Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs

Clotilde Lagier-Tourenne[1,2,9], Magdalini Polymenidou[1,2,9], Kasey R Hutt[2,3,9], Anthony Q Vu[2,3], Michael Baughn[4], Stephanie C Huelga[2,3], Kevin M Clutario[1,2], Shuo-Chien Ling[1,2], Tiffany Y Liang[2,3], Curt Mazur[5], Edward Wancewicz[5], Aneeza S Kim[5], Andy Watt[5], Sue Freier[5], Geoffrey G Hicks[6], John Paul Donohue[7], Lily Shiue[7], C Frank Bennett[5], John Ravits[4], Don W Cleveland[1,2] & Gene W Yeo[2,3,8]

FUS/TLS (fused in sarcoma/translocated in liposarcoma) and TDP-43 are integrally involved in amyotrophic lateral sclerosis (ALS) and frontotemporal dementia. We found that FUS/TLS binds to RNAs from >5,500 genes in mouse and human brain, primarily through a GUGGU-binding motif. We identified a sawtooth-like binding pattern, consistent with co-transcriptional deposition of FUS/TLS. Depletion of FUS/TLS from the adult nervous system altered the levels or splicing of >950 mRNAs, most of which are distinct from RNAs dependent on TDP-43. Abundance of only 45 RNAs was reduced after depletion of either TDP-43 or FUS/TLS from mouse brain, but among these were mRNAs that were transcribed from genes with exceptionally long introns and that encode proteins that are essential for neuronal integrity. Expression levels of a subset of these were lowered after TDP-43 or FUS/TLS depletion in stem cell–derived human neurons and in TDP-43 aggregate–containing motor neurons in sporadic ALS, supporting a common loss-of-function pathway as one component underlying motor neuron death from misregulation of TDP-43 or FUS/TLS.

ALS is an adult-onset neurodegenerative disorder that is characterized by premature degeneration of motor neurons, resulting in a progressive, fatal paralysis. Approximately 10% of disease instances are familial, with dominant mutations identified in several genes (reviewed in ref. 1). A seminal discovery was the finding that the primarily nuclear RNA- and DNA-binding protein TDP-43 is mislocalized in the cytoplasm of affected cells in essentially all sporadic ALS patients, where it forms insoluble, ubiquitinated inclusions[2,3]. TDP-43 pathology is also a prominent feature of frontotemporal lobar degeneration (FTLD), which is now recognized as being genetically, clinically and pathologically linked to ALS (reviewed in ref. 4).

The discovery that dominant point mutations in TDP-43 cause some instances of inherited ALS[5–7] and FTLD was rapidly followed by the identification of mutations in another RNA- and DNA-binding protein, FUS/TLS, that were responsible for an additional portion of inherited ALS[8,9] and rare cases of FTLD. Cytoplasmic accumulation of FUS/TLS in the nervous system of patients with FUS/TLS mutations[8,9], and in some instances of FTLD in the absence of mutations[10–12], disrupts its normal nuclear localization[13]. Whether the loss of FUS/TLS nuclear function or an acquired toxicity contributes to neuronal dysfunction in disease is unknown.

Both TDP-43 and FUS/TLS have been associated with multiple nuclear and cytoplasmic steps of RNA processing (reviewed in ref. 14). Comprehensive maps of TDP-43–binding sites in RNAs of normal mouse[15] or diseased human[16] CNS have been determined. Reducing TDP-43 expression in the adult nervous system alters pre-mRNA splicing of >900 genes, revealing an essential role for TDP-43 in sustaining the levels of RNAs derived from very long intron–containing pre-mRNAs that are important for neuronal function[15], some of which are reduced in human disease[16].

Using similar high-throughput sequencing and computational approaches, we identified ~5,500 RNA targets of FUS/TLS in both mouse and human brain. The expression levels of 610 genes were dependent on FUS/TLS, and only 112 of the genes were also dependent on TDP-43. Systematic comparison between the TDP-43– and FUS/TLS-dependent targets identified RNAs with exceptionally long introns and multiple binding sites for both TDP-43 and FUS/TLS. Depletion of FUS/TLS or TDP-43 in human neurons differentiated from pluripotent stem cells confirmed the downregulation of long intron–containing TDP-43 and FUS/TLS targets. Several of these common targets were substantially reduced at the protein level in affected post-mortem neurons of patients with sporadic ALS, thereby identifying a common pathogenic pathway in motor neurons from the misregulation of either TDP-43 or FUS/TLS.
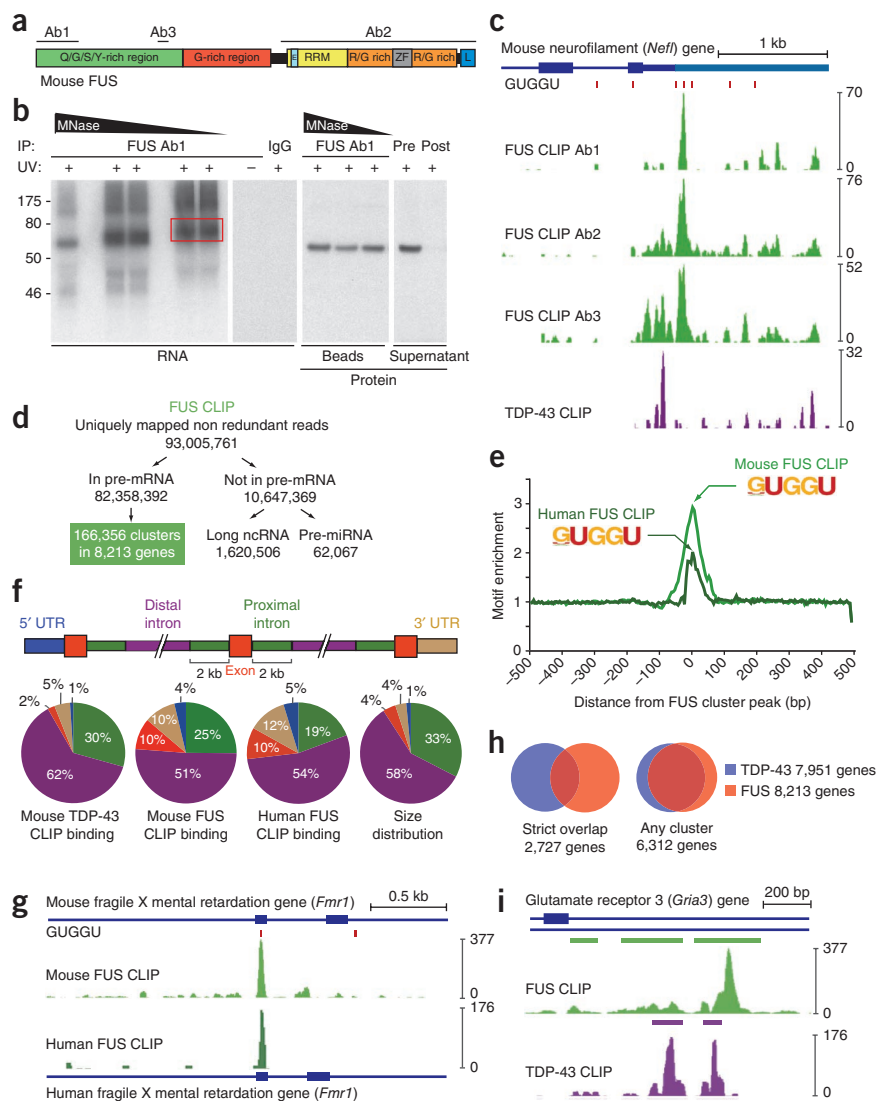
RESULTS
RNA targets of FUS/TLS in mouse and human brain
FUS/TLS protein–RNA complexes cross-linked *in vivo* by ultraviolet light were efficiently immunoprecipitated using antibodies to three

Figure 1 FUS/TLS RNA targets in mouse and human brain. (**a**) FUS/TLS protein domains used as antigens to generate antibodies Ab1, Ab2 and Ab3. Q/G/S/Y, glutamine, glycine, serine, tyrosine; G, glycine; E, nuclear export signal; RRM, RNA recognition motif; R/G, arginine/glycine; ZF, zinc finger; L, nuclear localization signal. (**b**) Autoradiograph of FUS/TLS protein–RNA complexes from mouse brain immunoprecipitated with Ab1 and trimmed with increasing concentrations of micrococcal nuclease (MNase) (first panel). Complexes highlighted by the red box were used for sequencing. Beads coated with IgG antibodies did not detect protein-RNA complexes (second panel). Immunoprecipitated FUS/TLS-RNA complexes migrated at the expected FUS/TLS mobility (third panel), and no FUS/TLS remained after immunoprecipitation (fourth panel). (**c**) FUS/TLS (green) and TDP-43 (purple) binding to low molecular-weight neurofilament subunit (*Nefl*) RNA. Vertical red lines show the positions of GUGGU motifs. The scale bar represents the read coverage per base. (**d**) Flow chart illustrating reads analyzed from three CLIP-seq experiments to define FUS/TLS clusters. (**e**) Positional distribution of the GUGGU motif in FUS/TLS CLIP clusters in human and mouse brain. (**f**) Percentages of TDP-43 and FUS/TLS CLIP clusters in pre-mRNAs regions as defined in the top panel. (**g**) FUS/TLS binding in human and mouse brain in orthologous exon 5 of the *Fmr1* RNA. (**h**) Venn diagrams showing genes with TDP-43 and FUS/TLS CLIP clusters overlapping by at least one nucleotide (left) or genes with both TDP-43– and FUS/TLS-binding sites (right). (**i**) Overlapping TDP-43 and FUS/TLS clusters in neighboring, but distinct, intronic positions in the *Gria3* pre-mRNA.

distinct regions of FUS/TLS (**Fig. 1a,b** and **Supplementary Fig. 1a,b**), resulting in a substantial reduction of FUS/TLS from mouse brain lysates (**Fig. 1b** and **Supplementary Fig. 1c**). Each antibody precipitated complexes with the expected mobility for a single FUS/TLS molecule bound to RNA, as well as some lower mobility complexes (**Supplementary Fig. 1a,b**). No complexes were immunoprecipitated when using nonspecific IgG or in the absence of ultraviolet cross-linking (**Fig. 1b**). After size selection, sequencing of RNAs immunoprecipitated with each antibody produced 12,344,760, 32,427,431 and 37,586,201 nonredundant reads that mapped to pre-mRNAs encoded in the mouse genome (University of California, Santa Cruz, reference mm9; **Supplementary Fig. 1d**). Previous TDP-43 CLIP-seq (cross-linking and immunoprecipitation, followed by high-throughput sequencing) libraries[15] were re-sequenced and reads were mapped onto the mm9 genome (**Supplementary Fig. 1d**).

Positions of mapped reads revealed similar binding sites for RNAs precipitated by all three FUS/TLS antibodies, including, for example, in the mouse neurofilament (*Nefl*) pre-mRNA (**Fig. 1c**). Binding sites for FUS/TLS were quite distinct from those obtained for TDP-43 (**Fig. 1c**). Genome-wide consistency of identified FUS/TLS-binding sites (**Supplementary Fig. 1e**) prompted the merging of the data sets from the three antibodies. Most (88.6%) of the >93 million combined nonredundant mapped reads represented the sense strand of the pre-mRNA in known genes (**Fig. 1d**). The majority of the remainder

mapped to annotated long noncoding RNAs (lncRNAs; **Fig. 1d** and **Supplementary Fig. 1f**). Using a cluster-finding algorithm with gene-specific thresholds that account for pre-mRNA length and variable expression levels[15,17], we identified 166,356 FUS/TLS-binding sites (with minimally 28 reads each and an average of 92 reads) in 8,213 mouse RNAs (**Fig. 1d**).

