Europe PMC Funders Group
Author Manuscript
Nat Neurosci. Author manuscript; available in PMC 2011 October 01.

Published in final edited form as:
Nat Neurosci. 2011 April ; 14(4): 452–458. doi:10.1038/nn.2778.

Characterising the RNA targets and position-dependent splicing regulation by TDP-43; implications for neurodegenerative diseases

James R. Tollervey[1,*], **Tomaž Curk**[2,*], **Boris Rogelj**[3,*], **Michael Briese**[1], **Matteo Cereda**[1,4], **Melis Kayikci**[1], **Tibor Hortobágyi**[3], **Agnes L. Nishimura**[3], **Vera Župunski**[3,5], **Rickie Patani**[6], **Siddharthan Chandran**[6], **Gregor Rot**[2], **Blaž Zupan**[2], **Christopher E. Shaw**[3], and **Jernej Ule**[1]

[1]MRC Laboratory of Molecular Biology, Hills Road, Cambridge, CB2 0QH, United Kingdom

[2]Faculty of Computer and Information Science, University of Ljubljana, Tržaška 25, SI-1000, Ljubljana, Slovenia

[3]MRC Centre for Neurodegeneration Research, King's College London, Institute of Psychiatry, De Crespigny Park, London, SE5 8AF, United Kingdom

[4]Scientific Institute IRCCS E. Medea, Via don L. Monza 20, 23842 Bosisio Parini (LC), Italy

[5]Faculty of Chemistry and Chemical Technology, University of Ljubljana, Aškerčeva 5, SI-1000, Ljubljana, Slovenia

[6]MRC Laboratory for Regenerative Medicine, Dept of Clinical Neurosciences, Robinson Way, Cambridge CB2 0SZ, United Kingdom

Abstract

TDP-43 is a predominantly nuclear RNA-binding protein that forms inclusion bodies in frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS). The mRNA targets of TDP-43 in the human brain and its role in RNA processing are largely unknown. Using individual-nucleotide resolution UV-crosslinking and immunoprecipitation (iCLIP), we demonstrated that TDP-43 preferentially binds long clusters of UG-rich sequences *in vivo*. Analysis of TDP-43 RNA binding in FTLD-TDP brains revealed the greatest increases in binding to MALAT1 and NEAT1 non-coding RNAs. We also showed that TDP-43 binding on pre-mRNAs influences alternative splicing in a similar position-dependent manner to Nova proteins. In addition, we identified unusually long clusters of TDP-43 binding at deep intronic positions downstream of silenced exons. A significant proportion of alternative mRNA isoforms regulated by TDP-43 encode proteins that regulate neuronal development or are implicated in neurological diseases, highlighting the importance of TDP-43 for splicing regulation in the brain.

TAR DNA binding protein (TDP-43) is a predominantly nuclear protein reported to regulate transcription, alternative splicing and RNA stability[1,2]. TDP-43 has two RNA-recognition motif (RRM) domains, which bind UG repeat sequences with high affinity[3]. TDP-43 mislocalisation and aggregation is implicated in the pathogenesis of amyotrophic lateral

*equal contribution

Author contributions J.R.T. carried out TDP-43 iCLIP, microarray and PCR experiments, M.B. carried out CELF2 iCLIP, T.C. and G.R. mapped the iCLIP sequence reads to genome, evaluated random barcodes, determined crosslink clusters and annotated the data, T.C. analyzed the reproducibility, sequence and positioning of TDP-43 crosslink sites and performed gene ontology analysis, B.R., A. L.N. and V.Ž. prepared RNA from knockdown cells and brain tissue, T.H. collected the brain tissue, M.C. and M.K. analyzed splice-junction microarray data and generated the RNA splicing map, R.P. prepared the embryonic stem cells, S.C, S.E.S, B.Z. and JU supervised the project, J.R.T., T.C., B.R., C.E.S and J.U. prepared the manuscript.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts
IDT Exhibit 1029.0001

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

sclerosis (ALS), frontotemporal lobar degeneration (FTLD-TDP). Mutations in the gene encoding TDP-43 are associated with 1–4% of familial and sporadic ALS and rare FTLD cases, and confer toxicity[4,5]. The pathology of ~60% of FTLD and ~90% of ALS cases are characterised by the formation of cytoplasmic inclusion bodies in neurons and astroglia that are rich in ubiquitinated, hyperphosphorylated C-terminal fragments of TDP-43[5,6]. Neurons containing cytoplasmic TDP-43 inclusions often have dramatically reduced TDP-43 staining in the nucleus[5,6]. Thus, TDP-43 inclusions might have a toxic gain of function in the cytoplasm and/or sequester nuclear TDP-43, disrupting its role in RNA processing. TDP-43 cleavage products found in the cytoplasmic inclusions have been shown to alter splicing of a TDP-43 regulated exon in minigene studies, indicating that they are capable of interfering with wild-type TDP-43 function[7]. The toxicity of TDP-43 overexpression is dependent on retaining its RNA binding function in yeast[8], drosophila and chick[9].

