PNAS

Supplementary Information for

Targeted RNA editing in brainstem alleviates respiratory dysfunction in a mouse model of Rett Syndrome.

John R. Sinnamon[a], Michael E. Jacobson[a], John F. Yung[a], Jenna R. Fisk[a], Sophia Jeng[b], Shannon K. McWeeney[b-e], Lindsay K. Parmelee[f], Chi Ngai Chan[f], Siu-Pok Yee[g] and Gail Mandel[a*]

[a] Vollum Institute, Oregon Health and Science University, Portland, OR, 97239, USA

[b] Knight Cancer Institute, Oregon Health and Science University, Portland, OR, 97239, USA

[c] Division of Bioinformatics and Computational Biology, Oregon Health and Science University, Portland OR, 97239, USA

[d] Department of Medical Informatics and Clinical Epidemiology, Oregon Health and Science University, Portland OR, 97239, USA

[e] Oregon Clinical and Translational Research Institute, Oregon Health and Science University, Portland OR, 97239, USA

[f] Integrated Pathology Core, Oregon National Primate Research Center, Beaverton, OR, 97006, USA

[g] Department of Cell Biology, University of Connecticut Health Center, Farmington, CT, 06030, USA

* **Correspondence**: Gail Mandel, Vollum Institute, Oregon Health and Science University, Portland, OR, 97239, USA
Email: mandelg@ohsu.edu

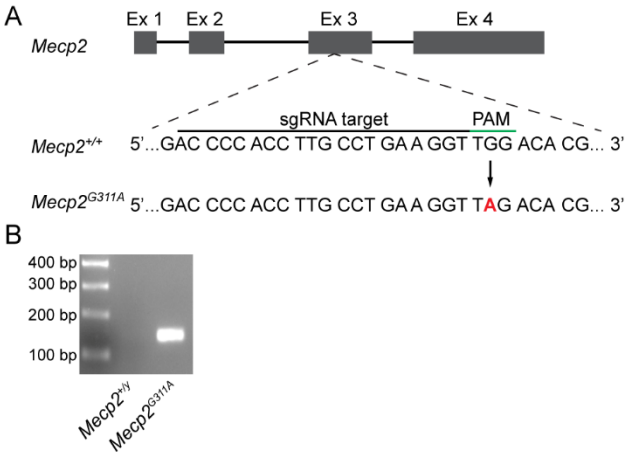
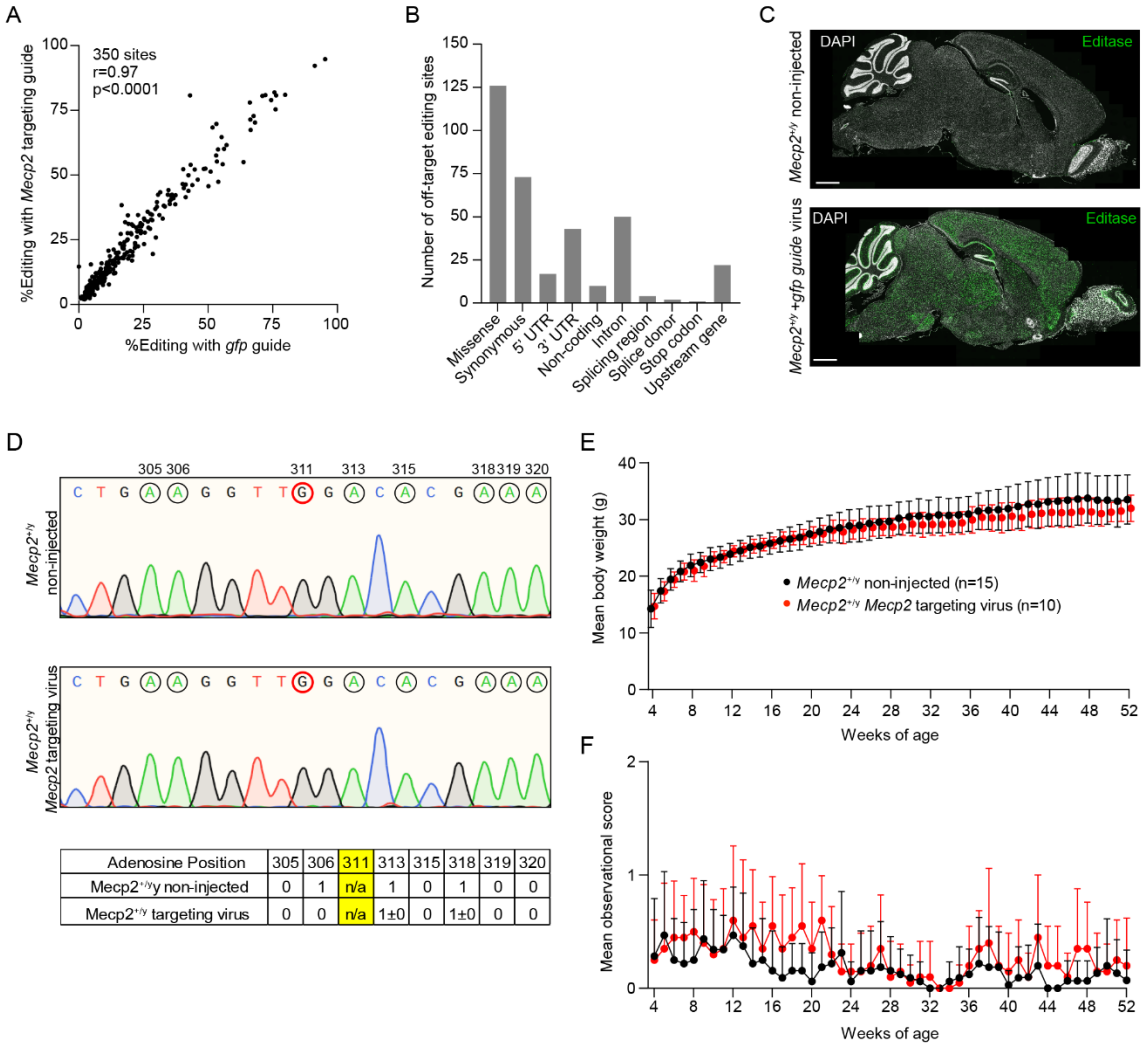


