SUPPLEMENTAL FIGURES

Figure S1. ***TAF15 HITS-CLIP of human brains and mouse neurons, Related to Figure 1***

A, B. Specificity of anti-TAF15 antibody, A300-308A**.** Immunoblots (**A**) of protein extracts of human brain lysate (lanes 1 and 2; two different concentrations) and of embryonic stem cell (ES)-derived neurons (lane 3) using A300-308A. Immunoprecipitates (**B**) of human brain lysate were silver stained (top) and immunoblotted (bottom) with A300-308A. NRS; non-immune serum (negative control). The ~66 kDa band (black box) was submitted for mass-spectrometry revealing its identity as TAF15 (TAF15 peptides recovered: TGKPMINLYTDK, GEATVSFDDPPSAK, and RSGGGYGGDRSSGGGYSGDRS).

C. Specificity of HITS-CLIP assayed by Crosslinking-Induced Mutation Site (CIMS) analysis (CIMS). The rate of deletions is plotted relative to the 5´ end of each CLIP tag for mouse (left panel) and human (right panel). Note that CLIP tags harbor significantly more deletions than RNA-Seq tags (left panel), as typically seen in HITS-CLIP.

D. (ES)-derived neurons schematic. Mouse neurons were differentiated as described in Experimental Procedures. Representative phase-contrast image of mouse neurons (Day 9 of differentiation) and immunoblot showing reduction of TAF15 after TAF15 knockdown. GAPDH served as loading control is shown. Days of differentiation utilized for various assays are indicated.

Figure S2. ***Reproducibility of HITS-CLIP and SSD-RNA-Seq, Related to Figures 1 and 2***

A-C. Scatter plots of the number of tags per gene for TAF15-CLIP of (**A**) human, (**B**) mouse, and (**C**) mouse controls (CTRL) and TAF15 knockdowns (KD) RNA-Seq libraries. Pearsons' correlation coefficient is noted on the graphs.

Figure S3. ***Correlation between TAF15 and FUS targets and impact of TAF15 on transcripts and proteins related to neurological disorders, Related to Figures 1 and 2***

A. Scatter plot of mouse genes targeted by TAF15 compared to those targeted by FUS.

B. Impact of TAF15 on transcripts of genes-related to neurological disorders. Representative gels from RT-PCRs are shown for CTRL and TAF15 KD.

C. Impact of TAF15 on TDP-43 and FUS proteins. Immunoblot analysis of cultured mouse neurons after TAF15 knockdowns (KD), compared to control knockdowns (CTRL).

Figure S4. ***Examples of differentially included exons of mRNAs coding for glutamergic and GABAergic receptors upon TAF15 knockdown or upon enforced expression of TAF15 wild-type or ALS variant R408C, in mouse neurons, Related to Figure 3***

A. Shown are schematics (constitutive exon: black box; intron: black line; differentially included exon: red box) with primers (arrows) used for PCR and position of TAF15 peaks (blue boxes). Quantitation and representative gels from RT-PCRs are shown for control (CTRL) and TAF15 knockdowns (KD).

B. Mouse neurons were transduced with FLAG-tagged TAF15 wild-type, or TAF15-R408C (ALS variant) or with empty vector (CTRL) and transduced proteins were detected with anti-FLAG immunoblot.

C. Quantitation and representative gels from RT-PCRs are shown for neurons with enforced expression of TAF15 wild-type (WT), TAF15 ALS-variant (R408C) or controls (CTRL) transduced with empty vector.

SUPPLEMENTAL TABLES

Table S1. ***TAF15 top ranked hexamers in human and mouse CLIPs***

Table S2. ***Top targeted genes of TAF15 mouse CLIPs***

Table S3. ***Common targets between mouse TAF15 gene targets and FUS targets.***

Table S4. ***TAF15 differentially regulated transcripts and exons***

SUPPLEMENTAL TEXT

HITS-CLIP in Human Brain and Mouse Neurons

We first tested the specificity of the polyclonal anti-TAF15 antibody, A300-308A; as shown in **Figure S1A**, the antibody recognized a single protein migrating at ~66 kDa on immunoblots of human brain and mouse neurons, whose identity was confirmed to be TAF15 by immunoprecipitation and mass spectrometry (**Figure S1B**).