To date, only a handful of RNAs regulated by TDP-43 have been identified. In order to understand the function of TDP-43, and its dysfunction in disease, we have comprehensively identified the RNA sites interacting with TDP-43 in human brain and characterised changes in RNA following reduced TDP-43 expression. In order to identify RNA targets on a global scale, we used individual-nucleotide resolution CLIP (iCLIP), which captures truncated cDNAs and quantifies them using a random barcode contained within the primer used for reverse transcription[10]. Duplicate sequences sharing the same random barcode were removed from analysis in order to avoid any apparent binding-site bias due to PCR artefacts.

In this study, we sought to identify the RNAs bound by TDP-43 in the healthy and FTLD brain tissue containing TDP-43 inclusions (FTLD-TDP). Although this was the first study of protein-RNA interactions in post-mortem human brain, we identified a high number of RNA binding sites using these samples. Most binding mapped to introns, long ncRNAs and intergenic transcripts, which also had the greatest enrichment of UG-rich motifs. We characterised the changes in alternative splicing following TDP-43 knockdown in neuroblastoma cells and identified splicing changes in 158 alternative cassette exons. Analysis of TDP-43 iCLIP crosslink sites in the vicinity of these exons provided insights into the mechanisms of TDP-43 splicing regulation. The exons regulated by TDP-43 are enriched in genes involved in neuronal development and those implicated in a range of neurodegenerative diseases.

RESULTS

Comparison of TDP-43 binding between healthy and FTLD-TDP brains

TDP-43 in complex with its direct RNA targets was purified from UV-crosslinked cells and analysed on SDS-PAGE gel. The size of the dominant TDP-43-RNA complex present in the low and high RNase conditions corresponded to a single TDP-43 molecule bound to the RNA, however a complex corresponding to two TDP-43 molecules bound to the RNA was also present, particularly in the low RNase experiment (Fig. 1a). This result agreed with the finding of a past study that TDP-43 binds RNA as a homodimer[2].

iCLIP experiments were performed with human cortical tissue from human post-mortem brain samples of three cognitively normal (healthy) and three sporadic FTLD-TDP patients, as well as from embryonic stem cells (hES) and SH-SY5Y neuroblastoma cells (Supplementary Table 1). In total, 17 million sequence reads aligned as single hits to the human genome by allowing a maximum of one nucleotide mismatch. After eliminating PCR amplification artefacts by removing sequences that truncated at the same nucleotide in the genome and shared the same random barcode, 3,7 million reads representing unique cDNA molecules were identified (Supplementary Table 2).

IDT Exhibit 1029.0002

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Genomic annotation of the cDNA sequences showed that the majority of TDP-43 binding occurred in introns, 74% in the cell lines and 58% in brain samples (Fig. 1b). The decreased proportion in brain samples, and corresponding increase in exonic sequences, may be a result of a greater sensitivity of intronic RNA to degradation in postmortem samples. Approximately 5% of iCLIP cDNAs mapped to long ncRNAs in all samples (Fig. 1c). Increased binding of TDP-43 to snoRNAs, snRNAs, rRNAs and telomeric ncRNAs was seen in FTLD-TDP (Fig. 1c). In total, 2139 ncRNA genes contained a crosslink site in control and 2702 in FTLD-TDP experiments.

The high proportion of intronic cDNA sequences indicated that the iCLIP data mainly represent nuclear RNA targets of TDP-43. To further evaluate this, we quantified the crosslinked complexes that were obtained using the iCLIP procedure from nuclear and cytoplasmic fractions of healthy and FTLD-TDP brain tissue. Only 6.5% of the total TDP-43-RNA complexes were obtained from cytoplasmic fractions of brain tissue, with insignificant difference between healthy and FTLD-TDP tissues (Supplementary Fig. 1b). iCLIP analysis of the RNAs bound by TDP-43 in the two fractions in SH-SY5Y cells showed 34% 3′UTR binding in the cytoplasm, 3.2% in the nuclear fraction and 3.8% in total cell extract (Supplementary Fig. 1d, Fig. 1b). As a control, we analysed crosslink complexes with TIAL1, a protein that is known to shuttle between nucleus and cytoplasm. 31.5% of total TIAL1-RNA crosslinked complexes were obtained from cytoplasmic fractions of healthy and FTLD-TDP brain tissue (Supplementary Fig. 1b). Thus, even though the TIAL1 analysis demonstrates that iCLIP can efficiently purify proteins bound to cytoplasmic RNAs, our analysis indicates that over 93% of TDP-43 iCLIP data presented in this study represents interactions with nuclear RNAs both in healthy and FTLD-TDP brain tissue.

