SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
MOLECULAR NEUROSCIENCE
NON-CODING RNA'S
GENE REGULATION
MRNA

Received
5 April 2012

Accepted
9 July 2012

Published
24 July 2012

Correspondence and requests for materials should be addressed to G.S. (sobueg@med.nagoya-u.ac.jp) or K.O. (ohnok@med.nagoya-u.ac.jp).

* These authors contributed equally to this work.

Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions

Shinsuke Ishigaki[1,4]*, Akio Masuda[2]*, Yusuke Fujioka[1], Yohei Iguchi[1], Masahisa Katsuno[1], Akihide Shibata[2], Fumihiko Urano[3], Gen Sobue[1,4] & Kinji Ohno[2]

[1]Department of Neurology, [2]Division of Neurogenetics, Center for Neurological Diseases and Cancer, Nagoya University Graduate School of Medicine, Nagoya, Japan, [3]Program in Gene Function and Expression, University of Massachusetts Medical School, Worcester, MA, [4]CREST, Japan Science and Technology Agency, Kawaguchi, Japan.

FUS is an RNA-binding protein that regulates transcription, alternative splicing, and mRNA transport. Aberrations of FUS are causally associated with familial and sporadic ALS/FTLD. We analyzed FUS-mediated transcriptions and alternative splicing events in mouse primary cortical neurons using exon arrays. We also characterized FUS-binding RNA sites in the mouse cerebrum with HITS-CLIP. We found that FUS-binding sites tend to form stable secondary structures. Analysis of position-dependence of FUS-binding sites disclosed scattered binding of FUS to and around the alternatively spliced exons including those associated with neurodegeneration such as *Mapt*, *Camk2a*, and *Fmr1*. We also found that FUS is often bound to the antisense RNA strand at the promoter regions. Global analysis of these FUS-tags and the expression profiles disclosed that binding of FUS to the promoter antisense strand downregulates transcriptions of the coding strand. Our analysis revealed that FUS regulates alternative splicing events and transcriptions in a position-dependent manner.

Amyotrophic lateral sclerosis (ALS) is one of the most devastating neurodegenerative disorders characterized by loss of motor neurons in the spinal cord and motor cortex. Approximately 10–20% of ALS patients have a family history. Among them several genes have been identified as a cause or risk-factor for ALS. Interestingly, recently identified ALS-related genes including *TARDBP* encoding the TAR DNA-binding protein (TDP-43) and *FUS* encoding the fused in sarcoma are RNA-binding proteins that regulate RNA metabolisms including gene transcription, RNA splicing, and mRNA transport[1,2]. Mutations in *FUS* have been identified in familial ALS (ALS6). In addition, aberrations of FUS have been linked to the pathogenesis of familial and sporadic ALS as well as frontotemporal lobar degeneration (FTLD)[3,4]. While it remains elusive how aberrations of FUS are linked to the pathogenesis of ALS and FTLD, several lines of evidence suggest that the loss-of-function mechanism is, at least partly, involved in FUS-associated neurodegeneration. First, in autopsied spinal cords and brains of sporadic and familial ALS as well as of FTLD, FUS redistributes into the cytoplasm from the nucleus and accumulates in ubiquitin-positive inclusions[3–5]. Second, artificially introduced mutant FUS proteins in cultured cells are distributed to the cytoplasm from the nucleus[6–9]. This is in contrast to the physiological condition where FUS is localized in the nucleus. Finally, the loss of FUS directly leads to neuronal cell death in drosophila[10] and zebrafish[11].

FUS is a multifunctional protein that is involved in several steps of gene expression regulation especially for transcription and RNA splicing. FUS belongs to the FET family of RNA-binding proteins, which includes FUS, Ewing's sarcoma (EWS), and TATA-binding protein-associated factor (TAFII68)[12]. FUS is present in polymerase II transcription complexes that function in the transcription process[13]. Interestingly, FUS inhibits the acetyltransferase activities of CREB-binding protein (CBP) and p300 on cyclin D1 (*Ccnd1*) in HeLa cells. Both CBP and p300 are co-activators of multiple classes of signal-dependent transcription factors and the TLS/CBP/p300 interactions result in the inhibition of histone-acetyltransferase (HAT) activities followed by repression of transcription[10]. Moreover, FUS is involved in the splicing machinery by cooperating with other splicing factors like serine/arginine-rich (SR) proteins (SC35, SRp75, and TLS-associated SR protein), SRm160, and PTB[14,15]. These observations suggest that compromised effects of FUS on transcription and alternative splicing could lead to neuronal cell degeneration in FUS-associated ALS and FTLD.

In an effort to understand the global roles of FUS on RNA metabolisms in neuronal cells, we analyzed exon arrays of primary cortical neurons after knocking down *Fus*. We also identified FUS-binding RNA segments in the mouse brain using the high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP)[16]. Our analysis revealed the global profiles of FUS-mediated regulations of transcription and pre-mRNA splicing, and the position dependence of FUS-binding sites in regulations of transcription and alternatively splicing.

Results

Silencing of *Fus* in primary cortical neurons. In an effort to identify global profiles of FUS-mediated gene expression and alternative splicing in neurons, we introduced lentivirus expressing shRNA against Fus into primary cortical neurons derived from E15 mouse embryos (Fig. 1a). To exclude possible off-targeting effects, we used two different shRNA, shFus1 and shFus2, and the experiments were performed in triplicate for shFus1 and shFus2 (Fig. 1b). The expression levels of *Fus* were suppressed by more than 80% in both shFus1 and shFus2 by real-time qPCR (Fig. 1c). Immunohistochemistry also showed that the protein levels of FUS were markedly decreased in primary cortical neurons infected with shFus1 and shFus2 (Fig. 1d).

We analyzed gene expression and alternative splicing of *Fus*-silenced primary cortical neurons using the Affymetrix Mouse Exon 1.0 ST Array (GEO accession number, GSE36153). Scatter plots of fold-changes (FC) of gene expressions in cortical neurons showed a correlation coefficient of 0.69 (data not shown). Filtering the gene-level signal intensities with the t-test p-value ≤ 0.1 increased the correlation coefficient to 0.96 (Supplementary Fig. S1a). We further restricted our analysis to 183 genes that had the t-test p-value ≤ 0.05 and FC ≤ 0.67 or $1.5 \leq$ FC for both shFus1 and shFus2. The 21 most altered genes were validated by real-time qPCR and were all indeed differentially expressed in shFus1 and shFus2 (Supplementary Fig. S1c).