We also validated the specificity of TAF15 HITS-CLIP by Crosslinking-Induced Mutation Site (CIMS) analysis as described in (Zhang and Darnell, 2011) and (Vourekas et al, 2012). CIMS is based on slippage of the reverse transcriptase at or near the site of the crosslinked amino-acid-RNA adduct resulting in nucleotide deletion at the site of the crosslink. We found that CLIP-tags harbor over 10 times more deletions than RNA-seq tags strongly supporting that our HITS-CLIP analyses can reliably identify *in vivo* TAF15 mRNA targets (**Figure S1C**). Furthermore, Pearson's correlation of human brain samples substantiated the high correlation among all samples (*R* between 0.83-0.94; **Figure S2A**). Notably, the correlation between the libraries that were generated from the same sample but from slightly different portions of the membrane (2L and 2H; 3L and 3H), were also very high (*R* of 0.91 and 0.90 respectively; **Figure S2A).** Furthermore, Pearson's correlation between the two mouse samples was 0.99 (**Figure S2B**), indicating that mouse TAF15 CLIP was remarkably reproducible.

The Impact of TAF15 on the transcriptome of Mouse Neurons

Pearson's correlations of the reads from the three CTRL and the three TAF15 KD libraries ranged from 0.74 to 0.92 (**Figure S2C**) indicating that the knockdowns and RNA-Seq experiments were highly reproducible. Pearson's correlation between the average number of TAF15 CLIP-tags and the mean normalized expression levels of the transcripts in the RNA-Seq control samples did not show any correlation (R=0.025) indicating that the number of TAF15 binding sites is not simply a function of RNA expression levels.

To assess whether the changes in splicing of the differentially regulated exons were directly related to TAF15 binding, we evaluated the presence of TAF15 peaks in exons and within a 2Kb window upstream and downstream of each exon. We found that 74% of changed exons contained TAF15 binding sites within this region. In contrast 62% of unaffected exons had CLIP-tags associated with them, indicating that exons whose inclusion changes after TAF15 knockdown are significantly more targeted by TAF15 (Fisher exact test, p-value < 10^{-16}). Furthermore, we did not identify TAF15 CLIP-tags in the vicinity of 26% of changed exons suggesting that the change in splicing of these exons may reflect an indirect consequence of TAF15 knockdown. We also observed that the majority of changed upregulated exons (82.6%) and dowregulated exons (85.6%) do not belong to changed genes.

EXTENDED EXPERIMENTAL PROCEDURES

HITS-CLIP Library Preparation from Human Brains and Cultured Mouse Neurons

TAF15 CLIPs were performed using a previously developed protocol (Chi et al., 2009) with modifications as listed in (Vourekas et al., 2012). Temporal cortices from three normal human brains were used. Briefly, each harvested tissue was homogenized on ice in a Pyrex tissue grinder using 5 ml of ice cold HBSS (GIBCO) with mild pipetting. The suspension was placed in a 10 cm plate. Protein-RNA crosslink was achieved by UV irradiating the cell suspension at 254 nm (400mJ/cm2) in Stratalinker 1800 (Stratagene), on ice. The cells were pelleted at 2,000 rpm for 10 min at 4°C and divided in several tubes. The cell pellets were frozen in liquid nitrogen and kept at -80°C until use. For Immunoprecipitations, the pellet was resuspened in 500 μl 1x PXL buffer with protease inhibitors (Roche) in the presence of 2U of RNasin per μL of lysate. The lysate was incubated with DNase I. After treatment, the lysate was incubated in the absence or presence of limited amounts of RNase T1 (Roche 10109193001). RNase T1 treatment reduces the size of crosslinked RNAs to sizes compatible with Illumina sequencing. For beads preparation, 50 μL of protein A Dynabeads (Invitrogen 100-02D) was rotated with 10 μg of TAF15 antibodies in 1x PXL buffer at 4°C for ~3 hr. TAF15 antibody (A300-308A) proved compatible with the use of 1x PXL for cell lysis and immunoprecipitation, yielding highly specific radioactive signal (**Figure 1B**). Labelling of the 3´adapter and ligation to immunoprecipitated RNA CLIP-tags, SDS-PAGE, and RNA extraction were performed. For isolating TAF15-RNA complexes, the membrane was cut directly at the main signal (1L, 2L, and 3L, **Figure 1B**) and up to 5-7 kDa higher (2H and 3H, **Figure 1B**). RNA extraction, 5´ adaptor ligation, RT-PCR and re-PCR amplifications were performed. Reaction products from two PCR amplifications were analyzed on 3% Metaphor 1x TAE/EtBr gels. Gels were extracted with QIAquick Gel Extraction Kit.

