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice (8–9 weeks). Sequencing was again performed with a high read count to achieve sufficient resolution and was performed using three mice as biological replicates to account for variability (Supplementary Table 1). Consistent with previous reports[15], Fl-H1.0 was distributed throughout the genome but depleted around transcriptional start sites (TSS) of transcribed genes (Fig. 2a). To look for changes in genome-wide distribution of Fl-H1.0, we computed Spearman correlation coefficients for the normalized ChIP-Seq reads between replicates. We observed high correlation between replicates regardless of the genotypes (Supplementary Fig. 4a), indicating that MeCP2 depletion had little impact on Fl-H1.0 distribution. Consistently, principal component analysis showed no separation between genotypes (Supplementary Fig. 4b).

To clarify the relationship between H1 and MeCP2, we computed correlations between Fl-H1.0 and MeCP2 ChIP-Seq data from the Cam-tTA:Fl-H1.0 mice, which were genetically identical to *fh-Mecp2*$^{+/y}$ mice. If MeCP2 competes with H1, we would expect to see a negative correlation between MeCP2 and Fl-H1.0, which should shift towards a positive correlation upon MeCP2 depletion. Contrary to our expectation, however, the correlations between MeCP2 and Fl-H1.0 were positive in both *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice (Supplementary Fig. 4a), and the correlation coefficient did not change when MeCP2 was depleted. This result suggested that Fl-H1.0 and MeCP2 occupied similar genomic regions, while the distribution of Fl-H1.0 was unaffected by the presence of MeCP2. To further validate our finding, we pooled the three replicates for each genotype together. The correlation for Fl-H1.0 between *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ was again very high (Fig. 2b), and the correlation between Fl-H1.0 and MeCP2 ChIP-Seq reads were also consistently high

in both *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice (Fig. 2c, d). We also performed correlation analysis using larger (1 Mb and 10 Mb) and smaller (1 kb and 0.2 kb) window sizes and obtained consistent results (Supplementary Fig. 4c).

To examine more local effects on Fl-H1.0, we examined Fl-H1.0 binding in the regions where MeCP2 is enriched, which included heterochromatic repeat elements and MeCP2 peaks identified from ChIP-Seq. We detected no change in Fl-H1.0 in either region (Fig. 3a, b). The results from the repeat elements were confirmed by ChIP-qPCR in the frontal cortex and hippocampus (Fig. 3c, d). To examine Fl-H1.0 in euchromatic regions and to correlate the results with gene transcription, we mapped Fl-H1.0 reads around TSS of genes, which were grouped according to the RNA expression levels[20]. Consistent with previous reports[15], the magnitude of depletion around TSS correlated with the gene expression (Fig. 3e). However, this correlation was observed similarly in *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice (Fig. 3e, Supplementary Fig. 4d). We also examined Fl-H1.0 distribution around the TSS of those genes differentially expressed in the cortical layer 5 excitatory cells of *Mecp2*$^{-/y}$ mice[20], and found no change (Fig. 3f and Supplementary Fig. 4e, f). Taken together, ChIP-Seq data revealed no change in Fl-H1.0 distribution upon MeCP2 depletion. This was further validated by the window analysis, in which we compared averaged enrichment for Fl-H1.0 in 200-bp windows using each biological replicate. The result showed Fl-H1.0 enrichment was comparable in 99.10% of the regions at a p-value threshold of 0.001 (uncorrected).

In conclusion, we found that H1.0 and MeCP2 tend to bind to similar genomic regions but, interestingly, we detected little change in H1.0 distribution upon MeCP2 depletion, suggesting that MeCP2 works independently of competition with linker H1.0. High-depth sequencing using multiple biological replicates was crucial to reach this conclusion since both MeCP2 and H1.0 bind the DNA broadly throughout the genome. Because ChIP-Seq has a limitation in detecting mild changes that occur diffusely, we may have missed subtle changes. However, the lack of major difference in Fl-H1.0 binding pattern in *Mecp2*-null brains (Fig. 2 and 3) and the lack of change in the MeCP2 distribution upon mild overexpression of H1.0 (Fig. 1) using both ChIP-Seq and ChIP-qPCR clearly showed that there is no direct MeCP2-H1.0 competition in the cortical excitatory neurons. Our findings argue for the need for additional studies to pinpoint the molecular mechanism by which MeCP2 alters chromatin architecture.

Online Methods

Animals

Mice were housed in an AAALAS-certified Level 3 facility on a 14 hour light cycle. The mice were weaned at 21 days after birth, and housed with 3–5 littermates per cage. Fl-H1.0 mice were generated in our lab (see below) and maintained in FVB background. Cam-tTA mice [129S6.Cg-Tg(Camk2α-tTA)1Mmay/JlwsJ] were purchased from Jackson Laboratory and maintained in 129S6 background [21]. *Mecp2*–null (Guy et al., 2001) and Mecp2-Tg3 lines [22] were maintained in 129S6 and FVB background, respectively. Male Cam-tTA heterozygous mice were mated with female Fl-H1.0 heterozygous mice, and the male offspring with 129/FVB hybrid background was used for analysis. To study Fl-H1.0 in *Mecp2*$^{-/y}$ mice and their *Mecp2*$^{+/y}$ littermate, we first mated Cam-tTA heterozygous male

mice with female *Mecp2*$^{-/+}$ mice to obtain females heterozygous for both Cam-tTA and *Mecp2*. These females (129S6 background) were crossed with male Fl-H1.0 mice (FVB background) to obtain male Cam-tTA:Fl-H1.0:*Mecp2*$^{+/y}$ and Cam-tTA:Fl-H1.0:*Mecp2*$^{-/y}$ mice (129/FVB hybrid background). These mice were scored based on six phenotypic features while the experimenter was blinded to the genotypes [23]. Other experiments were performed without blinding to genotypes. No randomization was performed but all the experiments were conducted using appropriate littermate controls. All procedures to maintain and use these mice were approved by the Institutional Animal Care and Use Committee for Baylor College of Medicine and Affiliates.