FUS/TLS targets were also identified by CLIP-seq from human brain cortex using FUS/TLS antibody 1 (Ab1), which efficiently immunoprecipitated human FUS/TLS (**Supplementary Fig. 2a,b**). The human FUS/TLS protein–RNA complexes showed similar migration patterns (**Supplementary Fig. 2a**) with those obtained from mouse brain (**Supplementary Fig. 1a**). After mapping to the human genome (University of California, Santa Cruz, reference hg18), 44,490 FUS/TLS-binding sites in pre-mRNAs from 5,594 human genes were identified from 9,373,763 nonredundant mapped reads (**Supplementary Fig. 2c**). Consistent with our findings in mouse, only a small fraction of reads (16%, 1,784,317 reads) mapped outside of annotated protein coding genes (**Supplementary Fig. 2c**). Notably, the majority of the FUS/TLS RNA targets in mouse brain (69%) were also targets in human brain, representing a highly significant overlap ($P < 10^{-37}$), illustrating that the FUS/TLS protein–RNA interactomes are conserved in both human and mouse.

Using the HOMER algorithm, we identified a significantly enriched primary RNA sequence motif GUGGU ($P < 10^{-3451}$; similar to the GGUG motif reported from *in vitro* RNA binding[18]) that was present in ~60% of mouse and human FUS/TLS clusters (**Fig. 1e**). Although this motif often coincided with the peak of FUS/TLS binding (**Fig. 1e**), its presence was neither necessary nor sufficient for FUS/TLS binding (**Fig. 1c**). The distribution of mouse FUS/TLS-binding sites along different pre-mRNA regions was visibly distinct from that of TDP-43 sites (**Fig. 1c,f**). FUS/TLS binding was also observed on small nuclear (snRNA) transcripts (**Supplementary Fig. 3a**). In particular, FUS/TLS CLIP-seq reads mapped to U1 snRNA harbored a high degree of sequence polymorphisms that reflected cross-linking at GUGG sequences defining the position of FUS/TLS binding (**Supplementary Fig. 3b,c**).

Mouse and human brain FUS/TLS sites showed similar distribution patterns in pre-mRNA transcripts, with the majority being localized in introns. The distal intronic region, defined as the sequences located at least 2 kb from the nearest exon-intron junction, accounted for the highest fraction of FUS/TLS-binding sites in accordance with the relative size of distal intronic regions in pre-mRNA (**Fig. 1f**). Notably, there was also a substantial enrichment for binding in exonic and untranslated regions (UTRs) compared to what would be expected from their contribution to pre-mRNA sequences (**Fig. 1f**). One example of exonic binding is the conserved interaction in mouse and human brain of FUS/TLS with exon 4 of the *FMR1* gene, which is mutated in fragile X mental retardation and adult onset ataxia/tremor syndrome (**Fig. 1g**).

FUS/TLS binding was found in the 3′ UTR of 4,272 mouse brain mRNAs, and gene ontology analysis identified enrichment for components of the synapse and molecules residing in neuronal projections (corrected *P* values: $P = 2.1 \times 10^{-5}$ or $P = 7.2 \times 10^{-7}$, respectively). This binding preference suggests that FUS/TLS is involved in the transport and localized translation of these RNAs, consistent with the role of FUS/TLS in dendritic spine formation[19]. We identified FUS/TLS binding to the 3′ UTR of several transcripts involved in the pathogenesis of ALS and other neurodegenerative disorders, including the genes encoding superoxide dismutase 1 (*SOD1*), the light (**Fig. 1c**), medium and heavy chains of neurofilament (*NEFL*, *NEFM* and *NEFH*), the glutamate transporter EAAT2 (*EAAT2*), ubiquilin 1 and 2 (*UBQLN1* and *UBQLN2*), and the microtubule-associated protein tau (*MAPT*) (**Supplementary Table 1**).

In the mouse brain, 6,312 genes contain both TDP-43 (ref. 15) and FUS/TLS clusters, with binding sites overlapping by at least one nucleotide in only 2,727 genes (**Fig. 1h**), consistent with a small fraction of TDP-43 and FUS/TLS proteins being present in the same complex[20]. In some instances, these overlapping or neighboring TDP-43–FUS/TLS clusters were characterized by distinct, but closely positioned, peaks of sequence reads, as exemplified by binding on the glutamate receptor 3 (*Gria3*) gene (**Fig. 1i**). In other cases, TDP-43– and FUS/TLS-binding sites were perfectly superimposed, with genome-wide analysis revealing a 2.2-fold enrichment of TDP-43 binding immediately surrounding (±50 bases) the center of FUS/TLS clusters and a 1.8-fold enrichment vice versa (**Supplementary Fig. 4a,b**). In contrast, FUS/TLS and TDP-43 CLIP reads were not enriched around the sites of an unrelated RNA-binding protein, Nova[21] (**Supplementary Fig. 4c**). A stronger enrichment for localization in 3′ UTRs was observed for overlapping TDP-43– and FUS/TLS-binding sites compared with either TDP-43 or FUS/TLS alone (**Supplementary Fig. 4d**). RNA targets with overlapping TDP-43– and FUS/TLS-binding sites include many genes that are crucial for proper neuronal function, including neurexin 1 and 2 (*NRXN1* and *NRXN2*), *UBQLN2*, *FUS/TLS*, amyloid beta precursor protein (*APP*) and ataxin 10 (*ATXN10*).

FUS/TLS and TDP-43 have distinct binding patterns

The FUS/TLS-binding mode across its target pre-mRNAs was markedly different from that of TDP-43. In contrast with the more discrete 'sharp' binding peaks[15,16] of TDP-43, FUS/TLS sites extended along broad regions of its pre-mRNA targets. This was exemplified by FUS/TLS binding on its own pre-mRNA in a region that was previously annotated as either an intron or an alternative 3′ UTR (**Fig. 2a**). This region has an abundance of RNA-seq reads and a high level of mammalian sequence conservation. We identified the presence of a FUS/TLS isoform that retained intron 7, but not intron 8, in both mouse and human brain, using primers specific for intron 7 and exon 9 (**Supplementary Fig. 5a**). This isoform is likely subjected to degradation by nonsense-mediated decay. FUS/TLS interacted with the entirety of this alternative 3′ UTR or retained intronic region, with a high density of CLIP-seq reads at the junction between exon 7 and intron 7 of *FUS/TLS* pre-mRNA (**Fig. 2a** and **Supplementary Fig. 5a**). Notably, TDP-43 also bound the same region, but at sites distinct from FUS/TLS. Binding of FUS/TLS to its own pre-mRNA, in conjunction with the production of an isoform that is a nonsense-mediated decay substrate, most likely reflects an autoregulatory mechanism similar to that of TDP-43 (ref. 15). We also observed a similar FUS/TLS binding pattern on a highly conserved alternative 3′ UTR or retained intronic region of the Ewing sarcoma breakpoint region 1 gene (*Ewsr1*) pre-mRNA (**Supplementary Fig. 5b**). This suggests potential cross-regulation between Ewsr1 and FUS/TLS, both closely related proteins, as has been observed for other heterogeneous nuclear ribonucleoparticle protein members[17].

To investigate whether FUS/TLS interacts with other alternative 3′ UTRs or retained introns, we first determined the distribution of wild-type mouse brain RNA-seq read density in annotated 3′ UTR exons. Reasoning that intronic regions containing RNA-seq reads above the median value of 3′ UTR read density are likely retained in mature mRNA, we identified 1,866 previously unannotated candidate 3′ UTR or retained intronic regions (**Supplementary Fig. 5c**). Consistent with our observations for *FUS/TLS* and *EWSR1*, more than 40% of these candidate regions had FUS/TLS CLIP clusters, exhibiting about fourfold more binding sites per region than expected (**Supplementary Fig. 5d**).

The most distinguishing feature of FUS/TLS binding was observed on genes with long introns (>100 kb), with FUS/TLS cluster density being substantially higher at the beginning of long (**Fig. 2b–d**), but not short (**Supplementary Fig. 6a**), introns, and a gradual decrease toward the 3′ end, producing a sawtooth-like pattern across the pre-mRNA. This pattern was evident in both mouse and human FUS/TLS CLIP-seq data, as illustrated by the neuroligin 1 (*Nlgn1*) gene (**Fig. 2b**), and contrasted sharply with the CLIP-seq patterns of other RNA-binding proteins, such as TDP-43 and RBFOX2 in mouse brain (**Fig. 2d**). A sawtooth-like binding pattern, which has been proposed to mark nascent transcripts[22], was observed in long intron–containing genes in our RNA-seq data from mouse brain (**Supplementary Fig. 6b**), consistent with FUS/TLS deposition along long introns during transcription elongation (**Fig. 2c**). The FUS/TLS-binding enrichment at the beginning of introns is likely a result of the higher prevalence of nascent transcripts containing this intronic region (**Supplementary Fig. 6b**) and not of an increased density of the FUS/TLS-binding motif at the 5′ end of introns (**Fig. 2e**).