A similar CLIP procedure was performed for cultured mouse neurons with few modifications. Briefly, two biological replicates of cultured neurons (~$2.5x10^7$ cells/per sample) were used to prepare TAF15 libraries. For beads preparation, we used 25 μl of protein A Dynabeads (Invitrogen 100-02D) with 5 μg of TAF15 antibodies in 1x PXL buffer and rotate at 4°C for ~3 hr.

Solid Support Directional (SSD) RNA-Seq Library Preparation

We performed SSD-RNA-seq (Vourekas et al., 2012) from control and TAF15-knockdowns of cultured mouse neurons (three biological replicates from each condition). Knockdowns were performed as described in the siRNA silencing section. Total RNA was isolated using TRIZOL reagent (Invitrogen), and treated with DNase I. The RiboMinus Eukaryote kit (Invitrogen A10837) was used per manufacturer's instructions to remove the ribosomal RNA from total RNA.

Approximately 1 μg of ribosomal RNA-depleted total RNA (RiboMinus RNA) was obtained from each replicate, and was fragmented using metal ion hydrolysis (Ambion AM8740). After RNA extraction, 50 picmoles of RL3-biotin adapter (IDT) was ligated to the fragmented RiboMinus-RNA using T4 RNA ligase in 80 μL of reaction mixture (Fermentas) at 16°C overnight. The ligated RNA was captured using (100 μL) M280 Streptavidin Dynabeads (Invitrogen 112.05D) and rotated for 15 min at RT. The beads were washed and resuspended in in the RL5 (150 pmoles) ligation mixture (80 μL) at 16°C for 6 hr. For preparation of cDNA, the beads were aliquoted in three tubes; the Titan One Tube RT-PCR System (11888382001 Roche) was used with one aliquot of beads per reaction. Reverse transcription was performed at 50°C for 45 min, and at 55°C for 15 min. Two PCR amplifications and gel extractions were performed as in HITS-CLIP procedure.

Genomic Analysis of HITS-CLIP and RNA-Seq Tags

Illumina sequencing was performed at a maximum of 76-nucleotides (nt) read length for HITS-CLIP and 100nt for RNA-Seq. The 3′-adaptor was removed from the sequences using the cutadapt software (Martin, 2011) and a 0.25 acceptable error rate for the alignment of the adaptor on the tag. To eliminate reads in which the adaptor was ligated more than one time, adaptor removal was performed 3 times. For RNA-Seq experiments an additional filtering step was introduced to remove the pool of TAF15 siRNAs used for knockdowns. All reads shorter than 15nt were excluded from further analysis. Sequencing reads were aligned to the human (hg19) and mouse (mm9) genomes respectively using Burrows-Wheeler Aligner and setting the default program parameters allowing for 0.04 fraction of missing alignments given a 2% uniform base error rate. All aligned reads that were shorter than 20nt and aligned to more than 1 position on the genome and all reads mapping on repeat elements as defined by RepeatMasker were excluded from further analysis. The human (hg19) and mouse (mm9) genome sequences and annotations for protein-coding genes were obtained from the University of California, Santa Cruz (UCSC) Genome Browser. To validate the specificity of our CLIPs, for each nucleotide position relative to the 5' end of the read, the mean percentage of reads with deletion in that position relative to the total number of reads with length equal or higher than this position was plotted. Mean percentages were calculated within 3 distinct library groups, 3 Human HITS-CLIP, 2 Mouse HITS-CLIP, and 3 Mouse control RNA-Seq.

