Cell Reports

Supplemental Information

MBNL Sequestration by Toxic RNAs and RNA Mis-Processing in the Myotonic Dystrophy Brain

Marianne Goodwin, Apoorva Mohan, Ranjan Batra, Kuang-Yung Lee, Konstantinos Charizanis, Francisco José Fernández Gómez, Sabiha Eddarkaoui, Nicolas Sergeant, Luc Buée, Takashi Kimura, H. Brent Clark, Joline Dalton, Kenji Takamura, Sebastien Weyn-Vanhentenryck, Chaolin Zhang, Tammy Reid, Laura P.W. Ranum, John W. Day and Maurice S. Swanson

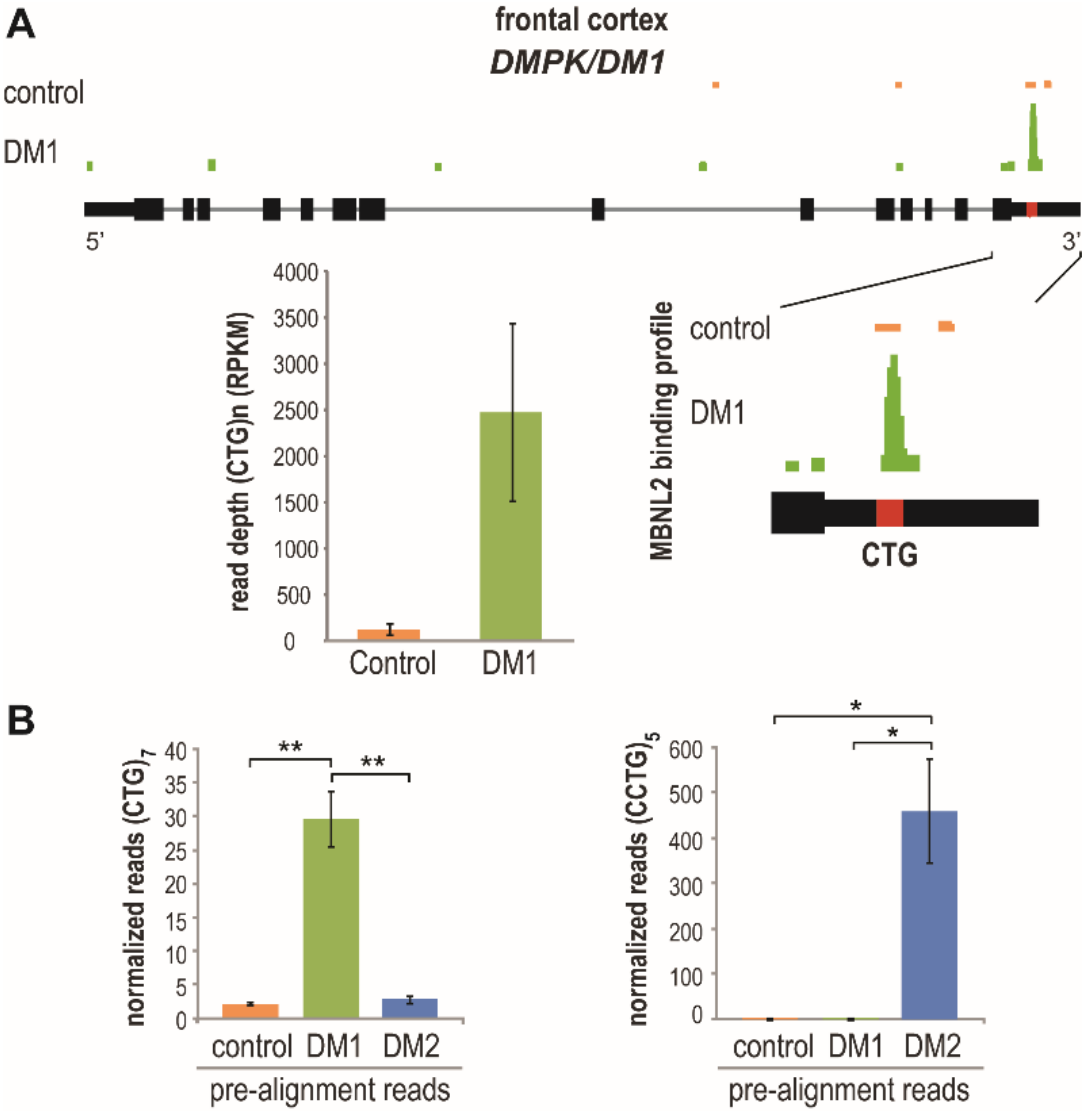

Figure S1. HITS-CLIP in DM1 and Mapping CTG/CCTG Repeats, Related to Figure 1

(A) MBNL2-CUG^{exp} interactions in DM1 frontal cortex. UCSC browser view of wiggle plots of MBNL2 CLIP binding profile in the *DMPK* reference gene (CTG repeat, red box; coding exons, thick black boxes; UTRs, thin black boxes; introns gray lines) for control (orange) and DM1 (green) patient frontal cortex. Zoomed-in view of *DMPK* terminal exon (bottom right) and quantification (bottom left) of average MBNL2 CLIP peak read depth (RPKM) in the *DMPK* CTG repeat region, showing enrichment (18-fold) over controls (n = 3 per group, data are reported ± SEM).

(B) Normalized $(CTG)_7$ (left) and $(CCTG)_5$ (right) pre-alignment reads (see Table S1 for details).