Fig. S1 Generation of a mouse model containing the mouse G-to-A Rett Syndrome patient mutation, *Mecp2*^{G311A}. (A) Schematic of the exon organization and CRISPR/Cas9 mutagenesis approach used to introduce the adenosine mutation (red) into the endogenous *Mecp2* gene (G311A; W104X knock-in). sgRNA, single guide RNA; PAM, protospacer adjacent motif. (B) Agarose gel showing amplification of the 146 base pair amplicon specific to the *Mecp2*^{G311A} site.



Adenosine Position	305	306	311	313	315	318	319	320
Mecp2$^{+/y}$ non-injected	0	1	n/a	1	0	1	0	0
Mecp2$^{+/y}$ targeting virus	0	0	n/a	1±0	0	1±0	0	0

Fig. S2. Testing off-target effects in mutant (A and B) and wild type mice (C-E). (A) Pairwise scatterplot of off-target editing percentages (350 sites) from the brainstem of *Mecp2*$^{G311A/y}$ mice injected with the indicated viruses. Each point represents an individual off-target editing event. X and Y axes are the mean percentages of editing from analysis of three mice with the indicated viruses. p values were determined using a non-parametric Spearman correlation. (B) Transcript-wide number and distribution of total off-target sites in the indicated regions, independent of viral condition. UTR, untranslated region; splicing region (off-target site within first 1-3 bases of the exon or 3-8 bases of the intron); upstream gene (located within 2kb upstream from promoter). (C) Representative confocal images of sagittal sections of a non-injected wild type *Mecp2*$^{+/y}$ mouse and an *Mecp2*$^{+/y}$ mouse injected with the *gfp* guide virus (same time of injection and with same amount of virus as for the studies of the *Mecp2*^{G311A} mice). Sections were stained for DAPI to label nuclei and immuno-labeled for Editase^{wt}. Scale bar = 1mm. (D) Results of Sanger sequencing analysis for bystander

editing efficiencies in the brainstem of wild type mice injected with the Editase^{wt} and *Mecp2* targeting guide. Left, representative sequencing electropherogram. Right, quantification of editing efficiencies in guide region. n = 3 mice each condition. (E) Body weight comparisons for *Mecp2*$^{+/y}$ mice injected with the *Mecp2* targeting virus and non-injected control *Mecp2*$^{+/y}$ mice. (F) Observational scores based on measurements of mobility, gait, general condition, hindlimb clasp and tremor. A score of 0 to 1 is common for wild type mice (1) out of 10 (lethality).