We also filtered the exon-level signal intensities with the t-test p-value ≤ 0.1, which gave rise to 3202 exons that were altered by both shFus1 and shFus2 with a correlation coefficient of 0.69 (Supplementary Fig. S1b). We chose 44 exons that were increased by 1.3-fold or more by both shFus1 and shFus2. Similarly, we chose 55 exons that were decreased by 1.5-fold or more by shFus1 and shFus2. We validated 17 exons with increased signals and 20 exons with decreased signals, and all showed altered splicing events by RT-PCR. Especially, we observed altered splicing events in genes associated with neuronal functions and neurodegeneration including *Mapt*, *Camk2a*, and *Fmr1* (Fig. 2 and Supplementary Fig. S2).

We next analyzed Gene Ontology (GO) terms of genes that were regulated by FUS using DAVID 6.7[17,18]. We found that genes for which FUS regulates the expression levels were enriched with GO terms of signaling cascades and metabolic processes. Similarly, genes for which FUS regulates alternative splicing events were enriched with GO terms of vesicle transport, neuronal impulse, and neuronal projection (Table1). FUS is thus likely to exert neuron-specific gene regulations by modulating alternative splicing events rather than gene expressions.

Global identification of FUS-binding RNA segments in mouse brain. We next identified FUS-binding RNA segments *in vivo* by HITS-CLIP (Fig. 1d). As the number of primary cortical neurons that we could obtain from E18 embryos was too small for the HITS-CLIP analysis, we used the mouse cerebrum. In the HITS-CLIP analysis, our first experiment yielded 47,647,724 CLIP tags of 50 nt, of which 17,703,863 were mapped to the mm9 genome using the default parameters with BioScope 1.3.1. A second CLIP experiment yielded 36,282,895 CLIP tags of 50 nt, of which 9,295,669 were mapped to the mouse genome. We only analyzed reads that were aligned uniquely in the genome and removed all potential PCR duplicates using the Avadis NGS software.

Mapping FUS CLIP-tags onto the ENSEMBL annotations revealed that a large fraction of FUS-binding regions were located in introns (Fig. 3a), as previously described[19]. As introns and interegenic regions are longer than the other regions, we normalized distribution of CLIP-tags for the length of each annotation, and found that FUS-binding regions were concentrated in 3′ UTRs as well as introns (Fig. 3b). When CLIP-tags were mapped to the relative positions of each gene, CLIP-tags were enriched close to the 3′ end of a gene and toward the 5′ end of a gene (Fig. 3c). In addition, CLIP-tags were enriched in genes with alternative transcription start/end sites compared to those with constitutive transcription start/end sites.

We next analyzed FUS-binding motifs with the BioProspector software[20], but detected variable motifs that widely varied even by adding, eliminating, or modifying combinations of parameters. Another motif analysis software, MEME[21], also predicted highly variable motifs. Hoell and colleagues reported that FUS binds to stem-and-loop structures with only three loosely conserved nucleotides at the boundaries of a stem and a loop[19]. We thus examined if our CLIP-tagged regions readily form secondary structures. We extracted 30-mer RNA stretches of CLIP-tagged regions and calculated a free energy, δG, of the most stable secondary structure using the mfold program[22]. As expected, the δG's of CLIP-tagged regions were -2.06 ± 3.10 kcal/mol (mean ± SD), whereas those of the controls were -1.50 ± 2.74 kcal/mol ($p < 0.0001$) (Fig. 3d). Stable structures with $\delta G \leq -14.0$ kcal/mol were observed in 28 CLIP-tagged regions and 3 control regions, and the predicted secondary structures are shown in Supplementary Fig. S4. Thus, our CLIP-tagged regions formed stable secondary structures compared to the control regions.

Analysis of position dependence of FUS-binding to splicing targets. FUS-tags were clustered in alternatively spliced splice sites rather than constitutively spliced splice sites, which suggests the essential roles of FUS on regulations of alternative splicing (Fig. 4a). Both exonic and intronic regions were similarly tagged by FUS, but regions around the 3′ and 5′ splice sites were less prominently recognized by FUS, as previously reported[19].

We next analyzed position dependence of FUS-binding to splicing targets and their effects on alternative splicing by comprehensively analyzing the exon array and HITS-CLIP. We analyzed positions of CLIP-tags of 37 FUS-responsive exons that we confirmed by RT-PCR: 17 exons were included and 20 exons were skipped by shRNA (Fig. 2, Supplementary Fig. S2). We combined these exons into a single composite pre-mRNA and made integrated RNA maps from our HITS-CLIP reads mapped to the corresponding genomic regions, as previously described[16,23,24]. The analysis revealed that FUS binding sites were scattered around the alternatively spliced exons (Fig. 4b). Among these, however, it was interesting to note that conspicuous bindings of FUS were observed at ~500 nt upstream of the 3′ end of the downstream intron (arrows in Fig. 4b). The similar peaks were observed when we extended our analysis to 78 FUS-responsive exons detected by the exon array analysis: 54 exons were skipped and 24 exons were included by shRNA (Supplementary Fig. S3). Bindings of FUS to the downstream introns were representatively observed in genes that we validated alternative splicing by RT-PCR (Fig. 2).

***Fus*-binding to the antisense strand at the promoter region downregulates gene expression.** We noticed that CLIP-tags were often located on the antisense strand in each promoter region. A previous report that FUS binds to noncoding RNA (ncRNA) and downregulates transcription of *Ccnd1*[10] prompted us to hypothesize that a transcript arising from the promoter antisense strand is the identity of the ncRNA recognized by FUS. We thus examined if binding of FUS to the promoter antisense strand downregulates

