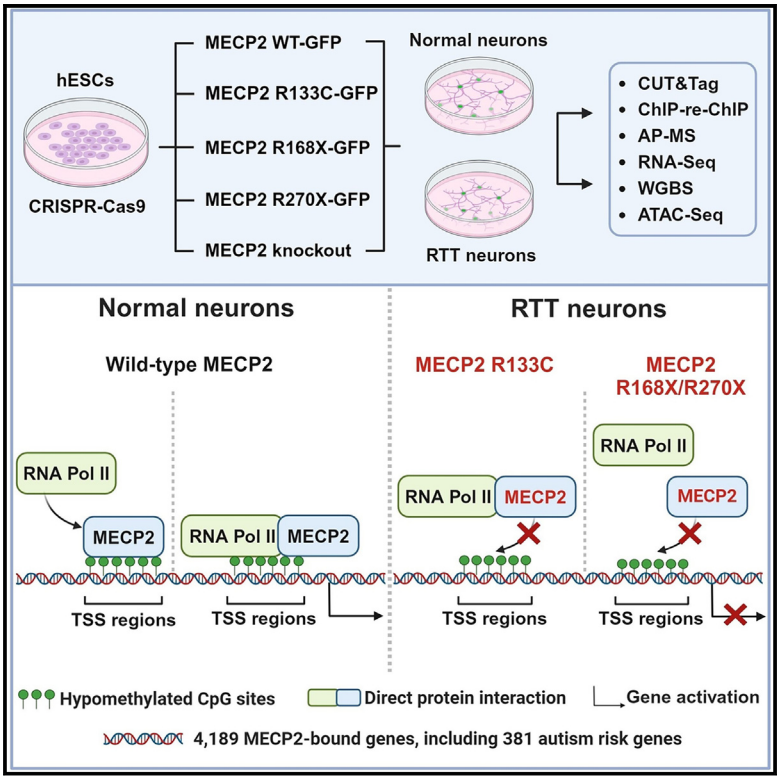
Neuron

Article

MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons

Graphical abstract



Highlights

- An integrated map of MECP2-binding sites in human normal and RTT neurons
- MECP2 and RNA Pol II co-occupy the TSSs
- A direct protein interaction between MECP2 and RNA Pol II
- MECP2 functions as a positive cofactor for RNA Pol II gene expression in neurons

Authors

Yi Liu, Anthony Flamier, George W. Bell, ..., Seychelle M. Vos, Richard A. Young, Rudolf Jaenisch

Correspondence

jaenisch@wi.mit.edu

In brief

Liu et al. utilized epigenomic, proteomic, and transcriptomic profiling to identify a conserved MECP2-RNA Pol II module critical for the regulation of neuronal gene expression, in particular to autism risk genes. Alteration to this module caused by *MECP2* patient mutations leads to transcriptional dysregulation associated with Rett syndrome (RTT).

Liu et al., 2024, Neuron *112*, 1–16
June 19, 2024
https://doi.org/10.1016/j.neuron.2024.04.007

CellPress

Neuron

CellPress

Article

MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons

Yi Liu,[1,7] Anthony Flamier,[1,5,6,7] George W. Bell,[1] Annette Jun Diao,[2] Troy W. Whitfield,[1] Hao-Che Wang,[1] Yizhe Wu,[1] Fabian Schulte,[1] Max Friesen,[1] Ruisi Guo,[1] Maisam Mitalipova,[1] X. Shawn Liu,[3] Seychelle M. Vos,[2,4] Richard A. Young,[1,2] and Rudolf Jaenisch[1,2,8,*]

[1]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[2]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[3]Department of Physiology and Cellular Biophysics, Columbia University Medical Center, New York, NY 10032, USA
[4]Howard Hughes Medical Institute, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[5]Present address: Department of Neuroscience, Université de Montréal, Montreal, QC H3C 3J7, Canada
[6]Present address: CHU Sainte-Justine Research Center, Montreal, QC H3T 1C5, Canada
[7]These authors contributed equally
[8]Lead contact
*Correspondence: jaenisch@wi.mit.edu
https://doi.org/10.1016/j.neuron.2024.04.007

SUMMARY

Mutations in the methyl-DNA-binding protein MECP2 cause the neurodevelopmental disorder Rett syndrome (RTT). How MECP2 contributes to transcriptional regulation in normal and disease states is unresolved; it has been reported to be an activator and a repressor. We describe here the first integrated CUT&Tag, transcriptome, and proteome analyses using human neurons with wild-type (WT) and mutant MECP2 molecules. MECP2 occupies CpG-rich promoter-proximal regions in over four thousand genes in human neurons, including a plethora of autism risk genes, together with RNA polymerase II (RNA Pol II). MECP2 directly interacts with RNA Pol II, and genes occupied by both proteins showed reduced expression in neurons with *MECP2* patient mutations. We conclude that MECP2 acts as a positive cofactor for RNA Pol II gene expression at many neuronal genes that harbor CpG islands in promoter-proximal regions and that RTT is due, in part, to the loss of gene activity of these genes in neurons.

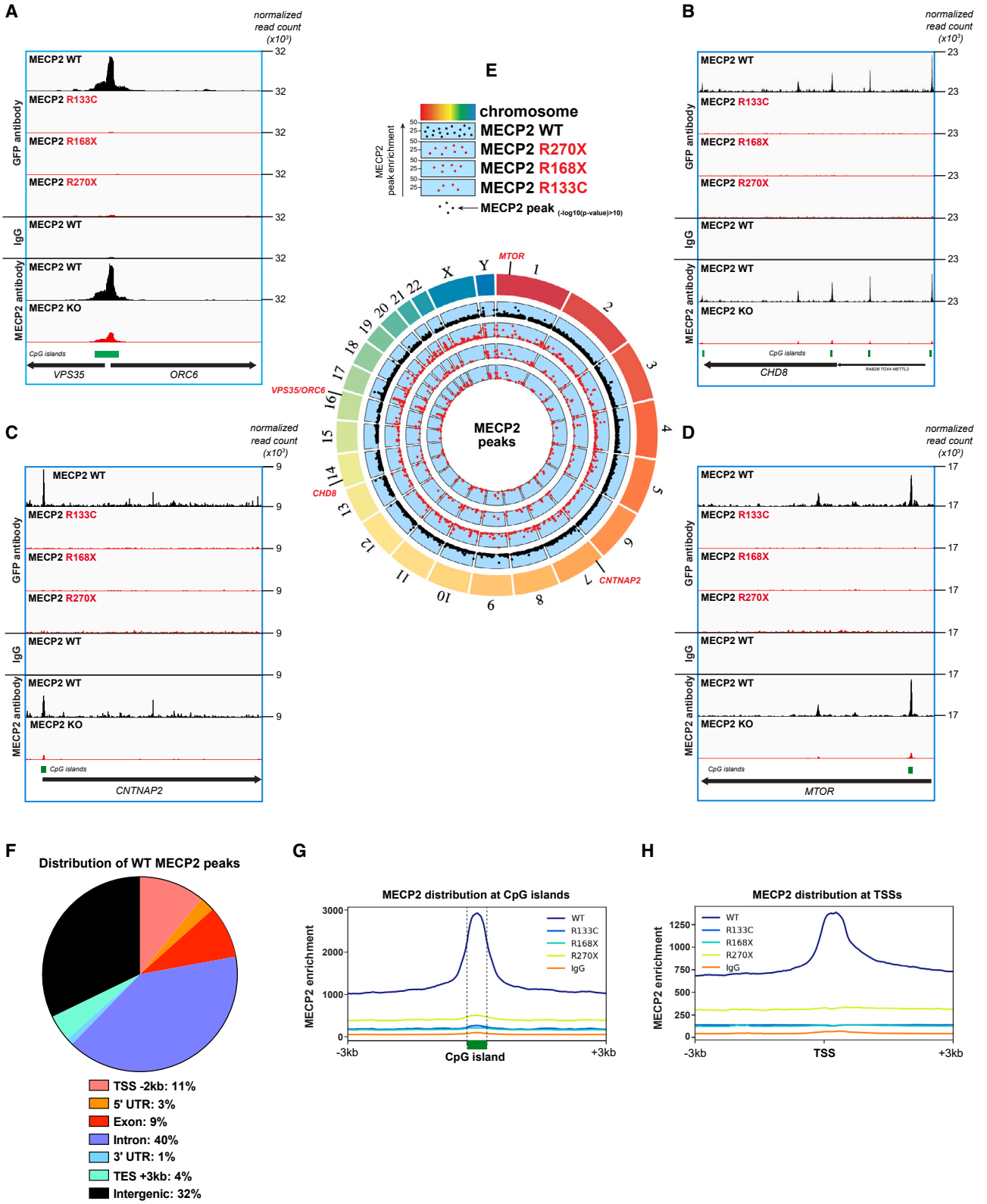
INTRODUCTION

Rett syndrome (RTT) is a progressive neurodevelopmental disorder characterized by severe mental disability and autism-spectrum-like behaviors that typically manifests in girls during early childhood and is caused by mutations in the X-linked methyl-CpG-binding protein 2 (*MECP2*) gene.[1–3] MECP2 is highly expressed in neurons,[4] and disruption of MECP2 in mouse brains causes RTT-like phenotypes, consistent with RTT being caused by MECP2 deficiency.[5–7] Despite significant efforts since the discovery that *MECP2* mutations cause RTT in 1999,[8] a clear consensus on how MECP2 dysfunction contributes to transcriptional regulation and disease development has not been achieved.

Although it is well established that MECP2 binds both 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC)-containing DNA,[9–14] it has proven challenging to identify the precise location of MECP2 binding throughout the human genome, likely due to limitations associated with reagents and technical processes involved in chromatin immunoprecipitation sequencing (ChIP-seq). Initial *in vitro* evidence suggested that MECP2 was strongly bound to methylated DNA and functioned as a transcriptional repressor based on the behavior of artificial expression constructs in cell lines.[15–17] However, *in vivo* studies on pre-symptomatic and post-symptomatic mutant mice indicated that MECP2 functioned as an activator of many genes, although it appeared to be necessary for repression of some genes.[18–25] Similarly, evidence for positive gene regulation came with reports of global transcriptional downregulation in mouse embryonic stem cell (mESC)- and human embryonic stem cell (hESC)-derived MECP2 mutant neurons.[26,27] Other studies proposed that MECP2 selectively repressed long genes with high gene body methylation.[28–34] These findings raised the perplexing question of how MECP2 could act as both a transcriptional repressor and activator. Because high-quality binding data are lacking, it has been difficult to ascertain whether the observed transcriptional changes are a direct or indirect effect of MECP2 deficiency.

It is important to note that most MECP2 genomic mapping studies to date were performed using adult mouse brains. A major technical challenge here is likely due to the resolution of ChIP-seq, which is not sufficient to deconvolute the specificity of

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007



(legend on next page)

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

MECP2 target sequences from a complex mixture of neuronal and non-neuronal cells.[4,12,13,35–37] Typical ChIP-seq methods involve cross-linking of proteins to DNA, chromatin fragmentation by sonication, immunoprecipitation with antibodies against target proteins, and release of bound DNA followed by library preparation for next-generation sequencing. In contrast, cleavage under targets and tagmentation (CUT&Tag) is an improved enzyme-tethering method with a higher signal-to-noise ratio.[38,39] In CUT&Tag, fresh cells are permeabilized to enable antibody binding to chromatin-associated factors, which allows tethering of protein A-Tn5 transposase. Upon activation, the transposase cuts chromatin close to the protein-binding sites, resulting in shorter lengths of DNA being sequenced. CUT&Tag enables the mapping of protein-DNA interactions under native conditions without cross-linking or IP, thereby providing high-quality profiles with low background and high sensitivity.[38,39] Thus, CUT&Tag appears to be an attractive solution for fine-resolution mapping of MECP2-chromatin interactions.

In this study, we use CUT&Tag to resolve high-resolution binding sites of MECP2 and construct a map featuring 4,189 high-confidence MECP2-bound genes in human neurons. By integrating proteomics, whole-genome bisulfite sequencing (WGBS), and RNA sequencing (RNA-seq), we identify a direct protein-protein interaction alongside genome-wide transcription start site (TSS) co-occupancy between MECP2 and RNA polymerase II (RNA Pol II) and also reveal MECP2's primary association with hypomethylated promoter-proximal regions to function as an activator. Notably, there is a positive correlation between MECP2 occupancy at TSSs and the expression levels of its target genes. In neurons with patient-derived mutations in *MECP2*, loss of MECP2 binding is associated with reduced RNA Pol II loading at TSSs, resulting in global gene downregulation. Overall, these findings reveal a previously undescribed mechanism by which MECP2 fine-tunes transcription through the regulation of RNA Pol II in human neurons.

RESULTS

High-resolution mapping of MECP2-binding sites in human neurons

To generate MECP2 hESC-derived reporter lines for wild type (WT) and various patient-derived mutations of *MECP2*, including R133C, R168X, and R270X, we used CRISPR-Cas9 to create *MECP2* alleles carrying the green fluorescent protein (GFP) sequences in the endogenous locus (Figures S1A and S1B). To efficiently induce neuronal differentiation, we introduced a doxycycline (DOX)-inducible NGN2 construct into the safe harbor *AAVS1*. Upon the addition of DOX, homogeneous populations of neurons were generated within 3 weeks from those MECP2 hESC-reporter lines.[40,41] All lines differentiated into neurons had the comparable expression of endogenous-tagged WT and mutant MECP2-GFP (Figures S1C and S1D) and expressed mature neuron markers, including MAP2, PSD95, NeuN, BRN2, and VGLUT1, and Ankyrin 2 (Figures S1E–S1O). Epigenomic profiling was performed with these NGN2-induced WT and mutant neurons.

To investigate the binding of MECP2 to the genome, we used the four GFP reporter lines (Figure S1A) and performed CUT&Tag on fresh NGN2-induced human neurons. After mapping the sequenced reads on the human genome, we found 67,685 high-confidence MECP2 peaks in WT neurons, 12,260 in R133C-neurons, 19,428 in R168X-neurons, and 22,923 in R270X-neurons ($q < 10^{-10}$ and enrichment vs. IgG > 5 using MACS2[42]). Figures 1A–1D showed the binding of MECP2 to five target genes (*CHD8*, *MTOR*, *CNTNAP2*, *VPS35*, and *ORC6*) in WT neurons, whereas this interaction was notably absent in both mutant neurons and the IgG control. A polyclonal antibody against the C terminus of MECP2 was also used for CUT&Tag, validated in western blot to selectively detect endogenous MECP2 in WT neurons (Figures S1P and S1Q). The representative results shown in Figures 1A–1D indicate that this MECP2 antibody recognized identical regions reported in WT MECP2-GFP reporter neurons but not in MECP2 knockout (KO) neurons. These data demonstrate the consistency of MECP2-binding sites in those target genes and also suggest that fusing GFP to the C terminus of MECP2 has no impact on its DNA binding. Additionally, we re-analyzed previous MECP2 ChIP-seq datasets of Lund human mesencephalic (LUHMES)-derived human neurons as a more relevant comparison with our CUT&Tag data.[43] Although the MECP2 ChIP signal was extremely broad across the whole genome, only 773 significant peaks were identified when using the least stringent peak-calling threshold (Figure S2A). In contrast, there were 123,538 MECP2 peaks obtained from CUT&Tag, recovering >30% of those ChIP-seq peaks using both stringent and relaxed conditions (Figures S2A–S2C). These analyses not only demonstrate certain conservation of MECP2-binding sites reported in two profiling methods but also indicate an improved signal-to-noise ratio using CUT&Tag.

Figure 1. High-resolution mapping of MECP2-binding sites in human neurons

(A–D) Genome browser tracks of CUT&Tag read coverage showing examples of MECP2-binding sites identified in the *VPS35*, *ORC6*, *CHD8*, *MTOR*, and *CNTNAP2* genes of WT, R133C, R168X, and R270X MECP2-GFP reporter neurons using a GFP antibody. CUT&Tag assays in control and MECP2 KO neurons were performed using a MECP2 antibody. CUT&Tag with IgG in WT MECP2-GFP reporter neurons was used as a negative control. $n = 3$ biologically independent samples per condition.

(E) Circos plot of the genomic distribution of significant MECP2 peaks ($q < 10^{-10}$) in WT, R133C, R168X, and R270X MECP2-GFP reporter neurons, identified by CUT&Tag using a GFP antibody. This Circos plot was generated by BioCircos.[44] $n = 3$ biologically independent samples per condition. (See Table S1 for a complete list of MECP2-binding sites.)

(F) MECP2 binding peaks in WT MECP2-GFP reporter neurons were categorized into intergenic regions, including 2 kb upstream of TSS (TSS −2 kb; pink), 3 kb downstream of transcription end site (TES +3 kb; green), and other intergenic regions (black), along with intragenic regions comprising 5′ UTR (orange), exon (red), intron (purple), and 3′ UTR (blue). MECP2 peaks were determined from three biologically independent experiments.

(G and H) MECP2 CUT&Tag metagene line plots showing read enrichment around CpG islands (G) and TSSs (±50 bp) (H) of all human genes in WT, R133C, R168X, and R270X MECP2-GFP reporter neurons: MECP2-WT (dark blue), MECP2-R133C (blue), MECP2-R168X (light blue), MECP2-R270X (green), and IgG (orange). CUT&Tag with IgG in WT MECP2-GFP reporter neurons was used as a negative control. $n = 3$ biologically independent samples per condition.