We also identified 843 FUS/TLS-binding sites on 71 of 234 literature-curated lncRNAs. Two-thirds (66%) of these contained
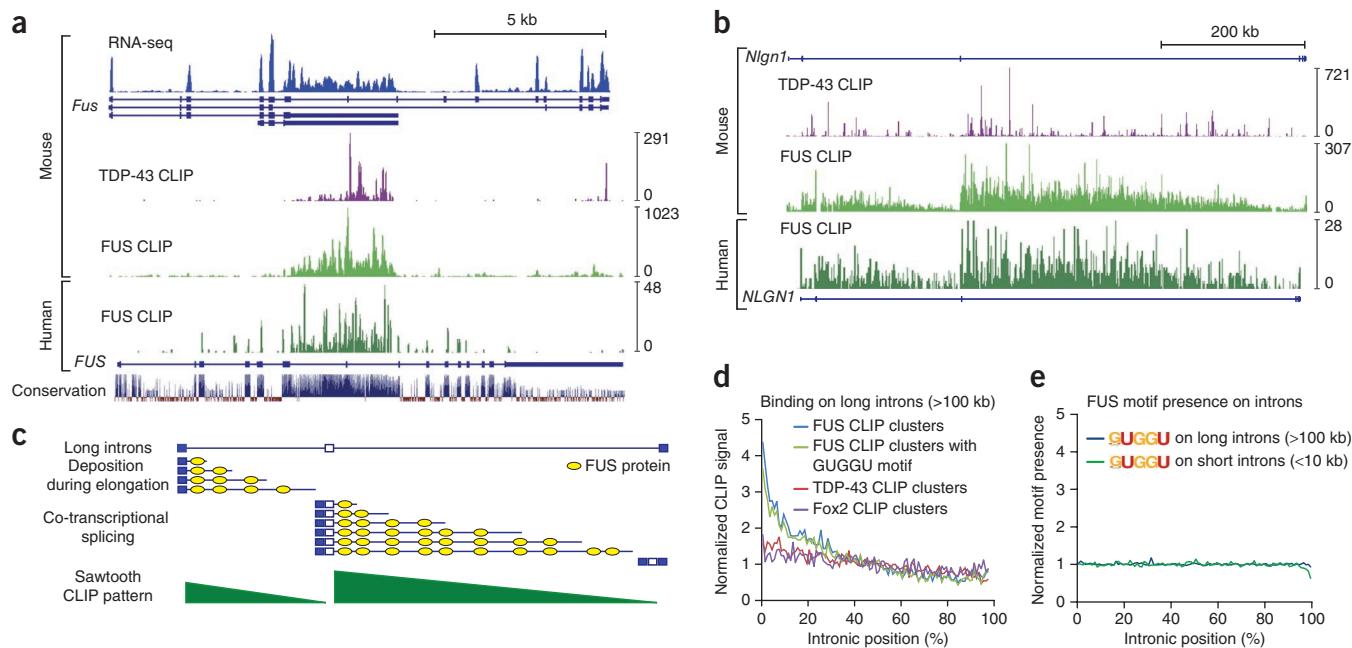
Figure 2 FUS/TLS binding patterns in mouse and human brain. (**a**) Mouse (light green) and human (dark green) FUS/TLS or TDP-43 (purple) binding to *FUS/TLS* RNA in a highly conserved region that represents either an alternative 3′ UTR, or a retained intron. (**b**) FUS/TLS bound to long transcripts such as neuroligin 1 (*Nlgn1*) in a characteristic sawtooth-like pattern. (**c**) Model for the deposition of FUS/TLS on long introns during transcriptional elongation and co-transcriptional splicing. A collection of nascent transcripts at different stages of elongation accumulated to produce a sawtooth-like pattern consistent with FUS/TLS deposition co-transcriptionally. (**d**) Graph showing enrichment of FUS/TLS, but not TDP-43 or RBFOX2, binding at the 5′ end of long introns. (**e**) Graph displaying a uniform density of FUS/TLS motif frequency across long and short introns measured by the signal-to-noise ratio (SNR) of the per gene fraction of clusters per 1% bin.

the FUS/TLS motif GUGGU. Similarly to what we observed for the protein-coding RNA targets, the position of FUS/TLS-binding sites on lncRNAs were distinct from those of TDP-43, as exemplified by binding on maternally expressed 3 (*Meg3*), imprinted in Prader-Willi syndrome (*Ipw*) or nuclear paraspeckle assembly transcript 1 (*Neat1*) lncRNAs (**Supplementary Fig. 7**).
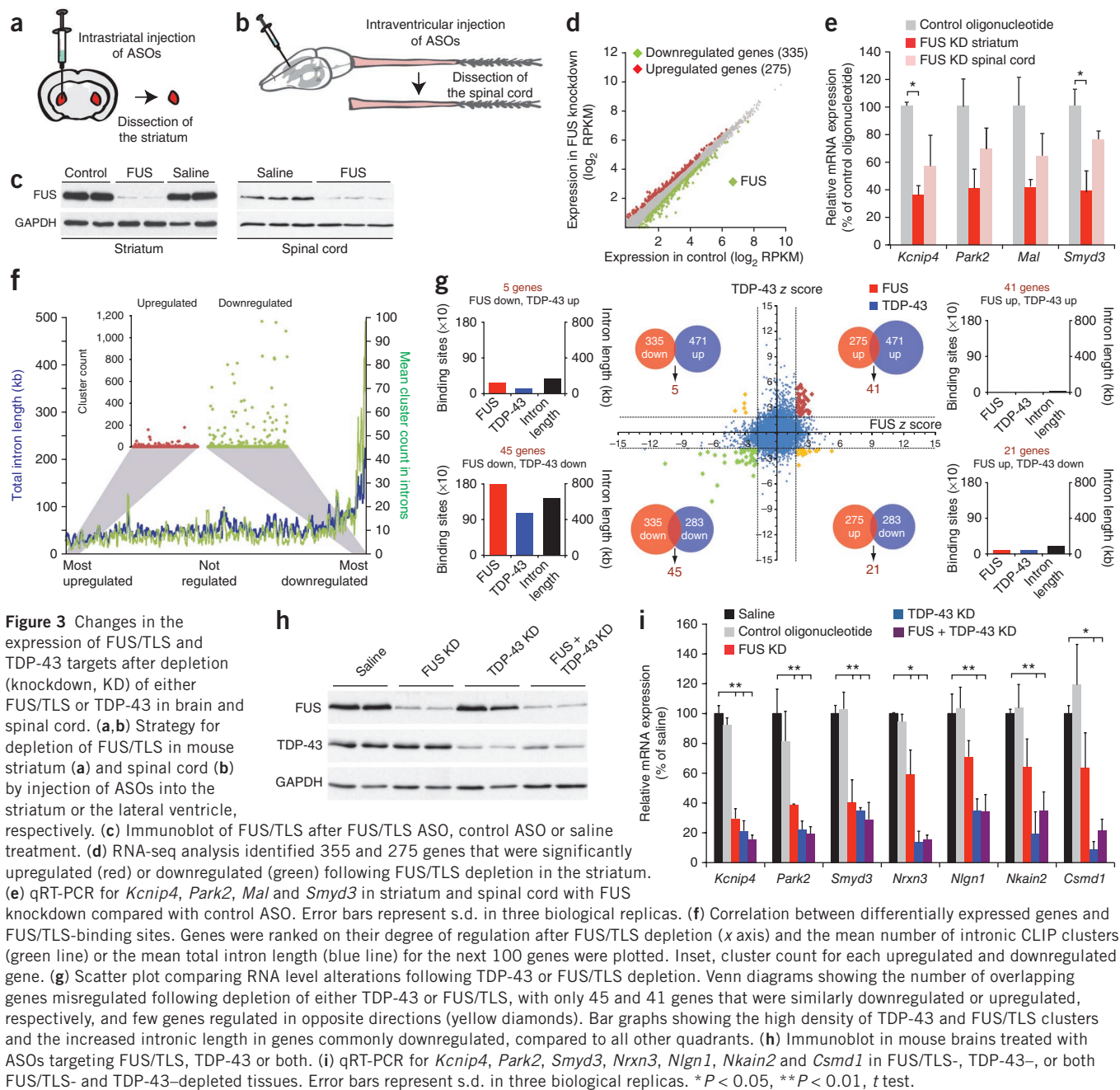
RNAs altered after *in vivo* FUS/TLS depletion

To determine expression and splicing changes after FUS/TLS depletion *in vivo,* we used single-stranded antisense oligonucleotides (ASOs) that direct catalytic degradation of RNA through the action of endogenous RNase H. An ASO specific for FUS/TLS, a control ASO without complementary sequences in the mouse genome, or saline were delivered into the adult CNS of mice by either stereotactic intrastriatal injection (**Fig. 3a**) or bolus injection into the lateral ventricle (**Fig. 3b**). The latter allows ASO distribution throughout the CNS, including all levels of the spinal cord[23,24]. Bolus injection or 2-week infusion of a FUS/TLS-specific ASO, but not a control ASO or saline, depleted FUS/TLS protein and RNA to ~10 or ~20% of normal levels, respectively, in the striatum and spinal cord (**Fig. 3c** and **Supplementary Fig. 8a**). No acute abnormal phenotype was observed in mice with marked reduction of FUS/TLS after 2 weeks of treatment.

FUS/TLS function in mRNA metabolism was assessed with strand-specific RNA-seq analysis of striata from three mice following ASO-mediated FUS/TLS depletion. For each sample, an average of ~26 million nonredundant reads uniquely mapped to the mouse mm9 genome (**Supplementary Fig. 8b**). Expression levels for each annotated protein-coding gene were determined by the number of mapped reads per kilobase of exon, per million mapped reads (RPKM). The combined sequence reads from the biological

replicates treated with FUS/TLS or control ASO (**Supplementary Fig. 8c**) were used to identify mRNAs encoded by 13,627 annotated protein-coding genes that satisfied the threshold of 0.5 RPKM in either condition. Pooling of reads per condition was supported by expected groupings of the independent biological replicates using hierarchical clustering of RPKM values of the differentially expressed genes (**Supplementary Fig. 8d**). The RPKM ratio of the gene encoding FUS/TLS confirmed a reduction of *Fus/Tls* (*Fus*) RNA to ~10% of normal levels (**Supplementary Fig. 8e**). FUS/TLS depletion affected a set of RNAs that was almost completely distinct from those dependent on TDP-43, as determined by comparison (**Supplementary Fig. 8d**) of similarly obtained RNA expression profiles from four biological replicates following ASO-mediated TDP-43 depletion in the striatum[15].

Statistical comparison of RPKM values between FUS/TLS- and control ASO–treated samples revealed that 275 genes were upregulated (defined by z score >1.96, $P < 0.05$) and 355 genes were downregulated (z score <−1.96, $P < 0.05$) following FUS/TLS depletion (**Fig. 3d** and **Supplementary Tables 2** and **3**). Similar expression changes were observed in the different biological replicates (**Supplementary Fig. 9**). Only four of the upregulated genes were changed by more than twofold. In contrast, 61 genes were downregulated more than twofold (**Supplementary Tables 2** and **3**). Reduced levels after FUS/TLS depletion were confirmed for selected RNAs by quantitative reverse transcription PCR (qRT-PCR) of striatal and spinal cord RNA samples (**Fig. 3e**). Notably, the degree of downregulation correlated with the level of FUS/TLS depletion in striatum versus spinal cord (**Supplementary Fig. 8a**), as exemplified by the genes encoding the potassium channel interacting protein 4 (*Kcnip4*), parkin (*Park2*), myelin and lymphocyte protein (*Mal*), and the histone methyltransferase Set and Mynd domain-containing protein 3 (*Smyd3*) (**Fig. 3e**).