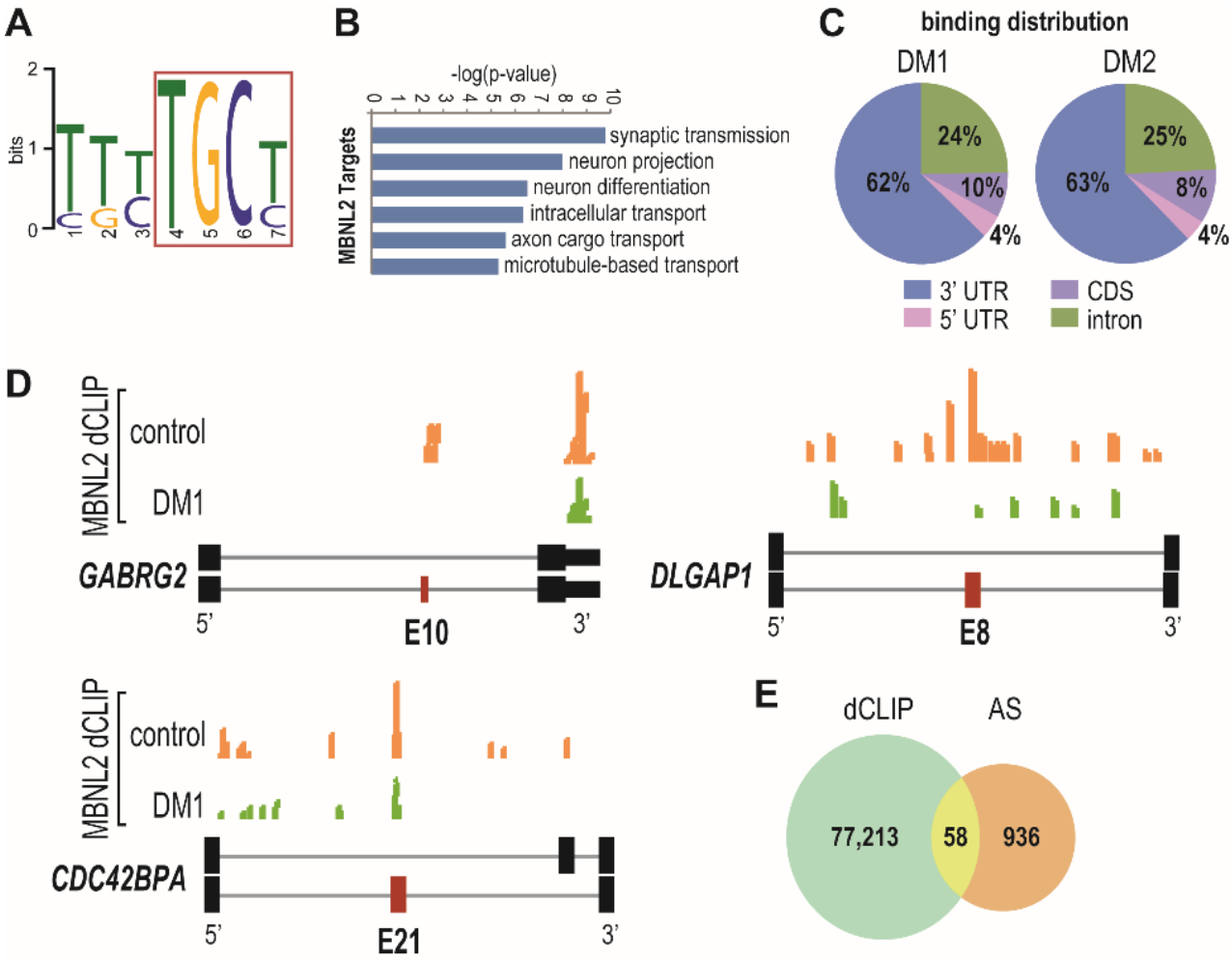

Figure S2. Altered MBNL2 Binding in the DM Brain, Related to Figure 2

(A) Preferred MBNL2 binding site motif in human brain with YGCY motif highlighted (red box).

(B) Gene ontology analysis of MBNL2 binding targets in the human brain showing principle biological processes.

(C) Pie charts of distribution of MBNL2 binding sites identified by differential CLIP (dCLIP) analysis for DM1 versus control (left) and DM2 versus control hippocampus.

(D) UCSC genome browser view of MBNL2 dCLIP binding profiles showing examples (*GABRG2*, *DLGAP1*, *CDC42BPA*) of reduced binding near mis-spliced exons in the DM1 brain versus controls (red box, alternative exon, thick black boxes, flanking exons; gray lines, introns) (n=3).

(E) Intersection of dCLIP (significance > 0.95, differential > |2.0| and AS regions (FDR <0.05, dI > |0.1|) where AS region = AS exon + flanking introns + flanking constitutive exons to find the overlap between differential MBNL2 binding and AS changes in DM1.