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

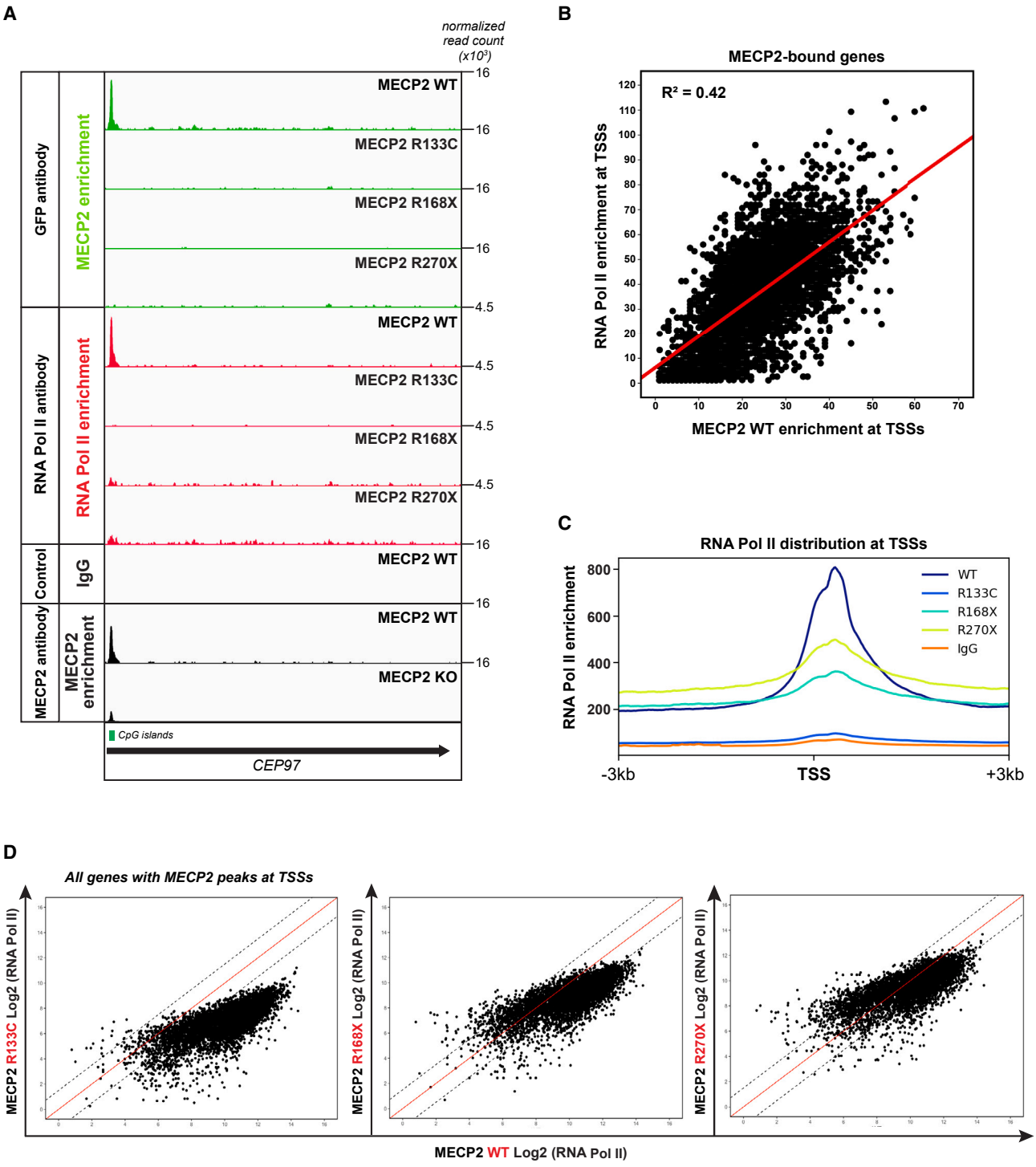


Figure 2. Genome-wide co-occupancy of MECP2 and RNA Pol II

(A) Genome browser tracks of CUT&Tag read coverage showing MECP2 and RNA Pol II peaks identified in the *CEP97* gene of WT and mutant MECP2-GFP reporter neurons, with IgG as a negative control. MECP2 peaks of *CEP97* in control and MECP2 KO neurons were performed using a MECP2 antibody. $n = 3$ biologically independent samples per condition.

(B) Correlation between the levels of RNA Pol II enrichment and MECP2 binding peak values at TSSs (±50 bp) in WT MECP2-GFP reporter neurons ($R^2 = 0.42$). $n = 3$ biologically independent samples per condition.

(legend continued on next page)

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

By plotting the distribution of MECP2 peaks on the human genome, we noticed a significant decrease of MECP2 binding in mutant neurons (Figure 1E), suggesting that these RTT-causing mutations result in reduced protein association with the genome. Furthermore, Figure 1F demonstrates that 47% of MECP2-binding sites in WT neurons were mapped between genes (intergenic), whereas 53% were intragenic, and of these, 11% were located at 2 kb upstream of TSSs and 3% were within the 5′ UTR. For those MECP2-binding sites in WT neurons, we observed a marked enrichment at CpG islands (Figures 1A–1D and 1G), indicating that the presence of CpG islands may determine MECP2-DNA interaction. When examining the overlapping and surrounding regions of MECP2 peaks in annotated genes of the human genome, we observed that MECP2-binding sites in WT neurons were mostly enriched at TSSs coincident with high CpG island density (Figures 1G and 1H). However, this TSS occupancy was reduced in all mutant neurons (Figures 1G and 1H), suggesting that RTT-causing mutations compromise the binding of MECP2 to these target genes. Based on the MECP2-TSS enrichment, we identified 4,189 genes that were significantly bound by MECP2 in WT neurons (Table S1), and this list of genes was used for the subsequent analyses. Additionally, CUT&Tag profiling of NGN2 neurons from different MECP2-GFP clones (WT, R133C, R168X, and R270X-clone2) obtained through independent gene editing presented similar WT MECP2 binding sites, which were significantly reduced in all three mutant neurons (Figure S2D). Overall, those results indicate that these CUT&Tag data consistently pinpoint high-confidence MECP2-binding sites and further verify the negative effect of RTT mutations on MECP2-DNA binding.

Genome-wide co-occupancy of MECP2 and RNA Pol II

To correlate the occupancy of MECP2 with gene expression, we performed CUT&Tag on WT and mutant MECP2-GFP reporter neurons using specific antibodies against RNA Pol II to examine their enrichment in the vicinity of genes occupied by MECP2. We noticed that MECP2 and RNA Pol II presented a similar genomic binding pattern around TSSs in WT neurons (Figures 1H and 2C). As a specific example, MECP2 occupancy co-localized with RNA Pol II at the TSS of the *CEP97* gene (Figure 2A). Consistently, we observed a positive correlation between MECP2 and RNA Pol II binding at TSSs (Figure 2B). In mutant neurons, loss of MECP2 at TSSs was associated with the reduced enrichment of RNA Pol II (Figures 2A and 2C). Additionally, CUT&Tag analysis of NGN2 neurons from different MECP2-GFP clones (WT, R133C, R168X, and R270X-clone2) derived from independent gene editing further validated the diminished binding of MECP2 and RNA Pol II in all three mutant neurons (Figures S2D and S2E). Furthermore, we measured the fold-change in RNA Pol II enrichment at TSSs between WT and mutant neurons for the genes bound by MECP2. This revealed that over 80% of these genes had a reduced RNA Pol II binding to TSSs in all mutant neurons as compared with WT (Figure 2D). In contrast, only a small subset of MECP2-bound genes was associated with increased enrichment of RNA Pol II at TSSs in the absence of MECP2 (Figure 2D). To further validate these data, RNA Pol II ChIP assay was performed with WT and mutant MECP2-GFP reporter neurons. Based on the MECP2 and RNA Pol II CUT&Tag data in WT neurons, we designed primers targeting their TSS-bound regions of the *CHD8*, *CNTNAP2*, *MTOR*, and *CEP97* genes (Figures S3A–S3D). The ChIP-qPCR results indicated a consistent reduction of RNA Pol II binding at the TSSs of these representative genes in all three mutant neurons (Figures S3E–S3H), reinforcing the negative impact of *MECP2* patient mutations on RNA Pol II TSS occupancy of their target genes. Following the RNA Pol II ChIP in WT MECP2-GFP reporter neurons, we performed a second MECP2 ChIP alongside IgG as a negative control. Our sequential ChIP-qPCR analysis confirmed the co-occupancy of MECP2 and RNA Pol II at TSS-proximal regions of the *CHD8*, *CNTNAP2*, *MTOR*, and *CEP97* genes (Figures S3I–S3L), consistent with their CUT&Tag data shown in Figures S3A–S3D. Overall, these findings revealed a significant concordance of MECP2 and RNA Pol II binding to TSSs in human neurons, and this association was lost upon disruption of MECP2 function.

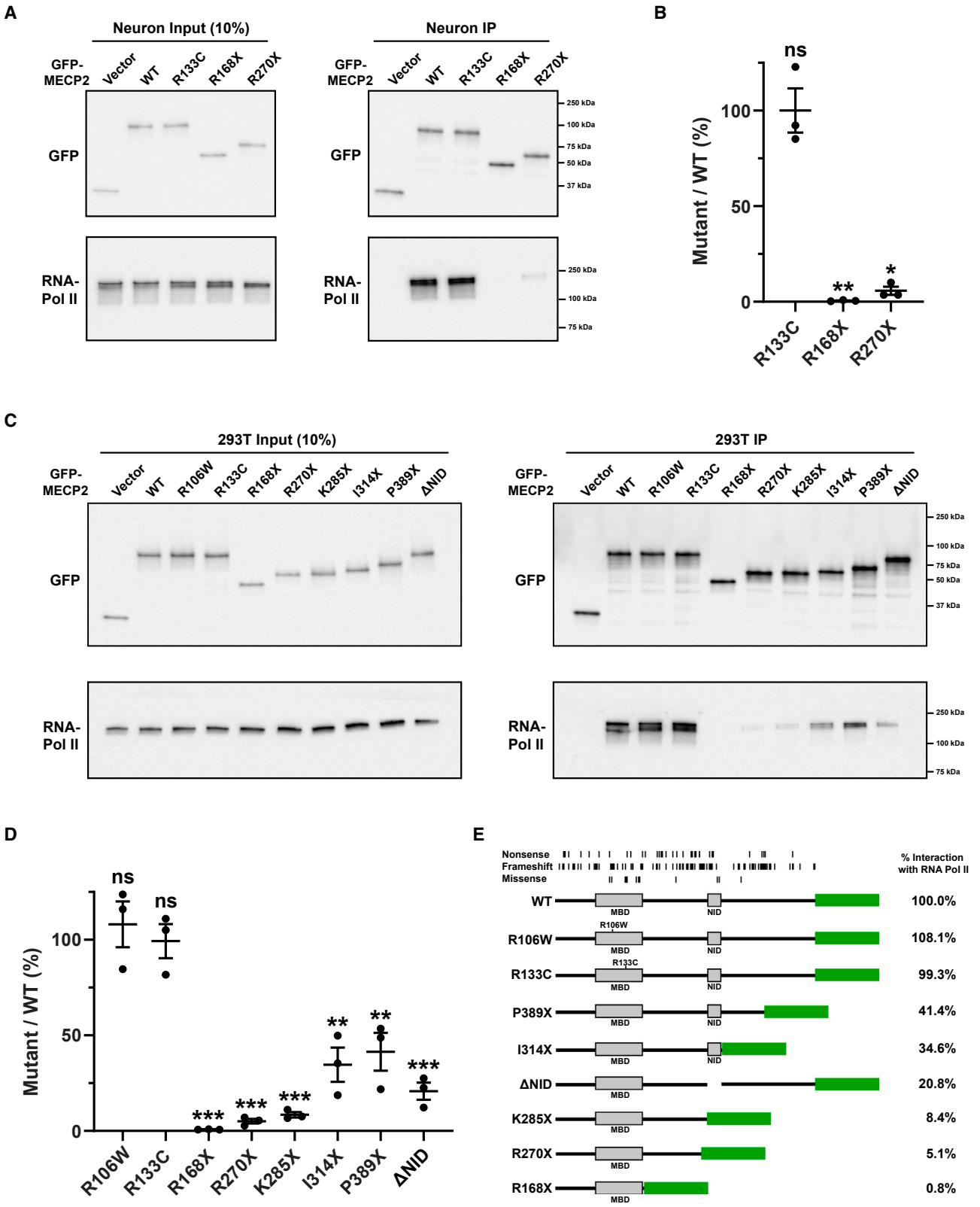
To search for characteristic features of MECP2 target genes, we used RNA-seq data from WT and MECP2 KO neurons to define a list of differentially expressed genes (Figure S6A; Table S1) and further examined a possible correlation of MECP2 and RNA Pol II levels at TSSs with transcriptional regulation. Alongside high RNA Pol II occupancy, genes activated by MECP2 (downregulated in MECP2 KO neurons) were associated with increased MECP2 enrichment at TSSs in WT neurons (Figures S2F and S2G) in contrast to genes repressed by MECP2. We also observed that MECP2-repressed genes displayed increased MECP2 binding within gene bodies in WT neurons (Figures S2F and S2G), implying an additional mechanism associated with MECP2's repressor function. Furthermore, the levels of MECP2 and RNA Pol II at TSSs in unaffected genes (no significant changes in KO neurons) were reduced in comparison with MECP2-activated genes but remained comparable to those observed in MECP2 repressed genes (Figures S2F and S2G). We speculate that those unaffected genes might not be measurably affected in MECP2 mutant neurons due to the involvement of other MECP2-binding cofactors or additional regulatory mechanisms. In summary, our findings are consistent with MECP2 in human neurons acting as both a transcriptional activator on the majority of MECP2-bound genes and a repressor on a subset of MECP2-bound genes through its co-occupancy with RNA Pol II at TSSs.

MECP2 is physically associated with the transcription apparatus in human neurons

The evidence that RNA Pol II occupancy was generally reduced at loci where mutant MECP2 occupancy was reduced led us to

(C) RNA Pol II CUT&Tag metagene line plots showing read enrichment around TSSs (±75 bp) of all human genes in WT and mutant MECP2-GFP reporter neurons: MECP2-WT (dark blue), MECP2-R133C (blue), MECP2-R168X (light blue), MECP-R270X (green), and IgG (orange). CUT&Tag with IgG in WT MECP2-GFP reporter neurons was used as a negative control. $n = 3$ biologically independent samples per condition.
(D) Scatter plots of RNA Pol II enrichment between WT MECP2-GFP reporter neurons and each mutant for all genes having a significant MECP2 peak at their TSSs (±75 bp) (see Table S1 for these MECP2-bound and regulated genes). $n = 3$ biologically independent samples per condition.

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007



(legend on next page)

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

investigate whether MECP2 and RNA Pol II might physically interact, either directly or through other components of the transcription apparatus. We conducted co-immunoprecipitation (co-IP) analysis in nuclear extracts from WT and mutant MECP2-GFP reporter neurons described in Figure S1A and found that endogenous WT MECP2 physically interacts with RNA Pol II, whereas R168X and R270X mutant MECP2 showed a marked reduction in their association with RNA Pol II (Figures 3A, 3B, and S4A), indicating that the NCoR/SMRT interaction domain[45] (NID) and C-terminal domain of MECP2 are involved in this association. We also observed that the interaction of R133C mutant MECP2 with RNA Pol II was comparable to that of WT (Figures 3A and 3B). However, this RTT missense mutation compromised DNA binding by MECP2, as shown in Figures 1A–1E and other studies,[11,46,47] which could contribute to the reduced RNA Pol II occupancy at TSSs observed in R133C mutant neurons (Figures 2C and 2D). These findings suggest that the MECP2-RNA Pol II interaction is independent of DNA, but MECP2 binding to target DNA is essential for maintaining normal levels of RNA Pol II genome occupancy. Additionally, we performed co-IP assays with mouse NGN2 neurons expressing endogenously WT and R168X mutant MECP2-GFP proteins. The results in Figure S4B confirmed that mouse RNA Pol II interacted with MECP2, whereas this association was disrupted by the R168X mutation, suggesting a conserved MECP2-RNA Pol II interaction between human and mouse neurons.

To further confirm their interaction, we performed affinity purification-mass spectrometry (AP-MS) using these MECP2-GFP reporter neurons. The Gene Ontology (GO) enrichment analysis consistently identified a group of proteins involved in the process of RNA-Pol-II-mediated transcription in WT neurons (Figure S5A). In contrast, this enrichment was reduced in two mutant neurons, which may be linked to their transcription defects (Figure S5A). Moreover, we identified ~78 differential protein interactions associated with R133C, R168X, and R270X mutant MECP2 in comparison with their WT (Figure S5B; Table S2). Of these, we noted diminished binding of RNA Pol II to R168X and R270X mutant proteins, but not to R133C (Figure S5B), and these results align with the co-IP data presented in Figures 3A and 3B. Additionally, previous proteomic studies have suggested a potential role of MECP2 in the regulation of RNA splicing.[48–50] Consistently, in MECP2 WT, R133C, and R168X mutant neurons, we noted a marked GO enrichment of MECP2-associated proteins involved in RNA splicing, whereas this enrichment was reduced in R270X mutant neurons (Figure S5A; Table S2). Additionally, we identified 20 out of the 78 differentially interacting proteins acting as MECP2-related splicing regulators (Figure S5B; Table S2). Given that splicing is often co-transcriptional,[51,52] this association implies a potential role of MECP2 to influence this co-transcriptional splicing through RNA Pol II.

To delineate the region in MECP2 required for the interaction with RNA Pol II, we shortened the NID and C-terminal domain of MECP2 to generate a set of truncation mutants, exogenously expressed in 293T cells for co-IP analysis. Our results indicated that the entire NID and C-terminal domain of MECP2 were necessary for its robust association with RNA Pol II (Figures 3C–3E). Although both R106W and R133C mutations have been shown to disrupt MECP2 binding to DNA,[11,46,47] the interaction of these two mutants with RNA Pol II was similar to that of WT (Figures 3C–3E). This result further confirms that MECP2's association with RNA Pol II is independent of its DNA binding. Additionally, the majority of *MECP2* mutations that cause RTT are truncating mutations that may alter their interaction with RNA Pol II (Figure 3E), which we speculate may help explain the broad transcriptional misregulation associated with various *MECP2* mutations.