Classifications in Functional Elements of the Genome

The number of tags which are contained within defined peaks and which overlap with functional elements of the genome (Intron, coding, 3′UTR, 5′UTR, non-coding, and intergenic), and

the percentages of these tags, which are contained within reproducible peaks, were counted. To assess for noncoding RNA binding, human and mouse lincRNAs were downloaded from NONCODE v3.0 and were filtered to keep only those that do not overlap with any genic region. The percentage of peaks overlapping these elements was measured. To capture potential effect from peaks not directly overlapping the lincRNAs but being in close proximity, a flanked region of 500nt on either side of the lincRNAs was used for overlap assessment.

Identification of Enriched Nmers in Peaks

To identify TAF15 binding motifs we estimated, independently for each library, the enrichment of every possible Nmer (4mers-7mers) within the top 5000 scored peaks that overlap UCSC annotated genes. Nmer enrichment was defined as the ratio of Nmer counts within these peaks over the number of times the Nmer was expected to be found based on the library nucleotide content. The ranked lists of all Nmers across all libraries were assigned p-values using the Rank Products method (Breitling et al., 2004). Significantly enriched Nmers with pfp<0.1 are shown in **Table S1**.

Gene Ontology Term Analysis of Targeted Genes

A list of genes was produced and sorted by the Rank Product of the number of HITS-CLIP tags mapping in peaks in the gene. An in-house method for GO enrichment based on Fisher's exact rank test was used against a background of genes that were present in both Human and Mouse CLIP datasets (at least 1 peak per gene) but were not highly targeted in any of the two species. Random permutations of the background list were also used to produce a better estimation of the probability of finding an enriched GO term by chance. For each GO term, ten thousand random sets of background genes with the same size as the original list were tested against a background of the union of the remaining background list and the original list. This allows for an estimation of the probability of a random background set being more significantly over-represented in the GO term up to 10^{-4}.

RNA-Seq and Functional Annotation Analyses

The estimation of expression levels for protein-coding genes was done using the number of RNA-Seq reads that mapped on the constitutively expressed exons of the gene. As constitutively expressed exons, we defined exons that were found in all known expressed isoforms of the gene downloaded from the UCSC Genome Browser. If a gene has both protein-coding and non protein-coding annotated transcripts only the protein coding transcripts are used for the definition of constitutive exons. For each gene and for each sample, the expression level was defined as the average number of reads mapped per Kb of constitutive exon. All gene expression levels per

sample were normalized using the upper quartile normalization factor, effectively dividing each expression level by the upper quartile of all expression levels. To identify differentially expressed genes, Student's t-test (N=3) was performed on the generalized logarithm of the gene expression levels (Vourekas et al, 2012). Genes significantly different between conditions (control and knockdowns) with $P<0.01$ were considered differentially expressed. For each exon, the number of reads per Kb that mapped on the exon was normalized using the per sample gene expression level quartile normalization factor. Differentially expressed exons were identified by the same method used for differentially expressed genes. Ribosomal RNAs and tRNAs were excluded from the analysis when calculating significantly changed genes or exons.

RT-PCRs

Total RNA was isolated using Trizol reagent according to manufacturer's protocol. cDNA was generated using random primers (IDT) and SuperScript III Reverse Transcriptase (Invitrogen). RT reaction was used as a template for transcript amplification in a 25 μl reaction with 25 pmol primer in a 1x GoTaq Mix. All samples were normalized to GAPDH. PCR amplification was performed at 94^0C for 3 min, 23-28 cycles at 94^0C for 30 s, 54^0C for 30 s, 72^0C for 30 s, 72^0C for 5 min. Amplified PCR products were resolved using 2% agarose/EtBr gels. Pictures were recorded by image quantity. Intensity ratio between products was averaged for three biological replicates per group. Differences between the control and the knockdown samples were assessed using Student's t-test. Primers were designed using Primer3 software (http://frodo.wi.mit.edu/) and sequences are available upon request.