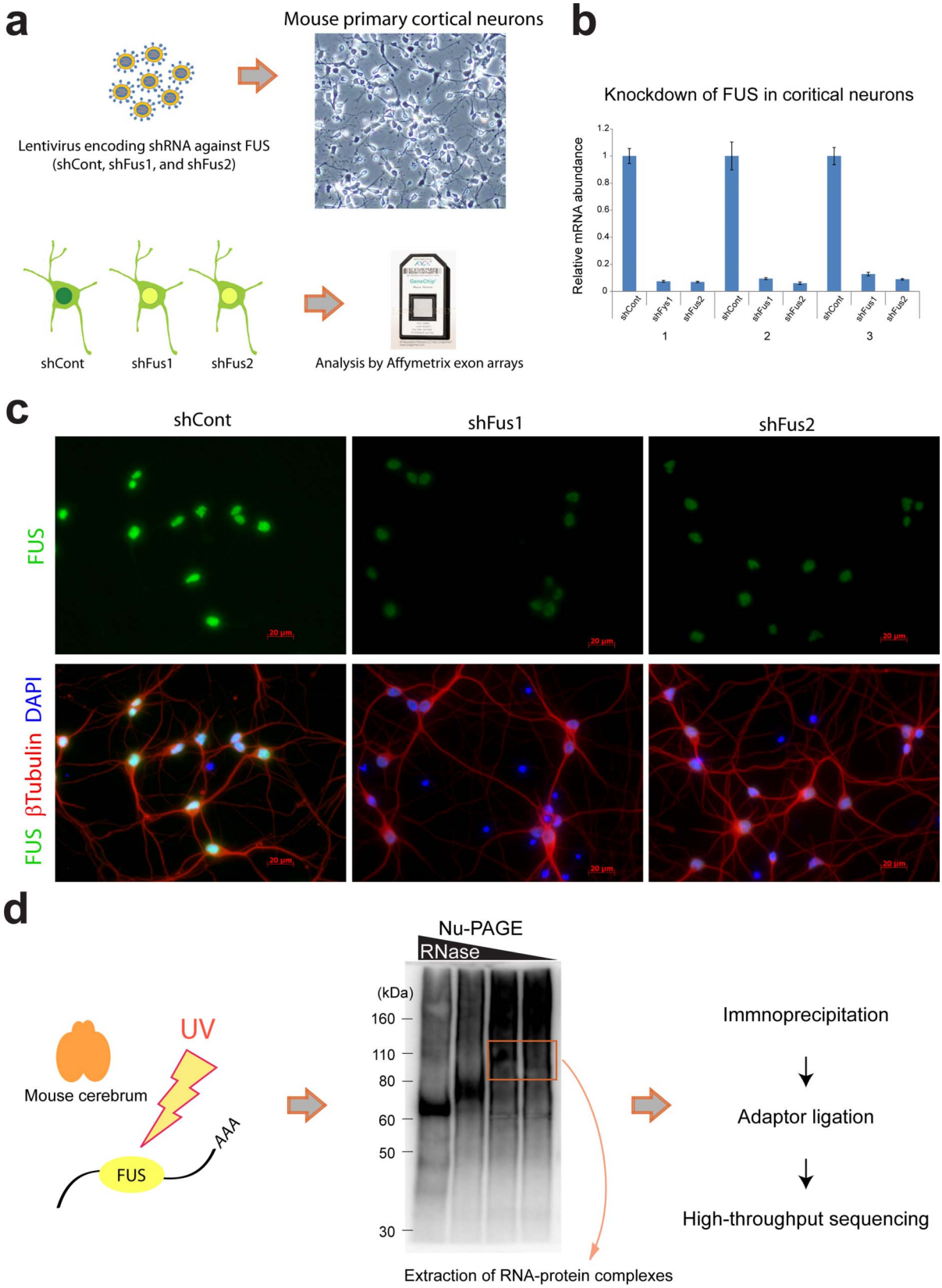

Figure 1 | Experimental schemes. (a) Mouse primary cortical neurons are prepared and infected with lentivirus expressing two different shRNA against FUS (shFus1 and shFus2) and control shRNA (shCont). Total RNA is isolated and analyzed by the Affymetrix Mouse Exon Array. (b) *Fus* is efficiently knocked down in primary cortical neurons, which is evaluated by real-time qRT-PCR. Bars indicate the mean and SD of three experiments. (c) Immunohistchimical analysis using anti-FUS antibody on primary cortical neurons silenced by shFus1, shFus2, and shCont. Cells are fixed and immunostained with anti-FUS antibody, anti-βTubulin antibody, and DAPI. (d) Mouse cerebrum derived from a 12-week-old C57Bl/6 mouse is UV-irradiated at 400 mJ and FUS-bound RNA segments are immuno-precipitated. High-throughput 50 bp single-end sequencing is performed using the SOLiD 3 sequencer.

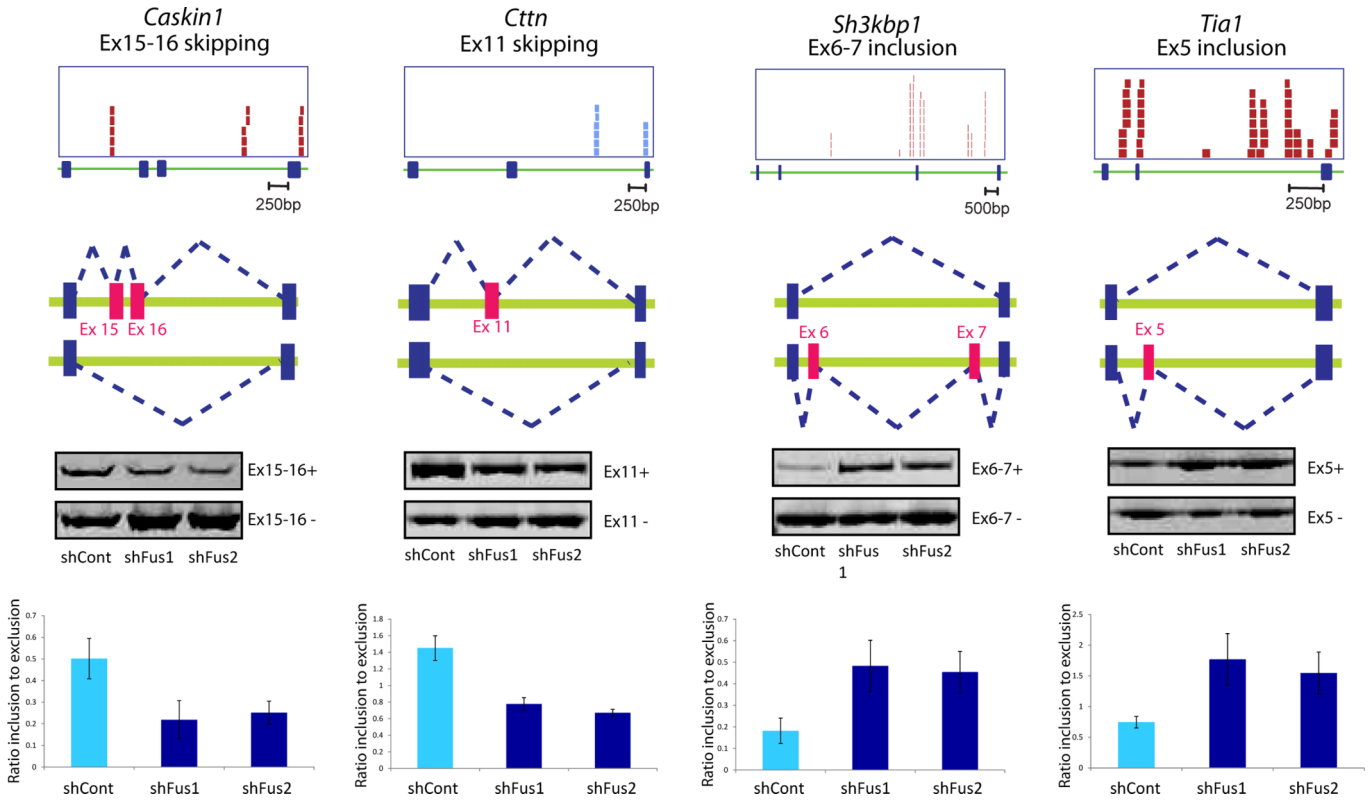

Figure 2 | Four representative FUS-mediated alternative splicing events. The top panels show the positions of CLIP-tags and exon-intron structures. The second panels represent schematic splicing changes mediated by FUS. shCont and shFus lead to the upper and lower splicing events, respectively. The third panels show representative RT-PCR of the indicated exons. The experiments are repeated in quadruplicate using four independent sets of samples. The last panels show densitometric quantification of RT-PCR (n = 4; mean and SD).