Figure 3 Changes in the expression of FUS/TLS and TDP-43 targets after depletion (knockdown, KD) of either FUS/TLS or TDP-43 in brain and spinal cord. (**a**,**b**) Strategy for depletion of FUS/TLS in mouse striatum (**a**) and spinal cord (**b**) by injection of ASOs into the striatum or the lateral ventricle, respectively. (**c**) Immunoblot of FUS/TLS after FUS/TLS ASO, control ASO or saline treatment. (**d**) RNA-seq analysis identified 355 and 275 genes that were significantly upregulated (red) or downregulated (green) following FUS/TLS depletion in the striatum. (**e**) qRT-PCR for *Kcnip4*, *Park2*, *Mal* and *Smyd3* in striatum and spinal cord with FUS KD knockdown compared with control ASO. Error bars represent s.d. in three biological replicas. (**f**) Correlation between differentially expressed genes and FUS/TLS-binding sites. Genes were ranked on their degree of regulation after FUS/TLS depletion (*x* axis) and the mean number of intronic CLIP clusters (green line) or the mean total intron length (blue line) for the next 100 genes were plotted. Inset, cluster count for each upregulated and downregulated gene. (**g**) Scatter plot comparing RNA level alterations following TDP-43 or FUS/TLS depletion. Venn diagrams showing the number of overlapping genes misregulated following depletion of either TDP-43 or FUS/TLS, with only 45 and 41 genes that were similarly downregulated or upregulated, respectively, and few genes regulated in opposite directions (yellow diamonds). Bar graphs showing the high density of TDP-43 and FUS/TLS clusters and the increased intronic length in genes commonly downregulated, compared to all other quadrants. (**h**) Immunoblot in mouse brains treated with ASOs targeting FUS/TLS, TDP-43 or both. (**i**) qRT-PCR for *Kcnip4*, *Park2*, *Smyd3*, *Nrxn3*, *Nlgn1*, *Nkain2* and *Csmd1* in FUS/TLS-, TDP-43–, or both FUS/TLS- and TDP-43–depleted tissues. Error bars represent s.d. in three biological replicas. *$P < 0.05$, **$P < 0.01$, *t* test.
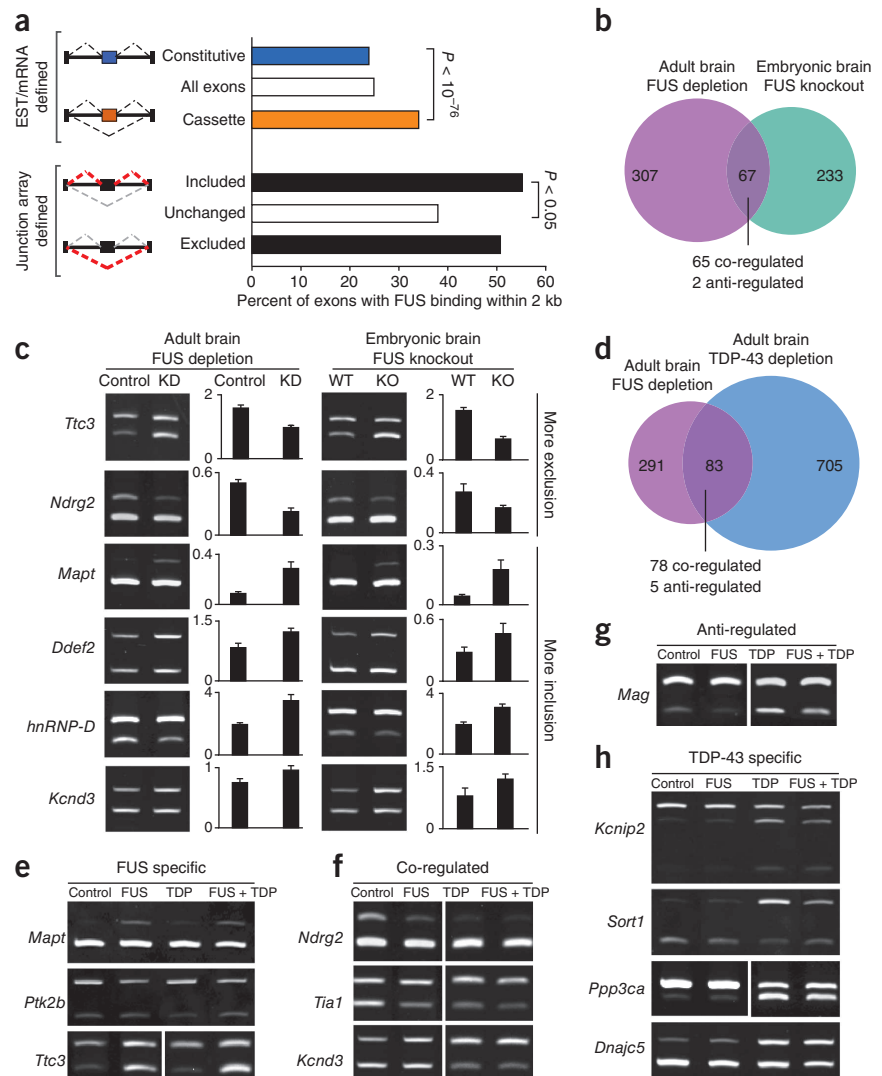
FUS and TDP-43 intersect in processing long RNAs

We previously established that TDP-43 is crucial for maintaining the mRNA levels of genes with exceptionally long introns, each of which contains numerous TDP-43–binding sites[15]. Applying the same analysis to RNAs affected by FUS/TLS depletion revealed that the genes that were downregulated the most were threefold longer (average of 158,997 bp) than those that were unaffected or upregulated (average of 51,687 bp, $P < 5.4 \times 10^{-9}$ by *t* test; **Fig. 3f**). Furthermore, an enrichment of FUS/TLS-binding sites was found in the most downregulated RNAs, with the hundred most reduced RNAs containing an average of 94 FUS/TLS-binding sites per pre-mRNA. Exceptionally long genes had more than 600 clusters each and include those encoding proteins with crucial functions in synaptic activity, such as *Nrxn1*, *Nrxn3*, *Nlgn1*, *Nlgn3* and *Kcnip4* (**Fig. 3f** and **Supplementary Tables 1** and **2**).

The prominence of FUS/TLS binding on downregulated RNAs was in contrast with a much lower density of binding on RNAs from upregulated genes, with an average of only six binding sites among the hundred most upregulated FUS/TLS targets. The infrequent binding sites and modest degree of upregulation are consistent with altered expression levels arising as an indirect consequence of FUS/TLS loss.

In contrast with the increased number of TDP-43–binding sites previously observed in the 3′ UTR of genes, such as progranulin (*Grn*), that are upregulated after TDP-43 reduction[15], no similar enrichment in FUS/TLS binding was seen in the 3′ UTR of upregulated FUS/TLS targets, consistent with the recent proposal that FUS/TLS does not affect the steady-state levels of mRNAs once they reach the cytoplasm[25].

Figure 4 FUS/TLS-dependent alternative splicing in mouse brain. (**a**) Schematics (left) and bar plots (right) displaying the percentage of alternatively excluded (cassette) or constitutively spliced, or all of the exons that contain FUS/TLS clusters within 2 kb (top). Bottom, percentage of exons that were excluded, unchanged or included following FUS/TLS depletion, as detected by splicing-sensitive microarray that contain FUS/TLS clusters in 2 kb. (**b**) Venn diagram showing the overlap of significantly changing alternative splicing events following FUS/TLS depletion in adult mouse brain and in embryonic brain from *Fus/Tls*[−/−] mice. The number of events changing in the same direction (co-regulated) and opposite direction (anti-regulated) are indicated. (**c**) Semi-quantitative RT-PCR analyses of alternative splicing changes following ASO-mediated FUS/TLS depletion in adult mouse brain and embryonic brain from *Fus/Tls*[−/−] mice compared with the respective controls. Graphs show the quantification (ratio of inclusion to exclusion) of three biological replicas per group; error bars represent s.d. (**d**) Venn diagram comparing the overlap of significantly changing alternative splicing events as a result of FUS/TLS depletion and previously published TDP-43 depletion in adult mouse brain. (**e**–**h**) Semi-quantitative RT-PCR validations of splicing events that changed only after FUS/TLS depletion (**e**), changed in similar direction when either FUS/TLS or TDP43 was depleted (**f**), changed in opposite directions when FUS/TLS or TDP-43 was depleted (**g**), or changed only after TDP-43 depletion (**h**).

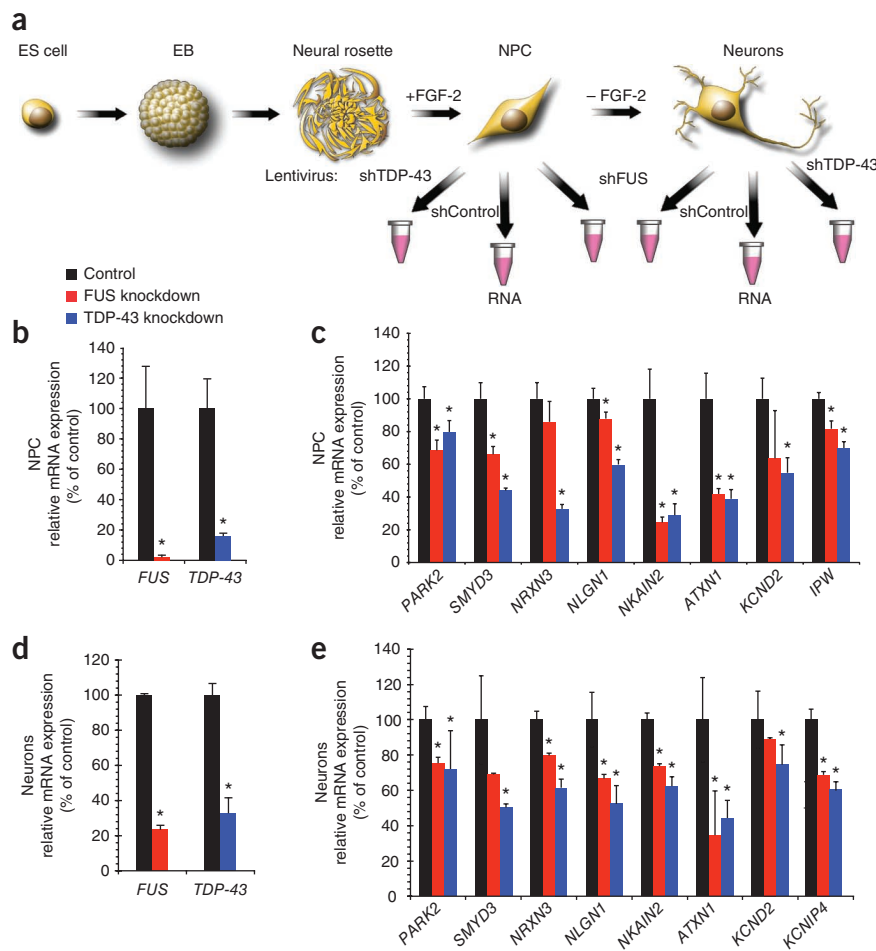## A small set of RNAs depends on both FUS/TLS and TDP-43

Systematic comparison of expression changes of all 13,627 expressed genes revealed highly distinct sets of RNAs that are dependent on FUS/TLS or TDP-43 (**Fig. 3g**). Only 41 and 45 genes were similarly upregulated or downregulated, respectively, in response to depletion of either protein, with downregulated genes bearing a marked enrichment for longer introns and binding sites for both TDP-43 and FUS/TLS (**Fig. 3g**).

To explore whether TDP-43 and FUS/TLS act synergistically on their targets, we depleted both simultaneously in the adult nervous system (**Supplementary Fig. 10a**). ASOs targeting both were injected and, after 2 weeks, RNA and protein for FUS/TLS and TDP-43 were reduced to 10% and 30% of normal levels, respectively, similar to the depletion achieved by injecting either ASO alone (**Fig. 3h** and **Supplementary Fig. 10b**). Levels of expression were confirmed by qRT-PCR for several long intron–containing genes that are down-regulated after either TDP-43 or FUS/TLS depletion, as well as in the samples with knockdown of both proteins (**Fig. 3i**). Of note, the selected transcripts were consistently more reduced following TDP-43 loss than following FUS/TLS loss. Depletion of both TDP-43 and FUS/TLS had no greater effect on the levels of expression of these long genes (which are enriched in genes that are important for synapse formation and activity; **Supplementary Table 4**), indicating that, even for these RNAs, TDP-43 and FUS/TLS act in parallel rather than synergistically.