To study the changes in TDP-43 binding to individual transcripts in FTLD-TDP, we analysed the proportion of cDNAs mapping to individual co-expressed genomic regions in the different brain samples. A significant (p value < 0.05 by Student's t-test, one-tailed, unequal variance) and greater than 0.025% change in proportion of cDNAs between control and FTLD-TDP brain was found in four ncRNAs, 3′ UTRs of seven transcripts and introns of 48 transcripts (Supplementary Table 3). The fact that introns had the greatest number of changes indicates that the changes detected by iCLIP mainly reflect changes in RNA interactions that occur in cell nuclei.

Four transcripts had a significant and greater than 0.1% change in proportion of cDNAs (p value < 0.05, Fig. 1d). Analysis of the abundance of these transcripts by RNA-seq and real-time PCR showed that the change in TDP-43 binding agreed with a similar extent of change in transcript expression (Fig. 1e, Supplementary Fig. 2). The most significant increase in TDP-43 binding in FTLD-TDP was observed in the nuclear paraspeckle assembly transcript 1 (*NEAT1*) and the metastasis associated lung adenocarcinoma transcript 1 (*MALAT1*, also known as *NEAT2*), which are transcribed from proximal loci on chromosome 11. Whereas *NEAT1* was previously identified as a 4 kb long ncRNA[11], we detected TDP-43 binding along the length of a 18kb ncRNA, with the peak binding in a tandem UG-repeat sequence close to the 3′ end (Fig. 1d, f). The primary binding sites in these two ncRNAs contained long UG repeats, and were identified in a highly reproducible manner (Fig. 1f, Supplementary Fig. 1f). The neurexin 3 (*NRXN3*) transcript and the glial excitatory amino acid transporter-2 (*EAAT2*, also *SLC1A2*) had significantly decreased cDNA proportion in FTLD-TDP (Fig. 1d, g). EEAT2 is necessary for glutamate clearance at the synapse, a process that prevents glutamate excitotoxicity[12].

TDP-43 binds to clusters of UG-rich sequences

In order to compare the RNA sequence specificity of TDP-43 in the different iCLIP experiments, we assessed the enrichment of all possible pentamers within 30 nt on either

IDT Exhibit 1029.0003

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

side of all crosslink sites. In all iCLIP data sets, the most significantly enriched pentamers were GUGUG and UGUGU (Fig. 2a, Supplementary Fig. 3a–c). The pentamer enrichment scores were correlated between healthy and FTLD-TDP brain (r=0.91, Fig. 2a).

We analysed the relationship between TDP-43 binding and UG repeats length by monitoring the frequency of crosslinking in the last ten nucleotides of UG tandem repeats compared to randomised positions. An example of a UG tandem repeat can be seen in the *NEAT1* ncRNA (Fig. 1f). We found that enrichment of TDP-43 crosslinking to UG repeats increased up to a repeat length of 15 nt, in a manner that was similar in healthy and FTLD-TDP brain (Fig. 2b, c). For comparison, we performed iCLIP with CELF2, which also binds UG-rich motifs in the human brain (Supplementary Fig. 3c). In contrast to TDP-43, binding of CELF2 to UG repeats increased only up to a length of 10 nt, and dropped once the repeats became longer than 15 nt (Fig. 2b). This indicates that TDP-43 has a stronger preference for tandem UG repeats compared to CELF2.