Generation of Flag-H1.0 Transgenic Mice

Mouse gene encoding H1.0 was directly cloned from wild-type C57/BL6 mouse genomic DNA by PCR using following primer sets to incorporate 3XFlag-tag to the N-terminus of H1.0:

Forward: 5'-ccgccggaattcaccatggactacaaagaccatgacggtgattataaagatcatgatatcgattacaaggatgacgatgacaag-accgagaactccacctccgc-3' and Reverse: 5'-ccgccgaagctttcacttcttcttgctggccct-3'. The PCR product was digested by restriction enzymes (EcoRI/HindII) and was inserted into p-TreTight vector (derived from Addgene vector 23966). The vector was purified by Nucleobond Xtra Midi Kit (Clontech), digested with BciVI, and the fragment was used for pronuclear injection according to standard procedures. The mouse strain was maintained in FVB background and stable germline transmission was confirmed for at least four generations. These Flag-H1.0 transgenic mice have been donated to and will be available from the Jackson Laboratory.

Gene Expression Analysis with RT-qPCR

Cortex from 8–9-week old mice were dissected and immediately frozen by liquid nitrogen. The frozen tissue was placed in 1 ml Purezol (Bio-Rad) on ice and immediately homogenized with Polytron homogenizer. The dissolved sample was diluted to 2X using Purezol, and 1 ml of the solution was processed by Aurum Total RNA Fatty and Fibrous Tissue kit (Bio-Rad) to collect RNA. DNA within the sample was removed with on-column DNase digestion which was included in the kit. First-strand cDNA was synthesized using M-MLV reverse transcriptase (Life Technologies). RT-qPCR was performed using Bio-Rad CFX96 Real-Time system. The primer sets were designed to amplify target genes using UCSC genome browser and Primer 3. The reactions were conducted in either duplicate or triplicate, and relative proportions of the cDNA were determined based on the threshold cycle (Ct). The results were averaged for each sample and normalized to the value of *gapdh*. Relative expression levels of the target genes were determined by normalizing the fold expression level in each sample to the average of WT controls. Statistical analyses were performed using delta Ct values over *gapdh* (i.e. $Ct_{gene} - Ct_{gapdh}$). Following PCR primers were used:

Gapdh: Forward 5'- ggagattgttgccatcaacga-3', Reverse 5'- tgaagacaccagtagactccacgac-3'

H1.0: Forward 5'- cacggaccaccccaagtattc-3', Reverse 5'- acccaccttgtagtggctct-3'

Flag-h1.0: Forward 5'- gactacaaagaccatgacggtga-3', Reverse 5'-acggttcttctctgcctggatag-3'

H1.c: Forward 5'- aaggtcaagagcgcgtctaa-3', Reverse 5'-tttcttggctgcaaccttct-3'

H1.x: Forward 5'- cgcgaagaaagtgaagaagg-3', Reverse 5'-gggcggatagggatagagac-3'

Mecp2: Forward 5'- acagcggcgctccattatc-3', Reverse 5'-cccagttaccgtgaagtcaaaa-3'

Histone Extraction and Western Blot

Cortex from 8–9-wk old mice were dissected and immediately frozen by liquid nitrogen. The frozen tissue was placed in 1 ml PBS on ice and gently homogenized using a dounce homogenizer. The samples were spun down at 6000 rpm for 5 min. The supernatant was discarded and 500 μl of ice-cold 0.4 M Sulfuric acid was added to the pellet. After overnight rotation at 4 ºC, the sample was centrifuged at 13,200 rpm for 10 min to collect supernatant, and 125 μl of trichloroacetic acid (TCA) was added. After 1 hour incubation, the precipitate containing histones was collected by spinning at 13,200 rpm, 4 ºC, for 30 min. The precipitate was dissolved in 2xSDS buffer (100 mM Tris (pH 6.8), 4% SDS, 20% glycerol), protein concentration was adjusted using Pierce BCA protein assay kit (Thermo Fisher) and was used for western blots using NuPAGE Bis-Tris gels. In the experiments where we extracted histones from the sorted nuclei, acid extraction was performed by overnight incubation with 0.4N sulfuric acid, followed by addition of the one tenth volume of 4N sodium hydroxide, and the sample was processed for western blot without precipitation by TCA. Relative protein levels of H1-isoforms, histone modifications, and MeCP2 were normalized to H3 levels. H3 levels were measured either by loading equivalent ratio of samples on a separate gel (Fig. 1f, Supplementary Fig. 3d) or by detecting both the target protein and H3 from the same gel (Supplementary Fig. 1). Full scan western blot images for each representative blot are shown in Supplementary Fig. 5. Following antibodies were used for western blot: HRP-conjugated anti-Flag (1:50,000, mouse monoclonal, Sigma, #A8592), anti-H1.0 (1:500, mouse monoclonal, Abcam, #ab11079), anti-H1.c (1:500, rabbit polyclonal, Proteintech, #19649–1-AP), anti-H1.x (1:500, rabbit polyclonal, Abcam, #ab31972), anti-H3 (1:5000, rabbit monoclonal, Cell Signaling, #4499P), anti-H3K9me3 (1:5000, rabbit polyclonal, Abcam, #ab8898), anti-H3K27me3 (1:5000, rabbit polyclonal, Millipore, #07–449) and anti-MeCP2 (1:5000, rabbit polyclonal, in house). Secondary antibodies conjugated with HRP were mouse anti-rabbit (Jackson ImmunoResearch Laboratories, 211–032-171) and donkey anti-mouse antibodies (Jackson ImmunoResearch Laboratories, 715–035-150). Immunoblot images were acquired with the ImageQuant LAS 4000 (GE Healthcare) and quantified using Image-Pro Analyzer 7 (Media Cybernetics).