MECP2 directly interacts with RNA Pol II

To further explore the basis for the interaction between MECP2 and RNA Pol II, we sought to determine whether there was a direct association between MECP2 and RNA Pol II using an *in vitro* pull-down experiment. RNA Pol II consists of 12 subunits (Rpb1-Rpb12) and forms a ~520-kDa complex. We isolated endogenous porcine (*S. scrofa*) RNA Pol II, which is nearly identical to human RNA Pol II.[53,54] Furthermore, we overexpressed recombinant GFP-tagged WT and mutant human MECP2 proteins in *E. coli* and purified them to homogeneity.[27] RNA Pol II and purified MECP2-GFP fusion proteins were mixed at physiologically relevant concentrations,[27] and the interaction between the proteins was assessed by immunoprecipitating GFP-tagged MECP2. We found that both WT and R133C MECP2 were associated with RNA Pol II at similar levels, whereas the association with RNA Pol II was notably reduced with R168X and R270X mutant proteins (Figure 4). It should be noted that no nucleic acids were included in these assays, indicating that the association is mediated by the proteins themselves. These data are consistent with our co-IP data using nuclear lysates from both human neurons and 293T cells (Figures 3A–3E). Overall, these protein-binding assays reinforce the notion that MECP2 directly

Figure 3. MECP2 is physically associated with transcription apparatus in human neurons

(A) Detection of endogenous-tagged MECP2-GFP WT, R133C, R168X, and R270X mutants co-immunoprecipitating with endogenous RNA Pol II in reporter human neurons. Neurons stably expressing GFP-empty (vector) were used as a negative control.

(B) The levels of RNA Pol II were normalized to IP MECP2 to assess the interaction between RNA Pol II and either WT or mutant MECP2. The ratio of normalized RNA Pol II for each mutant, compared with WT, was calculated as the percentage of mutant to WT ($n = 3$ per group, all error bars represent SEM, $^{**}p < 0.01$ and $^{*}p < 0.05$ in t test with Welch's correction).

(C) Detection of exogenously expressed WT MECP2-GFP and mutants co-immunoprecipitating with endogenous RNA Pol II in 293T cells.

(D) The levels of RNA Pol II were normalized to IP MECP2 to assess the interaction between RNA Pol II and either WT or mutant MECP2. The ratio of normalized RNA Pol II for each mutant, compared with WT, was calculated as the percentage of mutant to WT ($n = 3$ per group, all error bars represent SEM, p values are $^{**}p < 0.01$ and $^{***}p < 0.001$ in t test with Welch's correction).

(E) Schematic overview of human MECP2 domains with bar chart displaying *MECP2* coding mutations in patients with Rett syndrome found in the RettBASE database for each amino acid position. MBD, methyl-binding domain; NID, NCoR/SMRT interaction domain. The ratio of normalized RNA Pol II for each mutant, compared with the WT (D), was presented as the average percentage of mutant to WT.

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

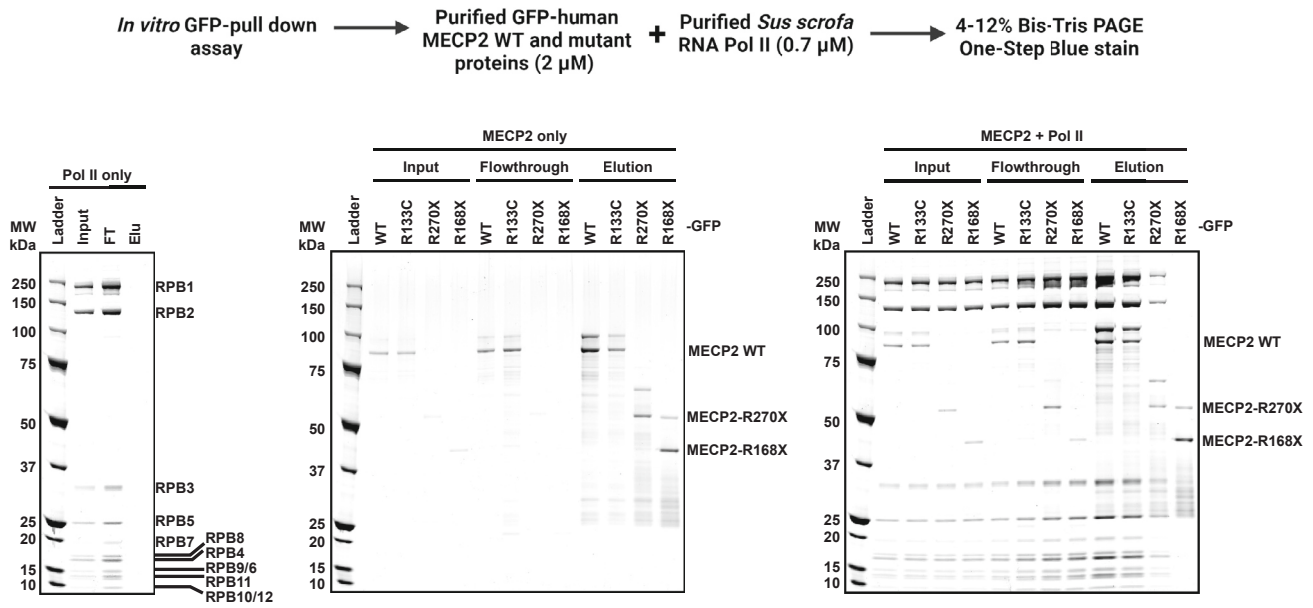


Figure 4. MECP2 directly interacts with RNA Pol II

In vitro pull-down assay using purified porcine (*S. scrofa*) RNA Pol II (0.7 μM) with recombinant GFP-MECP2 WT and mutant proteins (2 μM). Protein samples were run on a 4%–12% SDS-PAGE gel and stained with Coomassie blue.

associates with the RNA Pol II transcriptional apparatus to regulate gene expression in human neurons.

MECP2 TSS occupancy correlates with gene activation

To investigate the impact of MECP2 binding on transcriptional regulation, we conducted RNA-seq analyses comparing gene expression in mutant neurons to their isogenic controls. Initially, we focused on R168X, a severe and high-frequency RTT-causing mutation. As shown in Figure 5A, we observed a significant downregulation of global transcription in R168X mutant neurons as compared with WT. Consistent with a role in transcriptional activation, the results shown in Figures 2A–2D, 3A, and 3B indicate that the R168X mutation abolished the interaction of MECP2 with RNA Pol II and reduced its occupancy at TSSs in mutant neurons. When plotting differential expression of the 4,189 MECP2-bound genes, we found that the majority were highly expressed in WT neurons, whereas mutant neurons showed reduced transcription of around 88% of these genes (Figure 5B). About 10% of MECP2-bound genes were upregulated in mutant neurons, suggesting that MECP2 acts as a repressor of transcription on this fraction of genes (Figure 5B). Our data in Figure 2B revealed a significant correlation between MECP2 and RNA Pol II levels at TSSs. Furthermore, we observed that MECP2 occupancy at TSSs was positively correlated to the levels of gene expression (Figure 5D). Consistently, genes with higher MECP2 levels at TSSs were found to be more significantly downregulated in R168X mutant neurons (Figure 5C), suggesting that MECP2 TSS occupancy correlated with its role as a transcriptional activator. Additionally, we speculate that the modest *R* squared values observed in Figures 5C and 5D may help to explain the well-recognized role of MECP2 to slightly influence neuronal gene expression.[55]

To further confirm the extent of gene misregulation by loss of MECP2, we created MECP2 KO reporter hESCs (Figures S1P and S1Q) and generated their NGN2-derived neurons for RNA-seq analysis. Our data consistently revealed a marked reduction of global transcription in MECP2 KO neurons (Figures S6A and S6B). The magnitude of this gene downregulation corresponded with MECP2 binding levels at TSSs (Figure S6C), and a similar correlation was also observed in R133C mutant neurons (Figure S6D). For those dysregulated genes identified in our RNA-seq analysis, which were not bound by MECP2, were considered potential indirect targets of MECP2. As a comparison, we also integrated published microarray analysis of hESC-derived MECP2 KO neurons[26] along with RNA-seq datasets of NGN2-induced mouse R168X mutant neurons[27] and found that over 80% of MECP2-bound genes were consistently downregulated upon MECP2 disruption (Figures S7A–S7D), indicative of a highly conserved regulation of MECP2 in transcriptional activation.

DNA methylation status of MECP2-binding sites

To explore the methylation patterns of MECP2-bound genes, we performed WGBS on WT neurons to assess the levels of CpG and CA methylation (mCpG and mCA) within MECP2 binding regions. As demonstrated in Figures 2D, S2F, and S2G, MECP2 acts as an activator for over 80% of its target genes and, notably, these activated genes displayed higher levels of MECP2 and RNA Pol II binding at their TSSs compared with genes repressed by MECP2. We sought to determine whether this differential binding of MECP2 is dependent on mCpG, given that MECP2 was originally identified as a mCpG-binding protein.[9] However, we found that the mCpG levels of MECP2 binding regions at TSSs of MECP2-activated genes have a median value of 29%, which is significantly lower than the 89% observed in

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

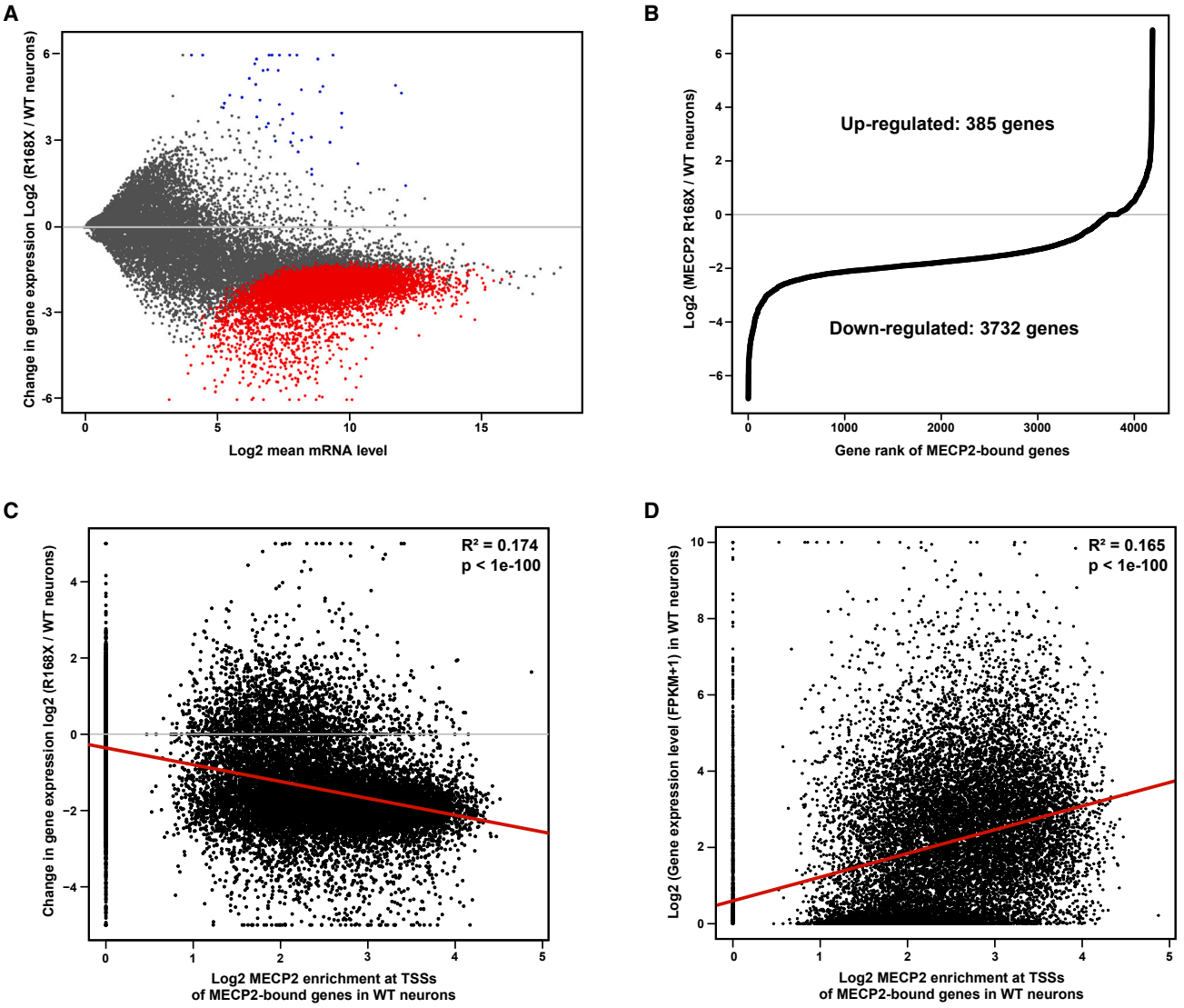


Figure 5. MECP2 TSS occupancy correlates with gene activation

(A) RNA-seq comparing MECP2-GFP WT and R168X mutant neurons. As compared with WT, upregulated genes in mutant neurons were marked in blue dots and downregulated genes were presented in red dots. Differentially expressed genes were highlighted in the plot as red/blue dots and determined by DESeq2 (two-tailed Wald test) with multiple test adjusted $p < 1 \times 10^{-5}$. (See Table S5 for the list of misregulated genes.) $n = 3$ biologically independent samples per condition.

(B) Plots showing the distribution of $\log_2$ fold changes ($\log_2$FCs) of all MECP2-bound genes. Classification of up- and downregulated genes in R168X MECP2-GFP mutant neurons was determined by positive and negative $\log_2$FCs, respectively. $n = 3$ biologically independent samples per condition.

(C) Correlation between MECP2 enrichment at TSSs in WT MECP2-GFP neurons and the differential gene expression in R168X MECP2-GFP mutant neurons ($R^2 = 0.174$, $p < 1 \times 10^{-100}$). $n = 3$ biologically independent samples per condition.

(D) Correlation between MECP2 enrichment at TSSs and gene expression levels (FPKM + 1) in human WT MECP2-GFP neurons ($R^2 = 0.165$, $p < 1 \times 10^{-100}$). $n = 3$ biologically independent samples per condition.

MECP2-repressed genes, suggesting that dense methylation does not correlate with MECP2 binding at TSSs (Figure 6A). This result is consistent with previous mapping of MECP2 in human SH-SY5Y and HCT116 cell lines, indicating that MECP2 binding to transcription regulatory regions is independent of DNA methylation.[56,57]

Previous ChIP-seq studies using mouse brains have shown that MECP2 preferentially represses long gene expression containing high levels of gene body mCA bound by MECP2.[28–34] During post-natal brain development, mCA levels have been shown to increase significantly in both human and mouse neurons.[58] A recent study indicates that neurons derived from hESCs using the conventional neuronal differentiation protocol display low mCA levels.[59] In contrast, our NGN2 induction appears to produce mature neurons with sufficient CA methylation that are comparable to those observed in human post-natal brain samples and mouse brains (Table S6). However, we observed that MECP2 binding regions in its repressed genes exhibited

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

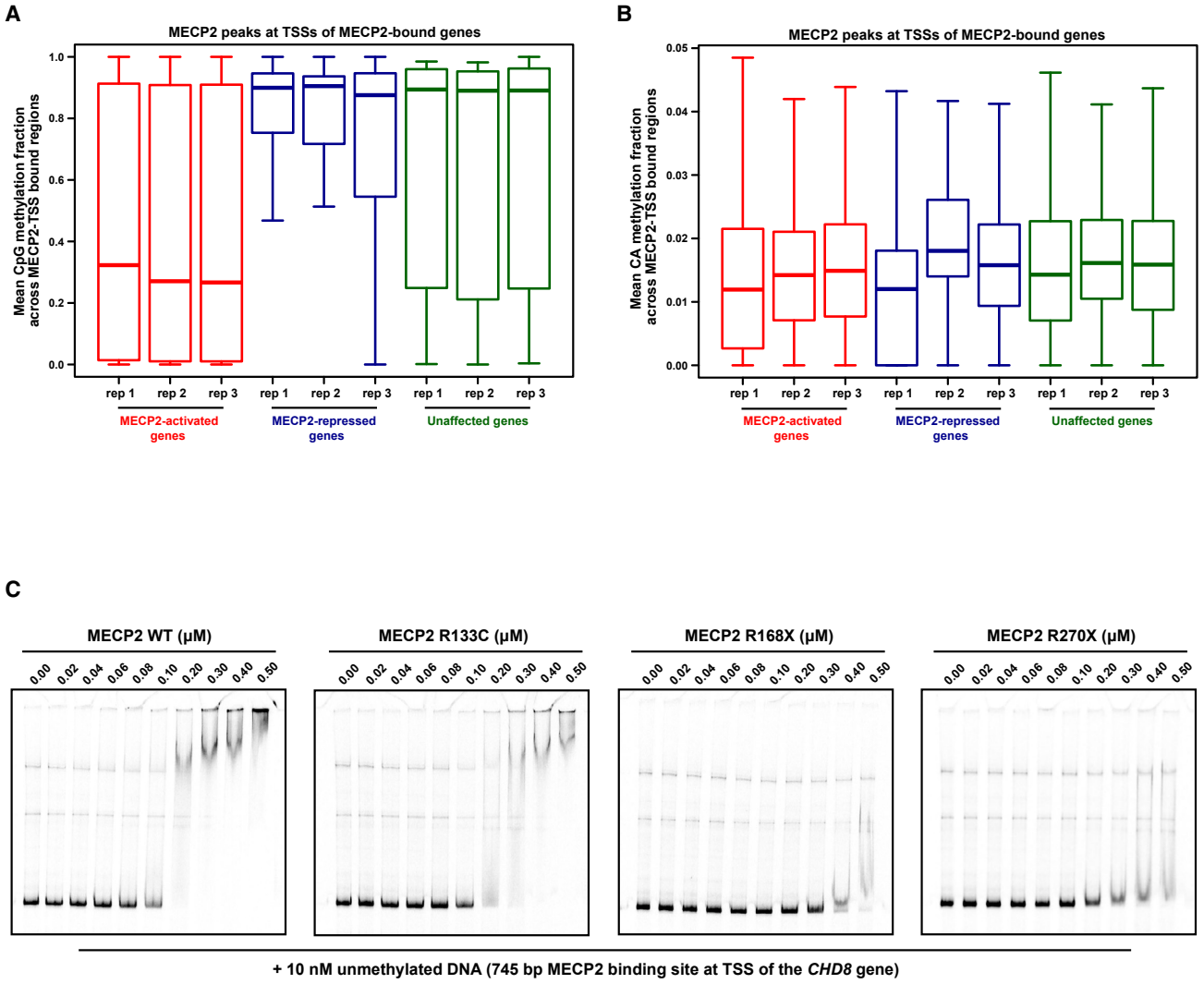


Figure 6. DNA methylation status of MECP2-binding sites

(A and B) The levels of mCpG and mCA in MEPC2 binding regions across TSSs of MECP2-activated, repressed, and unaffected genes (see Table S1 for the list of these genes). These differentially expressed genes bound by MECP2 were determined by RNA-seq data from WT and MECP2 KO neurons (Figure S6A; Table S1). For CpG methylation, MECP2-activated vs. repressed genes: $p = 4.8 \times 10^{-12}$; MECP2-activated vs. unaffected genes: $p = 4.8 \times 10^{-12}$; MECP2-repressed vs. unaffected genes: $p = 0.21$. For CA methylation, MECP2-activated vs. repressed genes: $p = 0.21$; MECP2-repressed vs. unaffected genes: $p = 0.99$. MECP2-activated genes showed a slight increase in CA methylation compared with unaffected genes ($p = 3.1 \times 10^{-4}$). ANOVA with post hoc Tukey HSD. Mean ± SEM. $n = 3$ biologically independent samples.