Proteins interacting with multiple identical sequences on the RNA often allow spacing between different binding sites; for instance, NOVA proteins bind clusters of a minimum of three YCAY tetramers, but can tolerate variable spacing between these[13]. To analyse if TDP-43 can similarly bind to more dispersed UG-rich motifs, we monitored the frequency of UGUGU motifs over a 200 nt region surrounding the TDP-43 crosslink sites compared to randomised positions. UGUGU was seven-fold enriched at the TDP-43 crosslink sites, and remained two-fold enriched even at a distance of 100 nts away from the crosslink sites (Fig. 2d, e). One cause for this phenomenon may be a tendency of TDP-43 to bind to long UG-rich RNA regions, such as the 90 nucleotide region in the 3′ UTRs of *TARDBP* mRNA, where TDP-43 binds to autoregulate its own mRNA levels (Fig. 2f)[14,15]. In contrast, even though UGUGU is four-fold enriched at the CELF2 crosslink sites, no such enrichment was seen at a distance of 100 nts (Fig. 2d). This indicates that TDP-43 has a unique capacity to recognise dispersed clusters of UG-rich motifs, or to spread its RNA binding to positions proximal to the UG-rich motifs.

To specifically analyse the high-affinity RNA binding sites of TDP-43, we grouped all data and searched for crosslink sites clustered with a maximum spacing of 15 nt that contained a significant cDNA count when compared to randomised positions (FDR < 0.05). 111,691 such crosslink clusters were identified. The reproducibility of crosslinking within these clusters between different samples increased with the number of iCLIP cDNA sequences that mapped to a cluster (Supplementary Fig. 3 d–g). For instance, 76% of clusters with summed cDNA counts of five or more in healthy brain were similarly bound in FTLD (Supplementary Fig. 3d). Clustered crosslink sites were highly enriched in UG tandem repeats, with increased enrichment at repeat lengths up to 30 nts (Fig. 2c). UGUGU was 18-fold enriched at the TDP-43 crosslink sites, and remained five-fold enriched even at a distance of 100 nts away from the crosslink sites (Fig. 2e). The increased incidence of UG-rich motifs at the crosslink clusters confirms that these clusters are the likely high-affinity binding sites of TDP-43.

To compare TDP-43 binding to different types of RNAs, we analysed the enrichment of UGUGU motif at the crosslink sites. The rRNAs, snRNAs, tRNAs and snoRNAs had the lowest UGUGU enrichment, indicating that crosslinking in these RNAs mainly reflected low-affinity, transient interactions (Supplementary Fig. 4). Three-fold enrichment was seen in 5′UTRs and ORFs, and four-fold UG enrichment was seen in 3′ UTRs, miRNA precursors and telomeric transcripts (Supplementary Fig. 4). Similar UG enrichment was seen in the 3′ UTRs when isolated with nuclear or cytoplasmic mRNA, indicating that TDP-43 binds to RNA in sequence-specific manner both in nucleus and the cytoplasm (Supplementary Fig. 1e). Seven-fold enrichment of UGUGU pentamers was seen in introns,

IDT Exhibit 1029.0004

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

long ncRNAs and intergenic RNAs, which together contain most TDP-43 binding (90% in cell lines and 76% in brain samples, Fig 1b, c). The high UG enrichment in intergenic RNAs is particularly interesting, indicating that high-affinity TDP-43 interactions occur with a large number of as yet non-annotated transcripts. Taken together, the iCLIP results indicate that TDP-43 binds pre-mRNAs and several types of ncRNAs in a sequence-specific manner.

The TDP-43 RNA splicing map

As the majority of iCLIP cDNAs mapped to introns, we reasoned that TDP-43 might have an important role in splicing regulation. Therefore, we used Affymetrix high-resolution splice-junction microarrays to evaluate splicing changes in TDP-43 knockdown SH-SY5Y cells. Analysis of splicing isoform reciprocity using the ASPIRE3 software identified splicing changes in 158 alternative cassette exons and 71 other types of splicing changes ($|\Delta Irank| \geq 1$) out of 30154 alternative exons that were detected by the microarray. 40 of the cassette exons were further assessed using RT-PCR and capillary electrophoresis. For 32 splicing events, primer pairs generated both isoforms, and 28 of these reproduced the direction of the observed splicing change, resulting in a 87% (28/32) validation rate of the microarray data (Supplementary Table 4).

To study how TDP-43 binds pre-mRNAs to regulate splicing, we visualised the positions of crosslink clusters in the form of an RNA splicing map, where the crosslink sites are plotted within 500 nt of the silenced (blue clusters) or enhanced (red clusters) cassette exons and the flanking exons (Fig. 3). The crosslink clusters identified three regions on the RNA splicing map where TDP-43 binding was mainly restricted to silenced or enhanced exons (Fig. 3). Binding in the region 1, 0–150 nt upstream or within the alternative exons, was present at 11 silenced and 1 enhanced exon (group S1). Binding in the region 2, 150–500 nt downstream of the alternative exons, was present at 7 silenced and 1 enhanced exon (group S2). Five of the silenced exons contained multiple crosslink clusters that spanned a sequence of 100 nt or longer in this region. This indicates that ‘deep’ intronic binding involving multiple dispersed binding sites might contribute to the ability of TDP-43 to silence exon inclusion. Finally, binding in the region 3, 0–150 nt downstream of the alternative exons (but not region 1 or 2), was present at four enhanced exons (group E).