Isolation of Neuronal Nuclei by Cell Sorting

Nuclei were isolated from frozen cortices as previously described [24] with a few modifications. Bilateral cortices were homogenized in 5 ml of ice-cold homogenization buffer (320 mM sucrose, 5 mM $CaCl_2$, 3 mM $Mg(Ac)_2$, 10 mM Tris (pH8.0), 0.1 mM EDTA, 0.1% Triton-X, 0.1 mM PMSF, 1x complete protease inhibitors (CPI, Sigma, #04693132001)). After incubation for 5 min, the lysate was mixed with 50% OptiPrep medium (50% OptiPrep, 320 mM sucrose, 5 mM $CaCl_2$, 3 mM $Mg(Ac)_2$, 10 mM Tris (pH8.0), 0.1 mM PMSF, and 1x CPI) and was gently loaded on top of 10 ml 29% iso-

osmolar OptiPrep solution. Samples were centrifuged at 10,000 × g, 4 °C, for 30 min. The nuclei pellet was resuspended in ice-cold nuclei suspension buffer (65 mM β-glycerophosphate (pH 7.5), 2 mM $MgCl_2$, 25 mM KCl, 340 mM sucrose, 5% glycerol, 0.1 mM PMSF, and 1x CPI), and incubated with Alexa 555-conjugated anti-NeuN antibody (1:500, mouse monoclonal, Millipore, #MAB377A5) and 2% normal goat serum for 1 hour at 4 °C. The sample was filtered through a cell strainer, diluted with the nuclei suspension buffer, and NeuN-positive nuclei were isolated using Cell Sorter SH800 (Sony). In order to set gatings, we first used FSC-A x BSC-A plot to separate larger particles from smaller debris. Next, we used FSC-A x FSC-H plot to remove duplicates. Finally, NeuN-positive population was identified based on the Alexa 555 fluorescent intensity. Fluorescence event histogram showed clear two peaks. NeuN-positive fraction was isolated by sorting 70–75% of the nuclei from the higher peak to increase specificity. We used the same gating threshold for the nuclei incubated with Alexa 555-labeled mouse IgG, and confirmed that the positive fraction was almost negligible.

Chromatin Immunoprecipitation

ChIP was performed as previously described [7] with several modifications. Frozen tissue was cross-linked by incubating in 1% PFA (Fisher, # 28906) in PBS for 10 min at room temperature. After quenching the fixation with 125 mM glycine in ice-cold PBS, the samples were homogenized in hypotonic solutions (10 mM Tris (pH 7.5), 0.5% Igepal CA-630, 10 mM NaCl, 30 mM $MgCl_2$, 1:1000 RNase cocktail (Ambion, #2286), 1 mM PMSF, and 1x CPI). Nuclei were collected at 3,000 rpm in a table-top centrifuge and resuspended in nuclei lysis buffer containing 50 mM Tris (pH 8), 10 mM EDTA, 1% SDS, 1 mM PMSF, and 1x CPI. The sample was incubated in the nuclei lysis buffer for 15 min on ice, and sonicated by Bioruptor Pico (Diagenode) for 7–10 cycles (30 sec off/ 30 sec on) to obtain DNA fragments of 100–500 bp. The chromatin samples were flash frozen in liquid nitrogen and stored at −80 ºC until needed. To perform ChIP, chromatin sample was first diluted based on their DNA concentration, further diluted in ChIP dilution buffer (0.01% SDS, 1.1% Triton-X, 1.2 mM EDTA, 16.7 mM Tris (pH 8.1), and 167 mM NaCl), and was prerecleared with Protein G Dynabeads (Invitrogen) for 1 hour. Approximately 1 μg of chromatin (in 1 ml volume) was incubated overnight at 4 °C with 1 μg of anti-Flag antibody (Sigma, #ab6556). An aliquot of 20% (vol/vol) of the precipitated chromatin was stored as input. The next day, 40 μl of Protein G Dynabeads were added, and the sample was rotated for 3 hours at 4 °C. The beads were then washed in 700 μl each of low-salt buffer (0.1% SDS, 1% Triton-X, 2 mM EDTA, 20 mM Tris (pH 8.1), and 150 mM NaCl), high-salt buffer (0.1% SDS, 1% Triton-X, 2 mM EDTA, 20 mM Tris (pH 8.1) and 500 mM NaCl), LiCl wash buffer (250 mM LiCl, 1% Igepal-CA630, 1 mM EDTA and 10 mM Tris (pH 8.1)) and twice with TE+NaCl buffer (10 mM Tris HCl, 1 mM EDTA and 50 mM NaCl), for 5 min each at room temperature. The beads were eluted twice in 250 μl of elution buffer (1%SDS, 100 mM $NaHCO_3$) for 15 min each at 65 °C. Precipitated chromatin and input samples were reverse crosslinked and treated with proteinase K. ChIP-DNA and input-DNA were recovered using PCR Purification Kit (Qiagen), and were used for qPCR or made into a library for sequencing. ChIP by anti-MeCP2 antibody (Millipore, #17–10491) was performed following the same steps, using 5 μg of chromatin (in 0.5 ml volume), 7.9 μg of the antibody, and Protein A Dynabeads (Invitrogen). The ChIP-Seq library was prepared and

its quality was checked by Genomic and RNA Profiling Core at Baylor College of Medicine using ThruPLEX DNA-seq kit (Rubicon). Sequencing was performed by Illumina HiSeq 2500 instrument following manufacturer's protocol. Following primers were used for ChIP-qPCR experiments:

Major satellite repeat: 5'- catccacttgacgacttgaaaa-3', 5'- gaggtccttcagtgtgcattt-3'.

Bdnf: 5'- ggattcccttcatcctcagat-3', 5'- ccaaagagtaagtgtgcccttc-3'.

Myc-4583: 5'- gccctttctttctggaggtc-3', 5'- gcccagaacctcactgtctc-3'.

Myc-3673: 5'- gtcatcaggctggggtacag-3', 5'- taggtgtgggaacccaggta-3'.

Myc-838: 5'- tgtagaccggcagagactcc-3', 5'- ggctctttcccctgtaggtc-3'.

Myc+1170: 5'- acccgtccctgagcctttta-3', 5'- gcgggggtcaggcttaaat-3'.

Gapdh (TSS): 5'-gtgcagtgccaggtgaaaat-3', 5'- agcatccctagacccgtacagt-3'.

Gapdh (promoter): 5'- gcagagattatgaaccccactc-3', 5'- cctgcagacccccttagtgtagt-3'.

Minor satellite repeat: 5'-catggaaaatgataaaaacc-3', 5'catctaatatgttctacagtgtgg-3'.

LINE L1: 5'- tttgggacacaatgaaagca-3', 5'- ctgccgtctactcctcttgg-3'.

