Supplementary information.

Figure S1. Brain morphology of *NMF291*$^{-/-}$ mice, Related to Figure 1.

(A-D) Sagittal sections of one-month-old wild type (+/+, A) and *NMF291*$^{-/-}$ (B) cerebella were stained with hematoxylin and eosin. Cerebellar lobules are indicated by Roman numerals (B). Higher magnification images of lobule IX from wild type (C) and *NMF291*$^{-/-}$ (D) cerebella.

(E-H) Sagittal sections of the hippocampus of 4-month-old wild type (E) and *NMF291*$^{-/-}$ (F) mice stained with hematoxylin and eosin. Higher magnification images of the dentate gyrus region from the wild type (G) and *NMF291*$^{-/-}$ (H) hippocampus. Examples of pyknotic nuclei are labeled by arrows (H). Scale bars: (B), 500 μm; (D), 50 μm; (F), 50 μm; (H), 25 μm.

Figure S2. Genetic mapping of the *NMF291* mutation and conservation of *Rnu2-6 – Rnu2-10* snRNAs, Related to Figure 2.

(A) Haplotypes of recombinant chromosomes from affected mice demonstrated that the *NMF291* mutation is located between *D11Mit200* and *D11Mit199*. The number of mice obtained for each haplotype is shown.

(B) The sequences of *Rnu2-6 – Rnu2-10* snRNAs are identical except for a single nucleotide polymorphism (filled circle) at nucleotide 170 in *Rnu2-6* RNA. The *NMF291* 5-nucleotide deletion is boxed and the U2/U6 Helix Ia, II, and III, and U2 branch site recognition sequences (BSRS) are labeled. The region corresponding to the DNA oligonucleotide probe used for northern blot analyses is underlined.

Figure S3. Spatial and temporal expression of the mutant U2 snRNAs and incorporation of mutant RNA into cerebellar U2 snRNPs, Related to Figure 3.

(A) Allele specific real-time PCR primers (upper panel) for detecting total U2 snRNA (U2F & U2R) and mutant U2 snRNA (Mutant U2 ASP & U2R) were validated in P30 wild type (+/+), *NMF291*$^{+/-}$, and *NMF291*$^{-/-}$ cerebellar cDNA (lower panel). To generate relative U2 levels, we first normalized total and mutant U2 expression levels to GAPDH levels. Then we defined the wild type (+/+) normalized total U2 level and *NMF291*$^{-/-}$ normalized mutant U2 level as arbitrary unit one for relative total and mutant U2 expression level, respectively. Note that the relative total U2 snRNA levels in the P30 wild type, *NMF291*$^{+/-}$, and *NMF291*$^{-/-}$ cerebellum are similar, and that the mutant allele specific primers preferentially (~30 fold) amplify mutant U2 in *NMF291*$^{-/-}$ cerebellar cDNA relative to wild type U2 in wild type cerebellar cDNA.

(B) Incorporation of mutant U2 snRNAs into U2 snRNPs in the *NMF291*$^{-/-}$ cerebellum. Immunoprecipitation (IP) experiments were performed using cerebellar supernatants from wild type and *NMF291*$^{-/-}$ mice. RT-PCR (primers; U2F & U2R, upper panel) detected the total U2 snRNAs present in the input supernatant and in the IP products of the Y12 (*Sm* Antigen) antibody, but not in IP products of the control (normal mouse IgG). Real-time PCR using mutant allele-specific primers demonstrated that the mutant U2 RNAs were present in the mutant input supernatant and Y12 IP products, but not in that of wild type (lower panel).

(C-D) Northern blot analysis of wild-type (+/+) and *NMF291*$^{-/-}$ (–/–) total RNA from P30 neuronal and non-neuronal tissues (C) or from the cerebellum of mice at different postnatal (P) ages (D). The U2 DNA oligo probe for northern blot is shown in Figure S2B. Ethidium bromide-stained RNA gels are shown as loading controls, and 5.8s and 5s rRNA, and tRNA are indicated. Cere, cerebellum; Hippo, hippocampus; S.C., spinal cord.

(E) Quantitative analysis of RNAse protection assays performed on *NMF291*$^{+/-}$ and *NMF291*$^{-/-}$ cerebellar RNA at different postnatal (P) ages (days after birth). Values are expressed as the percentage

of mutant/total U2 snRNA. Each genotype at each time point was supported by at least three independent mouse samples. *$p<0.001$, one-way ANOVA.

Figure S4. Granule neuron loss in hippocampal dentate gyrus region of mutant *Rnu2-8* transgenic mice, Related to Figure 4.

(A) Quantitative analysis of RNAse protection assays of Tg-MuU2 mice at different postnatal (P) ages (days after birth). Values are expressed as the percentage of mutant U2 of the total U2 snRNA. Each genotype at each time point was supported by at least three independent mouse samples. *$p<0.001$, one-way ANOVA.

(B-C) Hematoxylin and eosin stained coronal sections of the hippocampus from 3-month-old Tg-MuU2 mice shown at low (B) and higher (C) magnification. Pyknotic nuclei are indicated by arrows (C). Scale bars: (A), 75 μm; (B), 30 μm.

Figure S5. Expression of wild type U2 snRNAs decreases neuron death in aged Tg-268K4; *NMF291*$^{+/-}$ mice, Related to Figure 5.

(A) RNase protection assays were performed on cerebellar RNA from 15-month-old wild type (+/+), *NMF291*$^{+/-}$ (+/–), Tg-268K4; +/+, and Tg-268K4; *NMF291*$^{+/-}$ mice.