Calmodulin Binding Assay

The binding assay was performed as described in (Hisatsune et al., 1997) with few modifications. Briefly, cultured neurons were transiently transfected with control or TAF15 siRNAs. Cells were washed 3X with cold 1x PBS buffer and were collected in Buffer A (10 mM Tris (pH 7.5), 1% NP-40, 150 mM NaCl and complete EDTA-free protease inhibitors (Roche)) and lysed for 1 hr at 4^0C. The total lysate was centrifuged at 15,000 rpm for 30 min. Calmodulin-agarose beads (Sigma) were added to the lysate. The mixture was rotated for 4 hr at 4^0C. Samples were resolved in 4-12% NuPAGE and subjected to immunoblotting.

Cell Surface Biotinylation

The biotinylation assay was performed as described in ((Nishimoto et al., 2009). Briefly, after control and TAF15 knockdowns, cells were overlaid with 1x PBS buffer containing 1.0 mg/ml of sulfo-NHS biotin (Pierce) to crosslink surface proteins for 45 min at 4°C. The cells were washed

and lysed on ice in lysis buffer (50 mM Tris-HCl (pH 7.5), 1% TritonX-100, 0.5% sodium deoxycholate, 0.1% SDS, 100 mM NaCl, 1 mM EDTA) plus a protease inhibitor mixture (Roche). After centrifugation, the lysate was affinity-purified using streptavidin-agarose beads (GE healthcare) with rotation overnight at 4°C. Samples were resolved in 4-12% NuPAGE followed by immunoblotting. The total cell lysates and streptavidin-purified samples were probed for the intracellular protein GAPDH to control the possibility that the sulfo-NHS-biotin penetrates the cells during the biotinylation procedure. Only those experiment in which were negative for GAPDH control were used for analysis.

Plasmids and Lentiviral Transduction

TAF15 wild-type and TAF15-variant (R408C) entry vectors were kindly provided by A. Gitler's laboratory. PCR amplification was performed using the following primers: hTAF15 specific primers (Forward, hTAF15SpeIKozMFlagf 5'-CCCGGGACTAGTCACCATGGACTACAAGGACGACGATGACAAAATGTCGGATTCTGGAAG T-3'; Reverse, hTAF15MycNotIr 5'-CACGCGGCCGCCTACAGATCCTCTTTCTGAGATGAGTTTTTGTTCGTATGGTCGGTTGCGC-3'). PCR amplified fragments were cloned into the pEN-Tmcs entry vector and recombined using LR-clonase (Invitrogen) into pSLIK-Neo destination vector (Signaling-gateway). The pSLIK expression lentiviral vectors of TAF15 wild-type and TAF15-R408C were transfected along with packaging and pseudotyping plasmids into 293T cells using Lipofectamine 2000 reagent (Invitrogen) according to the manufacturer's protocol. The viral supernatant was collected 48 hr after transfection and concentrated by ultracentrifugation onto a 20% sucrose gradients using SW41 rotor (Beckman) at 20,000 rpm for 3 hr at 4 ºC. Viral pellets were resuspended in ADFNB medium and stored in aliquots at -80 ºC. For transduction, neuronal cells were mixed with the virus at a low MOI. To induce protein expression, 1μg/ml Doxycyline (Milipore) was added to the cells 24 hr after transduction and for five days. Cells were harvested for various assays.

SUPPLEMENTAL REFERENCES

Breitling, R., Armengaud, P., Amtmann, A., and Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett *573*, 83-92.

Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature *460*, 479-486.