SINE B1: 5'- gtggcgcacgcctttaatc-3', 5'- gacagggtttctctgtgtag-3'.

Primer sequences for *Major satellite repeat* and *Bdnf* are identical to those used in Chen et al[7]. Primer sequences for the *Myc* locus are from Cohen et al[16]. Those for *Minor satellite repeat, LINE L1*, and *SINE B1* are from Martens et al[25]. Enrichment over input (%) was plotted in the graphs. Statistical analyses were performed using delta Ct values (i.e. $Ct_{ChIP} - Ct_{input}$).

ChIP-Seq Analysis

***Data processing*:** A total of 24 ChIP libraries were sequenced using Illumina HiSeq 2500: 3 wild-type MeCP2 ChIP samples, 3 Cam-tTA:Fl-H1.0 MeCP2 ChIP samples, 3 *fh-Mecp2*$^{+/y}$ Fl-H1.0 ChIP samples, 3 *fh-Mecp2*$^{-/y}$ Fl-H1.0 ChIP samples, and 12 input libraries for respective ChIP samples. Each library was sequenced to a depth of ~90 million, 100bp, pair-end reads. Cutadapt version 1.7.1 was used to remove/trim adapter sequences. Trimmed paired mates were mapped to the mus musculus genome (mm10) using Bowtie2 with following parameters: -D 20 -R 3 -N 0 -L 20 -i S,1,0.50 (very-sensitive). The average mappability across all the samples was ~95%. SAMtools V0.1.19 and BEDtools v2.17.0[26] were used to convert Sequence Alignment/Map (SAM) files to Binary Alignment Map (BAM) and Browser Extensible Data (BED) files, respectively. Fl-H1.0 and MeCP2 ChIP-Seq samples (bam files) were normalized as the ratio to their respective inputs and converted to bigwig format using the bamCompare module in deepTools[27] (with parameters: ratio, --binsize 50, --smooth 150 and normalizeTo1x). Normalized bigwig files were visualized using UCSC genome browser. Genome level correlations between the data sets were

computed by binning the genome into 0.2 kb, 1kb, 10kb, 1Mb, and 10Mb windows using deepTools modules[27], computeMatrix and plotCorrelation.

***Methylation data*:** DNA methylation data from the frontal cortex was obtained from the GEO database (GSM1173786)[19] and the base level methylation was computed as the ratio of methylated Cytosines to total covered Cytosines. As the original data had been mapped to mm9, we re-mapped the data to mm10 coordinates using CrossMap 0.2.6[28] and converted the data to bigwig tracks using bedGraphToBigWig program in UCSC utilities. Line plots illustrating correlation/anti-correlation between mCG/mCH/unmethylated-C levels and MeCP2 ChIP-Seq data sets were plotted at 1kb resolution by sorting MeCP2 enrichment from low to high. Methylation and MeCP2 ChIP-Seq data sets were smoothed using a moving average window of 1 kb. Confidence intervals (95%) of MeCP2 ChIP-Seq data sets were computed from the biological replicates and plotted with the mean values of each genotype.

***Mecp2 binding around CG islands*:** Mouse (mm10) CG island coordinates were obtained from USCS genome browser and MeCP2 binding to CG islands was computed using ComputeMatrix and plotProfile modules of deepTools[27]. Significance of the difference between wild-type and Cam-tTA:Fl-H1.0 MeCP2 ChIP-Seq was computed using respective replicates.

***Peak calling*:** Peaks from the MeCP2 ChIP-Seq were called using MACS [29] with parameters --nomodel, --broad and --broad-cutoff 0.1. Aggregation plots of normalized Fl-H1.0 ChIP-Seq enrichment around MeCP2 peak regions (5 kb upstream and downstream of the peaks) were generated using ComputeMatrix and plotProfile modules of deepTools[27].

***Repeat quantification*:** To analyze Fl-H1.0 enrichment in repetitive elements, repeat annotations and respective sequences of mm10 were downloaded from UCSC table browser. Major satellite, minor satellite, L1 LINE and B1 SINE elements were quantified from *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ samples using Salmon[30].

Gene expression data: Microarray data from the pyramidal cells in motor cortex in *Mecp2*$^{+/y}$ and *Mecp2*$^{-/y}$ mice were acquired from GEO (GSE8720)[20]. The genes were categorized based either on their expression levels (top 20%, bottom 20%, and the rest) or on the expression level difference between *Mecp2*$^{+/y}$ and *Mecp2*$^{-/y}$ mice (top 100 genes up-regulated in *Mecp2*$^{-/y}$, top 100 genes down-regulated in *Mecp2*$^{-/y}$ mice, and the genes un-altered). The aggregation plots of Fl-H1.0 ChIP-Seq along the TSSs were generated with respect to these categories. Significance of the difference between *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ was computed using respective replicates.

Immunofluorescence

Brain sections were obtained as previously described[31]. The brains were fixed by transcardial perfusion of PBS-buffered 4% paraformaldehyde (PFA). Brains were removed, kept in 4% PFA overnight, cryoprotected in 25% sucrose solution, and frozen in optimal cutting temperature medium (O.C.T). Sagittal sections were obtained using a Leica

CM3050S cryostat at 25 μm thickness. The slices were incubated in a PBS-buffered blocking solution containing 2% normal goat serum and 0.3% Triton-X for 1 hour, followed by incubation in a primary antibody solution containing anti-MeCP2 antibody (1:1000, rabbit monoclonal, D4F3, Cell Signaling, Cat #3456), anti-Camk2a antibody (1:1000, rabbit monoclonal, Abcam, #ab52476), and anti-Flag antibody (1:1000, mouse monoclonal, M2, Sigma, #F1804) overnight at 4ºC. After three washes with PBS, the slices were incubated with secondary antibodies conjugated with either Alexa 488, 555 or 633 (1:1000, Invitrogen) overnight at 4ºC. After one wash wish PBS, PBS containing DAPI (2.5 μg/ ml) was applied for 5 min, followed by two additional washes with PBS. To obtain high-magnification images of Fl-H1.0, we used FITC-labeled anti Flag antibody (1:1000, mouse monoclonal, Sigma, #F4049) with antigen retrieval. The antigen retrieval was performed by incubating slices into the buffer containing 10 mM sodium citrate (pH 6) and 0.05% Tween 20, and heating them at 95ºC for 5 min. Stained brain sections were imaged using a Leica LCS SP8 confocal microscope with either 60x glycerol lens (N.A. 1.3) or 20x air lens (N.A. 0.7), and the images were processed using ImageJ and Image-Pro Analyzer 7.0 (Media Cybernetics). All image sets were acquired at a same laser / gain settings when comparing multiple genotypes, and they were processed using the same intensity thresholds.