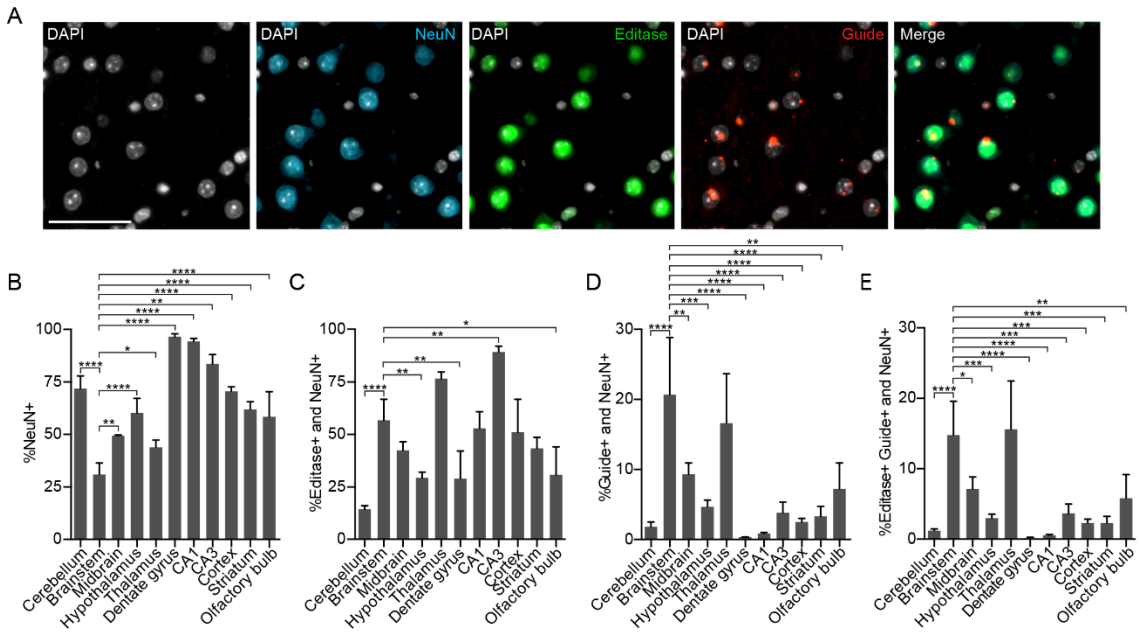
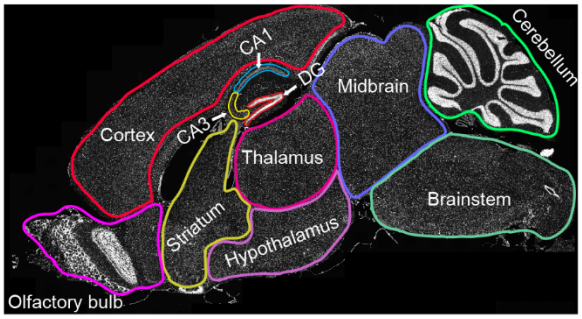


Fig. S3. Multiplex BaseScope detection of cells expressing Editase^{wt} and/or guide RNA in NeuN+ neurons across brain regions. (A) Representative images from the brainstem of a male *Mecp2*$^{+/y}$ mouse injected with the *gfp* guide virus. DAPI marks nuclei. Scale bars = 50 µm. Values are relative to DAPI (B) or to the total number of cells immuno-labeled with antibody to NeuN (C-E). n = 3 mice. Percentages are mean ± SD. *p<0.05, ** p< 0.01, *** p <0.001, **** p<0.0001. All comparisons lacking asterisks were not significant. Statistics were determined using a one-way ANOVA relative to brainstem values.