FUS/TLS mediates alternative splicing of its RNA targets

FUS/TLS, in association with other factors, has been proposed to participate in RNA splice site selection (reviewed in ref. 14). As an initial test, we used publicly available EST/mRNA transcript databases to determine that, relative to constitutive exons, FUS/TLS-binding sites are enriched in the proximity of alternatively included or excluded 'cassette' exons ($P < 10^{-76}$, χ^2 test; **Fig. 4a**). Furthermore, using splicing-sensitive arrays, we identified 374 splicing events in adult brain that changed following FUS/TLS ASO-mediated depletion (**Supplementary Fig. 11a** and **Supplementary Tables 5** and **6**). Notably, the number of FUS/TLS-dependent splicing events identified (374) is relatively small compared with the number of changes observed after depletion of other heterogeneous nuclear ribonucleoparticle proteins, including TDP-43 (refs. 15,17). However, more than 55% of the 114 cassette exons differentially included following FUS/TLS depletion have FUS/TLS-binding sites within 2 kb of the exon boundary ($P < 0.05$, compared to ~38% of unchanged exons; **Fig. 4a**), consistent with FUS/TLS having a direct role in repressing exon recognition. Although no strong rules for FUS/TLS splicing regulation were apparent, as in recent reports[26,27], position-specific preferences for FUS/TLS binding around the regulated exons versus those that remained unchanged were identified. A significant enrichment of FUS/TLS binding was found ~200 nucleotides upstream of exons skipped following

Figure 5 Alterations of TDP-43 and FUS/TLS human targets in neurons derived from ES cells. (**a**) Schematic for generation of embryoid bodies (EBs), NPCs and neurons from human ES cells, followed by transduction with lentiviruses encoding shRNAs targeting *TDP-43* (shTDP-43) or *FUS/TLS* (shFUS), or a control shRNA (shControl). (**b**) qRT-PCR validation of FUS/TLS or TDP-43 depletion in NPCs. (**c**) qRT-PCR of *PARK2*, *SMYD3*, *NRXN3*, *NLGN1*, *NKAIN2*, *ATXN1*, *KCND2* and *IPW* after TDP-43 or FUS/TLS depletion in NPCs. (**d**) qRT-PCR validation of *FUS/TLS* or *TDP-43* depletion in differentiated neurons. (**e**) qRT-PCR of *PARK2*, *SMYD3*, *NRXN3*, *NLGN1*, *NKAIN2*, *ATXN1*, *KCND2* and *KCNIP4* after either TDP-43 or FUS/TLS depletion in differentiated neurons. Error bars represent s.d. in at least two biological replicas in each group. *$P < 0.01$ by t test.

FUS/TLS depletion (that is, FUS/TLS promoted inclusion of the exon), and significant binding in multiple regions around exons included following FUS/TLS loss ($P < 0.05$ by χ^2 test; **Supplementary Fig. 11b**).

A role for FUS/TLS in alternative splicing was confirmed by analysis of *Fus/Tls*[−/−] mice[28], in which gene disruption produces a complete loss of full-length FUS/TLS (**Supplementary Fig. 11c**), accompanied by perinatal death. Using splicing-sensitive microarrays, we identified 300 splicing events in embryonic brains from *Fus/Tls*[−/−] mice (**Supplementary Fig. 11a**), 67 of which occurred in RNAs whose splicing was also altered by FUS/TLS depletion from adult brain. Despite developmental differences in the alternative splicing patterns of embryonic and adult brains, splicing of 65 of these 67 FUS/TLS-dependent RNAs was altered in the same direction after FUS/TLS loss in either embryonic or adult brains (**Fig. 4b**).

Semi-quantitative RT-PCR of selected RNAs confirmed these FUS/TLS-dependent splicing changes (**Fig. 4c** and **Supplementary Fig. 12**), including transcripts implicated in neurodegenerative diseases, such as the N-myc downstream-regulated gene 2 (*Ndrg2*), which is expressed in astrocytes[29] and mis-accumulated in Alzheimer's disease[30], and the microtubule-associated protein tau (*Mapt*) gene mutated in frontotemporal dementia[31]. Loss of FUS/TLS resulted in the inclusion of a 54-bp exon that is not annotated in the mouse genome, but corresponds to an alternatively included exon upstream of the classical exon 9 of the human *MAPT* gene (also reported as exon 10 on transcript variant 6 of the human *MAPT* gene (UCSC, hg18); **Fig. 4c**). Splicing arrays also identified increased inclusion of *Mapt* exons 3 and 4 (mm9 chr11: 104,143,794–104,156,172) and exon 10 (mm9 chr11: 104,171,848–104,182,664) in adult and embryonic brain after depletion of FUS/TLS protein, respectively (**Supplementary Tables 1** and **5**). Identification that FUS/TLS suppresses splicing of *MAPT* exon 10 is of particular relevance for disease pathogenesis, as mutations that increase the inclusion of exon 10 are linked to frontotemporal dementia[31].
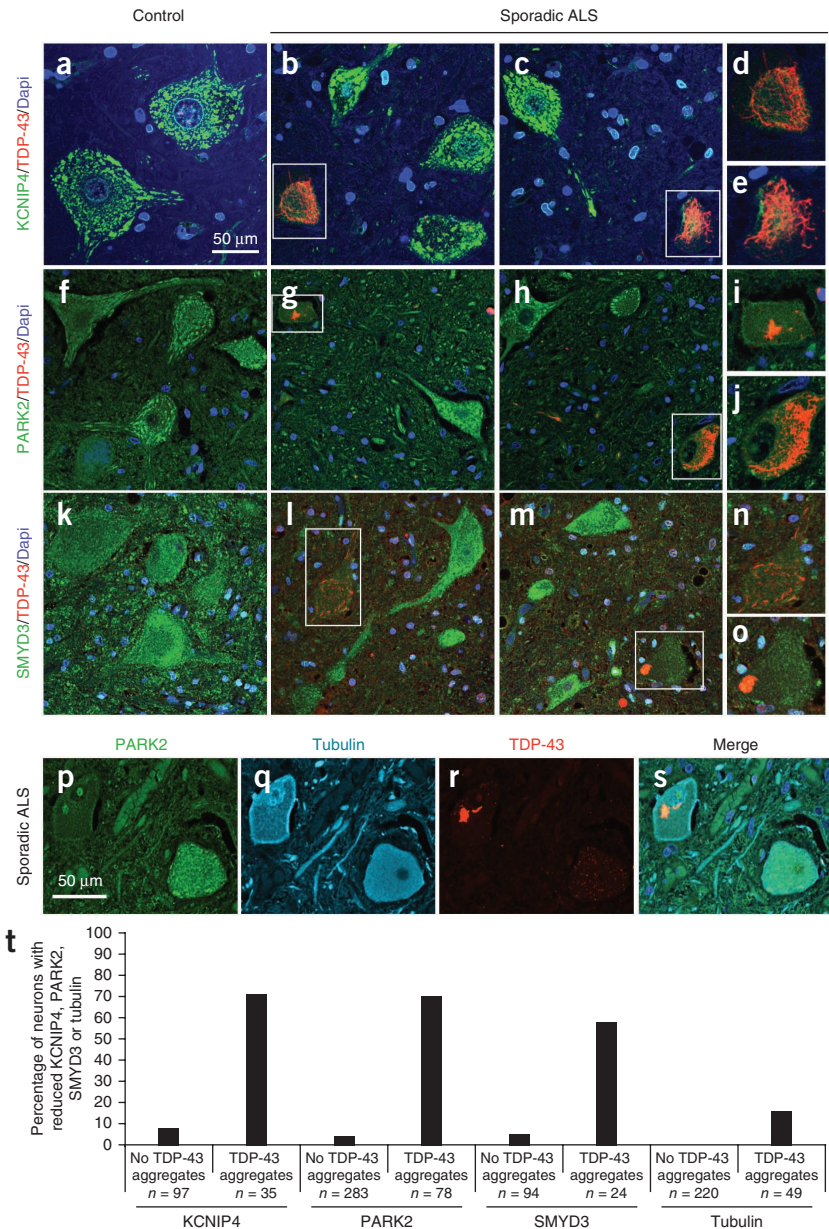
Comparison of the splicing changes after ASO-mediated depletion of either FUS/TLS or TDP-43 in adult brain revealed that TDP-43 affected a larger number (788) of splicing alterations than FUS/TLS

(374) (**Fig. 4d** and **Supplementary Table 6**). We identified 83 splicing events that were commonly regulated following FUS/TLS or TDP-43 depletion, 94% of which were differentially included or excluded in the same direction (**Fig. 4d**). Splicing alterations of selected TDP-43 and FUS/TLS RNA targets were validated by semi-quantitative RT-PCR (**Fig. 4e–h**). Exon 8 of *Ndrg2* showed increased exclusion after either TDP-43 or FUS/TLS depletion. Conversely, exon 5 of *Tia1*, whose encoded protein is a central component of RNA-containing stress granules, and exon 6 of the potassium voltage-gated channel gene (*Kcnd3*) were more included after loss of TDP-43 or FUS/TLS (**Fig. 4f**). FUS/TLS-dependent changes in splicing of *Mapt* were not found after the loss of TDP-43, despite TDP-43–binding sites within 2 kb of the affected exons (**Fig. 4e**). A rare example of differently regulated splicing events between FUS/TLS and TDP-43 is exon 12 of the myelin-associated glycoprotein gene (*Mag*), which encodes a protein that inhibits axonal regeneration through interaction with the Nogo receptor[32] (**Fig. 4g** and **Supplementary Table 6**).

Alterations of RNAs in human differentiated neurons

We tested the dependence of long intron–containing genes on TDP-43 and FUS/TLS in human neurons by differentiation of neural progenitor cells (NPCs) generated from human embryonic stem (ES) cells (**Fig. 5a**) and depletion of TDP-43 or FUS/TLS using lentiviral-encoded short hairpin RNAs (shRNAs). qRT-PCR revealed that human *FUS/TLS* (*FUS*) and *TDP-43* (*TARDBP*) mRNAs were reduced to 2 and 15% of their normal levels, respectively, when compared with NPCs treated with a scrambled shRNA (**Fig. 5b**). Of the target genes

Figure 6 Reduction of TDP-43 and FUS/TLS RNA targets in motor neurons from sporadic ALS patients. (**a–o**) Immunofluorescence localization in spinal cord autopsy samples from control, non-ALS individuals (**a,f,k**) or sporadic ALS (sALS) patients (**b–e,g–j,l–o**) for TDP-43 (red, **a–o**) and either KCNIP4 (green, **a–e**), Parkin (green, **f–j**) or SMYD3 (green, **k–o**). In both control and sALS motor neurons without TDP-43 aggregation, KCNIP4, Parkin and SMYD3 showed the expected, mainly cytoplasmic localization. Neurons bearing TDP-43 inclusions had markedly decreased levels of KCNIP4, Parkin or SMYD3. **d**, **e**, **i**, **j**, **n** and **o** are higher magnifications of the areas highlighted in **b**, **c**, **g**, **h**, **l** and **m**, respectively. (**p–s**) Triple immunofluorescence for images from a sALS spinal cord autopsy sample for Parkin (green, **p**), tubulin (blue, **q**) and TDP-43 (red, **r**). The merged images are shown in **s**. Notice that the cell on the left, which contained TDP-43 aggregates and had reduced Parkin levels, contained normal levels of tubulin. (**t**) Quantification of KCNIP4, Parkin, SMYD3 or tubulin levels in individual motor neurons from a total of 11 sALS patients and 3 control individuals revealed that the majority (~60–70%) of neurons with TDP-43 inclusions had reduced levels of KCNIP4, Parkin or SMYD3 staining. In contrast, the majority of TDP-43 inclusion–bearing cells (>80%) showed normal levels of tubulin staining, which is not a TDP-43 target and whose levels were not expected to be altered by TDP-43 loss.