Statistics Summary

Statistical tests for each assay were chosen based on their appropriateness for the assay. All statistical calculations were carried out using Graphpad Prism software with alpha level of 0.05. No statistical method was used to pre-determine sample size but our sample sizes are similar to or exceed previous publications[6,7]. For experiments comparing two groups, data distribution was examined by *F*-test. Two-tailed t-test was performed when the result was not significant, whereas non-parametric test was performed when the result from *F*-test was significant. For experiments comparing multiple groups, data distribution was assumed to be normal and this was not formally tested. Variance was estimated by standard errors of mean. Multiple comparisons were corrected by Tukey's post-hoc test. Statistical methods and values including all p values were described in figure legends.

Data Availability

All data sets in the manuscript will be available upon request. Data sets from MeCP2 and Fl-H1.0 ChIP-Seq were deposited to Gene Expression Omnibus database (GSE107533).

Life Sciences Reporting Summary

Additional information on the materials and methods can be found in the Life Sciences Reporting Summary.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was supported by NIH/NINDS 5R01NS057819 (H.Y.Z.), Japan Society for the Promotion of Science (A.I-I.), Genetically Engineered Mouse Core and Genomic and RNA Profiling Core at BCM, Neurovisualization

Core at the BCM Intellectual and Developmental Disabilities Research Center (NIH/1U54HD083092), NSF DMS#1263932 and CPRIT RP170387 (Z.L.). The authors declare no conflicts of interest. H.Y.Z. is an investigator with the Howard Hughes Medical Institute.

References

1. Amir RE et al. Nat Genet 23, 185–8 (1999). [PubMed: 10508514]
2. Lombardi LM, Baker SA & Zoghbi HY J Clin Invest 125, 2914–23 (2015). [PubMed: 26237041]
3. Nan X, Campoy FJ & Bird A Cell 88, 471–81 (1997). [PubMed: 9038338]
4. Lewis JD et al. Cell 69, 905–14 (1992). [PubMed: 1606614]
5. Chahrour M et al. Science 320, 1224–1229 (2008). [PubMed: 18511691]
6. Skene PJ et al. Mol Cell 37, 457–68 (2010). [PubMed: 20188665]
7. Chen L et al. Proc Natl Acad Sci U S A 112, 5509–14 (2015). [PubMed: 25870282]
8. Nikitina T et al. J Biol Chem 282, 28237–45 (2007). [PubMed: 17660293]
9. Ghosh RP, Horowitz-Scherer RA, Nikitina T, Shlyakhtenko LS & Woodcock CL Mol Cell Biol 30, 4656–70 (2010). [PubMed: 20679481]
10. Baker SA et al. Cell 152, 984–96 (2013). [PubMed: 23452848]
11. Kishi Y, Fujii Y, Hirabayashi Y & Gotoh Y Nat Neurosci 15, 1127–33 (2012). [PubMed: 22797695]
12. Catez F et al. Mol Cell Biol 24, 4321–8 (2004). [PubMed: 15121851]
13. Gemelli T et al. Biol Psychiatry 59, 468–76 (2006). [PubMed: 16199017]
14. Lu X et al. Science 340, 78–81 (2013). [PubMed: 23559249]
15. Cao K et al. PLoS Genet 9, e1003417 (2013). [PubMed: 23633960]
16. Cohen S et al. Neuron 72, 72–85 (2011). [PubMed: 21982370]
17. Lagger S et al. PLoS Genet 13, e1006793 (2017). [PubMed: 28498846]
18. Gabel HW et al. Nature 522, 89–93 (2015). [PubMed: 25762136]
19. Lister R et al. Science 341, 1237905 (2013). [PubMed: 23828890]
20. Sugino K et al. J Neurosci 34, 12877–83 (2014). [PubMed: 25232122]
21. Mayford M et al. Science 274, 1678–83 (1996). [PubMed: 8939850]
22. Collins AL et al. Hum Mol Genet 13, 2679–89 (2004). [PubMed: 15351775]
23. Guy J, Gan J, Selfridge J, Cobb S & Bird A Science 315, 1143–7 (2007). [PubMed: 17289941]
24. Swiech L et al. Nat Biotechnol 33, 102–6 (2015). [PubMed: 25326897]
25. Martens JH et al. EMBO J 24, 800–12 (2005). [PubMed: 15678104]
26. Quinlan AR & Hall IM Bioinformatics 26, 841–2 (2010). [PubMed: 20110278]
27. Ramirez F, Dundar F, Diehl S, Gruning BA & Manke T Nucleic Acids Res 42, W187–91 (2014). [PubMed: 24799436]
28. Zhao H et al. Bioinformatics 30, 1006–7 (2014). [PubMed: 24351709]
29. Zhang Y et al. Genome biology 9, R137 (2008). [PubMed: 18798982]
30. Patro R, Duggal G & Kingsford C biorxiv.org/content/early/2015/06/27/021592 (2015).
31. Ito-Ishida A, Ure K, Chen H, Swann JW & Zoghbi HY Neuron 88, 651–8 (2015). [PubMed: 26590342]