Brain Region	Mean cell #
Cerebellum	33456
Brainstem	26249
Midbrain	31338
Hypothalamus	13121
Thalamus	12955
Dentate gyrus	2375
CA1	1171
CA3	595
Cortex	39659
Striatum	20447
Olfactory bulb	31438

Fig. S4. Left: Schematic of brain regions selected for HALO analysis. Right: Mean cell numbers measured for the indicated brain regions based on two sections/region from three mice using HALO software.

	Sequence, 5′→ 3′
Genotyping primers	
Mecp2^{G311A} Fwd	CCC ACCTTGCCTGAAGGGTA
Mecp2^{G311A} Rev	CCTAGCCTTCCTACCCACCTG
Sex Fwd	ACCTTAAGAACAAGCCAATACA
Sex Rev	GGCTTGTCCTGAAAACATTTGG
Guide sequences	
Mecp2^{G311A} 2xBoxB guide	TCCTTTGTTTGGCCCTGAAAAAGGGCCCTTTCGTGTCCAACCTTCAGGCAAGGGGCCCTGAAAAAGGGCCGGTCATCATAC
gfp 2xBoxB guide	TCAGGGTAGTGGCCCTGAAAAAGGGCCAAGTGTTGGCCATGGAACAGGTAGTTTTCGGCCCTGAAAAAGGGCCTAGTGCAAAT
Amplification of *Mecp2* cDNA	
Mecp2 ATG Fwd	AACCCGTCCGGAAAATGGCC
Mecp2 3′UTR Rev	GGAAGCTTTGTCAGAGCCCTACCCATAAG
Sequencing primers for *Mecp2* RT-PCR	
Mecp2 554 Rev	CTCCTGGAGGGGCTCCCTCTC
Mecp2 914 Rev	GACCGTATGGAAGACTCCTTCA
Mecp2 1122 Rev	ACTGCTGCTGCGCCCCTT
Cloning primers	
Human U6 NdeI Fwd	GTGTCATATGCTTACCGTAACTTGAAAG
gfp 2xBoxB ApaI Rev	ATGCATCTAGAAAAAAATTTGCACTAGGCCC
Mecp2^{G311A} 2xBoxB XbaI Rev	ATGCATCTAGAAAAAAGTATGATGACCGGCCC
Editase KpnI Fwd	GCGCGTACCCACCATGGTGTACCCC
Editase EcoRI Rev	GCGCGAATTCTCAATGGTGATGGTGATGG
CAG ApaI Fwd	ATGCAGGGCCCACTAGTTATTAATAGTAATCAATTACGG
CAG KpnI Rev	ATGCAGGTACCGTTGGGCCGCAATTC
qPCR primers	
Mecp2 Fwd	CATACATAGGTCCCCGGTCA
Mecp2 Rev	CAGGCAAAGCAGAAACATCA
GAPDH Fwd	AGAAGGTGGTGAAGCAGGCA
GAPDH Rev	CGAAGGTGGAAGAGTGGGAG

Table S1. Guide RNA and primer sequences.