(C) EMSA of unmethylated DNA probes (MECP2-binding site at the TSS of the *CHD8* gene shown in Figure 1B; 10 nM) with increasing concentrations (0–0.5 μM) of recombinant human WT and mutant MECP2-GFP proteins.

similarly low mCA levels at both TSSs and gene bodies as compared with MECP2-activated genes (Figures 6B and S8A). This finding implies that CA methylation might not impact MECP2 binding to its regulated genes in human neurons. We also found that long genes were not preferentially repressed in MECP2 R168X mutant neurons (Figure S8B), which was confirmed using published microarray datasets of human MECP2 KO neurons[26] (Figure S8C). It is possible that these discrepant results may be due to the alternative regulatory mechanism, distinct chromatin profiling methods, or the differences between the human neuronal culture and mouse brains.

To verify these binding characteristics of MECP2, we utilized the electrophoretic mobility shift assay (EMSA) to examine the direct interactions between methylated and unmethylated DNA with recombinant WT and mutant MECP2 proteins. DNA fragment used in this assay corresponds to the MECP2-bound TSS region of the *CHD8* gene (*CHD8*-TSS), as shown in Figure 1B. The results indicate that WT MECP2 was capable of binding to unmethylated *CHD8*-TSS (Figures 6C and S9A), although its affinity for methylated DNA was increased at lower MECP2 concentrations (Figures S8D and S9B). Given that neuronal MECP2 is expressed at very high levels, near that of histone-octamers,[4] it is

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

CellPress

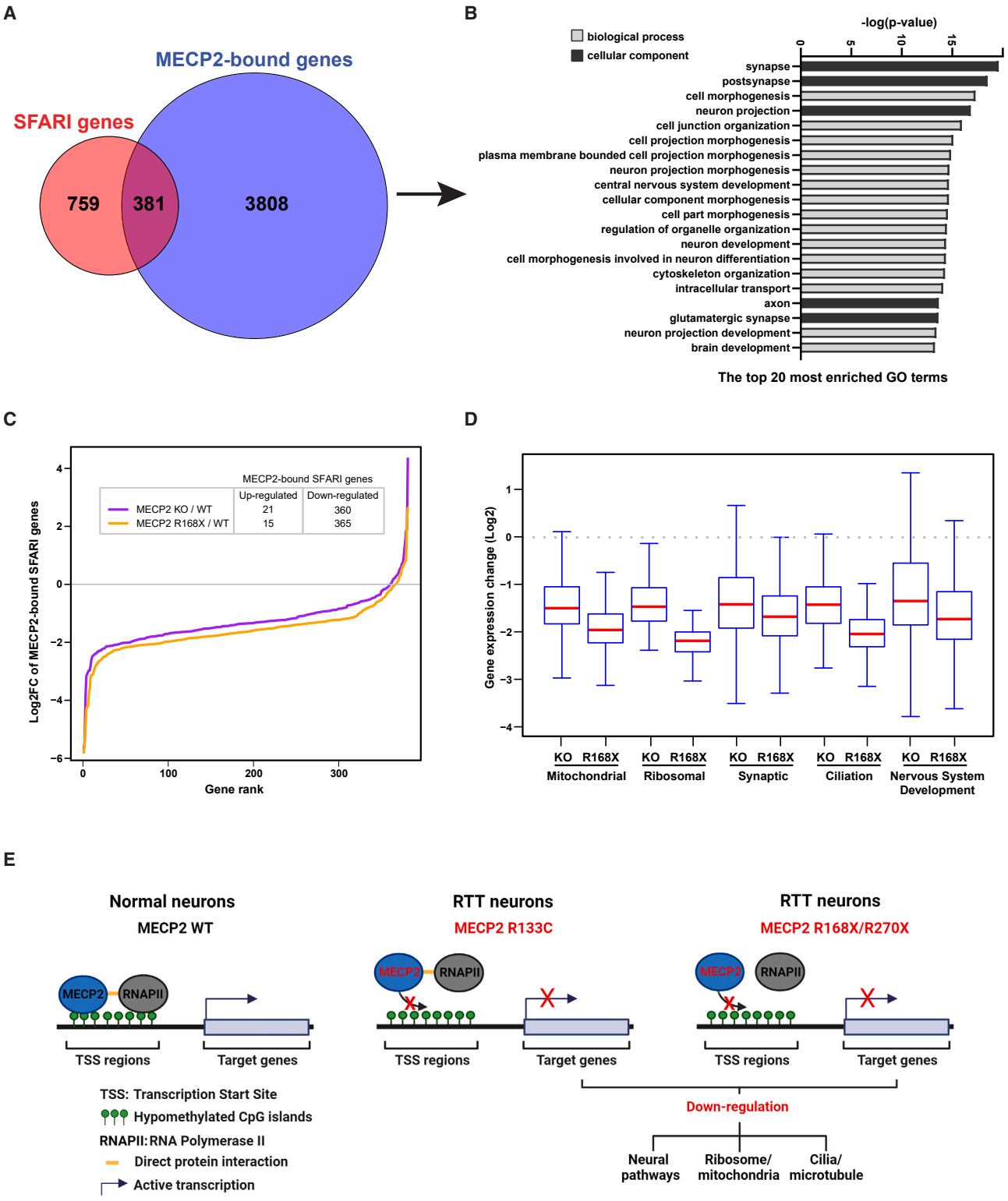


Figure 7. Functional characteristics of MECP2 target genes in Rett syndrome and autism

(A) Venn diagram comparing 4,189 MECP2-bound genes with 1,140 autism spectrum disorder (ASD) risk genes from the database of SFARI Gene 3.0 (see Table S7 for MECP2-bound SFARI genes).

(legend continued on next page)

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

plausible that MECP2 in human neurons exhibits binding affinities to both hypermethylated and hypomethylated DNA. We also observed that all three MECP2 mutants demonstrated reduced binding to both methylated and unmethylated *CHD8*-TSS (Figures 6C, S8D, S8E, S9A, and S9B), which is consistent with their CUT&Tag data presented in Figure 1B. In summary, these integrative analyses suggest that MECP2 has the capacity to directly bind the hypomethylated TSSs of its target genes, functioning as a transcriptional activator. Importantly, patient-derived mutations in *MECP2* compromise its interaction with DNA and affect its ability as a gene regulator.

Functional characteristics of MECP2 target genes in RTT and autism

To gain more insight into the function of MECP2-bound genes, we performed pathway analysis with the Simons Foundation Autism Research Initiative (SFARI) repository and found that MECP2 binds to 381 genes referenced by SFARI as implicated in autism susceptibility (Figure 7A; Table S7), and over 90% of these genes were downregulated in MECP2 KO and R168X mutant neurons (Figure 7C). As examples, MECP2 was associated with three master genes for autism shown in Figures 1B–1D. *CHD8* has been identified as a master gene for autism involved in modulating normal brain development.[60] The mTOR pathway has been directly correlated to disease-related cellular impairments observed in autism and RTT neurons.[26] CNTNAP2 KO mice were associated with stereotypic motor movements and behavioral inflexibility relevant to autism.[61] For all of these autism-related genes, we observed a global loss of MECP2 at their TSSs in mutant neurons (Figures 1B–1D) along with reduced transcription (Figure S10). These results support the role of MECP2 as an activator of target gene expression and may, in part, provide a previously unexplored mechanism for the complex autistic features of RTT.

Furthermore, GO enrichment analysis of those 4,189 target genes revealed that MECP2 primarily associates with genes related to synapse and axon functions, which are crucial for maintaining neuronal structure and normal brain development (Figure 7B). Consistently, our RNA-seq data indicate that downregulated genes in human RTT neurons are frequently involved in nervous system development, synapse formation, mitochondria, ribosome, and cilia formation (Figure 7D). These impaired biological processes have been identified as fundamental causes of RTT-related symptoms.[26,69] Furthermore, some evidence suggests that decreased MECP2 expression or *MECP2* mutations have been identified in autism patients.[1,70–74] It is plausible that MECP2 may broadly influence the expression levels of autism-associated genes, potentially contributing to the complex etiology of autism alongside its primary function in RTT.

DISCUSSION

We find that MECP2 and RNA Pol II co-occupy TSSs at a large population of active genes in human neural cells. MECP2 occupied 4,189 genes, the majority of which lost expression with RTT mutations. These findings address long-standing questions in the field about MECP2 genome occupancy and gene regulation, which are important for understanding disease development and considering therapeutic hypotheses.

At the mechanistic level, we have identified a direct protein interaction between MECP2 and RNA Pol II, and MECP2 binding positively correlates with RNA Pol II loading at TSSs and the expression levels of their target genes. These findings offer a previously unexplored mechanism to explain how and where MECP2 fine-tunes RNA-Pol-II-mediated transcription. We also find that the entire NID and C-terminal domain of MECP2 are required for its robust interaction with RNA Pol II. Further proteomics and structural analysis will be crucial to dissect more specific interaction modes. Moreover, the RNA Pol II interacting domain in MECP2 is also known to interact with other transcriptional factors, such as the NCoR co-repressor complex.[45] These associations with MECP2 might be either synergistic or competitive to influence the activator and repressor function of MECP2. We anticipate that further studies of the kinetic association and proximal mechanism of MECP2 with interacting partners at TSSs, gene bodies, or other genomic elements will be necessary to elucidate how MECP2 regulates target genes, while the involvement of other transcriptional regulators, post-translational modification of MECP2, DNA methylation dynamics, or genome organization in neurons might be subject to additional layers of complexities. To facilitate further studies extending our present results, we have developed the MECP2-NeuroAtlas portal (https://mecp2-neuroatlas.wi.mit.edu/) that includes additional assay for transposase-accessible chromatin with sequencing (ATAC-seq) and CUT&Tag for H3K4me3 and H3K9me3 in WT neurons, thereby providing streamlined access for these MECP2 multi-omics datasets with the broad spectrum of genomics, epigenetics, and proteomics.

(B) The top 20 most enriched GO terms ($p \leq 5.13 \times 10^{-14}$) of 4,189 MECP2-bound genes, black and gray boxes represent the GO terms belonging to cellular components and biological processes, respectively. This gene list enrichment analysis was performed with the portal of ToppGene Suite[62–65] (see Table S8 for the list of GO terms).

(C) Plots showing the distribution of $\log_2$FCs of 381 MECP2-bound SFARI genes (see Table S7 for this gene list). Classification of up- and downregulated genes in MECP2 KO and R168X MECP2-GFP mutant neurons was determined by positive and negative $\log_2$FCs, respectively. $n = 3$ biologically independent samples per condition.

(D) Boxplots of transcriptional changes in MECP2 KO and R168X mutant neurons in groups of genes related to synaptic, nervous system development, mitochondrial, ribosomal, and ciliation. These gene sets were sourced from previous studies[26,66–68] and listed in Table S9. $n = 3$ biologically independent samples per condition.

(E) Model of MECP2-mediated gene expression in human neurons. MECP2 directly interacts with RNA Pol II and co-occupies TSSs of their target genes to modulate active transcription. Loss-of-function mutations in *MECP2* compromise its DNA-binding capacity, leading to decreased RNA Pol II recruitment at TSS regions, although the R133C mutant MECP2 maintains its interaction with RNA Pol II. Therefore, MECP2 is required for the regulation of RNA Pol II recruitment, and alterations in this process lead to global transcriptional downregulation of genes critical for nervous system development, synapse formation, mitochondria, ribosome, and cilia functions, potentially contributing to the disease development.

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

To incorporate our conventional understanding of MECP2's function, we arrive at a model for gene activation that involves MECP2-mediated RNA Pol II binding at CpG-island-enriched TSSs to enable robust transcription and, consequently, MECP2 deficiency that alters this genomic association will cause global transcriptional repression of their target genes, leading to defects in neural pathways, mitochondria, ribosome, and ciliogenesis associated with RTT (Figure 7E). Reduced expression of the majority of genes is consistent with published results with MECP2 loss of function.[26,27] Our findings confirm that MECP2 represses a small subset of genes with higher mCpG levels but reduces MECP2 and RNA Pol II occupancy at TSSs. In contrast, genes activated by MECP2 exhibit lower mCpG levels and increased MECP2 and RNA Pol II occupancy at TSSs. Thus, our data suggest a DNA-methylation-independent binding mode of MECP2 at TSSs to modulate transcription, which is consistent with previous work in humans[56,57] but is not in agreement with the existing notion in mice. The observed differences might be due to the enhanced MECP2-binding patterns identified by CUT&Tag, which provides a superior signal-to-noise ratio for MECP2 chromatin profiling compared with ChIP-seq. Given the widespread and subtle impact of MECP2 loss on gene expression, it is essential to achieve high-resolution profiling to distinguish real binding events from background noise, thereby relating *in vivo* MECP2 binding behaviors with DNA methylation, chromatin, and gene regulation. Moreover, we need to consider the possibility of species-specific variations in MECP2-DNA bindings between humans and mice, as well as the heterogeneity of mouse brains that may influence the signal-to-noise ratio of MECP2 ChIP-seq data. These considerations emphasize the importance of comparative analyses using different chromatin profiling techniques in relevant mouse models, especially those with single-cell resolution. Taking into account the evidence presented here and in previous studies, it would be informative to re-evaluate the prevailing model of MECP2 as a transcriptional repressor.

At the physiological level, our global survey of MECP2-binding sites reveals a new list of target genes, with a prominent signature on 381 SFARI genes important for synaptic plasticity and neuronal activities. These include master genes in autism, such as *CHD8*, *MTOR*, and *CNTNAP2*, with MECP2 binding at their TSSs to activate gene expression. These findings provide compelling evidence for a previously unexplored mechanism of MECP2 in the regulation of autism risk genes. Although our findings shed light on the regulatory dynamics of MECP2, the functional outcomes of these interactions at cellular and behavioral levels remain to be deciphered. Future studies might explore how altering MECP2 binding affects neuronal morphology, synaptic plasticity, and, ultimately, behavioral phenotypes using mouse and human neuronal models of autism. In summary, by pinpointing genes targeted by MECP2 that are linked to autism and neuronal activities, we uncover potential pathways for new therapeutic directions. At the molecular level, we propose that innovative approaches to alter the genomic association between MECP2 and RNA Pol II, if designed to selectively influence the expression of their target neuronal genes, might offer therapeutic advantages for patients with RTT or autism.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - Cell culture
  - Genome editing
  - Neuronal differentiation
  - Co-Immunoprecipitation (Co-IP)
  - Western blot
  - CUT&Tag
  - Gene expression analysis
  - Whole-genome bisulfite sequencing (WGBS)
  - ChIP and Sequential ChIP assays
  - Quantitative RT-PCR (RT-qPCR)
  - Mass spectrometry
  - Protein purification
  - GFP-nanobody pull-down assay
  - Electrophoretic mobility shift assay (EMSA)
  - ELISA
  - Immunofluorescence microscopy
  - Statistics

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.neuron.2024.04.007.

ACKNOWLEDGMENTS

We thank members in the laboratories of Rudolf Jaenisch and Richard A. Young for discussions and Bingbing Yuan, Amanda Chilaka, Sumeet Gupta, Nancy M. Hannett, and Charles H. Li for technical advice and resources. We thank Andy Nutter-Upham for constructing the MECP2-NeuroAtlas portal. The graphical abstract and Figure 7E were created using BioRender.com. This work was supported by NIH grant 5R01MH104610 (to R.J.), Canadian Institutes of Health Research (CIHR) Fellowship (to Y.L.), and International Rett Syndrome Foundation (IRSF) Research Independence Award (to Y.L.). This research was supported by a generous gift from The Owens Family Foundation (to Y.L. and R.J.). S.M.V. is a Freeman Hrabowski Scholar of the Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

Conceptualization, Y.L., A.F., R.A.Y., and R.J.; methodology, software, and analysis, Y.L., A.F., G.W.B., and T.W.W.; investigation, Y.L., A.F., G.W.B., A.J.D., T.W.W., H.-C.W., Y.W., F.S., M.F., R.G., M.M., X.S.L., and S.M.V.; writing – original draft, Y.L., A.F., R.A.Y., and R.J.; visualization, Y.L.; supervision and funding acquisition, Y.L., R.A.Y., and R.J.

DECLARATION OF INTERESTS

R.J. is an advisor and co-founder of Fate Therapeutics and Fulcrum Therapeutics. A.F. is a co-founder and shareholder of StemAxon. R.A.Y. is a founder and shareholder of Syros Pharmaceuticals, Camp4 Therapeutics, Omega Therapeutics, Dewpoint Therapeutics, and Paratus Sciences.