To evaluate the proportion of TDP-43 iCLIP targets that changed splicing after knockdown, we analysed splicing of exons with crosslink clusters in the three regions defined by the RNA splicing map. We used RT-PCR to evaluate splicing of 41 putative alternative exons that had highest number of cDNAs mapping to these regions. 29% (12/41) of these exons showed no evidence of alternative splicing in SH-SY5Y cells (single band detected by RT-PCR). 79% (23/29) of the remaining exons changed splicing after knockdown, indicating that iCLIP data alone has high predictive value (Supplementary Table 4, Supplementary Fig. 5b). We also assayed splicing of 12 alternative exons that had UG-rich regions within 200 nt of the alternative exons, but were not identified either by microarray or iCLIP. A splicing change in 5 of these exons was validated by RT-PCR (Supplementary Table 4). Similar to the iCLIP binding, UG-rich motifs upstream of the exon or dispersed motifs downstream of the exon were associated with silencing, whereas proximal UG-rich motifs downstream of the exon were associated with enhancing function of TDP-43 (Supplementary Fig. 5c).

TDP-43 regulates splicing of mRNAs involved in neuronal development

TDP-43 crosslink clusters were also present in introns of 7,499 protein-coding genes and 3′ UTRs of 1,172 genes. 67% of the genes with 3′ UTR clusters (785/1,172) also had intronic clusters. For instance, two crosslink clusters were present in the intron downstream of the alternative exon 11 and three clusters in the 3′ UTR of myocyte enhancer factor 2D (*MEF2D*) mRNA (Fig. 4a). We identified a 29% decrease in inclusion of the *MEF2D* exon

IDT Exhibit 1029.0005

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

11 in TDP-43 knockdown SH-SY5Y cells (Fig. 4a). The additional binding sites in the 3′ UTR indicate that TDP-43 can remain associated with the *MEF2D* mRNA after splicing is finished and thereby regulate additional aspects of RNA processing.

Due to the strong sequence-specificity in binding to long ncRNAs (Supplementary Fig. 4), we also speculated that TDP-43 might regulate processing of such ncRNAs. iCLIP identified an alternative exon in the long ncRNA myocardial infarction associated transcript (*MIAT*). Eleven crosslink clusters are dispersed over the 1 kb region of both introns that flank this exon (Fig. 4b). We identified a 43% decrease in inclusion of the *MIAT* exon 10 in TDP-43 knockdown SH-SY5Y cells (Fig. 4b, Supplementary Fig. 5b).

We also identified a crosslink cluster in the *CFTR* pre-mRNA, which was the first identified RNA target of TDP-43[3]. Importantly, the crosslink overlapped with the UG-rich element upstream of exon 9, which is required for TDP-43-dependent splicing regulation of this exon (Supplementary Fig. 6). Past studies identified several candidate RNA targets of TDP-43, for which the direct *in vivo* interaction with TDP-43 has not been validated. Here, we identified TDP-43 binding sites in the FUS pre-mRNA[16], *hNFL* 3′ UTR[17], the *CDK6* intron and 3′ UTR[18], and the intronic region overlapping with the *mir558* gene[19] (Supplementary Fig. 6). Also of particular interest, we observed crosslink clusters in the mRNAs encoding HDAC6 and Casein kinase 1 and 2, which were reported to regulate phosphorylation and expression of TDP-43, respectively[20,21] (Supplementary Fig. 6).

We performed GO term analysis of genes regulated by TDP-43 at the level of alternative splicing, as determined by analysis of SH-SY5Y knockdown cells. This identified several terms related to development, including neural tube formation, organ morphogenesis and chordate embryonic development (Table 1). 15 of the exons regulated by TDP-43 are within genes encoding proteins with well-characterised functions in neuronal survival or development (Table 2). TDP-43 promotes inclusion of alternative exon 3 in Bcl-2 interacting mediator of cell death (BIM, also BCL2l11), and thereby suppresses production of the most cytotoxic BIM isoform (BIM_S)[22]. We have evaluated splicing of BIM in FTLD-TDP, and have found that the BIM_S isoform increases in FTLD-TDP tissue to a similar extent as in TDP-43 knockdown cells (Fig. 4c). This result indicates that aberrant splicing might contribute to the neurodegenerative process in human brain.