gene expression. We analyzed the fold-changes of gene expressions after silencing FUS in primary cortical neurons in the absence or presence of CLIP-tags on the antisense strand upstream of the transcription start sites, and found that the binding of FUS to the promoter antisense strand indeed increases expression of the target genes (Fig. 5a). Eighteen transcriptional start sites of 17 genes carried more than 250 CLIP tags on the promoter antisense strand, and we validated shFus-induced upregulation of transcriptions of four representative genes by realt-ime qPCR (Fig. 5b). The 17 genes, however, did not increase the expression levels more than 1.5-fold after silencing *Fus* according to the exon array analysis, and were not included in our initial validation of the eleven genes in which expression levels were increased with shFus (Supplementary Fig. S1c). On the contrary, the eleven genes scarcely had CLIP-tags on the promoter antisense strand, which suggests the presence of additional and possibly indirect mechanisms that regulate FUS-mediated gene expressions. We also found that genes in which FUS binds to the promoter antisense strand were enriched with GO terms related to reproductive process (Table 1).

Discussion

We globally analyzed FUS-mediated regulations of gene expressions and alternative splicing events using exon arrays. In addition, we globally mapped RNA-FUS interactions *in vivo* by HITS-CLIP and analyzed position dependence of FUS-binding to gene expressions and alternative splicing events.

Collation of the exon-level analysis of exon arrays and CLIP-tags enabled us to draw a normalized complexity map (Fig. 4b). The map disclosed scattered binding of FUS to the upstream and downstream introns with conspicuous binding peaks close to the 3′ end of the downstream intron. A similar peak close to the 3′ end of the downstream intron is observed with PTB and MBNL1 but not with CUGBP1[23,24]. The underlying mechanisms shared by these RNA-binding molecules, however, remain elusive.

Although we could not detect FUS-binding consensus motifs in our CLIP-tags, we found that FUS-binding regions readily form secondary structures. Our observation is consistent with previous reports that FUS binds to stem-and-loop structures[19] and that FUS has a weakly enriched motif endowed with G/C nucleotides, which is present in less than 10% of the FUS-bound sites[25]. *In vitro* SLELX analysis determined that FUS-binding motif is GGUG[26,27]. *Nd1-L*, however, has no GGUG motif but is able to specifically bind to FUS[28]. We assume that FUS binds to specific RNA targets with specific secondary structures, and simple analysis of primary sequences is unlikely to be sufficient.

Divergent transcriptions including bidirectional transcriptions at the promoter regions of protein-coding genes are widely recognized[29,30], although its biological significance is poorly understood. Collation of the gene-level analysis of exon arrays and CLIP-tags on the antisense strands at the promoter regions revealed that binding of FUS to the promoter antisense strand downregulates transcriptions of the coding sense strand (Fig. 5). FUS interacts with CBP and p300 in the presence of ncRNA and inhibits HAT followed by repression of transcription of *Ccnd1*, but the origin of ncRNA was not scrutinized[10]. Our global analysis suggests that binding of FUS to the promoter antisense strand and the subsequent downregulation of transcription of the sense-coding strand is likely to be instrumental in some but not all genes. Recently, Tan and colleagues reported that FUS binds to single-strand DNA at the promoter region, and up- or down-regulates transcriptions[31]. In our CLIP experiments, we needed to degrade DNA using DNase to make a precipitate with anti-FUS antibody. Lack of DNA in our RNA-protein complex was also indicated by that a high concentration of RNase decreased the molecular weight of the RNA-protein complex to that of FUS alone