(B) Quantitative analysis of RNase protection assays of cerebellar RNA from 15-month-old *NMF291*$^{+/-}$ and Tg-268K4; *NMF291*$^{+/-}$ (Tg-268K4; +/–) mice. Values are expressed as the percentage of mutant U2 of the total U2 snRNA. Each genotype was supported by at least three independent mouse samples. *$p<0.001$, paired t-test.

(C-F) Hematoxylin and eosin stained sagittal cerebellar sections from aged *NMF291*$^{+/-}$ (+/–; C, E) and Tg-268K4; *NMF291*$^{+/-}$ (Tg-268K4; +/–; D, F) mice. Higher magnification of lobule IV-V is shown in E, F. Note the decreased granule cell loss in the aged Tg-268K4; *NMF291*$^{+/-}$ cerebellum relative to that observed in that of *NMF291*$^{+/-}$ mice. Scale bars: (D), 500 μm; (F), 50 μm.

Figure S6. Exon array and RNA-Seq analysis, Related to Figure 7.

(A) RT-PCR analysis of alternative splicing events predicted to vary between wild type and *NMF291*$^{-/-}$ mice by exon array analysis. RT-PCR analysis was performed on cerebellar RNA obtained from four pairs of P30 wild type (+/+) and *NMF291*$^{-/-}$ (–/–) mice, independent of the mice used for the exon array experiment. The identity of RT-PCR products was confirmed by amplicon size and by DNA sequencing. Ten representative reactions from two pairs of wild type and *NMF291*$^{-/-}$ samples are shown.

(B-D) RT-PCR analysis of junctions differentially expressed between wild type and mutant cerebella as detected by RNA-Seq analysis. Splicing events in (B) were detected by RNA-Seq as unique junction reads only found in one genotype (median read count ≥ 3, and ratio of maxRPKM_25 ≤ 2). (C) Junction reads differentially expressed between the two genotypes (ratio of RJI ≥ 2, $p < 0.05$). (D) RT-PCR analysis of intron retention events that were predicted by RNA-Seq to increase in the mutant cerebellum. Note these introns were also retained in the wild-type cerebella as a minor form. RT-PCR analysis was performed on cerebellar RNA obtained from four pairs of P30 wild type (+/+) and *NMF291*$^{-/-}$ (–/–) mice, independent of the mice used for the RNA-Seq experiment. The identity of RT-PCR products was confirmed by amplicon size. Ten representative reactions from two pairs of wild type and *NMF291*$^{-/-}$ samples (B-D) are shown.

(E) cDNA used for RT-PCR validation of exon array and RNA-Seq analysis was free of genomic DNA contamination. GAPDH was amplified (40 PCR cycles) using a primer pair located within a single exon of the mouse GAPDH gene in the plus RT samples (upper panel), but not in the corresponding minus RT controls (lower panel). * indicates PCR primers (lower panel).

(F) U12-dependent splicing events were not affected in *NMF291*$^{-/-}$ cerebella. The splicing of five U12-type introns (Pessa et al., 2006) was examined in cerebellar cDNA generated from four pairs of P30 wild type and *NMF291*$^{-/-}$ mice. The length of *Drap1*, *Ipo4*, *Pex16*, *Psmc4*, and *smE* introns is 85, 117, 85, 182, and 419bps, respectively. Representative gel pictures of reactions from two pairs of wild type and *NMF291*$^{-/-}$ samples are shown.

(G) RNA-Seq analysis workflow. RNA-Seq libraries were prepared using cerebellar mRNA. Three *NMF291*$^{-/-}$ and two wild type (WT) mice were used as biological replicates. Libraries were sequenced on an Illumina GAIIX using a 76nt paired-end read strategy and reads were aligned to mouse genome reference via SpliceMap 3.3.6 allowing for a maximum of 4 mismatches. After bar code removal, the length of a mapped read for one end of the pair was 69bp. 53.48 ± 6.35 million reads per replicate were uniquely mapped to the mouse genome. (a) For junctions only found in one genotype, junctions with a median read number ≥ 3 were further analyzed. The maximal RPKM of 25bp upstream and downstream of the donor and acceptor site of the adjacent exons (maxRPKM_25) was computed to represent the isoform expression level. Junctions in which the ratio of the maxRPKM_25 ≥ 2 between the two genotypes were excluded to avoid potential false positives introduced by differential expression of RNA isoforms between WT and mutant mice. Using these filters, 52 and 1137 junctions were identified as only found in WT and mutant cerebella, respectively. (b) For junctions that were found in both genotypes, we further analyzed those with a median read number ≥ 3 for at least one genotype. Junction counts were normalized to the isoform expression (maxRPKM_25) to generate the Relative Junction Index (RJI), and the log value of the RJI was used for a two-tailed t-tests. 1636 junctions with a RJI ratio ≥ 2 between two genotypes and a p value < 0.05 were identified as differentially expressed between genotypes. (c) For analysis of intron retention, we identified introns with a maximal median of intron RPKM ≥ 2 for at least one genotype. The maximal RPKM of adjacent exons (maxRPKM_exon) was computed to represent expression levels of the transcript isoform, and the Relative Intron Retention Index (RII) was generated by normalizing the intron RPKM to the maxRPKM_exon. Two-tailed t-tests were performed on the log value of RII. 362 and 3978 introns were identified as differentially retained introns in WT and mutant, respectively, with an RII ≥ 2 between the genotypes and a p value < 0.05.

Movie 1.

Progressive ataxia in *NMF291*$^{-/-}$ mice, Related to Figure 1.

Videos of an *NMF291*$^{-/-}$ mouse taken at two-week intervals from 4-14 weeks of age demonstrate the progressive loss of motor control.

Extended Experimental Procedures.