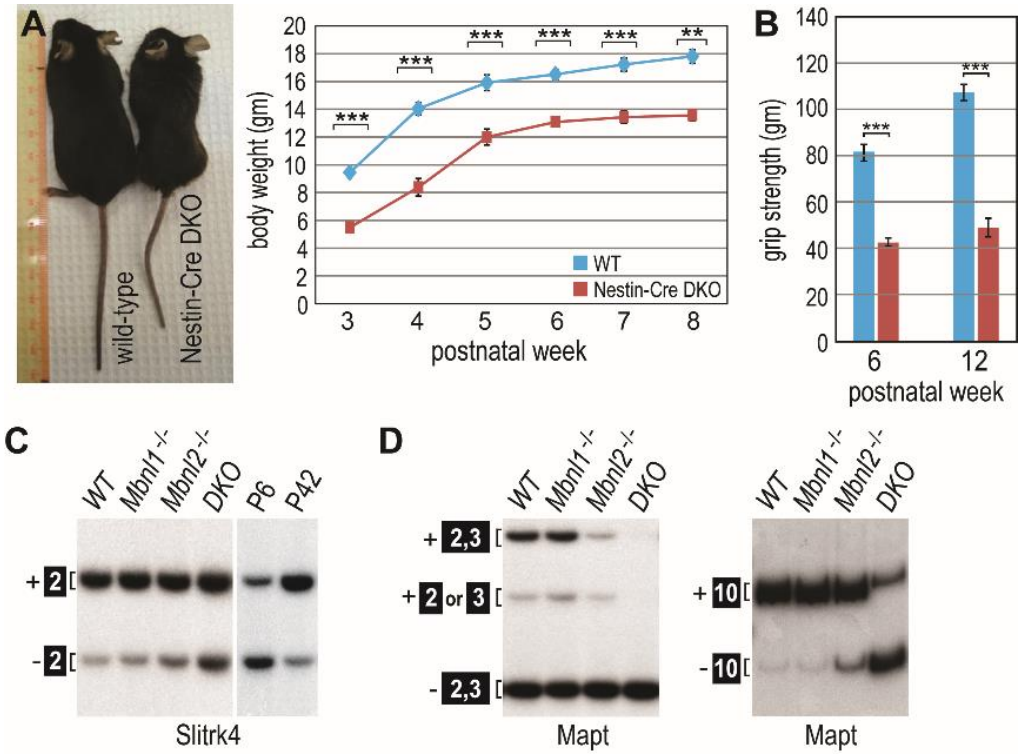

Figure S3. Phenotypic/Molecular Abnormalities in *Mbnl1*$^{-/-}$; *Mbnl2*$^{c/c}$; *Nestin-Cre*$^{+/-}$ DKO Mice, Related to Figures 3 and 4

(A) Smaller body size (left) and reduced body weight (right) of the *Nestin*-Cre DKO versus WT mice (n=21 for each group, data are reported ± SEM, ***p < 0.001, **p < 0.01).

(B) Grip strength analysis comparing *Nestin*-Cre DKO (red) and WT (blue) mice at postnatal weeks 6 and 12 (n ≥ 6 for each group, data are reported ± SEM, ***p < 0.001).

(C) RT-PCR splicing analysis of Slitrk4 in WT, *Mbnl1*$^{\Delta E3/\Delta E3}$, *Mbnl2* $^{\Delta E2/\Delta E2}$, and *Nestin*-Cre DKO brain. Developmental splicing pattern in WT mice at postnatal day (P)6 and P42 shown for comparison.

(D) Same as (C) except splicing analysis of Mapt exons 2 and 3 (left panel) and exon 10 (right).

Figure S4. MBNL2 Binding Sites in MAPT Intron 10, Related to Figure 5

MAPT gene sequence (exons 10 and 11, black boxes with white lettering) and the MBNL2 crosslinking sites (gray boxes) with YGCY motifs indicated (red lettering).