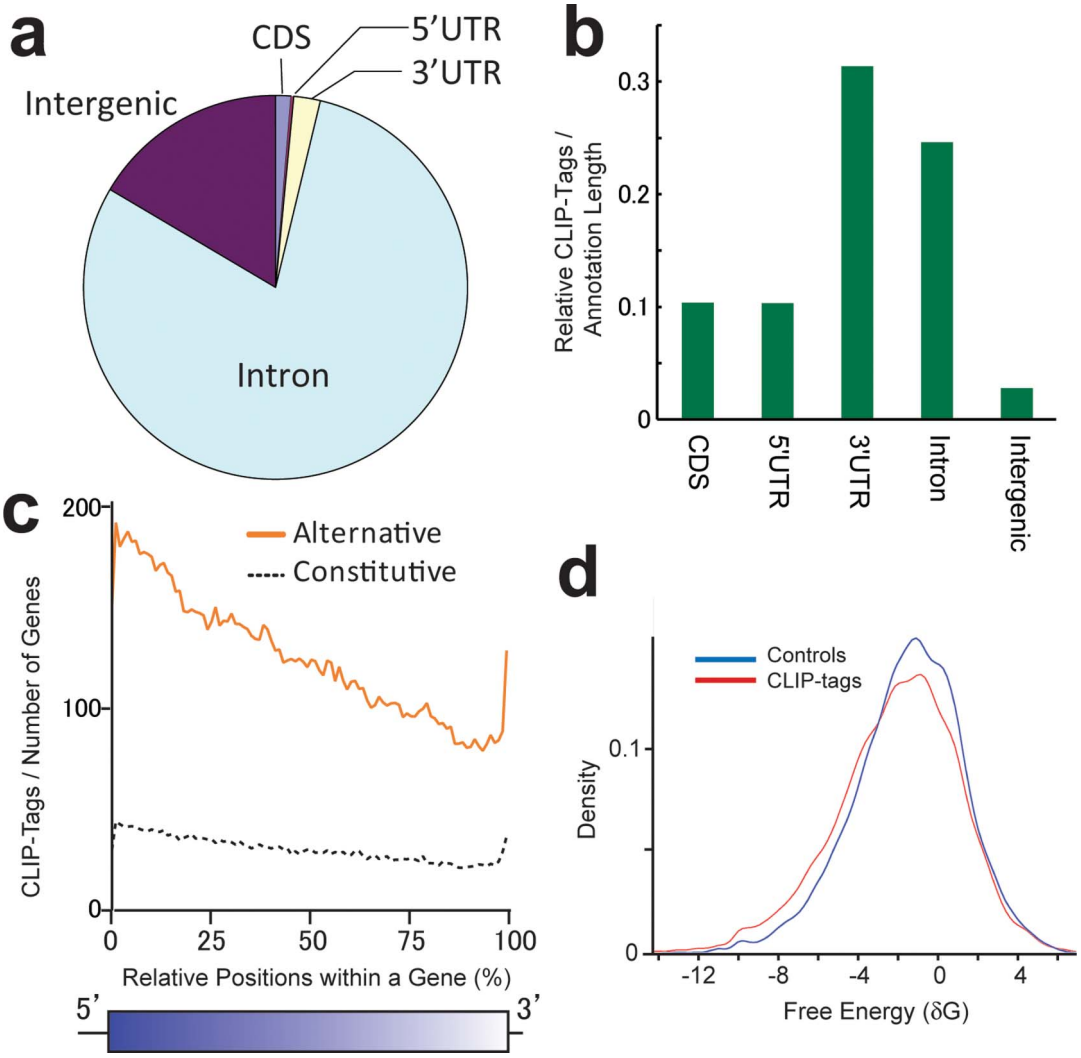

Figure 3 | Annotation mapping of FUS CLIP-tags. (a) Distributions of FUS CLIP-tags. Binding regions are mapped to CDS (coding sequence), 5′ and 3′ UTRs, introns, intergenic regions including tRNA and rRNA genes according to the ENSEMBL version e!61 annotation based on the mouse genome assembly NCBI build 37.1/mm9. Pie-charts show ratios of binding regions mapped to the indicated regions. (b) Distribution of FUS CLIP-tags normalized for the length of each annotation. (c) Distribution of FUS CLIP-tags mapped to the relative positions of each gene. The broken line indicates 12,508 genes with constitutive transcriptional start/end sites, and the solid line indicates 7,477 genes with alternative transcriptional start/end sites. (d) Probability density function of the minimum free energies, δG, of 30-mer stretches of CLIP-tagged regions. CLIP-tagged regions and controls are shown in red and blue lines, respectively.

(Fig. 1d). The underlying mechanisms shared by FUS-binding to single-strand DNA and to the promoter antisense RNA strand need to be further studied.

Involvement of TDP43 and FUS in the pathogenesis of ALS and FTLD suggests that the two diseases are likely to be caused by aberrations of RNA metabolisms[32]. Four additional RNA-binding proteins are causally associated with ALS: senataxin (*SETX*), angiogenin (*ANG*), elongation protein 3 (*ELP3*), and survival motor neuron (*SMN*)[33]. In the present study, we identified that FUS facilitates skipping of *Mapt* exon 10 in primary cortical neurons (Supplementary Fig. S2). *Mapt* encodes Tau protein and inclusion of exon 10 yields 4-repeat Tau (RD4), whereas skipping of exon 10 generates 3-repeat Tau (RD3). It has been reported that the RD4-to-RD3 ratio is increased in neurodegenerative disorders including PSP and FTLD[34,35]. Recently, a large hexanucleotide repeat expansion in intron 1 of *C9ORF72* has been reported in both familial ALS and familial FTLD[36,37]. Abnormally expanded repeats sequestrate MBNL1 in myotonic dystrophy[38], spinocerebellar ataxia type 8[39], and Huntington′s disease-like 2[40], and cause RNA gain-of-function pathologies, in which MBNL1 is sequestrated to abnormally expanded repeats, which compromises physiological functions of MBNL1. Thus, aberrations of RNA metabolisms are likely to be a common underlying mechanism shared by familial and sporadic ALS/FLTD. We hope that the global expression profiling and the global CLIP-mapping of FUS in our studies further facilitate discovery of the underlying pathophysiology leading to ALS/FLTD.

Methods

Lentivirus. We designed two different shRNAs against mouse *Fus* as well as a control shRNA. The targeted sequences were 5′-GCAACAAAGCTACGGACAA-3′ for shRNA/FUS1 (shFus1); 5′-GAGTGGAGGTTATGGTCAA-3′ for shRNA/FUS2 (shFus2); and 5′-AATTCTCCGAACGTGTCACGT-3′ for shRNA/control (shCont). These were cloned into a lentiviral shRNA vector (pLenti-RNAi-X2 puro DEST, w16-1, a kind gift from Dr. Eric Campeau at Resverlogix Corp.). Lentivirus was prepared following the Campeu's protocols[41]. Briefly, lentiviral particles were produced in HEK293T cells by transfection using Lipofectamine 2000 (Invitrogen). Lentivirus-containing supernatant was collected at 48 hours after transfection, and stored at -80°C. A titer of lentivirus was measured using NucleoSpin RNA Virus kit (Clontech Laboratories).

Primary cortical neurons. Mouse studies were approved by the Animal Care and Use Committee of the Nagoya University Graduate School of Medicine. Mouse fetal brain