Mice and mouse U2 snRNA genes

The *NMF291* mutant strain was derived from chimeric mice generated from 129S4/SvJae X C57BL/6J F1 ES cells treated with EMS (0.4 μg/ml) and used in a two-generation mating scheme as described (Munroe and Schimenti, 2009). To generate the mutant U2 snRNA transgenic mice, a DNA fragment containing the transcription unit of the mutant *Rnu2-8* and its upstream 780bp and downstream 569bp, was amplified by PCR from the *NMF291* homozygous genomic DNA (Forward, 5'-ctcttgcctagatgggcttg-3'; Reverse, 5'-tctctgcctctatggggttt-3'). The purified DNA fragment was microinjected into the pronuclei of C57BL/6J embryos and founders were identified using PCR. DNA sequencing of amplicons of this primer pair was performed to genotype *NMF291* mice. Genetic mapping of the *NMF291* mutation was performed using an intersubspecific intercross (B6; 129 NMF291 × CAST/Ei). The *NMF291* critical interval contains 5 mouse U2 snRNA genes. GenBank accession numbers for these genes are: *Rnu2-6* (JN863957), *Rnu2-7* (JN863958), *Rnu2-8* (JN863956), *Rnu2-9* (JN863959), and *Rnu2-10* (JN863960).

Histology, Immunohistochemistry and TUNEL Assay

Mice were intracardically perfused with Bouin's fixative or 4% paraformaldehyde for histology and immunohistology, respectively. Brains were postfixed and embedded in paraffin. Cleaved Caspase-3 (Cell Signaling) immunostaining and TUNEL assays (Roche) were performed as described previously (Zhao et al., 2005).

Immunoprecipitation of snRNPs from cerebellar extracts

Frozen mouse cerebella were homogenized with 1ml of reconstitution buffer (20mM Hepes-KOH pH 7.9, 50mM KCl, 5mM $MgCl_2$, 0.2mM EDTA, 5% glycerol) containing 0.01% NP40 as previously described (Gabanella et al., 2007). Homogenates were then passed through a 25G needle five times and

centrifuged 15 min at 10,000 rpm at 4ºC. Supernatants were collected for IP experiments and protein concentration was measured. Five µg of either the Sm Antigen (Y12) antibody (Lab Vision) or mouse normal IgG (Santa Cruz Biotechnology) conjugated to Dynabeads Protein G (Invitrogen) were incubated with 400µg of cerebellar extracts added to RSB-500 buffer (500mM NaCl, 10mM Tris-HCl pH 7.4, 2.5mM $MgCl_2$) containing 0.1% NP40, protease inhibitor cocktail (Roche), and phosphatase inhibitors (50mM NaF, and 0.2mM Na_3VO_4) at 4ºC for 8 hours. After five washes with the same buffer, bound RNAs were recovered from IP products by proteinase K treatment, phenol/chloroform extraction, and ethanol precipitation. Detection of total and mutant U2 snRNAs was performed by real-time PCR as described in Figure S3A.

RNase Protection, Northern Blot and RT-PCR Assays

Total RNA was extracted from mouse tissues and HEK293 cells using Trizol (Invitrogen) and treated with DNase I (Ambion). RT-PCR was performed on random-primed cDNA using primer sets described in Tables S2A, S2B and S2C. For the RPA probe, a DNA fragment covering 7-180 position (174nts) of the *Rnu2-8* transcription unit (191nts) was cloned into pCR2.1-TOPO (Invitrogen). The ^{32}P-radiolabeled RNA probe was generated by *in vitro* transcription of the linearized plasmid and gel purified. RPA analysis was performed as suggested by the manufacturer's instructions (Ambion). Briefly, total RNA was hybridized with radiolabeled RNA probe at 42˚C overnight. Non-hybridized RNA was digested by RNase A/T1, and the resulting protected RNA fragments were separated on a 10% denaturing polyacrylamide gel. Density of bands in the scanned X-ray films was determined by ImageQuant 5.2 software (Amersham). Statistical significance was determined by ANOVA and paired t- test (SPSS). For northern blots, total RNA was also separated on a 10% denaturing polyacrylamide gel. The membrane was hybridized with ^{32}P-end-labeled DNA oligonucleotides corresponding to U2 snRNA (5'-tatcagatattaaactgataagaacagatact-3') or 5s rRNA (5'- cctgcttagcttccgagatca-3').

Constructs and Mutagenesis

For expression of wild type and mutant U2 snRNA in the HEK293T cells, the mutant *Rnu2-8* DNA fragment used to generate transgenic mice and the corresponding wild type fragment were cloned into pCR2.1-TOPO. The Δ3, Δ6, and Δ9 U2 snRNA expression constructs were generated from the wild-type U2 construct by site-directed mutagenesis (Stratagene). The ISS- (pTN23) reporter was generated from the ISS+ (pTN24) plasmid (Nasim et al., 2002) by deleting the 25bp ISS using the same mutagenesis method. The L1CAM exon27-intron27-exon28 and exon28-intron28-exon29 cassettes were amplified from C57BL/6J genomic DNA and inserted in frame into pBPLUGA at the *Sal*I and *Bam*HI sites (Kollmus et al., 1996). Primers used for construct generation are described in the Table S2B.

Cell culture, Transfection and Luciferase Assays

HEK293T cells grown in DMEM medium containing 10% FBS, were used for transient transfection studies. Briefly, 3×10^5 cells/well were plated in a 12-well plate 24 hours before transfection with Lipofectamine 2000 (Invitrogen). 48 hours after transfection, cells were harvested in Passive Lysis Buffer (Promega). Luciferase and ß-galactosidase activities were measured as per manufacturer's instructions (Promega) in a multilabel counter. Statistical significance was determined by ANOVA analysis (SPSS).