DISCUSSION

Our results show that TDP-43 interacts with a diverse spectrum of RNAs with important functions in the brain. We identified binding of TDP-43 to non-coding RNAs, introns and 3′ UTRs of mRNAs, indicating a role for TDP-43 in coupling different aspects of gene expression. Even though only a minor portion of the protein was detected in the cytoplasm, we found a ten-fold increase of TDP-43 binding to 3′ UTR elements when comparing RNAs bound to cytoplasmic and nuclear protein. We also identified the changes in alternative splicing associated with reduced level of TDP-43 protein. Analysis of crosslink clusters in the regulated pre-mRNAs indicates that the splicing function of TDP-43 depends on the position where it interacts with pre-mRNAs.

We found that TDP-43 preferentially binds to UG tandem repeats or long clusters of UG-rich motifs. These results agree with the past *in vitro* studies which showed that RNA binding of TDP-43 increases with the length of the UG repeat[2,3]. iCLIP results also indicate that the increased affinity for long binding sites differentiates the RNA specificity of TDP-43 from CELF2, which recognises shorter clusters of UG-rich motifs. Since a single RRM domain generally recognises four nucleotides[23], this RNA specificity for long binding sites indicates that TDP-43 can cooperatively recognise multiple RNA binding sites and

IDT Exhibit 1029.0006

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

thereby achieve high-affinity RNA binding. The apparent RNA binding of TDP-43 in a homodimer state might contribute to its ability to recognise binding sites composed of multiple dispersed UG-rich motifs (Fig. 1a).

TDP-43 crosslink clusters were located upstream and within the silenced exons and immediately downstream of enhanced exons, which is also a characteristic of several other RBPs, including the Nova proteins[24]. In addition, dispersed clusters of TDP-43 crosslinking further downstream of exons were associated with splicing silencing. Similar mode of splicing silencing, involving multiple binding sites flanking an alternative exon, has also been reported for Sxl, hnRNP A1, PTB, hnRNP L and hnRNP C[10,25,26]. The glycine-rich C-terminal domain of Sxl and hnRNP A1 is required for the cooperative assembly of these proteins on multiple proximal binding sites and the associated splicing silencing[25]. Interestingly, TDP-43 also contains a glycine-rich C-terminus. Mutations in the TARDBP gene that were identified in ALS or FTLD patients are concentrated in the region encoding the glycine-rich C-terminus[5], indicating that the ability of TDP-43 for cooperative assembly on long RNA binding sites plays a role in the disease mechanisms.

The tandem UG repeat sequences recognised by TDP-43 correspond to the most common microsatellite (AC) in the human genome. Variation in the length of microsatellites can introduce heritable phenotypic variation[27] or lead to changes in transcription and alternative splicing[28]. Genes involved in organ morphogenesis and neurogenesis are enriched in the variable tandem repeats[29]. Similarly, the genes containing the exons regulated by TDP-43 are enriched for neural tube formation and organ morphogenesis. This is consistent with the finding that TDP-43 knockout is lethal for mice from embryonic day 3.5[30]. It could be speculated that the ability of TDP-43 to regulate splicing by binding the UG repeats in the human genome might contribute to variation in splicing regulation of genes involved in development.

TDP-43 regulates splicing of several transcripts encoding proteins involved in neuronal survival, as well as seven mRNAs encoding proteins with relevance for neurodegenerative diseases (Table 2). Delta-catenin (CTNND1) plays a role in the development and maintenance of dendritic spines[31,32]. Furthermore, CTNND1 interacts with presenilin-1 to inhibit production of Aβ peptide, the main constituent of amyloid plaques in the brains of Alzheimer's disease patients[33]. MEF2D is a transcription factor that was reported to play a role in neuronal survival in Parkinson's disease[34]. BIM is a core factor in the neuron-specific JNK-mediated apoptotic pathway, and loss of BIM protects mice from ischemia-induced damage[35,36]. The BIM splice isoforms regulated by TDP-43 have been functionally characterised; BIM_S antagonizes the pro-survival Bcl-2 family members more effectively than BIM_L, while BIM_{EL} is the least potent[22]. We found that production of the most cytotoxic BIM isoform (BIM_S) was increased in TDP-43 knockdown cells and in FTLD-TDP. However, analysis of other exons regulated by TDP-43 did not indicate a significant correlation with the splicing changes present in FTLD-TDP (unpublished observation).