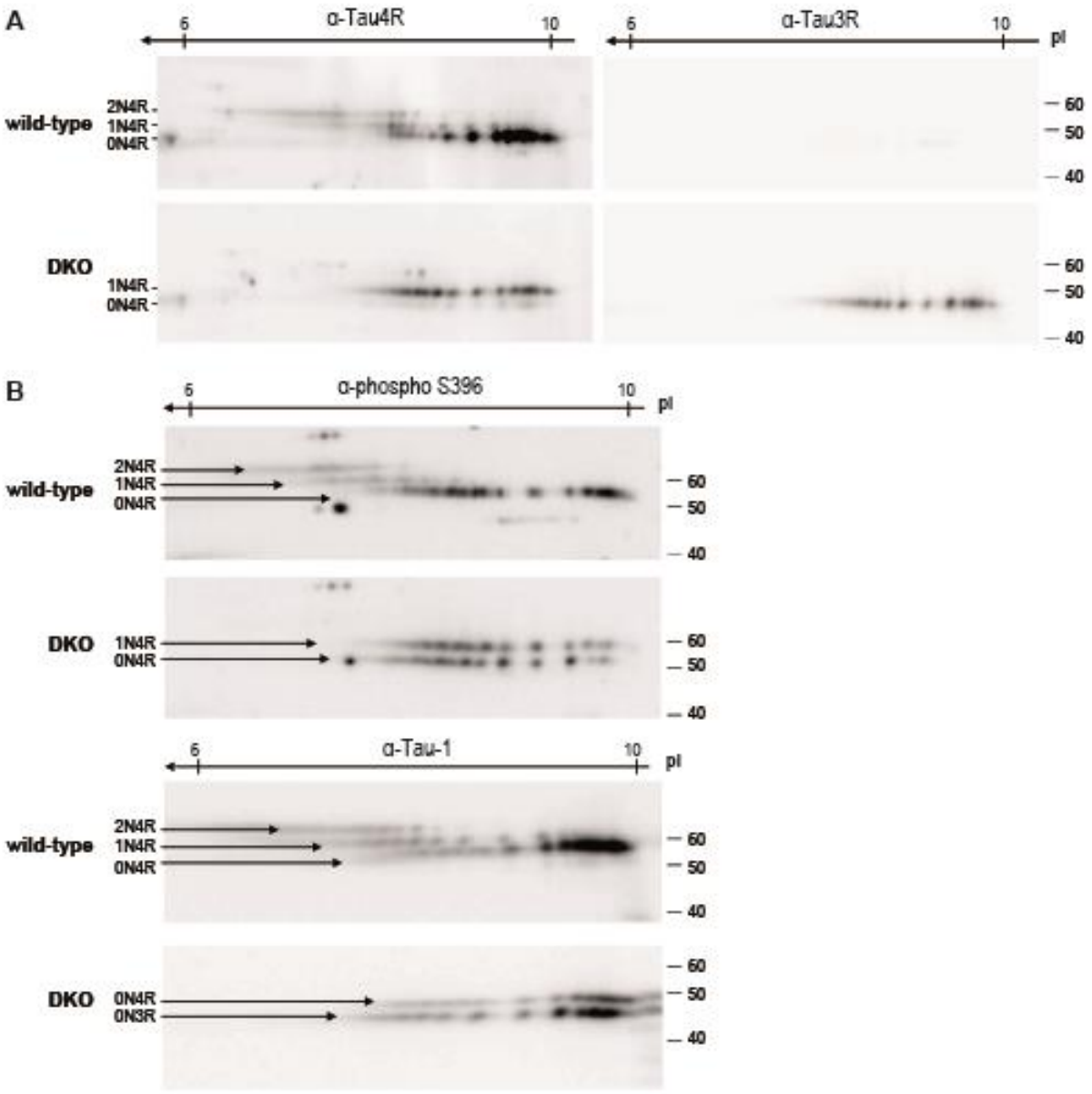

Figure S5. Altered Tau Expression in *Mbnl1*$^{-/-}$; *Mbnl2*$^{c/c}$; *Nestin-Cre*$^{+/-}$ Brain, Related to Figure 5

(A) Tau expression in *Nestin*-Cre DKO versus WT brain. Two-dimensional immunoblot analysis of Tau isovariants that include (2N4R, 1N4R, 0N4R, left) or exclude (0N3R, right) the forth microtubule binding domain. Tau 4R or Tau3R isoforms were detected with an exon 10 specific polyclonal antibody (α-Tau4R) or the Tau 3R mAb (α-Tau3R), which recognizes an epitope that overlaps exons 9/11 sequence. Note α-Tau3R does not detect tau proteins in wild-type brain extracts.

(B) Immunoblots of wild-type and *Nestin*-Cre DKO brains comparing abundance of tau isovariants phosphorylated at serine 396 (top panel) versus unphosphorylated (residues 198-205, bottom panel). Phosphorylated isovariants at Serine 396 (numbering according to the longest human tau isoform 441) were detected with the α-phospho S396 Tau antibody while unphosphorylated tau isovariants were detected with mAb Tau-1 antibody (α-Tau-1). Note that most basic isovariants, corresponding to the native protein, are lightly stained by α-phospho S396 while strongly detected by a-Tau-1 and vice versa for the acidic counterpart.