Functional GO Analysis

The gene lists of differentially expressed genes and exons identified by exon array, differentially expressed junctions found in one or both genotypes and retained introns identified by RNA-Seq, were uploaded into DAVID bioinformatics resources 6.7 (http://david.abcc.ncifcrf.gov/) (Huang da et al., 2009). The functional annotation chart and clustering analysis modules were employed for gene-term enrichment analysis.

Exon Array

Total RNA was isolated by TRIzolPlus (Invitrogen) and cDNA was synthesized, labeled and fragmented using the Affymetrix GeneChip WT Terminal Labeling kit. GeneChip® Mouse Exon 1.0 ST Arrays (Affymetrix) were hybridized with 2.3µg of biotin-labeled cDNA for 16 hours at 45°C. Post-hybridization staining and washing were performed according to manufacturer's protocols using the Fluidics Station 450 instrument (Affymetrix). Arrays were scanned with a GeneChipTM Scanner 3000 (Affymetrix) laser confocal slide scanner.

Affymetrix Expression Console Software was used for array quality assessment. Exon array data treatment was done by the EASANA® analysis system and interface visualization (GenoSplice Technology, www.genosplice.com), which is based on the FAST DB® annotations (de la Grange et al., 2007; de la Grange et al., 2005). Data were normalized by using quantile normalization method. Background correction was made using the anti-genomic probes. Probe selection was made as previously described (Clark et al., 2007; de la Grange et al., 2010). Briefly, only probes targeting exons annotated from full-length cDNA were kept for analysis. Among these pre-selected probes, bad quality probes (e.g. probes labeled as "cross-hybridizing" by Affymetrix) and probes with intensity signal too low compared to background probes with the same GC content were removed from analysis. Probes with a detection above background (DABG) p-value ≤ 0.05 in at least half of chips were considered for further statistical analysis as previously described (Clark et al., 2007; de la Grange et al., 2010). Exon intensities were normalized to the overall intensity of the corresponding gene for each replicate to obtain the normalized exon intensities (NI). The NI of each exon was then compared to that of the other genotype to give the fold change or splicing index. Paired statistical analyses were performed on gene signal intensities and on the splicing index using a Student's paired t-test. Results for both gene and exon levels were considered statistically significant at p-values < 0.05 and fold-change ≥ 1.5. Note that statistical corrections for sample size were not used to compute p-values, thus these values were only

used for relative ranking purposes. Primers used for exon array validation are described in the Table S2C. The GEO (Gene Expression Omnibus) accession number for our exon array data is GSE33069.

RNA-Seq Methods

Purification of mRNA, double-stranded cDNA preparation and ligation of barcode-containing adaptors were performed using the mRNA-Seq Sample Prep kit according to the manufacturer's protocol (Illumina). Size selection (300bp) was performed by agarose gel electrophoresis, and PCR was performed to enrich the adapter-modified DNA fragments. Validation of DNA fragment size was performed using an Agilent Technologies 2100 Bioanalyzer, and DNA concentration was determined by quantitative PCR, following the manufacturer's protocol (Kapa Biosystems). The DNA library was sequenced on an Illumina GAIIX using a 76nt paired-end read strategy following the manufacturer's protocol. The sequences were initially mapped to the mouse genome using CASAVA (Illumina) for assessment of sequence quality. The barcoded reads were separated allowing for 2 mismatches in the barcode tag, and analyzed for their base composition. Any bases at the start of the read deviating from an approximately uniform distribution were removed as they might represent adaptor sequence contamination. After pre-processing, one base was removed from the beginning of all the reads.

For de novo identification of junctions, the pre-processed reads were inputted to SpliceMap 3.3.6 (Au et al., 2010) and aligned to the mouse reference genome (mm9) (Figure S6G). For junctions found only in one genotype, the junctions where the median read number within a genotype was < 3 were filtered out. SpliceMap is not based on previous annotations and thus entire exons adjacent to junction reads were not readily defined. Furthermore the length of ~93.1% of exons in mouse genome is greater than 50bps (Sakharkar et al., 2005). Therefore, we recovered the RPKM (Mortazavi et al., 2008) value of 25bp, which is the minimal exon size SpliceMap 3.3.6 can utilize, upstream and downstream of the donor and acceptor sites of the junction. The larger RPKM value of the two sides was recorded to generate the

maxRPKM_25 to represent the isoform expression level. Any junctions in which the ratio of the maxRPKM_25 between the two genotypes was ≥ 2 were also removed. For junctions that appeared in both genotypes, we normalized the junction counts to the isoform expression (maxRPKM_25) to generate the Relative Junction Index (RJI). The log value of RJI was used for two-tailed t-tests. Note that statistical corrections for sample size were not used to compute p-values for RNA-Seq data, thus these values were only used for relative ranking purposes. To identify differentially expressed junctions with high reliability, we only included the junctions if the median count of junctions within the genotype with higher RJI value was ≥ 3, the ratio of RJI between two genotypes was ≥ 2, and the p value was < 0.05. For intron retention, we only considered reads for further analysis if the maximal median of intron RPKM within one genotype ≥ 2. The maximal RPKM of adjacent exons located to the 5' and 3' of the intron (maxRPKM_exon) was recovered for representing the isoform expression. The intron RPKM was normalized to the maxRPKM_exon to generate the Relative Intron retention Index (RII) and the log value of RII was used for two-tailed t-tests. To recover retained introns with high reliability, we only included the introns that the ratio of RII between two genotypes was ≥ 2 and the p value was < 0.05.