Received: November 17, 2023
Revised: March 8, 2024
Accepted: April 5, 2024
Published: May 1, 2024

REFERENCES

1. Chahrour, M., and Zoghbi, H.Y. (2007). The story of Rett syndrome: from clinic to neurobiology. Neuron *56*, 422–437. https://doi.org/10.1016/j.neuron.2007.10.001.
2. Lyst, M.J., and Bird, A. (2015). Rett syndrome: a complex disorder with simple roots. Nat. Rev. Genet. *16*, 261–275. https://doi.org/10.1038/nrg3897.
3. Ip, J.P.K., Mellios, N., and Sur, M. (2018). Rett syndrome: insights into genetic, molecular and circuit mechanisms. Nat. Rev. Neurosci. *19*, 368–382. https://doi.org/10.1038/s41583-018-0006-3.
4. Skene, P.J., Illingworth, R.S., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Andrews, R., and Bird, A.P. (2010). Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. Mol. Cell *37*, 457–468. https://doi.org/10.1016/j.molcel.2010.01.030.
5. Chen, R.Z., Akbarian, S., Tudor, M., and Jaenisch, R. (2001). Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. Nat. Genet. *27*, 327–331. https://doi.org/10.1038/85906.
6. Guy, J., Hendrich, B., Holmes, M., Martin, J.E., and Bird, A. (2001). A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat. Genet. *27*, 322–326. https://doi.org/10.1038/85899.
7. Gemelli, T., Berton, O., Nelson, E.D., Perrotti, L.I., Jaenisch, R., and Monteggia, L.M. (2006). Postnatal loss of methyl-CpG binding protein 2 in the forebrain is sufficient to mediate behavioral aspects of Rett syndrome in mice. Biol. Psychiatry *59*, 468–476. https://doi.org/10.1016/j.biopsych.2005.07.025.
8. Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U., and Zoghbi, H.Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat. Genet. *23*, 185–188. https://doi.org/10.1038/13810.
9. Lewis, J.D., Meehan, R.R., Henzel, W.J., Maurer-Fogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell *69*, 905–914. https://doi.org/10.1016/0092-8674(92)90610-o.
10. Meehan, R.R., Lewis, J.D., and Bird, A.P. (1992). Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Res. *20*, 5085–5092. https://doi.org/10.1093/nar/20.19.5085.
11. Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell *151*, 1417–1430. https://doi.org/10.1016/j.cell.2012.11.022.
12. Chen, L., Chen, K., Lavery, L.A., Baker, S.A., Shaw, C.A., Li, W., and Zoghbi, H.Y. (2015). MeCP2 binds to non-CG methylated DNA as neurons mature, influencing transcription and the timing of onset for Rett syndrome. Proc. Natl. Acad. Sci. USA *112*, 5509–5514. https://doi.org/10.1073/pnas.1505909112.
13. Lagger, S., Connelly, J.C., Schweikert, G., Webb, S., Selfridge, J., Ramsahoye, B.H., Yu, M., He, C., Sanguinetti, G., Sowers, L.C., et al. (2017). MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. PLoS Genet. *13*, e1006793. https://doi.org/10.1371/journal.pgen.1006793.
14. Tillotson, R., Cholewa-Waclaw, J., Chhatbar, K., Connelly, J.C., Kirschner, S.A., Webb, S., Koerner, M.V., Selfridge, J., Kelly, D.A., De Sousa, D., et al. (2021). Neuronal non-CG methylation is an essential target for MeCP2 function. Mol. Cell *81*, 1260–1275.e12. https://doi.org/10.1016/j.molcel.2021.01.011.
15. Nan, X., Campoy, F.J., and Bird, A. (1997). MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell *88*, 471–481. https://doi.org/10.1016/s0092-8674(00)81887-5.
16. Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. *19*, 187–191. https://doi.org/10.1038/561.
17. Nan, X., Ng, H.-H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386–389. https://doi.org/10.1038/30764.
18. Chahrour, M., Jung, S.Y., Shaw, C., Zhou, X., Wong, S.T., Qin, J., and Zoghbi, H.Y. (2008). MeCP2, a key contributor to neurological disease, activates and represses transcription. Science *320*, 1224–1229. https://doi.org/10.1126/science.1153252.
19. Ben-Shachar, S., Chahrour, M., Thaller, C., Shaw, C.A., and Zoghbi, H.Y. (2009). Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. Hum. Mol. Genet. *18*, 2431–2442. https://doi.org/10.1093/hmg/ddp181.
20. Jordan, C., Li, H.H., Kwan, H.C., and Francke, U. (2007). Cerebellar gene expression profiles of mouse models for Rett syndrome reveal novel MeCP2 targets. BMC Med. Genet. *8*, 36. https://doi.org/10.1186/1471-2350-8-36.
21. Kriaucionis, S., Paterson, A., Curtis, J., Guy, J., Macleod, N., and Bird, A. (2006). Gene expression analysis exposes mitochondrial abnormalities in a mouse model of Rett syndrome. Mol. Cell. Biol. *26*, 5033–5042. https://doi.org/10.1128/MCB.01665-05.
22. Nuber, U.A., Kriaucionis, S., Roloff, T.C., Guy, J., Selfridge, J., Steinhoff, C., Schulz, R., Lipkowitz, B., Ropers, H.H., Holmes, M.C., and Bird, A. (2005). Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome. Hum. Mol. Genet. *14*, 2247–2256. https://doi.org/10.1093/hmg/ddi229.
23. Tudor, M., Akbarian, S., Chen, R.Z., and Jaenisch, R. (2002). Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. Proc. Natl. Acad. Sci. USA *99*, 15536–15541. https://doi.org/10.1073/pnas.242566899.
24. Urdinguio, R.G., Lopez-Serra, L., Lopez-Nieva, P., Alaminos, M., Diaz-Uriarte, R., Fernandez, A.F., and Esteller, M. (2008). Mecp2-null mice provide new neuronal targets for Rett syndrome. PLoS One *3*, e3669. https://doi.org/10.1371/journal.pone.0003669.
25. Liu, S., Zheng, P., Wang, C.Y., Jia, B.B., Zemke, N.R., Ren, B., and Zhuang, X. (2023). Cell-type-specific 3D-genome organization and transcription regulation in the brain. Preprint at bioRxiv. https://doi.org/10.1101/2023.12.04.570024.
26. Li, Y., Wang, H., Muffat, J., Cheng, A.W., Orlando, D.A., Lovén, J., Kwok, S.M., Feldman, D.A., Bateup, H.S., Gao, Q., et al. (2013). Global transcriptional and translational repression in human-embryonic-stem-cell-derived Rett syndrome neurons. Cell Stem Cell *13*, 446–458. https://doi.org/10.1016/j.stem.2013.09.001.
27. Li, C.H., Coffey, E.L., Dall'Agnese, A., Hannett, N.M., Tang, X., Henninger, J.E., Platt, J.M., Oksuz, O., Zamudio, A.V., Afeyan, L.K., et al. (2020). MeCP2 links heterochromatin condensates and neurodevelopmental disease. Nature *586*, 440–444. https://doi.org/10.1038/s41586-020-2574-4.
28. Sugino, K., Hempel, C.M., Okaty, B.W., Arnson, H.A., Kato, S., Dani, V.S., and Nelson, S.B. (2014). Cell-type-specific repression by methyl-CpG-binding protein 2 is biased toward long genes. J. Neurosci. *34*, 12877–12883. https://doi.org/10.1523/JNEUROSCI.2674-14.2014.
29. Gabel, H.W., Kinde, B., Stroud, H., Gilbert, C.S., Harmin, D.A., Kastan, N.R., Hemberg, M., Ebert, D.H., and Greenberg, M.E. (2015). Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. Nature *522*, 89–93. https://doi.org/10.1038/nature14319.
30. Kinde, B., Wu, D.Y., Greenberg, M.E., and Gabel, H.W. (2016). DNA methylation in the gene body influences MeCP2-mediated gene repression. Proc. Natl. Acad. Sci. USA *113*, 15114–15119. https://doi.org/10.1073/pnas.1618737114.
31. Stroud, H., Su, S.C., Hrvatin, S., Greben, A.W., Renthal, W., Boxer, L.D., Nagy, M.A., Hochbaum, D.R., Kinde, B., Gabel, H.W., and Greenberg, M.E. (2017). Early-life gene expression in neurons modulates lasting epigenetic states. Cell *171*, 1151–1164.e16. https://doi.org/10.1016/j.cell.2017.09.047.

32. Renthal, W., Boxer, L.D., Hrvatin, S., Li, E., Silberfeld, A., Nagy, M.A., Griffith, E.C., Vierbuchen, T., and Greenberg, M.E. (2018). Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. Nat. Neurosci. *21*, 1670–1679. https://doi.org/10.1038/s41593-018-0270-6.

33. Boxer, L.D., Renthal, W., Greben, A.W., Whitwam, T., Silberfeld, A., Stroud, H., Li, E., Yang, M.G., Kinde, B., Griffith, E.C., et al. (2020). MeCP2 represses the rate of transcriptional initiation of highly methylated long genes. Mol. Cell *77*, 294–309.e9. https://doi.org/10.1016/j.molcel.2019.10.032.

34. Nettles, S.A., Ikeuchi, Y., Lefton, K.B., Abbasi, L., Erickson, A., Agwu, C., Papouin, T., Bonni, A., and Gabel, H.W. (2023). MeCP2 represses the activity of topoisomerase IIβ in long neuronal genes. Cell Rep. *42*, 113538. https://doi.org/10.1016/j.celrep.2023.113538.

35. Cohen, S., Gabel, H.W., Hemberg, M., Hutchinson, A.N., Sadacca, L.A., Ebert, D.H., Harmin, D.A., Greenberg, R.S., Verdine, V.K., Zhou, Z., et al. (2011). Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. Neuron *72*, 72–85. https://doi.org/10.1016/j.neuron.2011.08.022.

36. Baubec, T., Ivánek, R., Lienert, F., and Schübeler, D. (2013). Methylation-dependent and -independent genomic targeting principles of the MBD protein family. Cell *153*, 480–492. https://doi.org/10.1016/j.cell.2013.03.011.

37. Baker, S.A., Chen, L., Wilkins, A.D., Yu, P., Lichtarge, O., and Zoghbi, H.Y. (2013). An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. Cell *152*, 984–996. https://doi.org/10.1016/j.cell.2013.01.038.

38. Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat. Commun. *10*, 1930. https://doi.org/10.1038/s41467-019-09982-5.

39. Kaya-Okur, H.S., Janssens, D.H., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2020). Efficient low-cost chromatin profiling with CUT&Tag. Nat. Protoc. *15*, 3264–3283. https://doi.org/10.1038/s41596-020-0373-x.

40. Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron *78*, 785–798. https://doi.org/10.1016/j.neuron.2013.05.029.

41. Fernandopulle, M.S., Prestil, R., Grunseich, C., Wang, C., Gan, L., and Ward, M.E. (2018). Transcription factor-mediated differentiation of human iPSCs into neurons. Curr. Protoc. Cell Biol. *79*, e51. https://doi.org/10.1002/cpcb.51.

42. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

43. Cholewa-Waclaw, J., Shah, R., Webb, S., Chhatbar, K., Ramsahoye, B., Pusch, O., Yu, M., Greulich, P., Waclaw, B., and Bird, A.P. (2019). Quantitative modelling predicts the impact of DNA methylation on RNA polymerase II traffic. Proc. Natl. Acad. Sci. USA *116*, 14995–15000. https://doi.org/10.1073/pnas.1903549116.

44. Cui, Y., Chen, X., Luo, H., Fan, Z., Luo, J., He, S., Yue, H., Zhang, P., and Chen, R. (2016). BioCircos.js: an interactive Circos JavaScript library for biological data visualization on web applications. Bioinformatics Oxf. Engl. *32*, 1740–1742. https://doi.org/10.1093/bioinformatics/btw041.

45. Lyst, M.J., Ekiert, R., Ebert, D.H., Merusi, C., Nowak, J., Selfridge, J., Guy, J., Kastan, N.R., Robinson, N.D., de Lima Alves, F., et al. (2013). Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. Nat. Neurosci. *16*, 898–902. https://doi.org/10.1038/nn.3434.

46. Ghosh, R.P., Horowitz-Scherer, R.A., Nikitina, T., Gierasch, L.M., and Woodcock, C.L. (2008). Rett syndrome-causing mutations in human MeCP2 result in diverse structural changes that impact folding and DNA interactions. J. Biol. Chem. *283*, 20523–20534. https://doi.org/10.1074/jbc.M803021200.

47. Brown, K., Selfridge, J., Lagger, S., Connelly, J., De Sousa, D., Kerr, A., Webb, S., Guy, J., Merusi, C., Koerner, M.V., and Bird, A. (2016). The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. Hum. Mol. Genet. *25*, 558–570. https://doi.org/10.1093/hmg/ddv496.

48. Young, J.I., Hong, E.P., Castle, J.C., Crespo-Barreto, J., Bowman, A.B., Rose, M.F., Kang, D., Richman, R., Johnson, J.M., Berget, S., and Zoghbi, H.Y. (2005). Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. Proc. Natl. Acad. Sci. USA *102*, 17551–17558. https://doi.org/10.1073/pnas.0507856102.

49. Cheng, T.L., Chen, J., Wan, H., Tang, B., Tian, W., Liao, L., and Qiu, Z. (2017). Regulation of mRNA splicing by MeCP2 via epigenetic modifications in the brain. Sci. Rep. *7*, 42790. https://doi.org/10.1038/srep42790.

50. Jiang, Y., Fu, X., Zhang, Y., Wang, S.F., Zhu, H., Wang, W.K., Zhang, L., Wu, P., Wong, C.C.L., Li, J., et al. (2021). Rett syndrome linked to defects in forming the MeCP2/Rbfox/LASR complex in mouse models. Nat. Commun. *12*, 5767. https://doi.org/10.1038/s41467-021-26084-3.

51. Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. *18*, 637–650. https://doi.org/10.1038/nrm.2017.63.

52. Guo, Y.E., Manteiga, J.C., Henninger, J.E., Sabari, B.R., Dall'Agnese, A., Hannett, N.M., Spille, J.-H., Afeyan, L.K., Zamudio, A.V., Shrinivas, K., et al. (2019). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. Nature *572*, 543–548. https://doi.org/10.1038/s41586-019-1464-0.

53. Vos, S.M., Farnung, L., Urlaub, H., and Cramer, P. (2018). Structure of paused transcription complex Pol II–DSIF–NELF. Nature *560*, 601–606. https://doi.org/10.1038/s41586-018-0442-2.

54. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018). Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. Nature *560*, 607–612. https://doi.org/10.1038/s41586-018-0440-4.

55. Trostle, A.J., Li, L., Kim, S.Y., Wang, J., Al-Ouran, R., Yalamanchili, H.K., Liu, Z., and Wan, Y.W. (2023). A comprehensive and integrative approach to MeCP2 disease transcriptomics. Int. J. Mol. Sci. *24*, 5122. https://doi.org/10.3390/ijms24065122.

56. Yasui, D.H., Peddada, S., Bieda, M.C., Vallero, R.O., Hogart, A., Nagarajan, R.P., Thatcher, K.N., Farnham, P.J., and Lasalle, J.M. (2007). Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. Proc. Natl. Acad. Sci. USA *104*, 19416–19421. https://doi.org/10.1073/pnas.0707442104.

57. Lee, W., Kim, J., Yun, J.M., Ohn, T., and Gong, Q. (2020). MeCP2 regulates gene expression through recognition of H3K27me3. Nat. Commun. *11*, 3140. https://doi.org/10.1038/s41467-020-16907-0.

58. Lister, R., Mukamel, E.A., Nery, J.R., Urich, M., Puddifoot, C.A., Johnson, N.D., Lucero, J., Huang, Y., Dwork, A.J., Schultz, M.D., et al. (2013). Global epigenomic reconfiguration during mammalian brain development. Science *341*, 1237905. https://doi.org/10.1126/science.1237905.

59. Martin, S., Poppe, D., Olova, N., O'Leary, C., Ivanova, E., Pflueger, J., Dechka, J., Simmons, R.K., Cooper, H.M., Reik, W., et al. (2023). Embryonic stem cell-derived neurons as a model system for epigenome maturation during development. Genes *14*, 957. https://doi.org/10.3390/genes14050957.

60. Weissberg, O., and Elliott, E. (2021). The mechanisms of CHD8 in neurodevelopment and autism spectrum disorders. Genes *12*, 1133. https://doi.org/10.3390/genes12081133.

61. Peñagarikano, O., Abrahams, B.S., Herman, E.I., Winden, K.D., Gdalyahu, A., Dong, H., Sonnenblick, L.I., Gruver, R., Almajano, J., Bragin, A., et al. (2011). Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. Cell *147*, 235–246. https://doi.org/10.1016/j.cell.2011.08.040.

62. Chen, J., Xu, H., Aronow, B.J., and Jegga, A.G. (2007). Improved human disease candidate gene prioritization using mouse phenotype. BMC Bioinformatics *8*, 392. https://doi.org/10.1186/1471-2105-8-392.

63. Chen, J., Bardes, E.E., Aronow, B.J., and Jegga, A.G. (2009). ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic Acids Res. *37*, W305–W311. https://doi.org/10.1093/nar/gkp427.

64. Chen, J., Aronow, B.J., and Jegga, A.G. (2009). Disease candidate gene identification and prioritization using protein interaction networks. BMC Bioinformatics *10*, 73. https://doi.org/10.1186/1471-2105-10-73.

65. Kaimal, V., Bardes, E.E., Tabar, S.C., Jegga, A.G., and Aronow, B.J. (2010). ToppCluster: a multiple gene list feature analyzer for comparative enrichment clustering and network-based dissection of biological systems. Nucleic Acids Res. *38*, W96–W102. https://doi.org/10.1093/nar/gkq418.

66. Pirooznia, M., Wang, T., Avramopoulos, D., Valle, D., Thomas, G., Huganir, R.L., Goes, F.S., Potash, J.B., and Zandi, P.P. (2012). SynaptomeDB: an ontology-based KnowledgeBase for synaptic genes. Bioinformatics Oxf. Engl. *28*, 897–899. https://doi.org/10.1093/bioinformatics/bts040.

67. Pagliarini, D.J., Calvo, S.E., Chang, B., Sheth, S.A., Vafai, S.B., Ong, S.E., Walford, G.A., Sugiana, C., Boneh, A., Chen, W.K., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. Cell *134*, 112–123. https://doi.org/10.1016/j.cell.2008.06.016.