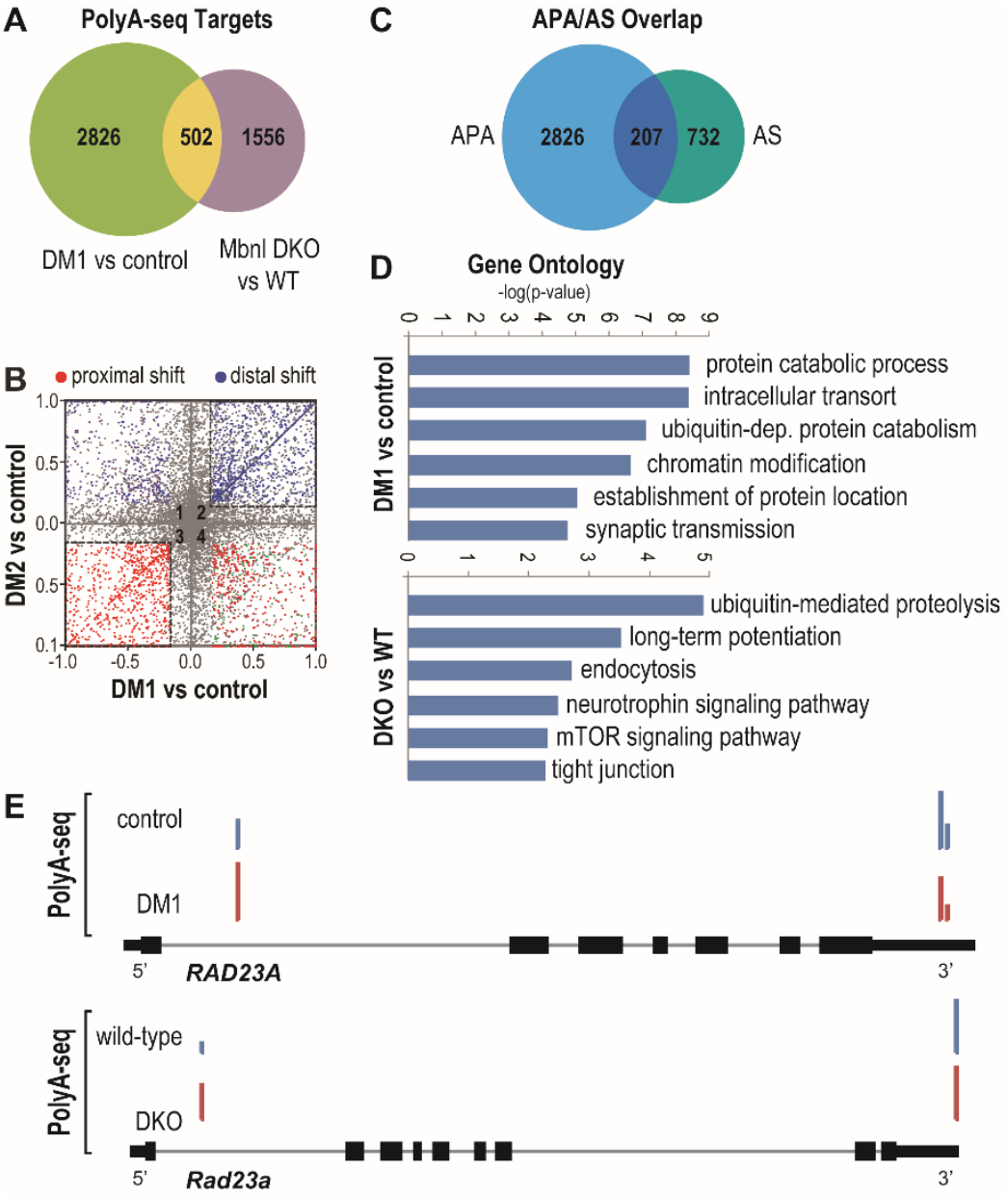

Figure S6. Alternative Polyadenylation Switches in the DM Brain, Related to Figure 6

(A) Venn diagram of gene overlap (502) with dysregulated APA in DM1 (total = 2,826 genes) and *Nestin*-Cre DKO brains (total = 1,556 genes).

(B) Scatter plot of correlations between pA site shifts (dI ≥ |0.15|, proximal; dI > 0, distal; FDR > 0.05 = unchanged/grey) in control versus DM1 and DM2. Squared quadrants (2 and 3) show APA shifts in the same direction for DM1 and DM2 while quadrant 1 (distal shifts in DM2, blue; proximal shifts in DM1, red) and quadrant 4 (distal in DM1, green; proximal in DM2, red) show shifts in the opposite directions between DM1 and DM2.

(C) Venn diagram showing overlap of genes misregulated in APA (FDR < 0.05, dI > |0.15|) and AS (FDR < 0.05, dI > |0.1|) in DM1 hippocampus.

(D) Gene ontology of dysregulated APA in DM1 brain versus controls (top) and *Nestin*-Cre DKO versus WT brain (bottom).

(E) UCSC browser view of the *RAD23A* gene showing wiggle plots of PolyA-seq data from control and DM1 (top) and WT and *Nestin*-Cre DKO brain (bottom) (n = 3).