reliably expressed in NPCs, long intron–containing pre-mRNAs for parkin (*PARK2*), SET and MYND domain-containing protein 3 (*SMYD3*), neurexin 3 (*NRXN3*), neuroligin 1 (*NLGN1*), Na$^+$/K$^+$ transporting ATPase interacting 2 or T-cell lymphoma breakpoint associated target 1 (*NKAIN2* or *TCBA1*), ataxin 1 (*ATXN1*), potassium voltage-gated channel (*KCND2*), and the noncoding RNA imprinted in Prader-Willi (*IPW*) were reduced after loss of either TDP-43 or FUS/TLS (**Fig. 5c**), confirming an intersecting and conserved role for both proteins in sustaining the levels of long intron–containing genes in human neuronal progenitors. Reduction of these pre-mRNAs targets and *KCNIP4* were also observed in neurons differentiated from human ES cells after FUS/TLS or TDP-43 were reduced to 23 and 33% of controls, respectively (**Fig. 5a,d,e**). Similarly, after differentiation into a pan-neuronal lineage of induced pluripotent stem (iPS) cells generated from an unaffected individual, endogenous TDP-43 was depleted using either of two independent lentiviral-encoded shRNAs (**Supplementary Fig. 13**). Assay of the resultant RNAs revealed downregulation of all seven of the long intron–containing genes that we examined (**Supplementary Fig. 13**), confirming a dependency of long pre-mRNA levels on TDP-43 in differentiated human neurons *in vitro*.

After identifying targets affected by the loss of TDP-43 or FUS/TLS, we tested mRNA levels of long intron–containing transcripts in primary fibroblasts from three symptomatic ALS patients: one carrying a *FUS/TLS*[R521G] mutation and two carrying the same *TDP-43*[G298S] mutation. One of the targets expressed in detectable levels in

fibroblast cells, *KCND2*, was substantially reduced in mutant fibroblasts compared with control fibroblasts from a non-ALS individual (data not shown). Although this preliminary evidence agreed with our hypothesis, the variability of gene expression amidst different fibroblast lines from other control individuals prevented us from drawing firm conclusions from these experiments and suggests that further studies using iPS cell–derived motor neurons may ultimately prove more conclusive.

Long pre-mRNAs altered in ALS patient motor neurons

Our data predict that the levels of this small set of common targets (**Supplementary Tables 2** and **3**) will be reduced in neurons with either FUS/TLS or TDP-43 disruption. Patients carrying FUS/TLS mutations are extremely rare, but, in the vast majority of sporadic ALS patients, cytoplasmic aggregates of TDP-43 accumulate in motor neurons, accompanied by a loss of nuclear staining[2,3]. To determine whether TDP-43– and FUS/TLS-dependent RNAs from genes with

long introns are affected in sporadic ALS, we initially reanalyzed expression alterations from a published microarray analysis of RNAs obtained after laser microdissection and pooling of approximately 1,000 motor neurons from 12 ALS patients[33]. Analysis of long intron–containing genes (including *PARK2*, *KCNIP4*, *SMYD3*, *NLGN1*, calcium-dependent secretion activator (*CADPS*), fibroblast growth factor (*FGF14*) and Ephrin A5 (*EFNA5*)) revealed a trend for reductions in each of these transcripts, although none reached statistical significance, with the exception of Ephrin A5 ($P = 0.0123$).

Given the well-recognized heterogeneity of TDP-43 pathology in patient motor neurons[34], we reasoned that a microarray-based analysis of altered expression of long intron–containing genes following loss of TDP-43 function might be confounded by pooling of RNAs from more or less affected neurons. Consequently, we assayed individual motor neurons in autopsy samples from 11 sporadic ALS patients and 3 control individuals by co-labeling for TDP-43 and either of three proteins (PARK2, KCNIP4 and SMYD3) encoded by long pre-mRNAs whose maturation is dependent on TDP-43 or FUS/TLS. Immunofluorescence staining of spinal cords was performed with either a polyclonal antibody specific for phospho–TDP-43 or a monoclonal antibody specific for the C-terminal part of TDP-43 that both recognize primarily TDP-43 inclusions (**Supplementary Fig. 14**). With the conditions used, neither of the TDP-43 antibodies tested produced a reproducible immunofluorescent staining of nuclear TDP-43. Nevertheless, a strong correlation was apparent between reduction in parkin, KCNIP4 or SMYD3 fluorescence and the presence of TDP-43 cytoplasmic inclusions (**Fig. 6** and **Supplementary Fig. 15**). Indeed, strong granular cytoplasmic staining of KCNIP4 was observed in 92% of motor neurons without identified TDP-43 aggregates in both control individuals and sporadic ALS patients. In contrast, 71% of motor neurons with TDP-43 aggregates were characterized by reduced KCNIP4 fluorescence (**Fig. 6a–e,t**). Similarly, parkin was strongly reduced in 70% of TDP-43 inclusion–containing motor neurons, whereas 95% of motor neurons without TDP-43 aggregates displayed strong cytoplasmic parkin levels (**Fig. 6f–j,t**). The nuclear and cytoplasmic protein SMYD3 was also reduced in 58% of motor neurons with TDP-43 inclusions, whereas 95% of non-affected motor neurons had normal staining for SMYD3 (**Fig. 6k–o,t**). Notably, tubulin, a protein that was not predicted to be regulated by TDP-43 or FUS/TLS, was not altered in the large majority (85%) of motor neurons carrying TDP-43 aggregates (**Fig. 6p–t**).

DISCUSSION

Our study provides a genome-wide, systematic comparison of the binding patterns and roles in gene regulation and alternative splicing in the CNS for TDP-43 and FUS/TLS, proteins that are involved in the pathogenesis of ALS and FTLD. We identified extensive FUS/TLS binding on thousands of mouse and human brain pre-mRNAs at positions that were largely distinct from those of TDP-43, including a sawtooth-like pattern on long pre-mRNAs. Our results also demonstrate that depletion of FUS/TLS leads to a broad misregulation of RNA processing, as the levels of >600 mRNAs were altered and the splicing patterns of >350 were changed. We found that FUS/TLS and TDP-43 affected the expression levels and processing of largely distinct sets of RNAs, and had only 86 and 78 gene regulation and splicing events in common, respectively.

Nevertheless, our evidence identifies intersection of the TDP-43 and FUS/TLS pathways in the regulation of a subset of transcripts that were downregulated after the loss of either TDP-43 or FUS/TLS (**Fig. 3**), harbor numerous intronic binding sites for both TDP-43 and FUS/TLS (**Figs. 2** and **3**), contain exceptionally long introns (**Figs. 2** and **3**), and encode protein products that are crucial for normal neuronal function

(**Supplementary Table 4**). These distinct attributes of target genes cannot be identifiable by similar analyses performed in non-neuronal cells. For example, none of the 45 RNAs that are downregulated following either FUS/TLS or TDP-43 depletion in the adult mouse brain are appreciably expressed (as revealed by RNA-seq data) in an immortalized cell line (293T cells) used in a previous analysis[35].

Analyses in progenitors and 1-month-old neurons derived from ES and iPS cells confirm the dependency of long intron–containing pre-mRNAs on TDP-43 or FUS/TLS in human cells (**Fig. 5** and **Supplementary Fig. 13**). Our findings predict that, in cells with functional interruption of either TDP-43 or FUS/TLS as a result of mutation and/or aggregation in disease, reduction of the levels of these target RNAs contributes to neuronal death in TDP-43 and FUS/TLS proteinopathies. We found that this is indeed the case in individual motor neurons in sporadic ALS spinal cords, where cytoplasmic mis-accumulation of TDP-43 is accompanied by loss of proteins encoded by long pre-mRNAs (**Fig. 6**). Our combined evidence supports that loss of function of TDP-43 or FUS/TLS contributes to ALS pathogenesis. Furthermore, given that the loss of either parkin or KCNIP4 individually leads to compromised neuronal function in human patients and mouse models[36–40], our finding that these targets are altered in affected motor neurons from ALS patients predicts that a combined reduction of multiple targets contributes to, rather than singlehandedly causes, neuronal death in TDP-43 and FUS/TLS proteinopathies.

The exact molecular mechanism by which TDP-43 or FUS/TLS binding on multiple intronic sites sustains the levels of these transcripts remains unknown and may or may not be identical for the two proteins. An attractive possibility is that binding of TDP-43 or FUS/TLS to long introns may serve as a protein scaffold that ensures the normal splicing of long introns, preventing unproductive splicing events. Alternatively, multiple TDP-43 and/or FUS/TLS molecules bound to these introns may facilitate transcription elongation, similar to what has been shown for another RNA-binding protein, SC-35 (ref. 41). In fact, our observation that FUS/TLS bound along long introns in a sawtooth-like pattern, as is expected for binding to nascent transcripts[22,42], supports the idea that it may be involved in transcriptional elongation.

A recent study found that proteins recruited to promoter regions can be co-transcriptionally deposited onto nascent mRNAs, thereby regulating decay of those RNAs by maintaining this interaction following mRNA transport to the cytoplasm[43]. Mindful of this, it is conceivable that a similar mechanism mediates the deposition of FUS/TLS on multiple sites of long intron–containing nascent RNAs. The very recent observation that FUS/TLS interacts with single-stranded DNA in promoter regions in human HeLa cells[44], in conjunction with the association of FUS/TLS with RNA polymerase II and the TFIID complex[45], supports this view. Alternatively, binding of FUS/TLS on ncRNAs with complementarity to promoter regions may be responsible for its recruitment there[46]. It will be necessary to determine whether the intersecting TDP-43 and FUS/TLS pathways that we uncovered are also affected by alterations in TAF15 (refs. 47,48) or EWSR1 (refs. 49,50), two additional RNA/DNA binding proteins that were recently shown to mediate ALS and FTLD.

METHODS

Methods and any associated references are available in the online version of the paper.

ARTICLES

AUTHOR CONTRIBUTIONS
C.L.-T., M.P., A.Q.V., M.B. and K.M.C. performed the experiments. K.R.H., S.C.H. and T.Y.L. conducted the bioinformatics analysis. S.-C.L. developed the polyclonal FUS/TLS-specific antibody Ab3. J.P.D. and L.S. conducted the preliminary splice-junction microarray analyses. G.G.H. provided the brain samples from the *Fus/Tls⁻/⁻* mice. M.P., C.L.-T, C.M., E.W., A.S.K., A.W., S.F. and C.F.B. conducted the ASO experiments. J.R. provided the ALS patient tissues. C.L.-T., M.P., K.R.H., G.W.Y. and D.W.C. designed the experiments. C.L.-T., M.P., K.R.H., A.Q.V., G.W.Y. and D.W.C. wrote the manuscript.