We also found that TDP-43 binds to long ncRNAs in highly sequence-specific manner both in control and FTLD-TDP samples. The greatest increase in TDP-43 binding in FTLD-TDP is seen in the *NEAT1* ncRNA. Expression of *NEAT1* and *MALAT1* significantly increases in FTLD-TDP, which might be the primary cause of the increased association of TDP-43 with these RNAs. *NEAT1* functions in paraspeckle assembly[11], and the *MALAT1* ncRNA, which recruits splicing factors to nuclear speckles and affects phosphorylation of SR proteins[37]. In conclusion, we have characterised the splicing function of TDP-43 and quantified the changes in its RNA binding in FTLD-TDP. Analysis of these RNAs in model organisms and in neurons containing TDP-43 inclusions will be required for full

IDT Exhibit 1029.0007

understanding of the mechanisms of TDP-43 loss- or gain-of-function effects in the neurodegenerative process.

Methods

iCLIP analysis

The iCLIP protocol was performed as described previously[10], with the following modifications. SH-SY5Y neuroblastoma or H9 human embryonic stem cells were irradiate once with 150 mJ/cm^2 in a Stratlinker 2400 at 254 nm, and brain tissue was dissociated in cold PBS and the suspension was crosslinked four times with 100 mJ/cm^2. TDP-43 was immunoprecipitated with protein A Dynabeads (Invitrogen) conjugated to rabbit-anti TDP-43 (Proteintech, 10782-2-AP). For iCLIP of CELF2, protein G Dynabeads conjugated to mouse anti-CELF2 (Sigma, C9367) were used. In both cases, the region corresponding to 55–100kDa complexes was excised from the membrane to isolate the RNA. High-throughput sequencing using Illumina GA2 was done using 54 or 72 cycles (Supplementary Table S1). The barcode sequences corresponding to the individual experiment were as described (Supplementary table 1). The random barcodes were registered and the barcodes were removed before mapping the sequences to the human genome sequence (version GRCh37/hg19) allowing one mismatch using Bowtie version 0.10.1 (command line: -a -m 1 -v 1).

Nucleocytoplasmic fractionation

50mg of crosslinked SH-SY5Y cells or brain tissue was resuspended in cold 1ml of cytoplasmic lysis buffer (50 mM Tris-HCl, pH 7.4, 10mM NaCl, 0.5% NP-40, 0.25% Triton X-100, 1mM EDTA, 0.5% RNAsin). After rotation in cold room for 5 minutes, homogeniser was used to complete the lysis. After centrifugation at 4°C, 3000 × g for 5 minutes, supernatant was collected and centrifugation was repeated at 4°C, 10000 × g for 10 minutes. Supernatant was collected, and 200ul of 0.5% SDS, 0.25% sodium deoxycholate, 0.5M NaCl was added; this was used as cytoplasmic fraction for CLIP or western blot analyses. The pellet (after 3000 × g spin) was resuspended in 1ml of CLIP Lysis Buffer[10], sonicated and used as nuclear fraction for CLIP or western blot analyses. The radioactive CLIP protein-RNA complexes on nitrocellulose membrane were quantified using phosphoimager.

Sequence analyses

Analysis of reproducibility of crosslink sites was done as described previously[10]. Identification of the significant iCLIP crosslink clusters and z-score analysis of enriched pentamers was done as described previously[38]. For analysis of crosslinking to UG repeats of different lengths, the number of UG repeats in transcribed genome with a crosslink site in at least one of the last ten nucleotides is determined for each data set. Enrichment represents the ratio in the number of UG repeats with a crosslink site in real data compared to randomized data. The pentamer with highest z-scores in TDP-43 iCLIP (UGUGU) was used to assess the positions at which the TDP-43 binding sites were enriched relative to the position of crosslink sites. The genomic sequences at positions within 100 nts of the iCLIP crosslink site were filtered to retain only the sequence containing the pentamer. The sequences were aligned at each iCLIP crosslink site, and the number of sequences containing the pentamer was determined at each nucleotide position between −100 to +100 nts. Enrichment represents the ratio in the number of UG repeats with a crosslink site in real data compared to randomized data.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

IDT Exhibit 1029.0008

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

siRNA knockdown

Acell Smartpool oligos against *TARDBP* (E-012394-00-0005, Dharmacon) were transfected according to the manufacturer's instructions (Acell, Dharmacon). Accell non-targeting pool oligos (D-001910-10-05, Dharmacon) were used as a control.