Table S1. HITS-CLIP Validation of MBNL2 Sequestration in Human Hippocampus, Related to Figures 1, 2, 5, 6

(A) Control MBNL2 significant peaks (hippocampus)

(B) Control MBNL2 significant peaks (frontal cortex)

(C) Control MBNL2 binding distribution (hippocampus)

(D) Control MBNL2 binding distribution (frontal cortex)

(E) Gene Ontology MBNL2 (hippocampus)

(F) Gene Ontology MBNL2 (frontal cortex)

(G) CLIP tag summary

(H) Repeat reads

(I) Fold change repeat reads

(J) DM1 MBNL2 significant peaks (hippocampus)

(K) DM2 MBNL2 significant peaks (hippocampus)

(L) Control, DM1 and DM2 CTG and CCTG Repeat Counts

Table S2. Differential CLIP Analysis, Related to Figures 2, 5

(A) Control versus DM1 (hippocampus)

(B) Control versus DM2 (hippocampus)

(C) Control versus DM1 (frontal cortex)

(D) Genes with CAG/CTG and CAGG/CCTG repeats that show increased MBNL2 binding by dCLIP

Table S3. RNA-seq of human control versus DM1 brain, Related to Figures 4, 5

(A) Control versus DM1 frontal cortex (coverage ≥ 20, |dI| ≥0.1, FDR ≤ 0.05)

(B) Control versus DM1 frontal cortex (coverage ≥ 20, |dI| ≥0.2, FDR ≤ 0.05)

(C) Differential AS events with coordinate changes in MBNL2 binding (dCLIP)

Table S4. PolyA-seq Analysis of DM and Nestin-Cre DKO brain, Related to Figure 6

(A) Control versus DM1 (frontal cortex)

(B) Control versus DM2 (frontal cortex)

(C) Wild-type versus *Mbnl1*$^{-/-}$*; Mbnl2*$^{c/c}$*; Nestin Cre*$^{+}$ double knockout (frontal cortex)

(D) PolyA-seq genomic distribution

Table S5. Human Brain Autopsy Samples and Sequencing Records, Related to Figures 1, 2, 5, 6

(A) Sample information and genotype

(B) Sequencing records

(C) Tissue and RIN values

Table S6. Primers Used for Alternative Splicing and Alternative Polyadenylation Validations, Related to Figures 2-6

Supplemental Experimental Procedures

HITS-CLIP

Illumina sequencing FASTQ files were groomed to Sanger format, PCR duplicate reads collapsed, sequences trimmed (5 nt from the 5'end and 3 nt from the 3' end), 3'-adaptors clipped, and the resulting files were combined with quality scores and converted from BAM to SAM format. The processed sequencing files were mapped to the human genome (hg18 and hg19) using BWA (version 0.5.9-r16) and reads mapped to the same chromosomal location were collapsed. For visualization in the UCSC Genome Browser, wiggle plots were generated with conversion to units of reads per kilobase per million (RPKM). Wiggle plots were also generated from BED files representing three biological replicates per group (control, DM1, DM2). Significant peaks (>3 overlapping CLIP tags) were ranked by peak height and intervals were annotated to determine the binding distribution. For CTG and CCTG repeat counts, pre-alignment reads with contiguous CTG (2-9) or CCTG (2-6) repeats were counted and normalized to the total number of sample reads and are listed (Table S1L) as repeat containing/100,000 reads.

Differential CLIP

After MA plot normalization and processing, the following code and modified parameters were used to define differential binding: perl /path/to/dCLIP1.6/bin/dCLIP.pl -f1 dCLIP1.5/path/RBP_disease.sam -f2 dCLIP1.5/path/RBP_control.sam -m1 5 -m2 5 -temp test -dir dCLIP1.5/path/output -mut "T2C, del" -filter 20. (-f1, SAM file for disease; -f2, SAM file for control; -m1, minimum number of reads in a cluster for disease (5 reads); -m2, minimum number of reads in a cluster for control (5 reads); -mut, CIMS designated as T>C substitution or deletion ("T2C, del"); -filter, minimum sum of reads per cluster from both conditions (20 reads). Output files were generated including summary.bed files with coordinates of all identified clusters; .bedgraph files for genome browser visualization of normalized CLIP clusters, and .txt files containing coordinates, significance, and differential values of all identified clusters. To generate annotated .txt files, the .txt files were intersected using the Bedtools software tool -intersectBed with hg19 Refseq bed file (from the UCSC Genome Browser). Annotated files were filtered by threshold

differential value and events with greater than 0.95 probability to identify significant differential binding events.