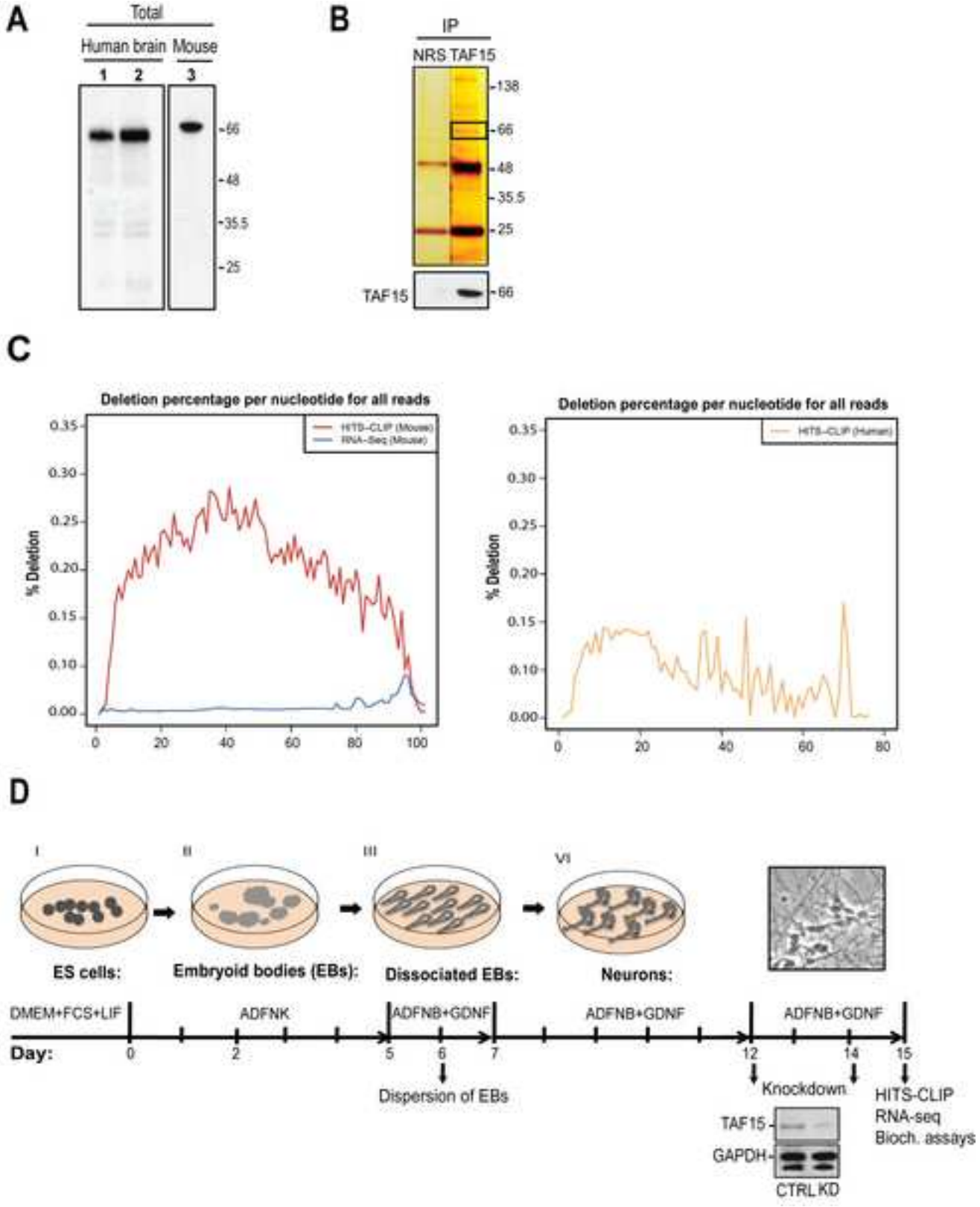
Hisatsune, C., Umemori, H., Inoue, T., Michikawa, T., Kohda, K., Mikoshiba, K., and Yamamoto, T. (1997). Phosphorylation-dependent regulation of N-methyl-D-aspartate receptors by calmodulin. J Biol Chem *272*, 20805-20810.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, North America *17*.