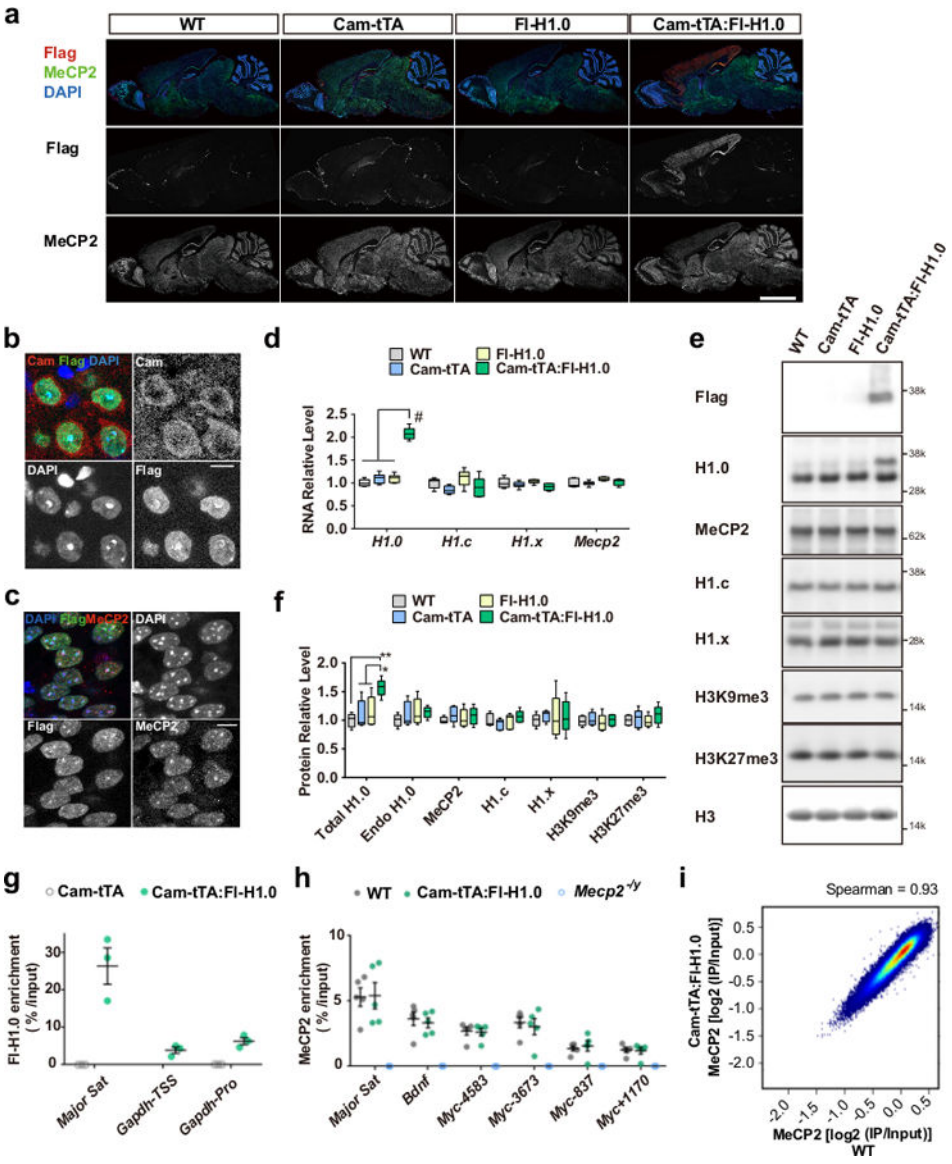

Figure 1.

Generation of mice expressing Flag-H1.0 in forebrain excitatory neurons.

a) Immunofluorescence images showing Fl-H1.0 expression in Cam-tTA:Fl-H1.0 mice (8–9 weeks). Scale bar 3 mm. The result was reproduced in four mice per genotype.

b, c) Fl-H1.0 was expressed in the nuclei and was enriched in DAPI-foci of layer 2 Camk2α + cortical cells (b) and of hippocampal CA1 pyramidal cells (c) in Cam-tTA:Fl-H1.0 mice (8–9 weeks). Scale bars 10 μm. The results were confirmed in four mice.

d) RNA expression levels of various genes in the cortex. N = 5 (WT), 6 (Cam-tTA), 6 (Fl-H1.0), 5 (Cam-tTA:Fl-H1.0) mice from at least two litters (8–9 weeks). One-way ANOVA for *H1.0*: $p < 0.0001$, F(3, 18) = 36.65. Tukey's multiple comparison for *H1.0*: $p = 0.26$ (WT vs Cam-tTA); $p = 0.26$ (WT vs Fl-H1.0); $p < 0.0001$(WT vs Cam-tTA:Fl-H1.0); $p > 0.9999$ (WT vs Fl-H1.0); $p < 0.0001$ (Cam-tTA vs Cam-tTA:Fl-H1.0); $p < 0.0001$(Fl-H1.0 vs Cam-

tTA:Fl-H1.0). One-way ANOVA for others: $p = 0.052$, F(3, 18) = 3.12 (*H1.c*); $p = 0.11$, F(3, 18) = 2.32 (*H1.x*); $p = 0.090$, F(3, 18) = 2.53 (*Mecp2*).

e) Western blot showing Fl-H1.0 expression in the cortex of Cam-tTA:Fl-H1.0 mice (8–9 weeks). Blot using anti-H1.0 antibody showed both endogenous and Flag-tagged H1.0. As the blot for H1.x resulted in a background band, we analyzed H1.x-corresponding band around 28 kDa. The experiment was repeated five times using independent mice with similar results.

f) Relative protein expression levels of H1.0 and other chromatin components, which were first normalized to H3 levels and then by the average of WT mice. N = 5 mice (per genotype) from at least two litters. One-way ANOVA for total H1.0: $p = 0.00224$, F(3, 16) = 7.466. Tukey's multiple comparison for total H1.0: $p = 0.88$ (WT vs Cam-tTA); $p = 0.68$ (WT vs Fl-H1.0); $p = 0.0024$ (WT vs Cam-tTA:Fl-H1.0); $p = 0.98$ (WT vs Fl-H1.0); $p = 0.011$ (Cam-tTA vs Cam-tTA:Fl-H1.0); $p = 0.023$ (Fl-H1.0 vs Cam-tTA:Fl-H1.0). One-way ANOVA for others: $p = 0.54$, F(3, 16) = 0.75 (endogenous H1.0); $p = 0.67$, F(3, 16) = 0.53 (MeCP2); $p = 0.3613$, F(3, 16) = 1.14 (H1.c); $p = 0.96$, F(3, 16) = 0.090 (H1.x); $p = 0.84$, F(3, 16) = 0.2851 (H3K9me3); $p = 0.59$, F(3, 16) = 0.66 (H3K27me3).