68. Gupta, G.D., Coyaud, É., Gonçalves, J., Mojarad, B.A., Liu, Y., Wu, Q., Gheiratmand, L., Comartin, D., Tkach, J.M., Cheung, S.W., et al. (2015). A dynamic protein interaction landscape of the human centrosome-cilium interface. Cell *163*, 1484–1499. https://doi.org/10.1016/j.cell.2015.10.065.

69. Frasca, A., Spiombi, E., Palmieri, M., Albizzati, E., Valente, M.M., Bergo, A., Leva, B., Kilstrup-Nielsen, C., Bianchi, F., Di Carlo, V., et al. (2020). MECP2 mutations affect ciliogenesis: a novel perspective for Rett syndrome and related disorders. EMBO Mol. Med. *12*, e10270. https://doi.org/10.15252/emmm.201910270.

70. Lam, C.W., Yeung, W.L., Ko, C.H., Poon, P.M., Tong, S.F., Chan, K.Y., Lo, I.F., Chan, L.Y., Hui, J., Wong, V., et al. (2000). Spectrum of mutations in the MECP2 gene in patients with infantile autism and Rett syndrome. J. Med. Genet. *37*, E41. https://doi.org/10.1136/jmg.37.12.e41.

71. Carney, R.M., Wolpert, C.M., Ravan, S.A., Shahbazian, M., Ashley-Koch, A., Cuccaro, M.L., Vance, J.M., and Pericak-Vance, M.A. (2003). Identification of MeCP2 mutations in a series of females with autistic disorder. Pediatr. Neurol. *28*, 205–211. https://doi.org/10.1016/s0887-8994(02)00624-0.

72. Nagarajan, R.P., Hogart, A.R., Gwye, Y., Martin, M.R., and LaSalle, J.M. (2006). Reduced MeCP2 expression is frequent in autism frontal cortex and correlates with aberrant MECP2 promoter methylation. Epigenetics *1*, e1–e11. https://doi.org/10.4161/epi.1.4.3514.

73. Monteggia, L.M., and Kavalali, E.T. (2009). Rett syndrome and the impact of MeCP2 associated transcriptional mechanisms on neurotransmission. Biol. Psychiatry *65*, 204–210. https://doi.org/10.1016/j.biopsych.2008.10.036.

74. Wen, Z., Cheng, T.L., Li, G.Z., Sun, S.B., Yu, S.Y., Zhang, Y., Du, Y.S., and Qiu, Z. (2017). Identification of autism-related MECP2 mutations by whole-exome sequencing and functional validation. Mol. Autism *8*, 43. https://doi.org/10.1186/s13229-017-0157-5.

75. Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., Alagappan, R., Frampton, G.M., Xu, P., Muffat, J., et al. (2010). Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. Cell *141*, 872–883. https://doi.org/10.1016/j.cell.2010.04.010.

76. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

77. Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800. https://doi.org/10.1371/journal.pone.0021800.

78. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

79. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics Oxf. Engl. *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

80. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.

81. Zheng, Y., Ahmad, K., and Henikoff, S. (2020). CUT&Tag Data Processing and Analysis Tutorial.

82. Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H., and Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. *47*, W171–W174. https://doi.org/10.1093/nar/gkz365.

83. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550. https://doi.org/10.1186/s13059-014-0550-8.

84. Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics Oxf. Engl. *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656.

85. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics Oxf. Engl. *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352.

86. Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. Bioinformatics *27*, 1571–1572. https://doi.org/10.1093/bioinformatics/btr167.

87. Concordet, J.P., and Haeussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. Nucleic Acids Res. *46*, W242–W245. https://doi.org/10.1093/nar/gky354.

88. Lovén, J., Orlando, D.A., Sigova, A.A., Lin, C.Y., Rahl, P.B., Burge, C.B., Levens, D.L., Lee, T.I., and Young, R.A. (2012). Revisiting global gene expression analysis. Cell *151*, 476–482. https://doi.org/10.1016/j.cell.2012.10.012.

89. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics Oxf. Engl. *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

90. Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906. https://doi.org/10.1038/nprot.2007.261.

91. Demichev, V., Messner, C.B., Vernardis, S.I., Lilley, K.S., and Ralser, M. (2020). DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. Nat. Methods *17*, 41–44. https://doi.org/10.1038/s41592-019-0638-x.

92. Schulte, F., Hasturk, H., and Hardt, M. (2019). Mapping relative differences in human salivary gland secretions by dried saliva spot sampling and nanoLC-MS/MS. Proteomics *19*, e1900023. https://doi.org/10.1002/pmic.201900023.

93. Kirchhofer, A., Helma, J., Schmidthals, K., Frauer, C., Cui, S., Karcher, A., Pellis, M., Muyldermans, S., Casas-Delucchi, C.S., Cardoso, M.C., et al. (2010). Modulation of protein properties in living cells using nanobodies. Nat. Struct. Mol. Biol. *17*, 133–138. https://doi.org/10.1038/nsmb.1727.

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
RNA polymerase II CTD repeat YSPTSPS antibody - ChIP Grade	Abcam	Cat# ab26721; RRID:AB_777726
GFP - ChIP Grade	Abcam	Cat# ab290; RRID:AB_303395
MECP2 - ChIP Grade	Diagenode	Cat# C15410052, RRID:AB_3095788
H3K4me3 - CUT&Tag grade	Active Motif	Cat# 39160; RRID:AB_2615077
H3K9me3- CUT&Tag grade	Active Motif	Cat# 39161; RRID:AB_2532132
Rabbit IgG - ChIP Grade	Cell Signaling	Cat# 2729, RRID:AB_1031062
GFP nanobody	This study	N/A
MAP2	Cell Signaling	Cat# 8707; RRID:AB_2722660
GAPDH	Cell Signaling	Cat# 2118; RRID:AB_561053
Rpb1 CTD (4H8) Mouse mAb	Cell Signaling	Cat# 2629; RRID:AB_2167468
ChromoTek GFP-Trap magnetic agarose beads	Proteintech	Cat# gtma-20; RRID:AB_2631358
Bacterial and virus strains		
LOBSTR cells	Cheeseman Lab (WI/MIT)	N/A
Critical commercial assays		
CUT&Tag-IT™ Assay Kit - Cells	Active Motif	53160
YourSeq (FT & 3'DGE) Strand-Specific mRNA Library Prep	Active Motif	23001
ChIP-IT® Express Enzymatic	Active Motif	53009
Re-ChIP-IT®	Active Motif	53016
Pico Methyl-Seq Library Prep Kit	Zymo Research	D5455
GFP ELISA Kit	Abcam	ab171581
NEBuilder HiFi DNA Assembly Master Mix	NEB	E2621S
RNeasy Plus Mini kit	QIAGEN	74134
P3 Primary Cell 4D X Kit L	Lonza	197184
Pierce BCA Protein Assay Kit	Thermo Scientific	23225
QIAEX II Gel Extraction Kit	QIAGEN	20021
ZymoTaq DNA Polymerase	Zymo Research	E2002
5-Methylcytosine dNTP Mix	Zymo Research	D1030
Chemicals, peptides, and recombinant proteins		
Poly-D-lysine	Sigma	P7405
Laminin	Gibco	23017015
Matrigel	Corning	354234
fetal bovine serum	HyClone	SH30396.03
L-glutamine	Gibco	25030
ROCK inhibitor	Tocris	1254
N2 supplement	Gibco	17502048
Doxycycline	Sigma	D9891
Non-essential amino acids	Gibco	11140050
B27 supplement	Gibco	17504044
BDNF	PeproTech	450-02
NT-3	PeproTech	450-03
Puromycin	Sigma	P7255

(*Continued on next page*)

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CloneR™2	STEMCELL Technologies	100-0691
mTeSR™ Plus	StemCell Technologies	100-0276
Accutase	StemCell Technologies	07920
BrainPhys™ Neuronal Medium	StemCell Technologies	05790
DAPI	Thermo Scientific	62248
Vectashield mounting medium	Vector Laboratories	H-1700
Recombinant *S. scrofa* RNA Pol II	Vos et al.[53,54]	N/A
Recombinant WT MECP2-GFP	Li et al.[27]	N/A
Recombinant R133C MECP2-GFP	Li et al.[27]	N/A
Recombinant R168X MECP2-GFP	Li et al.[27]	N/A
Recombinant R270X MECP2-GFP	Li et al.[27]	N/A
Experimental models: Cell lines		
HEK293T	ATCC	Cat# CRL-3216; RRID:CVCL_0063
WIBR1	Lengner et al.[75]	NIHhESC-10-0074; RRID:CVCL_9765
WIBR1 knock-in cell line with endogenous-tagged WT MECP2-GFP and DOX-inducible NGN2 at *AAVS1* locus	This study	N/A
WIBR1 knock-in cell line with endogenous-tagged R133C MECP2-GFP and DOX-inducible NGN2 at *AAVS1* locus	This study	N/A
WIBR1 knock-in cell line with endogenous-tagged R168X MECP2-GFP and DOX-inducible NGN2 at *AAVS1* locus	This study	N/A
WIBR1 knock-in cell line with endogenous-tagged R270X MECP2-GFP and DOX-inducible NGN2 at *AAVS1* locus	This study	N/A
WIBR1 MECP2 knock-out cell line with DOX-inducible NGN2 at *AAVS1* locus	This study	N/A
WIBR1 knock-in cell line with DOX-inducible NGN2 at *AAVS1* locus	This study	N/A
Recombinant DNA		
px330-mCherry	Addgene	98750
pX330-sgAAVS1	Addgene	85802
pUCM-AAVS1-TO-hNGN2 donor vector	Addgene	105840
pET His6 TEV LIC cloning vector (1B)	Addgene	29653
pUC19-donor vectors for MECP2-GFP WT, R133C, R168X or R270X	This study	N/A
pCL-CBFH-MECP2-GFP_WT, R106W, R133C, R168X, R270X, or P389X expression vectors	Li et al.[27]	N/A
pCL-CBFH-MECP2-GFP_K285X, I314X, or ΔNID expression vectors	This study	N/A
Software and algorithms		
GraphPad Prism	GraphPad Software	https://www.graphpad.com/scientific-software/prism/ RRID:SCR_002798
Fiji image processing package	Schindelin et al.[76]	http://fiji.sc RRID:SCR_002285
REVIGO	Supek et al.[77]	http://revigo.irb.hr/ RRID:SCR_005825
Bowtie2	Langmead and Salzberg[78]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml RRID:SCR_016368

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MACS2	Zhang et al.[42]	https://github.com/macs3-project/MACS RRID:SCR_013291
BEDTools suite	Quinlan et al.[79]	https://bedtools.readthedocs.io/en/latest/content/bedtools-suite.html RRID:SCR_006646
deepTools	Ramírez et al.[80]	https://deeptools.readthedocs.io/en/latest/ RRID:SCR_016366
CUT&Tag Data Processing and Analysis Tutorial	Zheng et al.[81]	https://yezhengstat.github.io/CUTTag_tutorial/ RRID:SCR_025059
CHOPCHOP	Labun et al.[82]	https://chopchop.cbu.uib.no/ RRID:SCR_015723
BioCircos	Cui et al.[44]	https://github.com/cran/BioCircos RRID:SCR_025055
DESeq2	Love et al.[83]	https://bioconductor.org/packages/release/bioc/html/DESeq2.html RRID:SCR_015687
featureCounts	Liao et al.[84]	https://subread.sourceforge.net/featureCounts.html RRID:SCR_012919
SAMtools	Li et al.[85]	https://www.htslib.org/ RRID:SCR_002105
bismark	Krueger and Andrews[86]	https://www.bioinformatics.babraham.ac.uk/projects/bismark/ RRID:SCR_005604
Trim Galore!	N/A	https://github.com/FelixKrueger/TrimGalore RRID:SCR_011847
Deposited data		
MECP2 and RNA Pol II CUT&Tag sequencing	This study	GEO: GSE230716
WGBS	This study	GEO: GSE230716
RNA-Seq	This study	GEO: GSE230716
MECP2 mass spectrometry proteomics data	This study	MassIVE database: ID: MSV000093210
Original data	This study	Mendeley data: https://doi.org/10.17632/g3tk3tdbwj.1
MECP2 NeuroAtlas portal	This study	https://mecp2-neuroatlas.wi.mit.edu/

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Rudolf Jaenisch (jaenisch@wi.mit.edu).

Materials availability

All unique/stable reagents generated in this study are available from the lead contact upon reasonable request with a completed Materials Transfer Agreement.

Data and code availability

Relevant data supporting the findings of this study can be found within the paper. CUT&Tag, WGBS, RNA-seq, and ATAC-Seq datasets generated for this paper have been deposited in NCBI's Gene Expression Omnibus (GSE230716). Proteomics data have been deposited in MassIVE database via FileZilla (ID: MSV000093210). The MECP2 NeuroAtlas portal, which incorporates genome browsers with bar graphs and boxplots to display CUT&Tag, DNA methylation, RNA-Seq, AP-MS, and ATAC-Seq data, is accessible at (https://mecp2-neuroatlas.wi.mit.edu/). Original images for EMSAs, Western blots, and imaging data are available through Mendeley data at (https://data.mendeley.com/datasets/g3tk3tdbwj/1). These data are publicly available as of the date of publication.

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

METHOD DETAILS

Cell culture

Human embryonic stem cell (hESC) line WIBR1 was maintained in mTeSR Plus media (STEMCELL Technologies 100-0276) on Matrigel (Corning 354234) coated tissue culture plates under the previously described culture conditions.[75] Karyotype analysis was performed by the Cell Line Genetics Laboratory (Madison, WI), confirming that hESC lines maintained genetic stability. HEK293T cells were purchased from ATCC (CRL-3216) and cultured in DMEM supplemented with 10% fetal bovine serum (HyClone SH30396.03), 100 U ml-1penicillin-streptomycin (Gibco 15140), 2 mM L-glutamine (Gibco 25030) following the ATCC guidelines. V6.5 mouse ES cells were cultured as previously described.[27] All cell lines were tested regularly for mycoplasma and were found to be negative.

Genome editing

The CRISPR-Cas9 system was applied to create genetically modified hESC lines. Target-specific sequences were cloned into a pX330 plasmid (Addgene #98750) containing a single guide RNA (sgRNA) backbone carrying a codon-optimized version of Cas9 and mCherry tag. For the generation of the MECP2-WT, R133C, R168X, or R270X mEGFP endogenously tagged lines, homology-directed repair arms were cloned into the pUC19 vector (NEB N3041S) using NEBuilder HiFi DNA Master Mix (NEB E2621S). The design of the CRISPR-Cas9 targeting strategy for those hESC reporter lines was summarized in Figure S1A. The homology repair template contained mEGFP cDNA sequence or R133C point mutation flanked on either side by 800 bp homology arms amplified from genomic DNA using PCR. Knockout (KO) of the *MECP2* gene was achieved by CRISPR-Cas9 using a single sgRNA to induce indel mutations. The following sgRNA sequences with PAM sequence in parentheses were used for CRISPR-Cas9 targeting:

sgRNA_MECP2_C-term: GTAAAGTCAGCTAACTCTCT(CGG)
sgRNA_MECP2_R168X: gTGGTTTCTGCTCTCGCCGGG(AGG)
sgRNA_MECP2_R270X: gTCAGGCCATTCCCAAGAAAC(GGG)
sgRNA_MECP2_R133C: gAAAAGCCTTTCGCTCTAAAG(TGG)
sgRNA_MECP2_KO: GGAGCGCCGCTGTTTGGGGG(AGG)

To generate genetically modified cell lines, hESCs were digested by Accutase treatment, single cells were counted (1 million cells), and nucleofected (program CB-150) with 2 μg Cas9 plasmid with sgRNAs and 4 μg pUC19 vector carrying homology repair templates using the P3 Primary Cell 4D X Kit L (Lonza 197184), per manufacturer's recommendations. Transfected cells were cultured in mTeSR Plus media with the p160-Rho-associated coiled kinase (ROCK) inhibitor (10 μM; Tocris 1254). At 48 hours post-transfection, single mCherry-positive cells were sorted into a Matrigel-coated 96-well plate. These single cells were cultured in mTeSR Plus media containing 10% CloneR™2 supplement (STEMCELL Technologies 100-0691) for 4 days and then switched to mTeSR Plus media alone for another 14 days of expansion until cell colonies appeared. QuickExtract (Biosearch Technologies QE09050) was used to prepare genomic DNA for PCR to test for the presence of a transgene of interest. At least one hundred colonies were screened for successful targeting using PCR genotyping and Sanger sequencing to verify the R133C point mutation (Figure S1B) and indel mutations of *MECP2* knockout (Figure S1P) as well as Western blot/ELISA to confirm the insertion of GFP (Figures S1C and S1D). The following primers were used for PCR genotyping:

MECP2-GFP_ for: GAGCGCAAAGACATTGTTTC
MECP2-GFP_rev: CAGTTAATCGGGAAGCTTTG
MECP2-GFP_R168X_ for: AGGGAAAAGCCTTTCGCTCT
MECP2-GFP_R168X_rev: AGTTTGAAAAGGCATCTTGA
MECP2-GFP_R133C_ for: TTTGTCAGAGCGTTGTCACC
MECP2-GFP_R133C_rev: AGGGATGTGTCGCCTACC
MECP2-GFP_R270X_ for: ACCACATCCACCCAGGTCAT
MECP2-GFP_R270X_rev: CAGATCGGATAGAAGACTCC
MECP2-knockout_ for: CACTTGTTCTGCAGACTGGC
MECP2-knockout_rev: CCTGCTACCTTGAGAACTTC

In addition, we used CRISPOR[87] to predict off-target sequences for all edited cell lines. Based on their CFD scores, we examined the top 10 potential off-targets and verified that the clones used in this study had no off-target mutations. Details of the off-target sequence analyses, including primers, can be found in Table S10.