Supplemental methods

Generation of *Mecp2*$^{311G>A}$ knock-in Mouse Strain using CRISPR/Cas9. The database https://chopchop.cbu.uib.no/ was used to scan *Mecp2* to identify sites to generate the Mecp2$^{311G>A}$;MeCP2^{W104X} knockin mutation. The *Mecp2* sgRNA (5'- ACC CCA CCT TGC CTG AAG GT) was prepared by in vitro transcription using MEGAshortscript T7 kit (Life Technologies) and a gBlock fragment from IDT as template. The gBlock fragment contains the T7 promoter followed by MeCP2 sgRNA sequence, crRNA, connecting loop and tracRNA. Transcribed sgRNA was then purified using a MEGAclear kit (Life Technologies). To prepare sgRNA/Cas9 ribonucleoprotein, sgRNA was incubated with Cas9 protein (obtained from IDT) at a final concentration of 50 ng/ul and 150 ng/ul, respectively. A single-strand DNA donor (obtained from IDT), which was used as template for homology-directed repair to introduce the *Mecp2*$^{311G>A}$; MeCP2^{W104X} mutation, was mixed with the sgRNA/Cas9 ribonucleoprotein to the final concentration of 50 ng/ul. The donor sequence is: 5'- A*A*G* ACT TGC TCT TAC TTA CTT GAT CAA ATA TAC ATC ATA CTT TCC AGC AGA TCG GCC AGA CTT CCT TTG TTT AAG CTT TCG TGT C**T**A ACC TTC AGG CAA GGT GGG GTC ATC ATA CAT AGG TCC CCG GTC ACG GAT AAT GGA GCG CCG* C*T*G (**T** is the 311G>A knockin mutation and *denotes phosphorothioate linkages to prevent degradation). This mixture was then microinjected into the pronucleus of C57BL6 one-cell embryos and then transferred into pseudo pregnant females for subsequent development. Founder animals were initially identified by PCR using primer pair (nMC-W104XF: 5'- CCC ACC TTG CCT GAA GGG TA and nMC-I3R: 5'- CCT AGC CTT CCT ACC CAC CTG) that amplified a fragment of 146 bp specific to the *Mecp2*^{G311A};MeCP2 ^{W104X} knock-in mutation. Potential founders were further confirmed by PCR using primer pair (nMeCE3F: 5'- CTC AGG CTC TGC CCC AGC AG and nMC-I3R: 5'- CCT AGC CTT CCT ACC CAC CTG) to amplify a fragment of 237 bp spanning the knock-in mutation followed by sequencing of the PCR product to confirm their identity.

Generation of nuclear extracts for Western blotting. Frozen tissue was placed in ice cold PBS pH 7.4 containing 1mM PMSF and minced with razor blades. Minced tissue was placed in a dounce homogenizer in 1ml per 100mg tissue in hypotonic lysis buffer (20 mM HEPES pH 7.9 KOH, 10 mM KCl, 1mM $MgCl_2$, 1X Protease Inhibitor Cocktail (Complete EDTA-free; Roche cat# 11836170001) and 0.5mM DTT). Suspended samples were first homogenized using five strokes of the loose pestle. The cells swelled for 10 min on ice before being homogenized using ten strokes of the loose pestle followed by ten strokes using the tight pestle. The homogenate was pelleted by centrifugation at 500 x relative centrifugal field (r.c.f) at 4 °C for 5 min. The pellet containing the nuclei was then resuspended in hypotonic lysis buffer and filtered through a 40-µm cell strainer. Nuclei were then pelleted by centrifugation at 500 x r.c.f. and washed

twice with hypotonic lysis buffer. After the second wash nuclei were resuspended in whole cell lysis buffer (20 mM HEPES pH 7.9 KOH, 10 mM KCl, 150mM NaCl, 1mM $MgCl_2$, 1X Protease Inhibitor Cocktail (Complete EDTA-free, Roche, cat# 11836170001), 1% Triton X100, 0.1% SDS and 0.5mM DTT) containing 750 units per ml of benzonase nuclease (Sigma-Aldrich, cat# E1014). Lysates were centrifuged at 14,000 × r.c.f. for 10 min at 4° C and the soluble fraction was used as the nuclear extract.

Plasmid constructs. All guide sequences and primers are included in Supplemental Table 1 and plasmids are available upon request. All constructs were verified by sequence analysis. Plasmid pGM1435 and plasmid pGM1437 contain the control *gfp* and *Mecp2* targeting guides, respectively. Each guide also contains two copies of the bacteriophage lambda BoxB RNA hairpin (2). Each plasmid encodes the ADAR2 wild type catalytic domain fused to two copies of the bacteriophage lambda 2xBoxB RNA binding domains (N-peptide; $Editase^{wt}$; (2). The *gfp* guide contains sequences complementary to a fragment of the humanized green fluorescent protein (Supplemental Table 1). A plasmid expressing the *gfp* guide under control of the human *U6* promoter (pGM1109) was produced by ligating synthetic oligonucleotides with BsaI over-hangs, representing the forward and reverse sequence of the guide, into the pENTR/*U6* polylinker backbone (pGM1099).