1. Da Cruz, S. & Cleveland, D.W. Understanding the role of TDP-43 and FUS/TLS in ALS and beyond. *Curr. Opin. Neurobiol.* **21**, 904–919 (2011).
2. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
3. Arai, T. *et al.* TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochem. Biophys. Res. Commun.* **351**, 602–611 (2006).
4. Polymenidou, M. *et al.* Misregulated RNA processing in amyotrophic lateral sclerosis. *Brain Res.* **1462**, 3–15 (2012).
5. Gitcho, M.A. *et al.* TDP-43 A315T mutation in familial motor neuron disease. *Ann. Neurol.* **63**, 535–538 (2008).
6. Kabashi, E. *et al.* TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat. Genet.* **40**, 572–574 (2008).
7. Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* **319**, 1668–1672 (2008).
8. Kwiatkowski, T.J. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205–1208 (2009).
9. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208–1211 (2009).
10. Neumann, M. *et al.* A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain* **132**, 2922–2931 (2009).
11. Woulfe, J., Gray, D.A. & Mackenzie, I.R. FUS-immunoreactive intranuclear inclusions in neurodegenerative disease. *Brain Pathol.* **20**, 589–597 (2010).
12. Munoz, D.G. *et al.* FUS pathology in basophilic inclusion body disease. *Acta Neuropathol.* **118**, 617–627 (2009).
13. Mackenzie, I.R. *et al.* Pathological heterogeneity in amyotrophic lateral sclerosis with FUS mutations: two distinct patterns correlating with disease severity and mutation. *Acta Neuropathol.* **122**, 87–98 (2011).
14. Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D.W. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.* **19**, R46–R64 (2010).
15. Polymenidou, M. *et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat. Neurosci.* **14**, 459–468 (2011).
16. Tollervey, J.R. *et al.* Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452–458 (2011).
17. Huelga, S.C. *et al.* Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep.* **1**, 167–178 (2012).
18. Lerga, A. *et al.* Identification of an RNA binding specificity for the potential splicing factor TLS. *J. Biol. Chem.* **276**, 6807–6816 (2001).
19. Fujii, R. *et al.* The RNA binding protein TLS is translocated to dendritic spines by mGluR5 activation and regulates spine morphology. *Curr. Biol.* **15**, 587–593 (2005).
20. Ling, S.C. *et al.* ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proc. Natl. Acad. Sci. USA* **107**, 13318–13323 (2010).
21. Zhang, C. *et al.* Integrative modeling defines the Nova splicing-regulatory network and its combinatorial controls. *Science* **329**, 439–443 (2010).
22. Ameur, A. *et al.* Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nat. Struct. Mol. Biol.* **18**, 1435–1440 (2011).
23. Smith, R.A. *et al.* Antisense oligonucleotide therapy for neurodegenerative disease. *J. Clin. Invest.* **116**, 2290–2296 (2006).
24. Kordasiewicz, H.B. *et al.* Sustained therapeutic reversal of Huntington's disease by transient repression of Huntingtin synthesis. *Neuron* **74**, 1031–1044 (2012).
25. Colombrita, C. *et al.* TDP-43 and FUS RNA-binding proteins bind distinct sets of cytoplasmic messenger RNAs and differently regulate their post-transcriptional fate in motoneuron-like cells. *J. Biol. Chem.* **287**, 15635–15647 (2012).
26. Ishigaki, S. *et al.* Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. *Sci. Rep.* **2**, 529 (2012).
27. Rogelj, B. *et al.* Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain. *Sci. Rep.* **2**, 603 (2012).
28. Hicks, G.G. *et al.* Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nat. Genet.* **24**, 175–179 (2000).
29. Okuda, T., Kokame, K. & Miyata, T. Differential expression patterns of NDRG family proteins in the central nervous system. *J. Histochem. Cytochem.* **56**, 175–182 (2008).
30. Mitchelmore, C. *et al.* NDRG2: a novel Alzheimer's disease associated protein. *Neurobiol. Dis.* **16**, 48–58 (2004).
31. Hutton, M. *et al.* Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* **393**, 702–705 (1998).
32. Liu, B.P., Fournier, A., GrandPre, T. & Strittmatter, S.M. Myelin-associated glycoprotein as a functional ligand for the Nogo-66 receptor. *Science* **297**, 1190–1193 (2002).
33. Rabin, S.J. *et al.* Sporadic ALS has compartment-specific aberrant exon splicing and altered cell-matrix adhesion biology. *Hum. Mol. Genet.* **19**, 313–328 (2010).
34. Bodansky, A. *et al.* TDP-43 and ubiquitinated cytoplasmic aggregates in sporadic ALS are low frequency and widely distributed in the lower motor neuron columns independent of disease spread. *Amyotroph. Lateral Scler.* **11**, 321–327 (2010).
35. Hoell, J.I. *et al.* RNA targets of wild-type and mutant FET family proteins. *Nat. Struct. Mol. Biol.* **18**, 1428–1431 (2011).
36. Kitada, T. *et al.* Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**, 605–608 (1998).
37. Itier, J.M. *et al.* Parkin gene inactivation alters behavior and dopamine neurotransmission in the mouse. *Hum. Mol. Genet.* **12**, 2277–2291 (2003).
38. Goldberg, M.S. *et al.* Parkin-deficient mice exhibit nigrostriatal deficits but not loss of dopaminergic neurons. *J. Biol. Chem.* **278**, 43628–43635 (2003).
39. Morohashi, Y. *et al.* Molecular cloning and characterization of CALP/KChIP4, a novel EF-hand protein interacting with presenilin 2 and voltage-gated potassium channel subunit Kv4. *J. Biol. Chem.* **277**, 14965–14975 (2002).
40. Massone, S. *et al.* RNA polymerase III drives alternative splicing of the potassium channel-interacting protein contributing to brain complexity and neurodegeneration. *J. Cell Biol.* **193**, 851–866 (2011).
41. Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S. & Fu, X.D. The splicing factor SC35 has an active role in transcriptional elongation. *Nat. Struct. Mol. Biol.* **15**, 819–826 (2008).
42. Wada, Y. *et al.* A wave of nascent transcription on activated human genes. *Proc. Natl. Acad. Sci. USA* **106**, 18357–18361 (2009).
43. Trcek, T., Larson, D.R., Moldon, A., Query, C.C. & Singer, R.H. Single-molecule mRNA decay measurements reveal promoter-regulated mRNA stability in yeast. *Cell* **147**, 1484–1497 (2011).
44. Tan, A.Y., Riley, T.R., Coady, T., Bussemaker, H.J. & Manley, J.L. TLS/FUS (translocated in liposarcoma/fused in sarcoma) regulates target gene transcription via single-stranded DNA response elements. *Proc. Natl. Acad. Sci. USA* **109**, 6030–6035 (2012).
45. Bertolotti, A. *et al.* EWS, but not EWS-FLI-1, is associated with both TFIID and RNA polymerase II: interactions between two members of the TET family, EWS and hTAFII68, and subunits of TFIID and RNA polymerase II complexes. *Mol. Cell Biol.* **18**, 1489–1497 (1998).
46. Wang, X. *et al.* Induced ncRNAs allosterically modify RNA-binding proteins in *cis* to inhibit transcription. *Nature* **454**, 126–130 (2008).
47. Ticozzi, N. *et al.* Mutational analysis reveals the FUS homolog TAF15 as a candidate gene for familial amyotrophic lateral sclerosis. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **156B**, 285–290 (2011).
48. Couthouis, J. *et al.* A yeast functional screen predicts new candidate ALS disease genes. *Proc. Natl. Acad. Sci. USA* **108**, 20881–20890 (2011).
49. Couthouis, J. *et al.* Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **21**, 2899–2911 (2012).
50. Neumann, M. *et al.* FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. *Brain* **134**, 2595–2609 (2011).

ONLINE METHODS

CLIP-seq library preparation and sequencing. Brains from 8-week-old female C57Bl/6 mice or frozen human autopsy cortical specimens were dissociated using a cell strainer of pore size 100 μm (BD Falcon) before ultraviolet cross-linking. CLIP-seq libraries were constructed and sequenced on the Hiseq 2000 platform for 36 cycles[15]. Three different antibodies to FUS/TLS were used: Ab1, a rabbit antibody (Bethyl Laboratories, A300-302A) specific for amino acids 1–50 of human FUS/TLS (per the company's description), which are almost identical between the human and the mouse FUS/TLS proteins, Ab2, a mouse monoclonal antibody specific for the C terminus of FUS/TLS (Santa Cruz Biotechnology, clone 4H11, sc-47711), and Ab3, a rabbit antibody to a peptide spanning amino acids 128–140. For each of the three CLIP-seq libraries, the brain of one mouse was used.

Injections of ASO in mice CNS and RNA-seq library preparation. All procedures were accomplished using a protocol approved by the Institutional Animal Care and Use Committee of the University of California at San Diego (Department of Health and Human Services, NIH Publication 86-23). Brains from *Fus/Tls*[−/−] mice, in which the *Fus/Tls* gene is interrupted[28], were collected at embryonic day 18.5 and RNAs and proteins were extracted using Trizol (Invitrogen). To deplete FUS/TLS *in vivo*, we performed stereotactic injections of either 2 or 10 μl of ASO solution, corresponding to a total of 75 or 300 μg of ASOs, respectively, into the striatum or the right ventricle of 8–10-week-old female C57Bl/6 mice[15]. Mice were monitored for any adverse effects for 2 weeks until they were killed, and striatum and adjacent cortex area or the spinal cord, respectively, were dissected and frozen in 1 ml of Trizol (Invitrogen). For the double knockdown experiment, 75 μg of TDP-43 ASO was mixed with or 50 μg of FUS/TLS ASO before injection into the striatum. A control group injected with the same dose (125 μg) of control ASO without any known target in the mouse genome was used for comparison. The ASOs are gapmers with capitalized nucleotides containing 2′-O-(2-methoxy)ethyl modifications with sequences as follows: CCTGGttatttcccaTGAGC (FUS/TLS); AAGGCttcatattgtACTTT (TDP-43). We sequenced 8 pmol of each constructed RNA-seq library[51] on the HiSeq 2000 platform for 100 cycles.

RT-PCR and qRT-PCR of RNA from mice samples. To validate alternative splicing, we carried out RT-PCR amplification (24–27 cycles) using poly-A–selected and reverse transcribed (Superscript III, Invitrogen) cDNA from three mice treated with either a control or FUS/TLS ASO. Isoform products were separated on 10% polyacrylamide gels and stained with SYBR gold (Invitrogen) and quantified with ImageJ software (US National Institutes of Health). Intensity ratios between products with included and excluded exons were averaged from three biological replicates per group. qRT-PCR for mouse *Tardbp* and *Fus/Tls* were performed using the Express One-Step SuperScript kits (Invitrogen) and thermocycler ABI Prism 7700 (Applied Biosystems), with the cyclophilin A (*Ppia*) gene as a normalization control. cDNA synthesis and amplification were performed using specific primers and 5′ FAM, 3′ TAMRA–labeled probes. qRT-PCR for all other genes tested was performed with 3–5 mice for each group (treated with saline, control ASO, or ASO to either FUS/TLS or TDP-43) and two technical replicates using the iQ SYBR green supermix (Bio-Rad) on the iQ5 multicolor real-time PCR detection system (Bio-Rad). Analysis was performed using the iQ5 optical system software (Bio-Rad; version 2.1). Expression values were normalized to at least two of the following control genes: *Actb*, *Actg1* and *Rsp9*. Expression values were expressed as a percentage of the average expression of saline-treated samples. Inter-group differences were assessed by two-tailed Student's *t* test. Primer sequences designed using Primer3 software (http://frodo.wi.mit.edu/primer3/) are available in **Supplementary Table 7**.