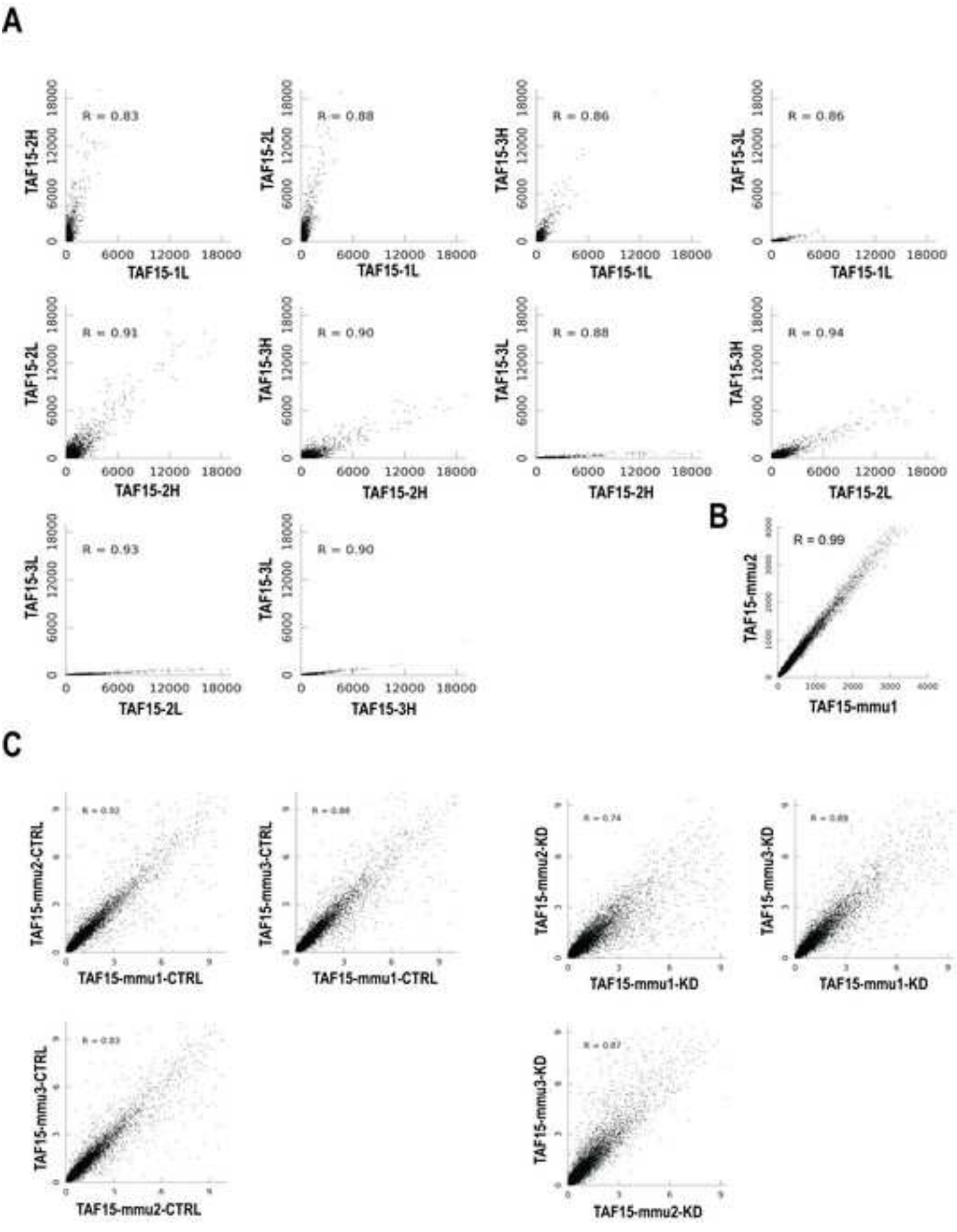
Nishimoto, T., Kihara, T., Akaike, A., Niidome, T., and Sugimoto, H. (2009). AMPA reduces surface expression of NR1 through regulation of GSK3beta. Neuroreport *20*, 161-165.

Vourekas, A., Zheng, Q., Alexiou, P., Maragkakis, M., Kirino, Y., Gregory, B.D., and Mourelatos, Z. (2012). Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. Nat Struct Mol Biol *19*, 773-781.