To generate pGM1435, we removed the *U6* promoter and CRISPR sequences from plasmid pX552 (RRID: Addgene_60958), by restriction digest with NdeI and ApaI, and replaced them with a *U6* promoter and hybrid 2X BoxB-GFP guide that was prepared by PCR amplification and restriction with NdeI and ApaI enzymes from plasmid pGM1109. The guide sequences are present as a single copy in pGM1435. The $Editase^{wt}$ cDNA was amplified by PCR from pGM1090 (2), restricted with KpnI and EcoRI and cloned into pGM1435. To achieve ubiquitous expression of the Editase *in vivo*, the *Synapsin I* promoter was removed from pGM1435 with ApaI and KpnI. This fragment was and replaced with the CAG (CMV enhancer, chicken ß-actin promoter) promoter by PCR amplification from pAAV CAG-NLS-GFP (RRID: Addgene_104061). A woodchuck hepatitis virus post-transcriptional regulatory element (WPRE), for increasing transgene expression from AAV, was then added to pGM1435. This was accomplished by removing the existing bovine growth hormone (bGH) poly(A) sequence, and adding bGH sequences that contained both a WPRE and the bGH poly(A) sequence from pAAV CAG-NLS-GFP, using EcoRI and BstxI enzymes. To create plasmid pGM1437, containing the $Mecp2^{G311A}$ targeting guide, the *U6* GFP 2xBoxB guide was removed from pGM1435 by NdeI and XbaI digestion and replaced with the *U6* promoter and $Mecp2^{W104X}$ guide PCR amplified with NdeI and XbaI overhangs.

Genomic DNA extraction and exome sequencing. Genomic DNA was extracted from tail biopsies and treated with RNase A using the Qiagen DNeasy Blood and Tissue kit (Qiagen, cat# 69504) according to the manufacturer's instructions. Mouse whole exome library preparation and sequencing were

performed by Genewiz. Libraries were prepared using an Illumina DNA prep kit with exome enrichment and sequencing was performed using 2x150 bp read length on an Illumina HiSeq 2500 at a depth of 12gb per sample.

Next Generation Sequencing. All RNA samples were tested for purity using a Bioanalyzer 2000 and had RIN values of >9.3. cDNA libraries were made by the OHSU Massively Parallel Sequencing Shared Resource using the TruSeq stranded total RNA with ribo-zero kit (Illumina). Library quality was assessed using a TapeStation 220 and followed by quantification using quantitative PCR. Libraries were sequenced using paired end sequencing on a NovaSeq 6000 using a 200 cycle S4 flow-cell kit containing 10 million clusters.

Bioinformatic analysis for RNA editing events. RNA-seq data was first trimmed using fastp (3) and aligned to the mm39 reference genome using the STAR protocol (4). The whole exome sequencing data was aligned to the same reference using bwa-mem (5). The RNA and DNA reads were then quantified using REDItools suite of python scripts (6). Specifically, the number of DNA base reads with an adenosine ('A') were compared to the number of RNA reads at the same position with a guanosine ('G'). Results were filtered by requiring 10 DNA-seq reads (6) and 5 RNA-seq reads. To predict the impact of RNA-editing and annotate the sites, we used Variant Effect Predictor (7). To determine whether RNA editing at sites identified in the samples from mice injected with the *Mecp2* targeting virus occurred at greater frequency compared to either the *gfp* guide virus or non-injected mutant control samples, we compared the proportion of reads with RNA base 'G', using a one-sided Pearson's Chi-squared test and a Benjamani-Hocberg correction (8). A 2% lower threshold of off-target editing was determined by comparing DNA bases with reads with an 'A' to the proportion of RNA reads corresponding to a 'C' or 'T,' which cannot be introduced by editing.

Observational scoring of wild type mice injected with the *Mecp2* targeting virus. Mice were scored weekly beginning at P28-P35, prior to injection, as described previously (9). Briefly, individual mice were placed onto a metal laminar flow hood for observation where they were accorded a score (0-2, 2 being the most severe) for each of the following phenotypes: mobility, gait, hindlimb clasp, tremor, and general condition. An animal was considered to have reached a humane endpoint when they lost 20% of their peak body weight, or exhibited a tremor or general condition score of 2. The scoring was performed blind to condition and by two independent people in the lab.