To associate junctions with genes, junctions were assigned to a given gene if the nearest exon to the donor and acceptor sequence belonged to the same gene in RefSeq annotations, or if the junction matched uniquely to a RefSeq annotated junction. To extract intron length information from RefSeq, annotations were downloaded from UCSC and the length between all annotated exons was recorded. Primers used for RNA-Seq validation are described in Table S2C.

Analysis of exon array and RNA-Seq data overlap

To analyze the overlap between exon array and RNA-Seq data, we recorded the exon array probe position, which is annotated by the EASANA visualization interface, and localized it in our RNA-Seq data set. For intron retention events detected by exon array, the fold change of the corresponding events

from RNA-Seq data was calculated by the ratio of Relative Intron retention Index (RII). For other alternative splicing events detected by exon array, the RNA-Seq junction reads spanning the array probe were recorded if the reads appeared in at least 2 replicates and the ratio of Relative Junction Index (RJI) was converted into the fold change. If multiple junctions spanned the exon probe, we did not include the event into the comparison analysis. Only 17 non-intron retention events identified by exon array passed these criteria, and our RNA-seq data set indicated 15/17 with the same up-/down regulation direction as detected by exon array. However, these events failed to meet the stringent thresholds set for differential expressed junctions.

Supplementary References:

Au, K.F., Jiang, H., Lin, L., Xing, Y., and Wong, W.H. (2010). Detection of splice junctions from paired-end RNA-seq data by SpliceMap. Nucleic acids research *38*, 4570-4578.

Clark, T.A., Schweitzer, A.C., Chen, T.X., Staples, M.K., Lu, G., Wang, H., Williams, A., and Blume, J.E. (2007). Discovery of tissue-specific exons using comprehensive human exon microarrays. Genome Biol *8*, R64.

de la Grange, P., Dutertre, M., Correa, M., and Auboeuf, D. (2007). A new advance in alternative splicing databases: from catalogue to detailed analysis of regulation of expression and function of human alternative splicing variants. BMC Bioinformatics *8*, 180.

de la Grange, P., Dutertre, M., Martin, N., and Auboeuf, D. (2005). FAST DB: a website resource for the study of the expression regulation of human gene products. Nucleic Acids Res *33*, 4276-4284.

de la Grange, P., Gratadou, L., Delord, M., Dutertre, M., and Auboeuf, D. (2010). Splicing factor and exon profiling across human tissues. Nucleic Acids Res *38*, 2825-2838.

Gabanella, F., Butchbach, M.E., Saieva, L., Carissimi, C., Burghes, A.H., and Pellizzoni, L. (2007). Ribonucleoprotein assembly defects correlate with spinal muscular atrophy severity and preferentially affect a subset of spliceosomal snRNPs. PLoS One *2*, e921.

Huang da, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc *4*, 44-57.

Kollmus, H., Flohe, L., and McCarthy, J.E. (1996). Analysis of eukaryotic mRNA structures directing cotranslational incorporation of selenocysteine. Nucleic Acids Res *24*, 1195-1201.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods *5*, 621-628.

Munroe, R., and Schimenti, J. (2009). Mutagenesis of mouse embryonic stem cells with ethylmethanesulfonate. Methods Mol Biol *530*, 131-138.

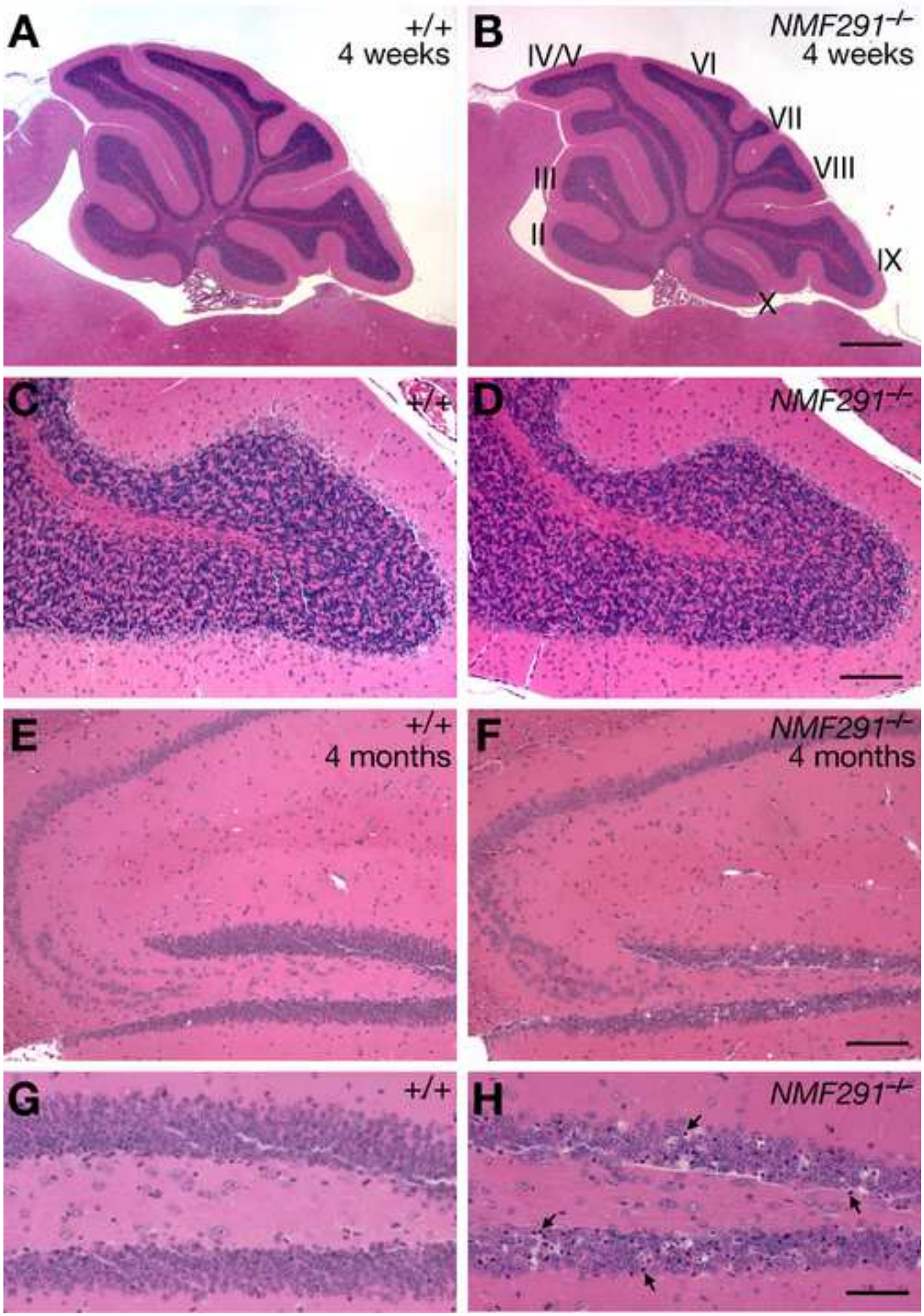
Nasim, M.T., Chowdhury, H.M., and Eperon, I.C. (2002). A double reporter assay for detecting changes in the ratio of spliced and unspliced mRNA in mammalian cells. Nucleic Acids Res *30*, e109.