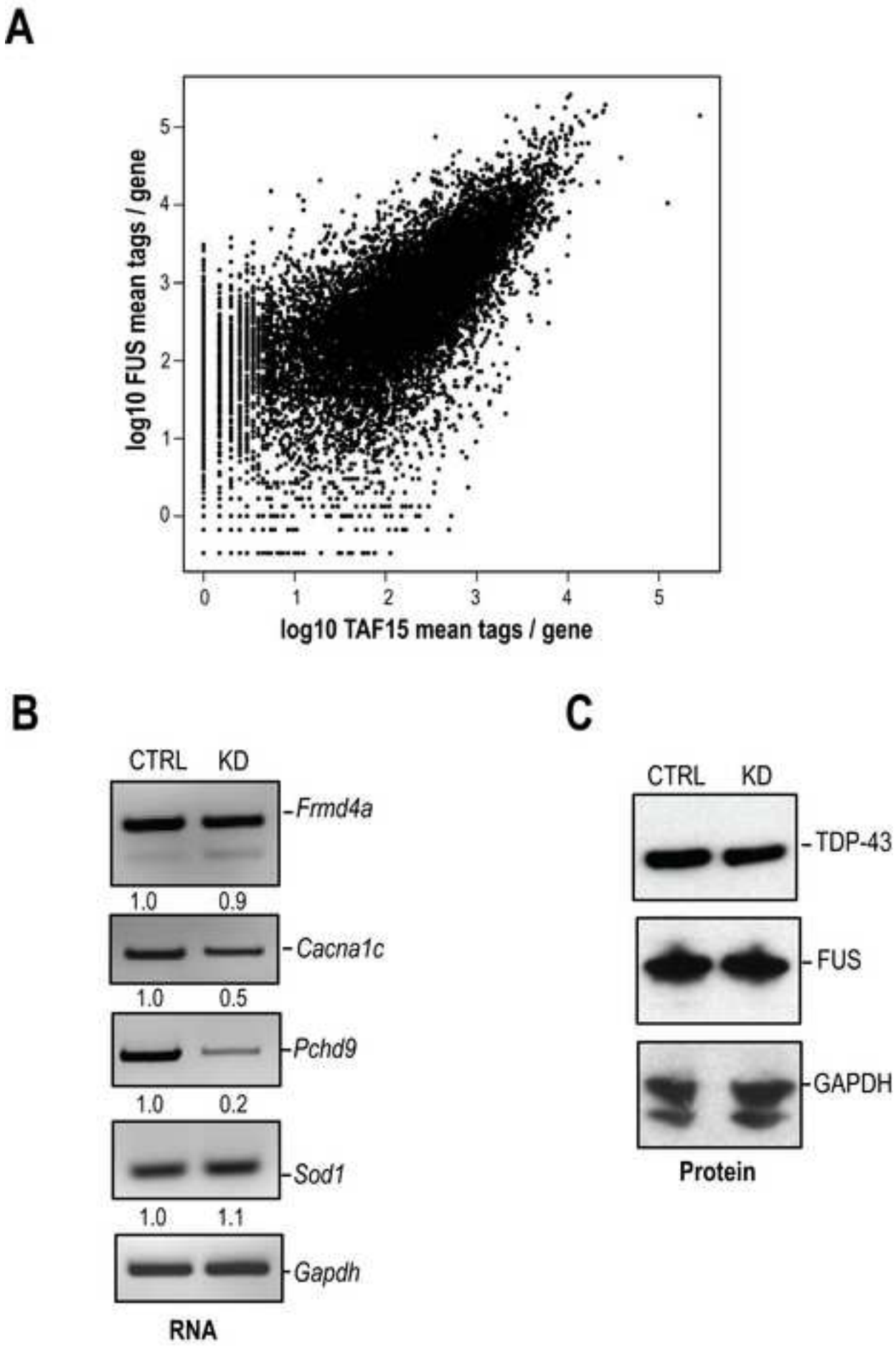
Supplemental Figure 1



Supplemental Figure 2



Supplemental Figure 3



Supplemental Figure 4