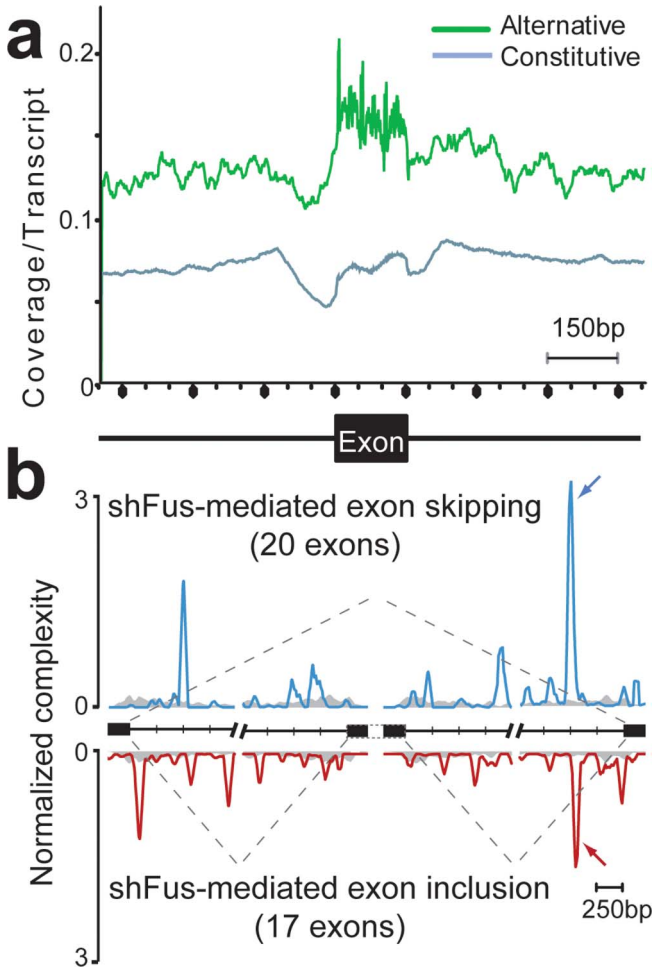

Figure 4 | Mapping of CLIP-tags on exon-intron structures. (a) Distribution of CLIP-tags on constitutively or alternatively spliced exons and the flanking intronic regions. The abscissa indicates an intron-exon-intron structure. The sizes of all the exons are normalized to 150 nucleotides. The number of exonic CLIP-tags is also normalized accordingly. Intronic CLIP-tags within 1,000 nucleotides upstream or downstream of exons are indicated. The number of CLIP-tags is normalized for the number of transcripts belonging to each category of constitutive and alternative exons. (b) Normalized complexity map of FUS-dependent splice sites. shFus-mediated alternative splicing events are compiled. Arrows point to conspicuous peaks at ~500 nt upstream of the 3′ end of the downstream intron. Shaded areas indicate an average of 100 sets of normalized complexity of 20 randomly selected constitutive exons.

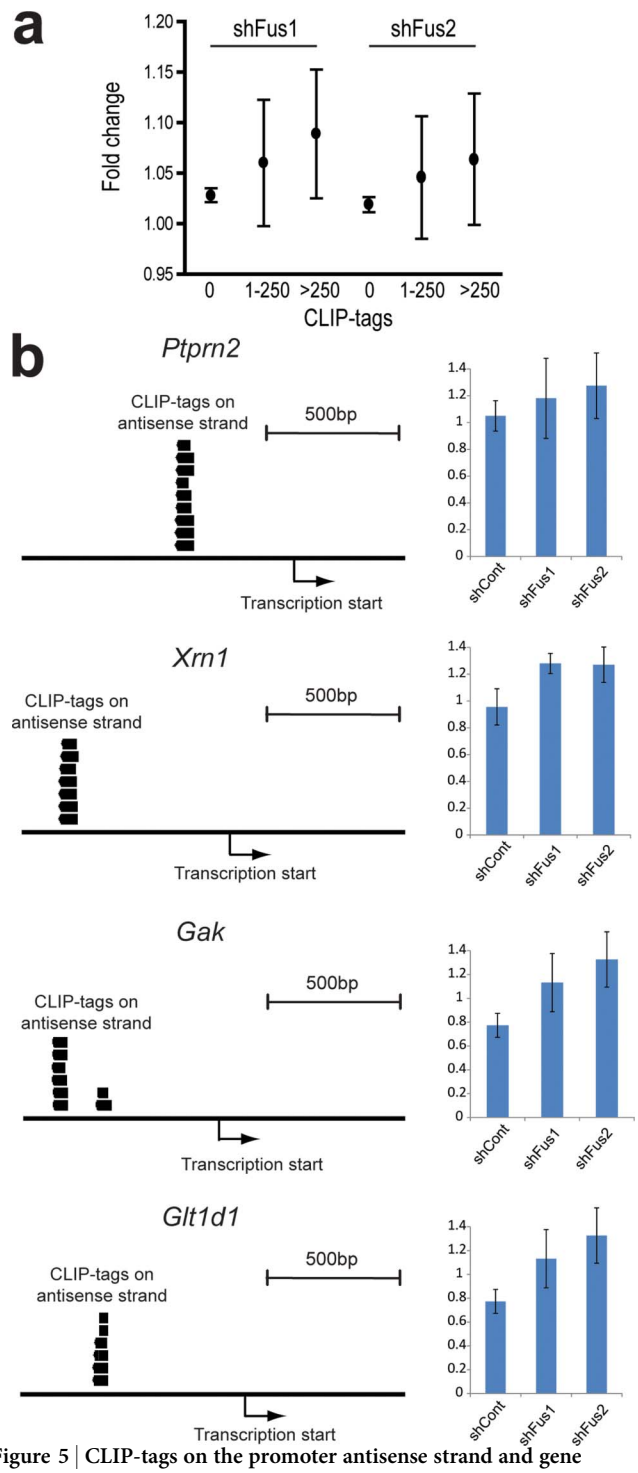

Figure 5 | CLIP-tags on the promoter antisense strand and gene expression profiles. (a) The numbers of CLIP-tags on the antisense strand at 1 to 700 nucleotides upstream of the transcription start sites of each gene are divided into three categories according to the partitioning functionality of the JMP 8.0 software. Fold-changes of gene expression levels with shFus1 and shFus2 are calculated for each category. Means and SEs are plotted. The numbers of transcription start sites are 2390 for no tag, 53 for 1-250 tags, and 18 for more than 250 tags. The number of CLIP-tags represents the total number of nucleotides covered by the tags. No statistical difference is observed for each dataset with the one-way ANOVA analysis. (b) Four representative genes, *Ptprn2*, *Xrn1*, *Gak*, and *Glt1d1*, for which more than 250 CLIP-tags are bound to the promoter antisense strand, are validated by real-time qPCR. Changes in gene expression levels in cortical neurons after knocking down *Fus* with shFus1 and shFus2 are indicated by means and SDs (n = 3).

was taken from C57BL/6 mouse embryos at E15. After removing meninges cortical tissue was dissociated into a single-cell suspension by Sumilon dissociation solution (Sumitomo Bakelite, Tokyo Japan). Cells were plated at a density of 1.5 x 10^6 cells in a 60 mm-well culture plate with the medium containing 0.5x Sumilon nerve-culture medium (Sumitomo Bakelite), 0.5x Neurobasal medium, 1% FBS, 0.5x B27 supplements (Invitrogen), 0.5x Glutamax, 5 μg/ml of BDNF, 5 μg/ml of CNTF, and 0.5% Pen-Strep. A day after plating (day 2), neurons were supplemented with 10 ng/μl of AraC and incubated overnight. On day 5, neurons were infected with 2x10^{10} copies/well (1.5x10^7 copies/μl) of lentivirus expressing shRNA against mouse *Fus* (shFus1 or shFus2) or control (shCont). After 4 hr of infection, the virus media was removed. Neurons were then cultured for 6 additional days, and were harvested on day 11 followed by RNA extraction and cDNA synthesis. Each knock-down experiment was performed in triplicate for each microarray analysis. For immunohistochemistry we used anti-FUS antibody (A300-293A, Bethyl Laboratories), anti-βTubulin antibody (TU20, Santa Cruz), and DAPI.