Neuronal differentiation

Neurons were derived from hESC reporter lines through the expression of neurogenin2 (NGN2), which is based on the differentiation protocol by Fernandopulle et al.[41] To integrate a doxycycline-inducible NGN2 cassette into the *AAVS1* safe harbor site, MECP2 hESC reporter lines were nucleofected with pUCM-AAVS1-TO-hNGN2 donor construct (Addgene #105840) and pX330-sgAAVS1 vector (Addgene #85802) using the method described in genome editing. Cells were sorted at 48 hours post-transfection for the expression of mCherry to enrich transfected cells. This population was allowed to expand on Matrigel-coated plates for 5 days before a 3-day

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

treatment with puromycin (1 μg/ml; Sigma P7255), followed by culturing in media without puromycin, allowing colonies to form by 7 days. According to protocols outlined in Fernandopulle et al.,[41] those cells with NGN2 transgene insertion at *AAVS1* will maintain stable expression of mCherry, and thus a second time sorting will isolate ~10,000 mCherry-positive cells plated in serial dilution in a 6-well plate and allowed to expand for 7 days before individual colonies were manually picked into a Matrigel-coated 24-well plate. The following clonal selection by PCR genotyping was carried out using protocols by Fernandopulle et al.,[41] and the inducible expression of NGN2 was confirmed by Western blot. Following neural differentiation protocols by Fernandopulle et al.,[41] hESCs were singularized by Accutase (STEMCELL Technologies 07920) treatment, replated in a Matrigel-coated 6-well plate, and cultured in neuronal induction medium with 2 μg/ml doxycycline (Sigma D9891) for 3 days. At day 3, neurons were dissociated with Accutase, replated in BrainPhys neuronal medium (STEMCELL Technologies 05790) supplemented (B27/BDNF/NT3), and maintained by half-medium change every three days for three weeks of differentiation to generate mature cortical neurons, which were characterized by the expression of multiple neuronal markers (Figures S1E–S1J), but were negative for the glia, astrocyte, and ES cell markers (Figures S1K–S1N). Additionally, we replaced mCherry with GFP in the pUCM-AAVS1-TO-hNGN2 donor construct. Using the same method, we generated control NGN2 neurons stably expressing GFP-empty, which were used as a negative control for Co-IP and AP-MS experiments shown in Figures 3A, 3B, S5A, and S5B. Additionally, mouse NGN2 neurons derived from endogenous-tagged WT and R168X MECP2-GFP mouse ES cells were generated as previously described.[27]

Co-Immunoprecipitation (Co-IP)

To examine the interaction between MECP2 and RNA polymerase II, HEK293T cells were transfected with expression constructs carrying GFP-tagged WT or mutant MECP2 transgenes (as shown in Figures 3C and 3D) using X-tremeGENE™ 9 DNA Transfection Reagent (Millipore Sigma). An empty GFP vector was used as a negative control. At 48 hours post-transfection, cells were harvested with cold 1x PBS and centrifuged for 5 min at 1000 rpm (4°C). Similarly, MECP2-GFP reporter neurons used for co-IP assays in Figures 3A, 3B, and S4B were cultured according to the method outlined in the neuronal differentiation section and harvested using cold PBS for the subsequent procedures. Next, cell pellets were lysed immediately with 300 ul hypotonic buffer (10mM HEPES-KOH, 1.5mM MgCl2, 10mM NaCl, 0.5mM DTT, 0.15% NP40, and Halt protease and phosphatase inhibitor cocktail (100x; Thermo Scientific)), incubated on ice for 5 min, and centrifuged for 8 min at 3000 rpm (4°C). The supernatant (cytosolic fraction) was removed, and pellets (nuclear fraction) were washed twice, resuspended in 300 ul lysis buffer (20mM HEPES-KOH, 1.5mM MgCl2, 150mM NaCl, 0.5mM DTT, 1% NP-40, Halt protease and phosphatase inhibitor cocktail (100x; Thermo Scientific), and 25U/ul Benzonase), and incubated on a rotator for 1 hour at 4°C. After centrifugation at 20,000xg for 10 min (4°C), the supernatant (nuclear fraction) was transferred into a fresh tube, and a fraction (10%) of those nuclear extracts was saved as the input before co-IP. ChromoTek GFP-Trap magnetic agarose beads (Proteintech Cat No. gtma) were equilibrated using ice-cold wash buffer (20mM HEPES-KOH, 1.5mM MgCl2, 150mM NaCl, 0.5mM DTT, 0.05% NP-40, Halt Protease and Phosphatase Inhibitor Cocktail (100x; Thermo Scientific)). The nuclear lysate was incubated with the equilibrated beads on an end-to-end rotator for 1 hour at 4°C. After incubation, the beads were pelleted and washed three times with wash buffer. Samples of both input and co-IP were mixed in 2x Laemmli sample buffer (Bio-Rad 161-0737) with 5% β-mercaptoethanol and boiled at 95°C for 5 min. Additionally, RNA Pol II Co-IP experiments in Figure S4A were performed using Universal Magnetic Co-IP Kit (Active Motif 54002) with an anti-RNA Pol II antibody (Cell Signaling #2629) and mouse IgG (Cell Signaling #5415) according to the manufacturer's instructions. Co-immunoprecipitated proteins were analyzed by Western blot using indicated antibodies.

Western blot

Cell lysates were prepared by resuspending cell pellets in 2x Laemmli sample buffer (5% β-mercaptoethanol and benzonase nuclease) and incubating at 4°C for 30 min. Lysates were then boiled at 95 °C for 5 min. Samples were run on a 4-20% precast polyacrylamide gel (Bio-rad; 4561094) at 80 V for 30 min, followed by 120 V until the dye front reached the bottom of the gel. Protein was transferred to a 0.2 μm PVDF membrane (Bio-rad 1704272) in transfer buffer (Bio-rad 10026938) at 25 V for 7 min using the Trans-Blot Turbo system (Bio-rad 1704150). After transfer, membranes were blocked with 5% non-fat milk in TBS for 1 hour at room temperature and then incubated with primary antibodies in 5% non-fat milk in TBST (0.1% Tween-20) overnight at 4°C. Primary antibodies were verified by their respective vendors. After washing three times with TBST for 10 min, membranes were incubated with horseradish peroxidase (HRP) conjugated-anti-Mouse IgG (GE Healthcare NXA931V) or anti-Rabbit IgG (GE Healthcare NA934V) secondary antibodies diluted in 5% non-fat milk in TBST for 1 hour at room temperature. After washing three times with TBST for 10 min, membranes were developed with Clarity Western ECL Substrate (Bio-rad 1705061) and imaged using the ChemiDoc MP System (Bio-rad 12003154). Rabbit monoclonal antibody used in this study was anti-GAPDH (dilution 1:1000; Cell Signaling #2118). Rabbit polyclonal antibodies used in this study were anti-RNA polymerase II CTD (dilution 1:1000; Abcam ab26721), anti-MECP2 (dilution 1:1000; Diagenode C15410052), and anti-GFP (dilution 1:1000; Abcam ab290). Mouse monoclonal antibody used in this study was anti-RNA Pol II antibody (dilution 1:1000; Cell Signaling #2629).

CUT&Tag

For each condition, NGN2-derived neurons were harvested from three independent differentiations, and the cell number was determined by Countess Automated Cell Counter (Invitrogen) as previously described.[26,27] 500,000 neurons per condition were centrifuged and processed for CUT&Tag using CUT&Tag-IT Assay Kit (Active Motif 53160) with an anti-GFP antibody (Abcam; ab290),

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

RNA Pol II antibody (Abcam ab26721), anti-MECP2 (Diagenode C15410052), H3K4me3 antibody (Active Motif 39160), H3K9me3 antibody (Active Motif 39162), and rabbit IgG (Cell Signaling #2729) according to the manufacturer's instructions. Libraries were pooled and sequenced on the NovaSeq 6000 (Illumina) platform, and the raw reads were processed following the protocol described by Y. Zheng, K. Ahmad, and S. Henikoff, which is accessible at https://yezhengstat.github.io/CUTTag_tutorial/. Paired-end reads (51x 51 bp) were mapped to the human (GRCh38) reference genome using the Bowtie2[78] aligner (version 2.4.2) with options: "--local --very-sensitive --no-mixed --no-discordant --phred33 -I 10 -X 700". The paired-end reads were also mapped to the *E. coli* (K-12 strain, MG1655 substrain) reference using the Bowtie2 options "--local --very-sensitive --no-overlap --no-dovetail --no-mixed --no-discordant --phred33 -I 10 -X 700". The *E. coli* reads were used to create output that was normalized to epitope abundance and was suitable for visualization (i.e. bedGraph and bigWig format files) in genome browsers. The mapped paired-end reads were processed using MACS2[42] (v. 2.2.7.1) with a $q < 10^{-10}$ threshold to identify protein binding or enrichment regions (peaks) for the MECP2 variants, RNA Pol II and histone modifications. Peak annotation was generated using the "intersect" function from the BEDTools suite[79] (v. 2.26.0). Density plots were generated using the computeMatrix function of deepTools[80] (v. 3.5.0).

Gene expression analysis

RNA-seq was used to investigate the differential expression of genes between MECP2 WT and mutant neurons. Total RNA was collected using the RNeasy Mini Plus kit (QIAGEN 74034). The amount of RNA samples was quantified using a Nanodrop spectrophotometer (Thermo Scientific). A fixed amount of ERCC RNA Spike-In (Invitrogen 4456740) was added to each sample in proportion to the number of cells for RNA-seq normalization as previously described.[26,27,88] Samples were treated with NEBNext Poly(A) mRNA Magnetic Isolation Kit (NEB E7490S) before library preparation using the YourSeq (FT & 3'DGE) Strand-Specific mRNA Library Prep (Active Motif 23001) for full-transcript (FT) sequencing. Libraries were qPCR'ed using KAPA qPCR library quant kit as per the manufacturer's protocol. The samples were loaded on the NovaSeq 6000 (Illumina) based on qPCR concentrations and sequenced on a paired-end 51x51 run.

Paired-end reads (51 x 51) were mapped with STAR (v 2.7.1a)[89] to a genome index created from the human (hg38) genome and the ERCC spike-in sequences, using Ensembl Release 106 annotations (with sjdbOverhang = 50). Gene counts were obtained for all human genes and ERCC spike-ins using featureCounts[84] and the same GTF file as for STAR and options "-p -s 1". Only protein-coding and lncRNAs with gene symbols were used for further analysis. Differential expression analysis was performed in R (v 4.2.1) using DESeq2 (v 1.36.0). Counts from ERCC spike-ins were used for normalization (applying estimateSizeFactors (dds, controlGenes = ERCC), independent filtering was set to false, and logFC shrinkage was performed with the ashr method. For scatterplot figures, the log2FC floor and ceiling were set to -6 and 6, respectively. For MA plots, differential expression was defined by an FDR $< 1 \times 10^{-5}$ and at least a 2-fold change. Previously published expression profiles were compared and visualized using public expression metrics from NCBI GEO (GSE50584, GSE139033). For mouse expression profiles, orthologous human genes were identified using Ensembl Release 106.

Whole-genome bisulfite sequencing (WGBS)

3-week differentiated endogenous-tagged WT MECP2-GFP NGN2-neurons (See neuronal differentiation method) were collected for the genomic DNA extraction using the Quick-gDNA MicroPrep (Zymo D3020). Bisulfite conversion and library preparation were performed using Pico Methyl-Seq Library Prep Kit (Zymo D5455) according to the manufacturer's instructions. Libraries were sequenced on NovaSeq 6000 (Illumina) on a paired-end 100x100 run with ~650-800 million reads for each replicate. Adapters were trimmed with Trim Galore! (v0.6.7), and reads were mapped to the bisulfite-prepared canonical chromosomes of the human genome (GRCh38/hg38) with bismark (v0.22.3) and deduplicated. The methylation status of each cytosine of each read was assayed with bismark_methylation_extractor, and methylation fraction and read coverage at each position were calculated with the bedtools suite (v2.29.2). Cytosines were partitioned into CpG and mCA contexts, and any sites with read coverage less than 5 were removed. The bedtools suite was used to further mean-summarize methylation fraction across genome features.

ChIP and Sequential ChIP assays

Chromatin immunoprecipitation (ChIP) assays were performed using ChIP-IT® Express Enzymatic Kit (Active Motif 53009). Neurons were fixed with 1% formaldehyde for 10 minutes, followed by quenching of the cross-linking with PBS containing glycine for 5 minutes. Neurons were then harvested and lysed according to the manufacturer's instructions. Specifically, nuclei were extracted using a dounce homogenizer (Active Motif 40401) and digested with the Enzymatic Shearing Cocktail (Active Motif 53009) for 15 minutes at 37°C. The reaction was quenched by the addition of cold EDTA. To assess the efficiency of digestion, the sheared chromatin samples were subjected to crosslink reversal, DNase-free Proteinase K digestion, and extraction by Phenol:Chloroform:Isoamyl Alcohol (Invitrogen 15593031) according to the manufacturer's instructions. The sheared DNA was then evaluated by electrophoresis for its size, ranging between 200 and 500 bp. Each sheared genomic DNA samples were incubated with either 2 μg of RNA Pol II (Abcam ab26721) or 2 μg of rabbit IgG (Cell Signaling #2729) as negative control for 4 hours at 4°C. After magnetic separation, the pull-down genomic DNA in the complex was subjected to reverse crosslinking followed by treatment with DNase-free Proteinase K. In line with ChIP assays, the following Sequential ChIP experiments were performed with Re-ChIP-IT® Kit (Active Motif 53016). Specifically, washed beads containing immunocomplexes from the first RNA Pol II ChIP were resuspended in Re-ChIP-IT elution buffer and subsequently purified through desalting columns according to the manufacturer's instructions. Each eluted RNA Pol II-DNA sample was

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

split into two equal aliquots and incubated with either 2 μg of MECP2 antibody (Diagenode C15410052) or 2 μg of rabbit IgG (Cell Signaling #2729) as negative control. Genomic DNA was then purified from the immunocomplexes by reverse crosslinking and treatment with DNase-free Proteinase K.

For ChIP and Sequential ChIP assays, precipitated DNA fragments were purified using Chromatin IP DNA Purification Kit (Active Motif 58002) in preparation for Quantitative real-time PCR (qPCR). The primers were designed to target the TSS-proximal regions of the *CHD8*, *MTOR*, *CNTNAP2*, and *CEP97* genes, which were bound by RNA Pol II and MECP2 as evidenced by their CUT&Tag data (Figures S3A–S3D). Fast SYBR™ Green Master Mix (Thermo Fisher 4385618) was used for qPCR reactions, and data recording was carried out on a Thermo Fisher Scientific QuantStudio 6 machine. Primers for ChIP-qPCR and Sequential ChIP-qPCR were listed below:

CHD8_for: TTCCAGCGAGCTAGTGTGAG
CHD8_rev: GTCGCCTACGTACCCTTCTC
MTOR_for: GGGCGGGGACTGTTCAAATA
MTOR_rev: TCACGACCGATTGGTTCTCC
CNTNAP2_for: GCAGCCTCTGCTCTTCTGAT
CNTNAP2 _rev: GAAGGCAAAGGAGGGGGAAG
CEP97_for: CCTACCCCGAGTCCTAAGGT
CEP97_rev: CGACCACAGATCCGCATCAA

Quantitative RT-PCR (RT-qPCR)

Total RNA was collected using RNeasy Plus Micro Kit (QIAGEN 74134) according to the manufacturer's protocol. A fixed amount of ERCC RNA Spike-In (Thermo Fisher 4456740) was added to each sample in proportion to the number of cells as previously described[26,27] and following the manufacturer's instructions. cDNAs were obtained by reverse transcription (RT) using qScript cDNA Supermix (QuantaBio 95048-500). RT-qPCR reactions were performed using Fast SYBR™ Green Master Mix (Thermo Fisher 4385618) and recorded on a Thermo Fisher Scientific QuantStudio 6 machine. The expression level of target genes was normalized to ERCC spike-in using the $2^{-\Delta CT}$ method as previously described.[26] Primer pairs against ERCC-00130 and 00096 spike-in standards were designed as previously described,[26] and primer pairs against *NANOG*, *OCT4*, *MAP2*, *PSD95*, *BRN2*, *VGLUT1*, *Ankyrin 2*, *OLIG2*, *GFAP*, and *NeuN* were verified in a previous study.[40] Primers for RT-PCR are listed below:

ERCC-00130_for: GGTTCAACCGCCTGACGAGTGCT
ERCC-00130_rev: CTGTAGCCGCTTGGCAGGACCA
ERCC-00096_for: TTTGCCGGCGCACAAGCACG
ERCC-00096_rev: TGAGTCCCTCAGCGTGACGCAGT
NANOG _for: GAAATACCTCAGCCTCCAGC
NANOG _rev: GCGTCACACCATTGCTATTC
OCT4_for: AGAACATGTGTAAGCTGCGG
OCT4_rev: GTTGCCTCTCACTCGGTTC
MAP2_for: CAGGAGACAGAGATGAGAATTCC
MAP2_rev: CAGGAGTGATGGCAGTAGAC
PSD95_for: AGTCAGAAATACCGCTACCAAG
PSD95_rev: CCGTTCACCTGCAACTCATATC
BRN2_for: AAAGTAACTGTCAAATGCGCG
BRN2_rev: GCTGTAGTGGTTAGACGCTG
VGLUT1_for: TCAATAACAGCACGACCCAC
VGLUT1_rev: TCCTGGAATCTGAGTGACAATG
Ankyrin 2_for: AGCACTCTTTCCCCAAACTC
Ankyrin 2_rev: TCTCTGATGTTTCTGTTCCTGG
OLIG2_for: GTTCTCCCCTGAGGCTTTTC
OLIG2_rev: AGATAGTCGTCGCAGCTTTC
GFAP_for: CCTCCAGCGATTCAACCTTT
GFAP_rev: GAAGCTCCAAGATGAAACCAAC
NeuN_for: CATAGAATTCAGGCCCGTAGAC
NeuN_rev: GTAGAGGGACGGAAAATTGAGG

Mass spectrometry

GFP pull-down assays were conducted using WT, R133C, R168X, and R270X MECP2-GFP reporter neurons following the protocol outlined in the Co-IP section. Neurons stably expressing GFP-empty were used as a negative control. The captured proteins on

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

magnetic beads were reconstituted in a 100 μL protein digestion buffer containing 10 mM TCEP, 40 mM CAA, and 100 mM TEAB. To ensure thorough protein solubilization, reduction of cysteine residues, and alkylation of cysteine residues, the samples underwent a 10-minute heat treatment at 70 °C. Following this, each sample received 1 μg of trypsin in 10 μL of 100 mM TEAB and was allowed to incubate overnight at 37 °C within an incubator. The following day, an additional 1 μg of trypsin in 1 μL of 100 mM TEAB per sample was introduced, and the digestions continued for an additional 4 hours at 37 °C. The resultant peptides were purified using Stage Tips with SDB-RPS sorbent, following the protocol developed by Rappsilber et al.[90]

LC-MS/MS data measurements were executed using a Vanquish Neo nanoLC system coupled with an Orbitrap Eclipse mass spectrometer, a FAIMS Pro Interface, and an Easy Spray ESI source, all provided by Thermo Fisher Scientific (Waltham, MA, USA). NanoLC separation utilized an Acclaim PepMap trap column (75μm x 2 cm) combined with an EasySpray ES802 column (75 μm x 25 mm, 100 Å) from Thermo Fisher Scientific. A 5 μL peptide extract was injected into the system. Peptide separation was conducted with a mobile phase consisting of 0.1% (v/v) formic acid in water (solution A) and 0.1% (v/v) formic acid in 80% (v/v) acetonitrile (solution B), flowing at 300 nL/min, while the column temperature was held constant at 40 °C. The column underwent an initial 2-minute conditioning with 3% solution B, followed by a gradual linear gradient up to 40% solution B over a 60-minute period. Any remaining peptides adhering to the C18 resin were afterward washed off with 95% solvent B for 11 minutes.