Antibodies for immunoblots. Immunoblots for FUS/TLS or TDP-43 were performed in a least three mice for each treatment group (**Fig. 3**). For primary antibodies, we used rabbit antibody to FUS/TLS (Bethyl Laboratories, A300-302A, 1:5,000), mouse antibody to FUS/TLS (Santa Cruz Biotechnology, clone 4H11, sc-47711, 1:500), rabbit antibody to amino acids 128–140 of FUS/TLS (serum diluted to 1:2,000), rabbit antibody to TDP-43 (Proteintech, 10782, 1:2,000), mouse DM1α antibody to tubulin (1:10,000) and mouse antibody to GAPDH (Abcam, AB8245, 1:10,000).

CLIP-seq and RNA-seq read mapping. Reads from CLIP-seq and RNA-seq libraries had adaptor sequences trimmed and homopolymeric runs removed before mapping to the repeat-masked mouse genome (mm9) using bowtie (version 0.12.2), with parameters −q −p 4 −e 100 −a −m 10 −best–strata. Reads with identical sequences were collapsed into a single read. Reads were allowed to map to multiple locations, but only the best scoring alignment was kept for downstream analysis. The entire trimmed library was mapped against reference snRNA sequences rather than the genome, and reads mapped to the U1 snRNA reference sequence were plotted as cumulative read coverage per nucleotide. Mismatch sites provided evidence of incorrect base incorporation during the reverse transcriptase step, as a result of protein adducts left by ultraviolet cross-linking[52].

CLIP-seq cluster identification, motif and pattern analysis. CLIP cluster identification was performed as previously described[15,17,53,54]. Significant clusters of reads were calculated using a Poisson distribution using two different frequencies to determine a *P* value. First, a transcriptome-wide frequency was calculated by dividing the total length of all pre-mRNAs by the total CLIP reads mapping in them. Second, a gene-specific frequency was calculated by dividing the size of a particular pre-mRNA by the total CLIP reads in it. A significant cluster had sufficient reads to satisfy a bonferroni-corrected *P* value $<10^{-4}$ for both frequencies. To determine CLIP clusters outside of annotated pre-mRNAs, we calculated a genomic frequency using genome size divided by total genomic CLIP reads, as well as a local frequency using a ±1-kb region surrounding a potential cluster divided by the total CLIP reads overlapping this region. We used 150-nt regions surrounding the center of CLIP clusters to identify strand-specific motifs against a nucleotide frequency–matched background of all pre-mRNAs, using the software HOMER[55]. To visualize sawtooth-like patterns, we determined intron distributions of CLIP clusters by assigning a cluster to a 1%-resolution intron bin dependent on the length of the intron. For each cluster, a score was added to this bin equal to the fraction this cluster represents versus all clusters on the entire gene to normalize for differing bin sizes. Motifs (defined from HOMER) and RNA-seq reads were analyzed the same way.

RNA-seq analysis and annotation of candidate intron retention events. Differentially expressed genes computed from RPKM[56] values were defined by a *z* score > 1.96 (*P* < 0.05). A subset of 17 genes was excluded from further analysis, as they also changed in control oligo–treated samples compared with saline-treated samples. To identify putative intron-retention/alternative 3′ UTR events, RPKMs of individual regions, not entire genes, were calculated for known 3′ UTRs and their distribution was plotted. The most frequent RPKM value was chosen as the cutoff (natural log RPKM > 2) that we used to distinguish expressed intronic regions (1,866 regions) from non-expressed ones.

Splicing-sensitive microarray analysis. Microarray analysis was performed as previously described, filtering events with a *q* value >0.05 and an absolute separation score > 0.5 (ref. 57). An RNA splicing map was produced as previously reported[17], overlapping FUS/TLS CLIP-seq clusters around splicing array–detected cassette exons promoted by FUS/TLS (skipped following knockdown, *n* = 63), repressed by FUS/TLS (included following knockdown, *n* = 114) or unchanged by FUS/TLS (*n* = 3,664).

Generation of NPCs and neurons from pluripotent stem cells. Human pluripotent embryonic stem cells (HUES6 line) and iPS cells were induced into neural progenitors and differentiated into mature neurons using a pan-neuronal protocol as previously described[58,59]. Stem cells grown on matrigel-coated plates (BD) with mTeSR1 growth media (Stem Cell Technologies) at 37 °C and 5% CO_2 were removed by treatment with collagenase IV (Sigma). Whole colonies were grown on ultra low–attachment plates in DMEM/F12 + GlutaMAX supplemented with N2 for 1 week as embryoid bodies, before being transferred onto polyornathine/laminin-coated plates and DMEM/F12 + GlutaMAX supplemented with N2 and FGF-2 (20 ng ml−1). After 1 week, neural rosettes were manually picked, enzymatically dissociated and re-plated in DMEM/F12 + GlutaMAX with N2, B27-RA and FGF-2. To differentiate the rosette derived NPCs into mature neurons, 10 μM ROCK inhibitor (Y27632) was added to the media for 2 d and cultured for an additional 4 weeks without FGF-2. To generate iPS cells, human fibroblasts from a non-ALS individual were subjected to infection by retroviruses harboring the reprogramming factors c-Myc, Klf4, Oct4 and Sox2 and replated

onto mouse embryonic fibroblasts after 5 d. Colonies with ES cell–like morphology, including well-defined borders, high nuclear to cytoplasmic ratio, and phase-bright characteristics were isolated after 4 weeks of reprogramming and propagated separately on Matrigel-coated plates and mTeSR1 growth media.

Lentiviral shRNA–mediated depletion of human TDP-43 and FUS/TLS. Lentiviral shRNA constructs (Open Biosystems) toward *TDP-43* (TRCN0000016038 and TRCN0000016041) and *FUS/TLS* (TRCN0000010450) in the pLKO.1 vector system was used to produce lentivirus as previously described[53]. The control virus used was pLKO.1 containing a shRNA targeting *GFP*. The efficacy of the lentivirus was tested by infection of NPCs at varying viral concentrations and qRT-PCR quantitation of *TDP-43* and *FUS/TLS* 5 d after infection.

qRT-PCR in human neuronal progenitor cells and differentiated mature neurons. qRT-PCR was performed on 7900HT Fast Real-time PCR system (Applied Biosystems) with Fast SYBR-Green PCR Master Mix (Applied Biosystems) in triplicate from polyA-selected, reverse-transcribed cDNA from control, FUS/TLS-depleted and TDP-43–depleted human neural progenitor cells and differentiated mature neurons. Gene expression values were normalized using either *SMYD5*, *BMI* or *GAPDH* levels.

Human tissue acquisition and processing. Human CNSs were obtained and archived via an institutional review board and Health Insurance Portability and Accountability Act-compliant process. The ALS tissues were obtained from patients who met the modified El Escorial criteria for definite ALS[60]. The diagnosis was confirmed post-mortem through histologic analysis of TDP-43 aggregation[2,3]. Mutations in *SOD1*, *C9orf72*, *TDP-43* and *FUS/TLS* genes were excluded. Control tissues were obtained from non-neurologically affected patients out of the hospital's critical care unit when life support was withdrawn, or patients on hospice. Autopsies were performed immediately following death, with an average post-mortem interval of 4 h. For histological studies, gross segments were fixed in 10% neutral buffered formalin (wt/vol) for 14 d, embedded in paraffin, and stored at 20–25 °C. For biochemical analysis, tissue segments were embedded in optimal cutting temperature medium (Tissue Tek, 4583) and frozen.

Immunofluorescence from human autopsy samples. We performed double immunofluorescence on formalin-fixed, paraffin-embedded lumbar spinal cord sections from SALS or control patients using an antibody to TDP-43 and one to either Parkin, KCNIP4, SMYD3 or tubulin. The slides were deparaffinized, rehydrated through a descending ethanol wash series, permeabilized by immersion in 0.2% Triton X-100 (vol/vol) for

10 min, and washed twice in phosphate-buffered saline (PBS). Antigen retrieval was performed in ScyTek Citrate Buffer pH 6.0 (CBB999) for 20 min at 125 °C in a Dako Pascal pressure chamber. After washing and blocking the slides were incubated overnight with one of the following primary antibodies: TDP-43 (CosmoBio, TIP-PTD-P01, phospho–TDP-43 rabbit polyclonal IgG, 1:50; in-house FL4 antibody, mouse monoclonal IgG, dilution of ascites 1:1,000), Parkin (R&D, AF1438, goat polyclonal IgG, 1:50), KCNIP4 (Sigma, HPA022862, rabbit polyclonal IgG, 1:200), SMYD3 (Abcam, ab16027, rabbit polyclonal IgG, 1:200) and tubulin (Sigma Aldrich, T8660, mouse monoclonal IgG, 1:1,000). Secondary antibodies (Alexa 488–conjugated goat antibody to rabbit, Cy3-conjugated donkey antibody to mouse, Cy3-conjugated donkey antibody to rabbit, Alexa 488–conjugated donkey antibody to goat or Cy5-conjugated donkey antibody to mouse) were diluted directly in PBS and applied at a 1:500 dilution ratio. Slides were stained using 10 µg ml^{-1} DAPI in ddH$_2$0 for 10 min, rinsed twice in ddH$_2$0, autofluorescence-masked in 0.1% Sudan Black (vol/vol, Fisher, BP109-10) dissolved in 70% ethanol for 10 s, and rinsed twice in PBS. Visualization and imaging was performed using a Nikon Eclipse Ti confocal microscope system. Motor neurons were counted and classified upon the presence or absence of TDP-43 inclusion and the qualitative difference of staining for Parkin, SMYD3, KCNIP4 or tubulin. The result presented in **Figure 6** is the average for each category of motor neurons determined by two independent investigators.

51. Parkhomchuk, D. *et al.* Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res.* **37**, e123 (2009).
52. Zhang, C. & Darnell, R.B. Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.* **29**, 607–614 (2011).
53. Yeo, G.W. *et al.* An RNA code for the FOX2 splicing regulator revealed by mapping RNA-protein interactions in stem cells. *Nat. Struct. Mol. Biol.* **16**, 130–137 (2009).
54. Zisoulis, D.G. *et al.* Comprehensive discovery of endogenous Argonaute binding sites in *Caenorhabditis elegans*. *Nat. Struct. Mol. Biol.* **17**, 173–179 (2010).
55. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
56. Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–628 (2008).
57. Sugnet, C.W. *et al.* Unusual intron conservation near tissue-regulated exons found by splicing microarrays. *PLOS Comput. Biol.* **2**, e4 (2006).
58. Yeo, G.W. *et al.* Alternative splicing events identified in human embryonic stem cells and neural progenitors. *PLOS Comput. Biol.* **3**, 1951–1967 (2007).
59. Marchetto, M.C. *et al.* A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143**, 527–539 (2010).
60. Brooks, B.R., Miller, R.G., Swash, M. & Munsat, T.L. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* **1**, 293–299 (2000).