Microarray analysis. Total RNA was extracted from primary cortical neurons by the RNeasy Mini kit (Qiagen). We confirmed that the RNA integrity numbers (RIN) were all above 7.0. We synthesized and labeled cDNA fragments from 100 ng of total RNA using the GeneChip WT cDNA Synthesis Kit (Ambion). Hybridization and signal acquisition of the GeneChip Mouse Exon 1.0 ST exon array (Affymetrix) were

Table 1 | Gene Ontology terms enriched in FUS-associated genes

GO ID	GO_TERM_BP_FAT	P-value
Gene expression		
GO:0007242	Intracellular signaling cascade	0.0083
GO:0006796	Phosphate metabolic process	0.0115
GO:0006793	Phosphorus metabolic process	0.0115
GO:0007265	Ras protein signal transduction	0.0163
GO:0046486	Glycerolipid metabolic process	0.0296
Alternative splicing		
GO:0016192	Vesicle-mediated transport*	0.0077
GO:0006887	Exocytosis	0.0081
GO:0019226	Transmission of nerve impulse*	0.0101
GO:0016044	Membrane organization	0.0188
GO:0048812	Neuron projection morphogenesis*	0.0283
Promoter antisense strand		
GO:0032504	Multicellular organism reproduction	0.0060
GO:0048609	Reproductive process	0.0060
GO:0006350	Transcription	0.0092
GO:0007276	Gamete generation	0.0145
GO:0043933	Macromolecular complex organization	0.0205

GO terms with the five best P-values in each category are indicated.
*GO terms with neuron-specific functionalities.

performed according to the manufacturer's instructions. Each array experiment was performed in triplicate. The exon-level and gene-level signal intensities were normalized by the RMA and iterPLIER methods, respectively, using the Expression Console 1.1.2 (Affymetrix). We followed the gene annotation of the ENSEMBL version e!61, which was based on the mouse genome assembly NCBI build 37.1/mm9. All microarray data were registered in the Gene Expression Omnibus with an accession number of GSE36153.

We compared the gene-level signal intensities of three controls treated with shCont and three samples treated either with shFus1 or shFus2 using the Student's t-test. Among the 21,603 genes on the mouse exon array, 1,814 genes had the t-test p-values ≤ 0.10 for both shFus1 and shFus2, and the correlation coefficient of the fold-changes became 0.963 between shFu1 and shFus2 (Supplementary Fig. S1a). By applying the t-test p-values ≤ 0.01 for both shFus1 and shFus2, we obtained 1,500 genes, the expression levels of which were altered by knocking down *Fus*. We also analyzed alternative splicing profiles by filtering the exon-level signal intensities with the t-test p-value ≤ 0.1, which gave rise to 3,202 exons that were altered by both shFus1 and shFus2 with the correlation coefficient of 0.925 (Supplementary Fig. S1b).

RT-PCR for alternative splicing analysis. Total RNA was isolated from cells using RNeasy Mini Kit (Qiagen) followed by DNaseI (Qiagen) treatment. cDNA was synthesized from 1 μg of total RNA with the Oligo-dT primer (Promega). Primers for each candidate exon were designed using Primer3 software (http://frodo.wi.mit.edu/primer3/input.htm). The sequences of primers were shown in Table S1. Semi-quantitative RT-PCR was performed using Ex Taq (Takara) at 25-30 cycles at 98°C for 10 sec, 60°C for 30 sec, and 72°C for 1 min. PCR products was electrophoresed on a 15% acrylamide gel and stained with ethidium bromide. The intensity of each band was measured by Multi Gage software (Fujifilm).

Real-time qPCR for gene expression analysis. Total RNA was isolated from cells using RNeasy Mini Kit (Qiagen) and reverse transcribed using 1 μg of total RNA with Oligo-dT primer. For the thermal cycle reaction, the CFX96 system (BioRad) was used at 95°C for 3 min, then 40 cycles at 95°C for 10 sec and 55°C for 30 sec.

The relative amount for each transcript was calculated by drawing a standard curve of cycle thresholds for serial dilutions of cDNA samples and normalized to the amount of β-actin. The PCR was performed in triplicate for each sample, after which all experiments were repeated twice. The sets of primers (Table S1) and iQ SYBR Green Supermix (BioRad) were used for real-time qPCR.

HITS-CLIP analysis. Mouse cerebrum derived from 12-week-old C57BL/6 mice was UV-irradiated at 400 mJ and CLIP was performed as previously described[24]. High-throughput 50-bp single-end sequencing was performed with the SOLiD 3 sequencer (Life Technologies) using one quad of a SOLiD sequencing slide for each sample. All HITS-CLIP data were registered in NCBI SRA. Sequences of RNA oligonucleotide adaptors and PCR primers used for HITS-CLIP analysis were as follows:

5′-RNA linker 5′-CCACUACGCCUCCGCUUUCCUCUCUAUGGGCAGUCG-GUGAU-3′

3′-RNA linker phospho-5′-AGAGAAUGAGGAACCCGGGGCAGUU-3′-amino

RT Primer 5′- CCCGGGTTCCTCATTCTCTCGCCTTGGCCGTACAG-3′

Forward PCR Primer 5′- CCACTACGCCTCCGCTTTCCTCTCTATG-3′

Reverse PCR Primer 5′-CTGCCCCGGGTTCCTCATTCT-3′

Bioinformatics analysis. SOLiD reads were mapped to the mouse genome (NCBI build 37.1/mm9) with default parameters using the BioScope 1.3.1 (Life Technologies). We next removed multiply aligned reads, unreliable reads, and PCR duplicates with Avadis NGS software 1.3 (Strand). We analyzed the uniquely aligned reads according to the ENSEMBL version e!61 gene annotations of the mouse genome (NCBI build 37.1/mm9) by writing and running Perl and Excel VBA programs, as well as by running BEDTools utilities[42].

We extracted CLIP-tagged regions using the enrichment functionality of Avadis NGS. Motifs of the tagged regions were analyzed by BioProspector[20] and MEME[21]. As these motif analysis tools gave rise to highly variable motifs, we extracted 17,385 30-mer RNA stretches of the CLIP-tagged regions by moving a 30-mer window every five nucleotides, and calculated the free energy, δG, of the most stable secondary structure using the mfold program[22]. We similarly calculated the free energies of control sequences for which nucleotide sequences were scrambled from the 30-mer stretches of the CLIP-tagged regions.

We identified 2,461 distinct transcription start sites in the 1,500 genes, the transcription levels of which were significantly altered by both shFus1 and shFus2. We counted CLIP tags on the antisense strand at 1 to 700 nucleotides upstream of each transcription start site. We analyzed the expression profiles of primary cortical neurons after knocking down *Fus* in relation to the number of CLIP tags on the antisense strand at the promoter region.

Normalized complexity maps of FUS-RNA interactions were generated as previously described[24]. For the control, normalized complexity map was similarly generated by analyzing 100 sets of 20 constitutive exons that were randomly selected from 118,969 constitutive exons in the mouse genome. To identify enriched Gene Ontology terms, we used the Database for Annotation, Visualization and Integrated Discovery (DAVID 6.7)[17,18].