Pessa, H.K., Ruokolainen, A., and Frilander, M.J. (2006). The abundance of the spliceosomal snRNPs is not limiting the splicing of U12-type introns. Rna *12*, 1883-1892.

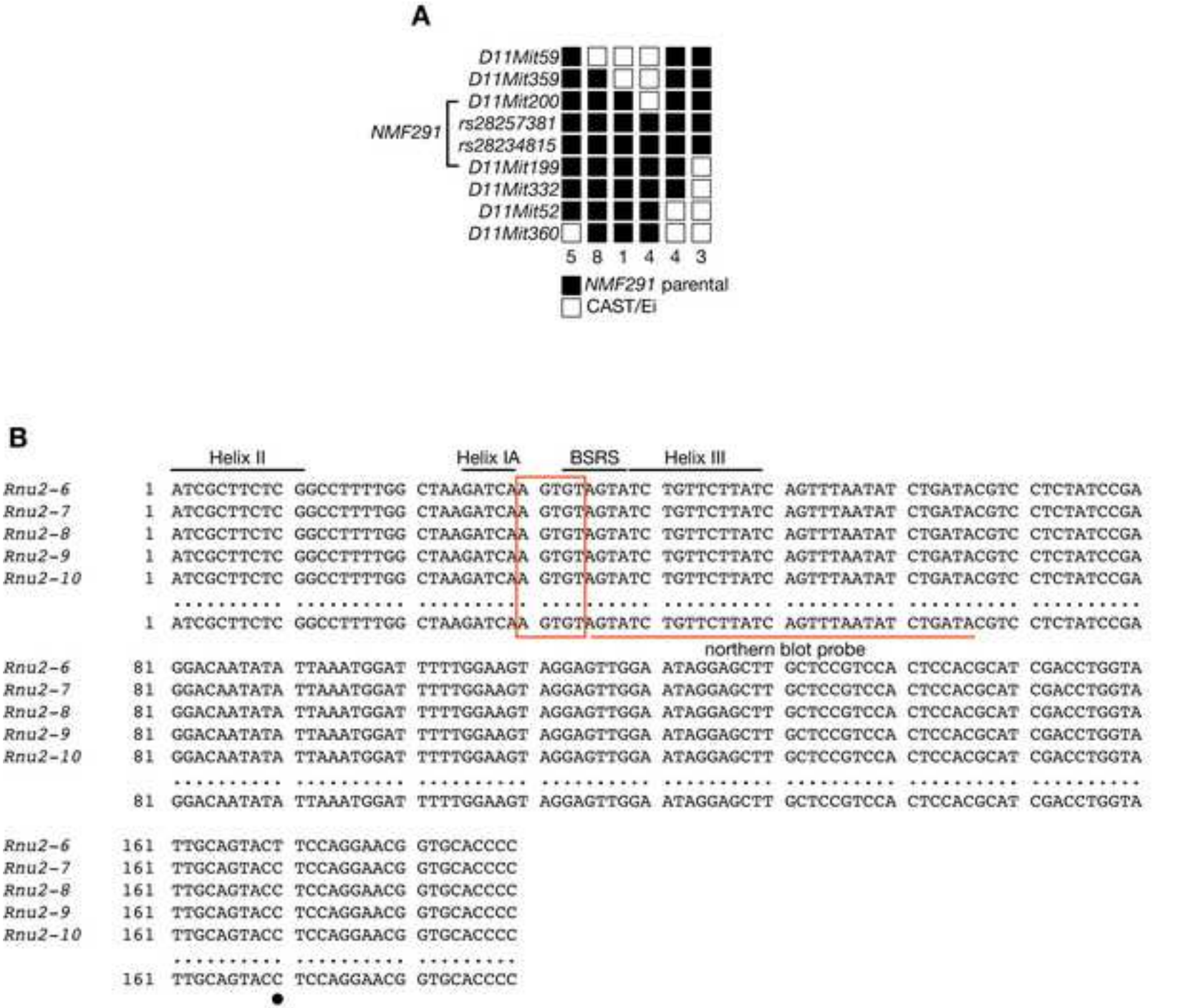
Sakharkar, M.K., Perumal, B.S., Sakharkar, K.R., and Kangueane, P. (2005). An analysis on gene architecture in human and mouse genomes. In Silico Biol *5*, 347-365.