Tissue Preparation for Two-dimensional Gel Electrophoresis

For 2D-GE, frozen brain tissue was added to 10X volumes of 2M thiourea, 8M urea and 1% SDS (w/v) and homogenized (Teflon) followed by further sonication at 60 Hz. A total amount of 30 µg of protein was precipitated with 10X volumes of ice-cold acetone for 20 min at -20°C and proteins were pelleted at 12,000 x *g* for 10 min at 4°C. The acetone supernatant was discarded, the pellet was dried under nitrogen and then resuspended in 2D buffer containing 8M urea, 2M thiourea and 4% CHAPS (Fernandez-Gomez et al., 2014).

IEF was achieved by applying a total voltage of 13,000 Vh. Immobiline DryStrips were then equilibrated in 6M Urea, 2% (w/v) SDS, 30% glycerol, 50 mM Tris-HCl pH 8.6 buffer with 1% DTT. A second equilibration bath was performed with 4.7% iodoacetamide in the same buffer without DTT. The strips were then loaded on the top of Criterion XT™ 12% gels and SDS-PAGE run according to the manufacturer's instructions (BioRad). Proteins were transferred to 0.45 µM nitrocellulose membranes (GE HealthCare), which were subsequently incubated with primary antibodies for 2 hr at RT or overnight at 4°C. The following primary antibodies were used; 1) a-TauCter rabbit antiserum (final dilution 1:4000 in TBS-T) raised against the carboxy-terminus of human tau (Sergeant et al., 2001); 2) Tau-1 mAb (Millipore, clone PC1C6, final dilution 1:8000) against a non-phosphorylated epitope comprising tau residues 198-205 (numbering according to the longest human brain tau 441 isoform); 3) rabbit pAb against phosphorylated serine 396 (Life Technologies, 44-752G); 4) mAb against the 3R Tau isoform (anti-Tau 3-repeat isoform RD3, 1:2000 in TBS-T) (clone 8E6/C11, Millipore); 4) rabbit pAb against the sequence encoded by exon 10 (1:2000 in TBS-T) (Sergeant et al., 1999). Incubation with the secondary antibody was preceded by 3 washes of 10 min in TBS-T. Immunoreactive complexes were detected using the ECL™ luminescence system (GE Healthcare) and exposed to Hyperfilm™ ECL (GE healthcare) or using the LAS-4000 digital imaging system (GE healthcare). Raw images were used for figure preparation.

PolyA-seq

Total RNA (4 µg) was used as starting template for library preparation and libraries were sequenced (Illumina). Sequencing files were processed and

mapped to the mouse (mm10) or human (hg19) genome using the OLego aligner (Wu et al., 2013), filtered to remove events resulting from internal oligo(dT) mis-priming to A-rich stretches, and converted to wiggle files (RPKM). Valid pA sites were identified and the reads were trimmed to the exact nucleotides representing the PAS. The peak height for each pA site was recorded and differential analysis was performed to identify significant APA changes between groups (FDR ≤ 0.05, dI ≥ I0.15I).

Supplemental References

Fernandez-Gomez, F.J., Jumeau, F., Derisbourg, M., Burnouf, S., Tran, H., Eddarkaoui, S., Obriot, H., Dutoit-Lefevre, V., Deramecourt, V., Mitchell, V.*, et al.* (2014). Consensus brain-derived protein, extraction protocol for the study of human and murine brain proteome using both 2D-DIGE and mini 2DE immunoblotting. J Vis Exp *86*, doi: 10.3791/51339.

Wu, J., Anczukow, O., Krainer, A.R., Zhang, M.Q., and Zhang, C. (2013). OLego: fast and sensitive mapping of spliced mRNA-Seq reads using small seeds. Nucl Acids Res *41*, 5149-5163.