g) Fl-H1.0 enrichment was measured by ChIP-qPCR using the frontal cortex of Cam-tTA:Fl-H1.0 and control (Cam-tTA) mice. Fl-H1.0 enrichment was very low to undetectable (0–0.1%) in Cam-tTA. N =3 mice per genotype (8–9 weeks).

h) ChIP-qPCR using frontal cortex showed MeCP2 enrichment was comparable between WT and Cam-tTA:Fl-H1.0 mice in various MeCP2 targets (N = 5 mice per genotype from at least two litters, 14–21 weeks). ChIP-qPCR from *Mecp2*$^{-/y}$ mice (N =3 mice, 8–9 weeks) resulted in very low enrichment (0.00–0.02%). Two-tailed t-test between WT and Cam-tTA:Fl-H1.0: $p = 0.98$, $t_8 = 0.022$ (*Major Sat*); $p = 0.79$, $t_8 = 0.28$ (*Bdnf*); $p = 0.90$, $t_8 = 0.13$ (*Myc*-4583); $p = 0.56$, $t_8 = 0.60$ (*Myc*-3673); $p = 0.65$, $t_8 = 0.47$ (*Myc*+1170). Two-tailed Mann-Whitney U test: $p = 0.55$ (*Myc*-837).

i) Scatter plot showing correlation for MeCP2 ChIP-Seq reads between WT and Cam-tTA:Fl-H1.0 mice using 264,922 data points pooled from the three mice per genotype. ChIP-Seq reads were normalized to input and averaged at 10 kb window bins.

Box-and-whisker plots show median, 25th and 75th percentile, and min and max values (d, f). Graphs with individual data show average ± sem (g, h). * $p < 0.05$, ** $p < 0.01$, # $p < 0.0001$.

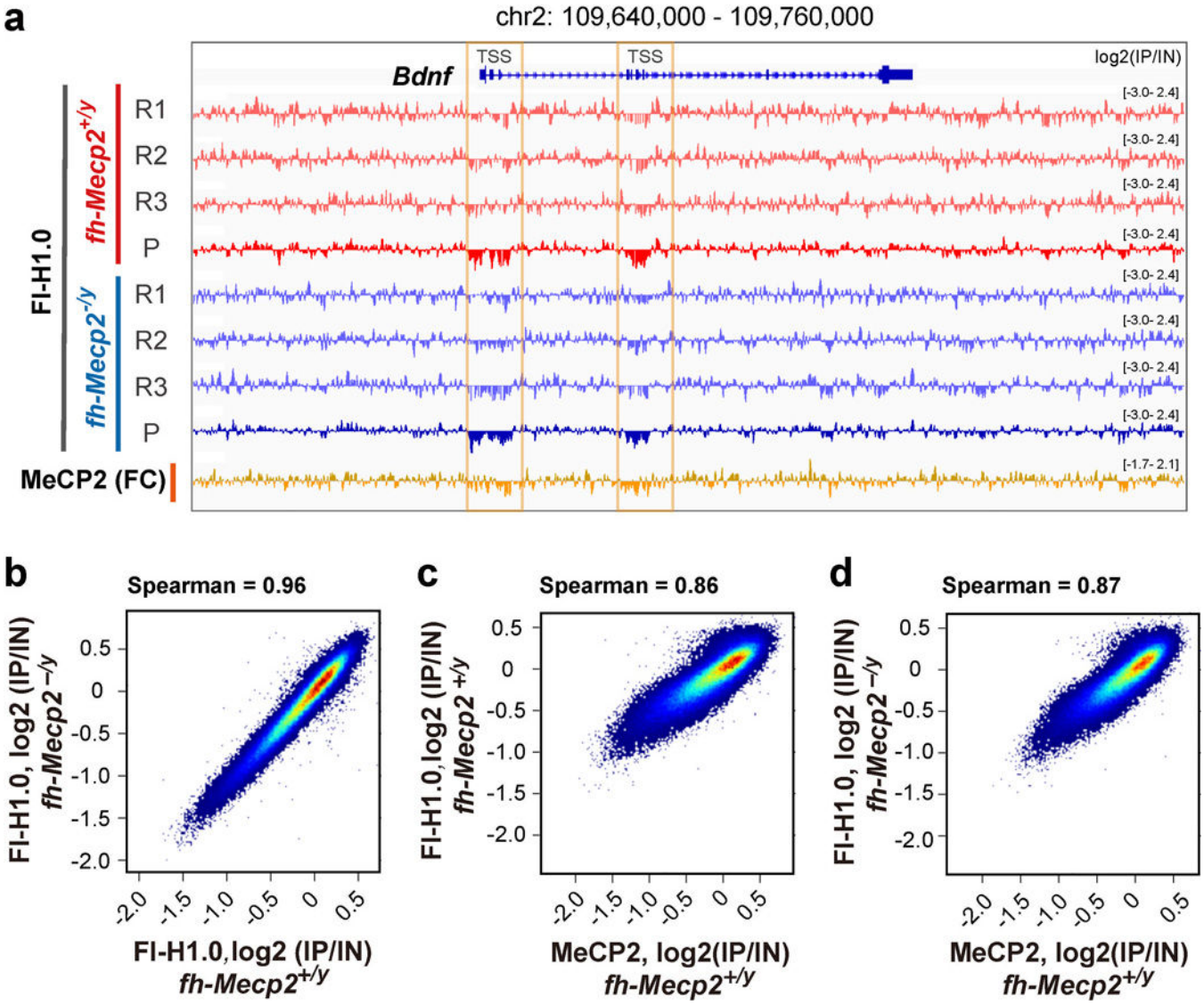

Figure 2.

ChIP-Seq revealed no changes in genomic distribution of Fl-H1.0 upon MeCP2 loss.

a) Representative tracks of Fl-H1.0 ChIP-Seq from individual replicates (R1–3) and pooled data sets (P) are shown together with MeCP2 ChIP-Seq from the frontal cortex (FC) of the Fl-H1.0 expressing mice. Note that Fl-H1.0 distribution is widely similar between replicates, and that Fl-H1.0 is depleted around two alternative start sites of *Bdnf* (orange box).

b) Scatter plot showing correlation for Fl-H1.0 ChIP-Seq reads between *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice (264,901 data points from pooled data).

c) Scatter plot showing correlation between MeCP2 and Fl-H1.0 ChIP-Seq from *fh-Mecp2*$^{+/y}$ mice (264,903 data points from pooled data).

d) Scatter plot showing correlation between MeCP2 and Fl-H1.0 ChIP-Seq from *fh-Mecp2*$^{-/y}$ mice (264,907 data points from pooled data).