1. Strong, M. J. & Volkening, K. TDP-43 and FUS/TLS: sending a complex message about messenger RNA in amyotrophic lateral sclerosis? *FEBS J* 278, 3569–3577 (2011).
2. Lagier-Tourenne, C. & Cleveland, D. W. Rethinking ALS: the FUS about TDP-43. *Cell* 136, 1001–1004 (2009).
3. Kwiatkowski, T. J., Jr. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* 323, 1205–1208 (2009).
4. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323, 1208–1211 (2009).
5. Deng, H. X. *et al.* FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. *Ann Neurol* 67, 739–748 (2010).
6. Bosco, D. A. *et al.* Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. *Hum Mol Genet* 19, 4160–4175 (2010).
7. Dormann, D. *et al.* ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *EMBO J* 29, 2841–2857 (2010).
8. Ito, D., Seki, M., Tsunoda, Y., Uchiyama, H. & Suzuki, N. Nuclear transport impairment of amyotrophic lateral sclerosis-linked mutations in FUS/TLS. *Ann Neurol* 69, 152–162 (2011).
9. Kino, Y. *et al.* Intracellular localization and splicing regulation of FUS/TLS are variably affected by amyotrophic lateral sclerosis-linked mutations. *Nucleic Acids Res* 39, 2781–2798 (2011).
10. Wang, X. *et al.* Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454, 126–130 (2008).
11. Kabashi, E. *et al.* FUS and TARDBP but not SOD1 interact in genetic models of amyotrophic lateral sclerosis. *PLoS Genet* 7, e1002214 (2011).
12. Guipaud, O. *et al.* An in vitro enzymatic assay coupled to proteomics analysis reveals a new DNA processing activity for Ewing sarcoma and TAF(II)68 proteins. *Proteomics* 6, 5962–5972 (2006).
13. Bertolotti, A., Lutz, Y., Heard, D. J., Chambon, P. & Tora, L. hTAF(II)68, a novel RNA/ssDNA-binding protein with homology to the pro-oncoproteins TLS/FUS and EWS is associated with both TFIID and RNA polymerase II. *EMBO J* 15, 5022–5031 (1996).
14. Meissner, M., Lopato, S., Gotzmann, J., Sauermann, G. & Barta, A. Proto-oncoprotein TLS/FUS is associated to the nuclear matrix and complexed with splicing factors PTB, SRm160, and SR proteins. *Exp Cell Res* 283, 184–195 (2003).
15. Yang, L., Embree, L. J., Tsai, S. & Hickstein, D. D. Oncoprotein TLS interacts with serine-arginine proteins involved in RNA splicing. *J Biol Chem* 273, 27761–27764 (1998).
16. Licatalosi, D. D. *et al.* HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456, 464–469 (2008).
17. Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4, 44–57 (2009).
18. Dennis, G., Jr. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* 4, P3 (2003).
19. Hoell, J. I. *et al.* RNA targets of wild-type and mutant FET family proteins. *Nat Struct Mol Biol* 18, 1428–1431 (2011).
20. Liu, X., Brutlag, D. L. & Liu, J. S. BioProspector: discovering conserved DNA motifs in upstream regulatory regions of co-expressed genes. *Pac Symp Biocomput* 127–138 (2001).
21. Bailey, T. L. & Elkan, C. The value of prior knowledge in discovering motifs with MEME. *Proc Int Conf Intell Syst Mol Biol* 3, 21–29 (1995).

22. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31, 3406–3415 (2003).
23. Xue, Y. *et al.* Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol Cell* 36, 996–1006 (2009).
24. Masuda, A. *et al.* CUGBP1 and MBNL1 preferentially bind to 3′ UTRs and facilitate mRNA decay. *Sci Rep* 2, 209 (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* Epub (2012).
26. Lerga, A. *et al.* Identification of an RNA binding specificity for the potential splicing factor TLS. *J Biol Chem* 276, 6807–6816 (2001).
27. Iko, Y. *et al.* Domain architectures and characterization of an RNA-binding protein, TLS. *J Biol Chem* 279, 44834–44840 (2004).
28. Fujii, R. & Takumi, T. TLS facilitates transport of mRNA encoding an actin-stabilizing protein to dendritic spines. *J Cell Sci* 118, 5755–5765 (2005).
29. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848 (2008).
30. Seila, A. C. *et al.* Divergent transcription from active promoters. *Science* 322, 1849–1851 (2008).
31. 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 U S A* Epub (2012).
32. Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D. W. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum Mol Genet* 19, R46–64 (2010).
33. van Blitterswijk, M. & Landers, J. E. RNA processing pathways in amyotrophic lateral sclerosis. *Neurogenetics* 11, 275–290 (2010).
34. Hong, M. *et al.* Mutation-specific functional impairments in distinct tau isoforms of hereditary FTDP-17. *Science* 282, 1914–1917 (1998).
35. Yoshida, M. Cellular tau pathology and immunohistochemical study of tau isoforms in sporadic tauopathies. *Neuropathology* 26, 457–470 (2006).
36. Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257–268 (2011).
37. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256 (2011).
38. Ranum, L. P. & Cooper, T. A. RNA-mediated neuromuscular disorders. *Annu Rev Neurosci* 29, 259–277 (2006).
39. Moseley, M. L. *et al.* Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat Genet* 38, 758–769 (2006).
40. Rudnicki, D. D. *et al.* Huntington's disease--like 2 is associated with CUG repeat-containing RNA foci. *Ann Neurol* 61, 272–282 (2007).
41. Campeau, E. *et al.* A versatile viral system for expression and depletion of proteins in mammalian cells. *PLoS One* 4, e6529 (2009).
42. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842 (2010).

Acknowledgements

This work was supported by Grants-in-Aid from the MEXT/JST, MEXT/ICBIT, CREST/JST, SRPBS/JST, and MHLW of Japan. We are grateful to Dr. Kunio Ihara at the Center for Gene Research of Nagoya University for the SOLiD sequencing analysis.

Author contributions

S.I., A.M., Y.F., Y.I., and M.K. performed the experiments. S.I., A.M., A.S., and K.O. analyzed the data. S.I., A.M., F.U., G.S. and K.O. prepared the manuscript. All authors reviewed the manuscript.

Additional information

Accession code: The HITS-CLIP and microarray data were registered in the NCBI Gene Expression Omnibus database with accession numbers GSE37190 and GSE36153, respectively.

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Ishigaki, S. *et al.* Position-dependent FUS-RNA interactions regulate alternative splicing events and transcriptions. *Sci. Rep.* **2**, 529; DOI:10.1038/srep00529 (2012).