Single Cell Analysis for MeCP2 protein repair. Acquisition of images. A Zeiss 710 laser scanning confocal microscope equipped with a 63x Plan-Apochromat objective was used to acquire single plane images from the tissue sections using Zen Digital Imaging Software (Zeiss, RRID: SCR_013672). For each image, the field size corresponded to 18211.5 μm^2 at a resolution of 1024x1024 pixels. Fluorescent images corresponding to staining with antibodies to HA (488 laser), MeCP2 (561 nm laser) and to DAPI (405 nm laser) were sequentially acquired.

The laser strengths for MeCP2 and DAPI were adjusted independently to fall within the non-saturating 0-255 range for wild type samples and were held constant for image acquisition for the corresponding *Mecp2*^{W104X} injected mice. The laser strength for HA was adjusted within the non-saturating 0-255 range for *Mecp2*^{W104X} injected mice and was held constant for wild type samples.

MeCP2 immunofluorescence quantitation. MeCP2 immunofluorescence within the heterochromatic foci in the nuclei of mice was quantitated using Image J (RRID: SCR_003070) (10). For wild type, mutant and virally injected mutant mice, we selected nuclei and measured 4x4 pixel arrays for all the foci, which varied between 1 and 4. In each case, we also measured background fluorescence corresponding to regions devoid of cell nuclei. Low background levels were comparable among the three conditions. For wild type mice, 50 nuclei were chosen randomly for the heterochromatin analysis of endogenous MeCP2 protein. The median value for all foci was determined after background subtraction. The same methodology was performed for mice injected with the *Mecp2* targeting virus except that 150 nuclei were measured. The measurements were used to bin cells into distinct repaired and non-repaired categories. Cells were considered repaired if the nuclear intensity was above background levels. For the specific experiments comparing the nuclear MeCP2 immunofluorescence between wild type mice and repaired cells in the mice injected with the *Mecp2* targeting virus, we measured MeCP2 levels in both the nucleoplasm and the heterochromatin puncta. These measurements were made in 44 cells from a single mouse representing each condition.

Measurement of the fraction of cells expressing HA-Editase^{wt} and exhibiting MeCP2 protein repair. Cell counts were performed using the ImageJ cell counter plugin as previously described (2, 10). Briefly, the proportion of HA immunolabeled cells was determined by counting the fraction of DAPI positive nuclei that were also HA positive. DAPI positive nuclei were considered repaired if there was nuclear MeCP2 protein and immunolabeling in heterochromatic foci.

Low magnification images of brainstem and cerebellum. A Zeiss 710 laser scanning confocal microscope equipped with a 40x Plan-Apochromat objective was used to acquire tiled single plane images from the tissue sections using Zen Digital Imaging Software (Zeiss, RRID: SCR_013672). The tiled images were then stitched using a 10% overlap with Zen Digital Imaging software.

SI References

1. J. Guy, J. Gan, J. Selfridge, S. Cobb, A. Bird, Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315**, 1143-1147 (2007).
2. J. R. Sinnamon *et al.*, Site-directed RNA repair of endogenous Mecp2 RNA in neurons. *Proc Natl Acad Sci U S A* **114**, E9395-E9402 (2017).
3. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884-i890 (2018).
4. A. Dobin *et al.*, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).

5. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **1303.3997** (2013).
6. C. Lo Giudice, M. A. Tangaro, G. Pesole, E. Picardi, Investigating RNA editing in deep transcriptome datasets with REDItools and REDIportal. *Nat Protoc* **15**, 1098-1131 (2020).
7. W. McLaren *et al.*, The Ensembl Variant Effect Predictor. *Genome Biol* **17**, 122 (2016).
8. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **57**, 289-300 (1995).
9. S. K. Garg *et al.*, Systemic delivery of MeCP2 rescues behavioral and cellular deficits in female mouse models of Rett syndrome. *J Neurosci* **33**, 13612-13620 (2013).
10. J. R. Sinnamon *et al.*, In Vivo Repair of a Protein Underlying a Neurological Disorder by Programmable RNA Editing. *Cell Rep* **32**, 107878 (2020).