Using the MS in positive mode, the ion source temperature was set to 305 °C, and ionized peptides were passed through the FAIMS Pro unit at -50 V. Mass spectra were collected in MS1 mode with a resolution of 120,000, spanning the mass range of m/z 350-2000. This was carried out using standard automatic gain control (AGC) settings and automatic injection time. For MS/MS fragmentation, a data-independent acquisition (DIA) mode was employed within the mass range of 375-1200 at a resolution of 30,000. The collision energy was set at 30%, and an AGC target of 1000% was applied, utilizing m/z 25 isolation windows with 0.5 m/z overlaps for the collection of MS2 spectra.

The analysis of proteome data was conducted using the DIA-NN 1.8.1 software platform.[91] LC-MS data files were processed and analyzed with the following parameters: FASTA database: Homo sapiens (UP000005640); FASTA digest for library-free search/library generation; Protease: Trypsin/P; missed cleavages: 1, N-term M excision; C carbamidomethylation; Peptide length range: 7-30; Precursor charge range: 1-4; precursor m/z range: 300-1800; fragment ion m/z range: 200-1800; Precursor FDR (%): 1.0; Use isotopologues; heuristic protein interference; no shared spectra; Protein interference: Genes; Neural network classifier: Single-pass mode; Quantitation strategy: Robust LC (high precision); Cross-run normalization: RT-dependents; Library generation: Smart profiling; Speed and RAM usage: optimized for optimal results. Within each experiment, reporter intensity values (protein peak areas) were normalized by calculating the median for each sample, equalizing the medians across all samples, then subtracting the AP-MS results from the negative control (GFP-empty), and these adjusted binding intensity for WT and mutant MECP2 were subsequently used for statistical analysis, following the methodology outlined by Schulte et al.[92] using the Real Statistics Package. The statistical analysis included the $\log_2$ transformation of protein quantities, the computation of $\log_2$ mean ratios for the experimental conditions R133C, R168X, and R270X against the experimental WT MECP2, as well as the computation of p-values (ANOVA) that were FDR-corrected by using the Benjamini-Hochberg procedure. Proteins with adjusted p-values ≤ 0.05 and Log_2 mean ratios $\leq / \geq 1$ were considered differentially regulated.

Protein purification

RNA Pol II was purified as previously described. Briefly, RNA Pol II was isolated from *S. scrofa thymus* (Pel-Freez Biologicals, Rogers, AR, USA) essentially as described.[53,54] We performed a final size-exclusion step using a Sephacryl S-300 16/60 column (GE Healthcare Life Sciences), equilibrated in 150 mM NaCl, 10 mM Na-HEPES (pH 7.25), 10 μM ZnCl2, and 10 mM DTT. Peak fractions were concentrated using 100 kDa molecular weight cutoff (MWCO) Amicon Ultra Centrifugal Filters (MerckMillipore), aliquoted, flash-frozen in liquid nitrogen, and stored at -80 °C.

GFP-MECP2 WT and mutant proteins were purified using our previously published protocol. In detail, human MECP2 cDNA was cloned into a modified version of a T7 pET expression vector, which was engineered to include sequences encoding an N-terminal 6×His followed by mEGFP alongside a 14-amino-acid linker sequence 'GAPGSAGSAAGGSG'. cDNA sequences, amplified by PCR, were inserted in-frame after the linker sequence using NEBuilder HiFi DNA Assembly Master Mix (NEB E2621S). Mutant cDNA sequences were generated by PCR and inserted into the same base vector as described above. All expression constructs were verified by Nanopore long-read whole plasmid sequencing. For protein expression, plasmids were transformed into *E. coli* LOBSTR cells (gift from I. M. Cheeseman) and grown as follows. A fresh colony was inoculated into LB medium containing kanamycin (35 μg/ml) and chloramphenicol (35 μg/ml) and grown overnight at 37 °C. Cells were diluted 1:30 in 500 ml pre-warmed LB with freshly added kanamycin (35 μg/ml) and chloramphenicol (35 μg/ml) and grown 1.5 h at 37 °C. To induce expression, IPTG was added to the bacterial culture at 1 mM final concentration, and growth continued for 4 h. Induced bacteria were then pelleted by centrifugation, and bacterial pellets were stored at -80 °C until ready to use. For protein purification, the 500-ml cell pellets were resuspended in 15 ml of lysis buffer (50mM Tris-HCl pH 7.5, 500 mM NaCl, and 1× cOmplete protease inhibitors) followed by sonication of ten cycles of 15 s on, 60 s off. Lysates were cleared by centrifugation at 12,000g for 30 min at 4 °C, added to 1 ml of pre-equilibrated Ni-NTA agarose, and rotated at 4 °C for 1.5 h. The slurry was centrifuged at 3,000 rpm for 10 min, washed with 10 volumes of lysis buffer, and proteins were eluted by incubation for 10 or more minutes, rotating with lysis buffer containing 50 mM imidazole, 100 mM imidazole, or 3× 250 mM imidazole followed by centrifugation and gel analysis. Fractions containing protein of the correct size were dialyzed against two changes of buffer containing 50 mM Tris-HCl pH 7.5, 125 mM or 500 mM NaCl, 10% glycerol, and 1 mM DTT at 4 °C. Protein concentration of purified proteins was determined using the Pierce BCA Protein Assay Kit (Thermo Scientific 23225). Recombinant

proteins were stored in 10% glycerol, 50 mM Tris-HCl pH 7.5, 125 mM or 500 mM NaCl, 1 mM DTT. To set up the pull-down and EMSA assay with purified GFP-MECP2 WT and mutant proteins, Amicon Ultra Centrifugal filters (30K MWCO; Millipore) were used to both concentrate proteins to the desired working concentrations and to exchange them to the required assay buffer.

GFP-nanobody pull-down assay

Anti-GFP nanobody was expressed using a modified version of a previously published sequence.[93] The gene was synthesized (Twist) and inserted into a modified pET28b vector containing an N-terminal 6xHis tag followed by a tobacco etch virus cleavage site (UC Berkeley MacroLab 1b vector, Addgene #29653). The plasmid was transformed into *E. coli* BL21 (DE3) RIL cells. The cells were grown at 37 °C in 2xYT medium until reaching an OD600 of 0.4 whereupon expression was induced by the addition of 0.5 mM IPTG. The cells were grown for another 3 hours at 37 °C and collected by centrifugation, followed by resuspension in Lysis 300 buffer (300 mM NaCl, 20 mM Na•HEPES pH 7.4, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM β-mercaptoethanol (BME), 0.284 μg/ml leupeptin, 1.37 μg/ml pepstatin A, 0.17 mg/ml PMSF, and 0.33 mg/ml benzamidine). Cells were lysed by sonication and cleared by centrifugation. The clarified lysate was applied to a 5 mL HisTrap HP column (GE Healthcare Life Sciences) equilibrated in Lysis 300 buffer. The column was washed with 10 column volumes of Lysis 300 buffer, followed by 4 column volumes of High Salt buffer (1000 mM NaCl, 20 mM Na•HEPES pH 7.4, 10% (v/v) glycerol, 30 mM imidazole pH 8.0, and 5 mM BME). The column was washed with 4 column volumes of Lysis 300 buffer, and the protein was eluted over a gradient with a buffer containing 300 mM NaCl, 20 mM Na•HEPES pH 7.4, 500 mM imidazole pH 8.0, 10% (v/v) glycerol, and 5 mM BME. Peak fractions containing anti-GFP nanobody were pooled, and protein purity was assessed by SDS–PAGE and Coomassie staining. Anti-GFP nanobody was concentrated with 30 kDa MWCO Amicon Ultra Centrifugal Filters (MilliporeSigma) and applied to a HiLoad S75 16/600 pg column (Cytvia) equilibrated in 300 mM NaCl, 20 mM Na•HEPES pH 7.4, 10% (v/v) glycerol, and 1 mM dithiothreitol (DTT). Protein purity was assessed by SDS–PAGE and Coomassie staining. Pure fractions with anti-GFP nanobody were concentrated with 30 kDa MWCO Amicon Ultra Centrifugal Filters (MilliporeSigma). Protein concentration was determined by measuring absorption at 280 nm and using the predicted extinction coefficient for the complex. Protein was aliquoted, flash-frozen, and stored at -80 °C.

Lyophilized CnBr-activated Sepharose 4B affinity beads (Cytiva) were activated by incubation at 25 °C for 5 min while rotating with activation buffer (10 mM HCl). The beads were washed 10 times with activation buffer. The resulting resin was washed 3 times with coupling buffer (500 mM NaCl, 100 mM NaHCO3). The resin was then incubated with anti-GFP nanobody (35 μM) for 2 hours at 25 °C, rotating. The coupling reaction was quenched by washing 3 times with 1 M Tris-HCl pH 7.9. The resin then underwent the following series of washes: 3 times with coupling buffer, 3 times with low pH buffer (500 mM NaCl, 100 mM acetic acid/NaOAc pH 4.0), 3 times with coupling buffer, and 4 times with phosphate-buffered saline (PBS). The beads were stored at 4 °C in PBS.

Purified GFP-MECP2 WT, R133C, R270X, or R168X proteins (2 μM) were incubated with 0.7 μM RNA Pol II in pull-down buffer (75 mM NaCl, 20 mM Na-HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT and 1 mM $MgCl_2$) for 20 min at 25 °C. The binding reactions were incubated with anti-GFP nanobody-conjugated beads for an additional 20 min. The beads were washed three times with pull-down buffer to remove unbound protein. The GFP-tagged protein was eluted from the beads by applying 100 mM glycine pH 2.0 to the beads. 20 μl of the eluted sample was neutralized with 100 mM Tris-HCl pH 7.5 and was applied to a 4-12% SDS–PAGE and stained with Coomassie blue.

Electrophoretic mobility shift assay (EMSA)

To test the binding affinity of MECP2 with DNA, recombinant MECP2-WT, R133C, R168X, or R270X proteins were isolated according to the protocol outlined in the protein purification section and then mixed with 10 nM methylated or unmethylated Cy5-labeled DNA at selected concentrations in buffer containing 10 mM HEPES-KOH pH 7.5, 12 mM Tris-HCl pH 7.4, 50 mM NaCl, 50 mM KCl, 5 mM MgCl2, 0.01 mM ZnCl2, 0.01% NP-40 (Sigma # 85124), 5% glycerol, and 0.5 mM DTT. The reactions were performed in a 20 μL volume and were incubated in the dark for 1 hr at room temperature to achieve equilibrium. The reactions were then loaded into 4-15% Criterion™ TGX™ Precast Midi Protein Gel (Bio-Rad #5671085) by mixing with a 2-fold sample volume of Native sample buffer (Bio-Rad #1610738) for 1 hr at 200V at 4°C. The gels were imaged using Typhoon FLA95 imager with the Cy5 fluorescence module. The DNA sequence used in this assay was the MECP2-bound TSS region of the *CHD8* genes as identified by MECP2 CUT&Tag assay (Figure 1B), which was amplified from the genomic DNA of human NGN2 neurons using oligonucleotide primers with 5′-Cy5 fluorophore modifications (MilliporeSigma; for: GACTGAGGGGCCTCCTAAGT; rev: CCTCTCCCCCGGCCTCTCACA AA). PCR was performed using ZymoTaq™ kit (Zymo E2002), and 5-Methylcytosine dNTP Mix (ZYMO D1030) was added to generate methylated DNA templates. Fluorescent PCR products were gel purified using the QIAEX II Gel Extraction Kit (QIAGEN 20021). The methylation of templates was validated by the methyl-specific restriction digestion using BseYI (NEB R0635S). The 745-bp methylated and unmethylated DNA sequence used in EMSA assays (Figures 6C, S8D, and S8E) is GACTGAGGGGCCTCCTAAGTGTATACAGCAGA GTATGTTCGGGCTGGAC CTAGATAGTCGGACAGAACACCGACAGGGTCTTCACATTCTTTAAGCACTCTCACCCAGTCCCGGAGT CCGAGCCAAGATGGCTTCCCAGAAGTAGGTAACTAAGTGTTGTTTTAGTTCACAGCGCCCCCAATGGCGGGGGTCAAAAAAAGATG CCGAGGGGCCTGGCCGCCTGTGAACTGAGTCGTCAGAGGCGAGCCGCCGCCAGATTTTATTTTCAATCAGGCATATTTGTGGTCG CCTACGTACCCTTCTCATCCTCCTAGGACCTTCCACCCTCAAGAGACTTCCTGGATGGCGCTCACACTAGCTCGCTGGAAGAGGC CGGGCTTCGGTGCTGCCATTGGCGGATGGGAGCGCAGGGGGGCGGAGCCTAGGGAGAATGGACTACTAGCAAGCTAGATTGAG GCGAGGGCTGGCCATGCTCCGCCGGAAGGCGCCTTGTTCTCGGAGAGGGTGAGTTGAAACGCTGCCTGGAAAGGAAGTACCAG GACTTGCACAGGTAGGCTTCCCGCTCGCACCCGCTCCTCAGTGCTGGGCCAGGAGTCAGCAAAAGCTGAGGCTCTCAGAGATTTG

Please cite this article in press as: Liu et al., MECP2 directly interacts with RNA polymerase II to modulate transcription in human neurons, Neuron (2024), https://doi.org/10.1016/j.neuron.2024.04.007

ATTTAAGTAACTCCTTCTCCTCTCTCTTCCACCCCCTGCTCGCCCCCCAGTTGTGAGGGATCAGGTGCGGGGTGATGCTCCGGAC TGAACCAGTCTCCTTTGTGAGAGGCCGGGGGGAGAGG. Additionally, we used similar methods to generate the 200-bp methylated and unmethylated DNA probes for EMSA assays shown in Figures S9A and S9B using oligonucleotide primers modified with 5′- Cy5 fluorophore (IDT; for: GGTAACTAAGTGTTG; rev: ATCCAGGAAGTCTCT). The 200 bp DNA sequence is GGTAACTAAGTGTTGT TTTAGTTCACAGCGCCCCCAATGGCGGGGGTCAAAAAAAGATGCCGAGGGGCCTGGCCGCCGTGAACTGAGTCGTCAGAGGCG AGCCGCCGCCAGATTTTATTTTCAATCAGGCATATTTGTGGTCGCCTACGTACCCTTCTCATCCTCCTAGGACCTTCCACCCTCAAG AGACTTCCTGGAT. This methylated template was verified by the methyl-specific restriction digestion using Fnu4HI (NEB Cat log: R0178S).

ELISA

The GFP ELISA kit (Abcam ab171581) was used to measure GFP levels in MECP2-GFP WT and mutant reporter neurons. 500,000 neuron aliquotes were determined by Countess Automated Cell Counter (Invitrogen) as previously described,[26,27] which were lysed and diluted for ELISA assay according to the manufacturer's protocol (Abcam ab171581). Final absorbance values were read at 450 nm using a Thermo Fisher Multiskan Go plate reader. The standard curve was derived from plotting a dilution series of GFP standards, which allowed calculating the concentration of GFP-tagged proteins. The normalized MECP2-GFP expression ratio for each mutant compared to the WT was calculated from nine biologically independent samples, each with two technical replicates. For statistical analysis, the technical replicates were averaged, as depicted in Figure S1D.

Immunofluorescence microscopy

Human NGN2 neurons were plated on glass coverslips coated with poly-D-lysine (100 μg/ml; Sigma P7405) and laminin (1 μg/mL; Gibco 23017015) following the published protocol.[41] For immunostaining, neurons were fixed with 4% paraformaldehyde in PBS, washed 3 times with PBS, and then permeabilized using 0.5% Triton X-100 in PBS for 10 min at room temperature. Next, neurons were washed three times with PBS, blocked for 1 h in 4% IgG-free BSA (Jackson ImmunoResearch 001-000-162) in PBS, and then stained overnight with the MAP2 antibody (dilution 1:200; Cell Signaling #8707) in 4% IgG-free BSA at 4°C in a humidified chamber. Subsequently, neurons were washed three times with PBS. Secondary antibodies were added to neurons in 4% IgG-free BSA and incubated for 1 h at room temperature. Neurons were then washed twice in PBS, and stained with DAPI (dilution 1:1000; Thermo Scientific 62248) in PBS for 5 min, and then mounted in Vectashield mounting medium (Vector Laboratories H-1700). Images were acquired with a Zeiss LSM980 Microscope with Airyscan detector using 63x objective and Zen Black software (ZEISS) at the W.M. Keck Microscopy Facility, MIT.

Statistics

Individual p-values and the number of replicates for statistical testing were indicated in corresponding figure legends. For *t*-tests, data were assumed to be normal. For RNA-Seq analysis, a two-sided Wald test was used to determine differentially expressed genes, and p-values were adjusted for multiple comparisons. Other statistical tests were performed as indicated in figure legends.