ChIP-Seq reads were normalized to input and averaged at 10 kb windows (b−d).

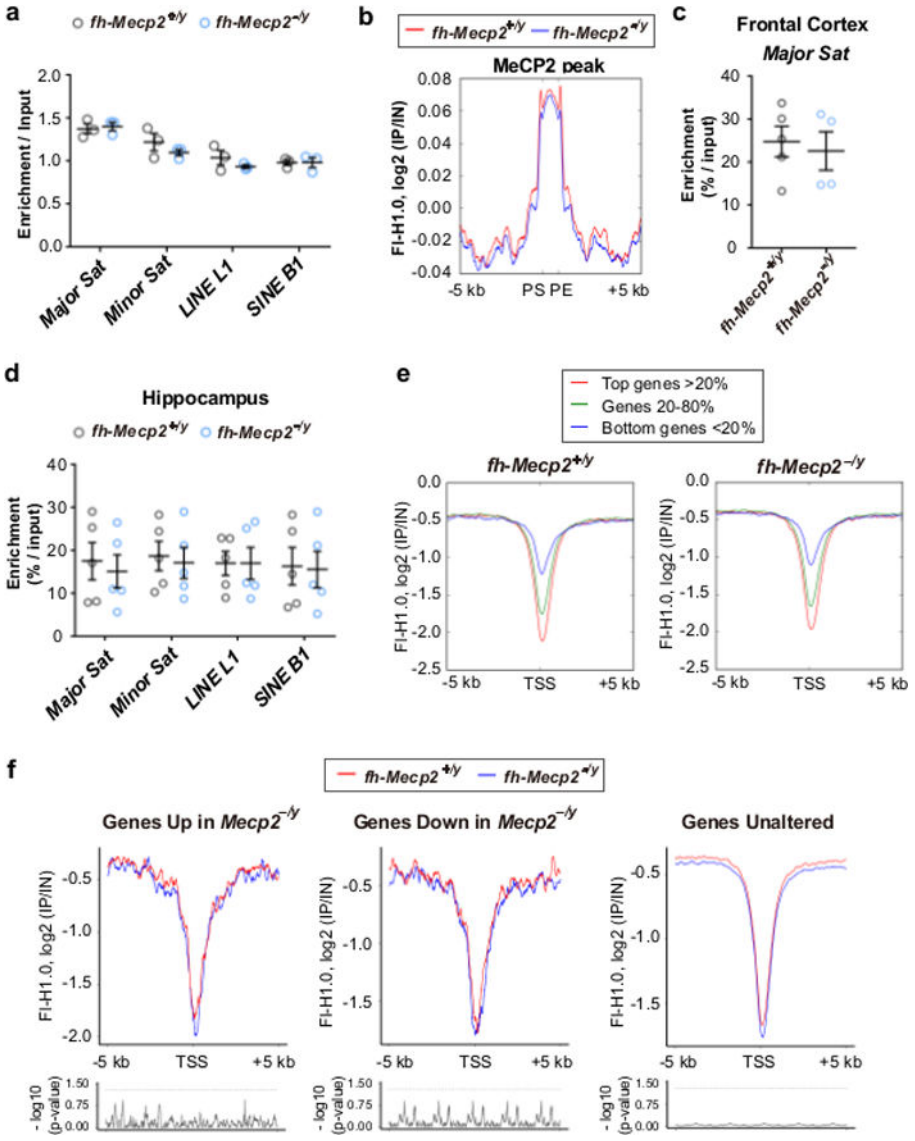

Figure 3.

MeCP2 depletion did not affect enrichment of Fl-H1.0 in the repeat sequences and around TSS.

a) Enrichment of Fl-H1.0 in the repeat elements was not affected in *fh-Mecp2*$^{-/y}$ mice. N = 3 mice from three litters. Two-tailed t-test: $p = 0.74$, $t_4 = 0.36$ (*Major Sat*); $p = 0.32$, $t_4 = 1.14$ (*Minor Sat*); $p = 0.29$, $t_4 = 1.21$ (*LINE L1*); $p = 0.97$, $t_4 = 0.040$ (*SINE B1*).

b) Enrichment of Fl-H1.0 was plotted around MeCP2 ChIP-Seq peaks obtained from the frontal cortex of Fl-H1.0 expressing mice. PS, peak start; PE, peak end.

c) Result from Chip-qPCR using the frontal cortex showed no difference in the Fl-H1.0 enrichment in major satellite repeats. N = 5 and 4 mice for *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice, respectively, from at least three litters. Two-tailed t-test: $p = 0.70$, $t_7 = 0.41$.

d) Results from Chip-qPCR using the hippocampi of Fl-H1.0 expressing mice. Enrichment of Fl-H1.0 in various repeats was not different between *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice.

N = 5 mice from at least three litters. Two-tailed t-test: $p = 0.71$, $t_8 = 0.38$ (*Major Sat*); $p = 0.72$, $t_8 = 0.37$ (*Minor Sat*); $p = 0.91$, $t_8 = 0.12$ (*LINE L1*); $p = 0.90$, $t_8 = 0.12$ (*SINE B1*)

e) Enrichment of Fl-H1.0 was aligned around the TSS of genes that were grouped into three based on expression levels.

f) Enrichment of Fl-H1.0 was aligned around the TSS of genes that were increased, decreased, and unaltered in *Mecp2*$^{-/y}$ excitatory neurons (top). P-values were obtained by running t-tests in every 100 bp windows between *fh-Mecp2*$^{+/y}$ and *fh-Mecp2*$^{-/y}$ mice and were plotted along the TSS' (bottom). Dotted lines indicate $p = 0.05$.

Graphs in (a), (c), and (d) show individual data with average ± sem.</sure_Caption>