Zhao, L., Longo-Guess, C., Harris, B.S., Lee, J.W., and Ackerman, S.L. (2005). Protein accumulation and neurodegeneration in the woozy mutant mouse is caused by disruption of SIL1, a cochaperone of BiP. Nat Genet *37*, 974-979.

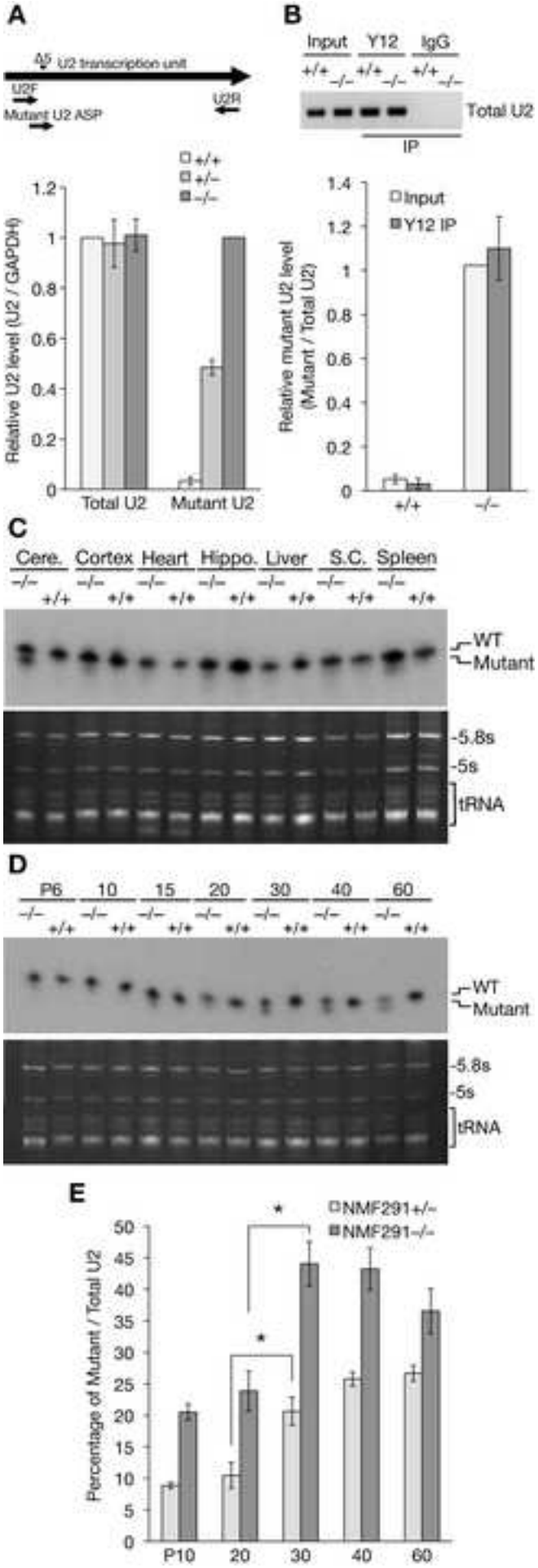
Supplemental Figure 1



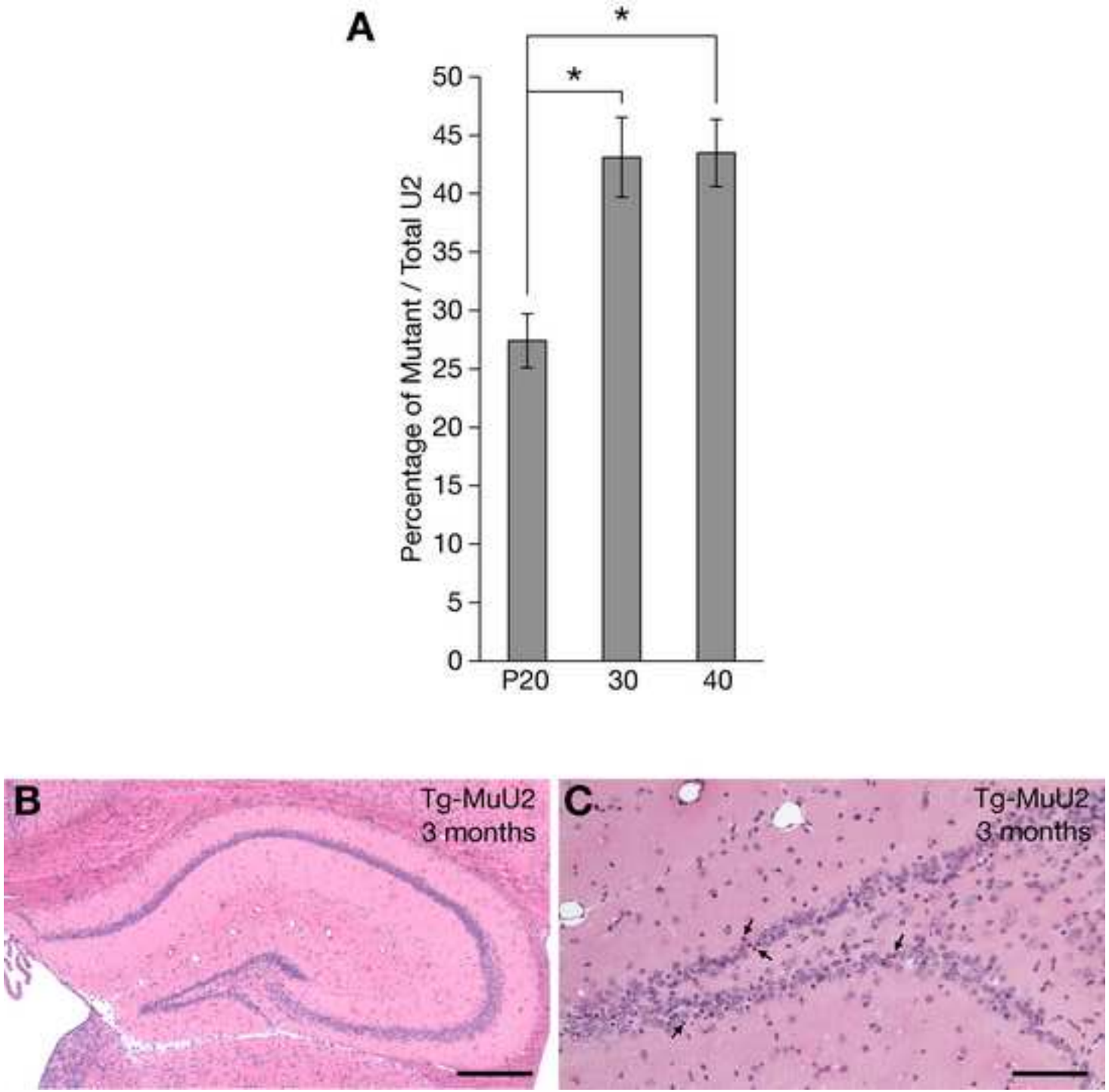
Supplemental Figure 2



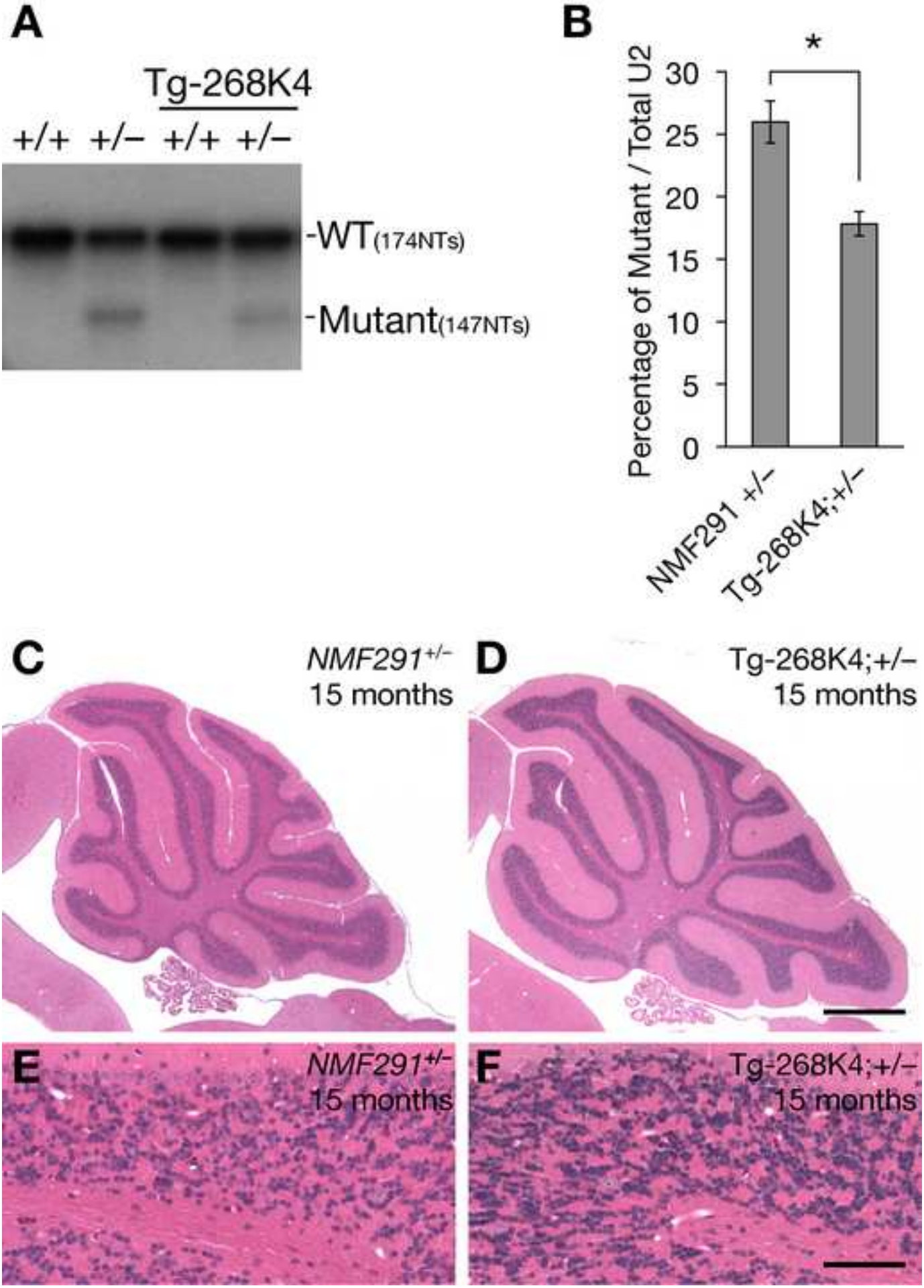
Supplemental Figure 3



Supplemental Figure 4



Supplemental Figure 5



Supplemental Figure 6