A-012394-14, TARDBP: 5′- GGCUCAAGCAUGGAUUCUA -3′

A-012394-15, TARDBP: 5′- GUCUCAAGUCAAAUGGAUU -3′

A-012394-16, TARDBP: 5′- GUGUUAAGUGAAAUGAUAC -3′

A-012394-17, TARDBP: 5′- GGGUGAUGUUCUAUUUACA -3′

Western blot

SH-SY5Y cells were lysed in iCLIP lysis buffer and sonicated. To digest DNA, 10 μl of DNase (Ambion) was added per 1 ml of lysate, and incubated at 37°C for 3 minutes. The samples were centrifuged and the supernatants were collected for Western analysis. The protein concentration was determined using Lowry's Assay (Bio-RAD). Equal amounts of protein were loaded on 4–12% Bis-Tris gels (Invitrogen) and transferred to a nitrocellulose membrane. The membrane was incubated with rabbit anti-TDP-43 antibody (Proteintech, 10782-2-AP) (1:1000 dilution). Rabbit anti-GAPDH (Cell Signalling) (1:5000 dilution) or anti-actin antibodies were used as a loading control.

Splice-junction microarray

A total of 6 samples were used, 3 from the siRNA control and 3 from KD. The high-resolution AltSplice splice-junction microarrays were produced by Affymetrix, the cDNA samples were prepared using the GeneChip WT cDNA Synthesis and Amplification Kit (Affymetrix) and data was analysed with the version 3 of ASPIRE (Analysis of SPlicing Isoform REciprocity)[10].

RT-PCR

Total RNA was extracted using RNasy Kit (Qiagen) and 200 ng of RNA was used for reverse transcription using Superscript III (Invitrogen) according to the manufacturer's instructions. For analysis of transcript levels, real-time PCR was performed using SYBR Green Fast PCR master mix (Applied Biosystems). For analysis of splicing, PCR was performed using Immomix (Bioline) using primers listed in Supplementary Table 4. The PCR products were visualised using QIAxcel capillary electrophoresis system (Qiagen). To calculate exon inclusion (I), the percentage of the peak representing exon inclusion was divided by the total percentage of peaks representing exon inclusion and skipping. Splicing change was calculated by subtracting the exon inclusion in the knockdown cells from the inclusion in wild-type cells (thus, a positive ΔI represents exons enhanced by TDP-43, and negative those silenced by TDP-43).

RNA splicing map

In order to analyze the impact of TDP-43 positioning on splicing regulation, we assessed the positioning of crosslink clusters in proximity of the exon/intron boundaries of alternative exons and flanking exons, including 50 nt of exonic and 500 nt of intronic sequence. All cassette exons with identified splicing change in knockdown SH-SY5Y cells and at least one significant crosslink cluster in the analysed regions are shown in the RNA splicing map.

IDT Exhibit 1029.0009

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors wish to thank Kathi Zarnack and Julian König for their advice and comments on the manuscript, Francisco Baralle, Yuna Ayalla and Andrea D.Ambrogio for providing HeLa knockdown RNA, and James Hadfield and the genomic team at CRI for Illumina sequencing. This work was supported by the European Research Council grant (206726-CLIP), the Medical Research Council (MRC), Slovenian Research Agency (P2-0209, J2-2197, L2-1112, Z7-3665), Wellcome Trust and Medical Research Council Strategic Grant Award (089701/Z/09/Z), Motor Neuron Disease Association, Heaton-Ellis Trust and Psychiatry Research Trust.

References

1. Buratti E, Baralle FE. Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. Front Biosci. 2008; 13:867–878. [PubMed: 17981595]
2. Kuo PH, Doudeva LG, Wang YT, Shen CK, Yuan HS. Structural insights into TDP-43 in nucleic-acid binding and domain interactions. Nucleic Acids Res. 2009; 37:1799–1808. [PubMed: 19174564]
3. Buratti E, Baralle FE. Characterization and functional implications of the RNA binding properties of nuclear factor TDP-43, a novel splicing regulator of CFTR exon 9. J Biol Chem. 2001; 276:36337–36343. [PubMed: 11470789]
4. Sreedharan J, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science. 2008; 319:1668–1672. [PubMed: 18309045]
5. Lagier-Tourenne C, Polymenidou M, Cleveland DW. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. Hum Mol Genet. 2010; 19:R46–64. [PubMed: 20400460]
6. Neumann M, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science. 2006; 314:130–133. [PubMed: 17023659]
7. Igaz LM, et al. Expression of TDP-43 C-terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies. J Biol Chem. 2009; 284:8516–8524. [PubMed: 19164285]
8. Johnson BS, McCaffery JM, Lindquist S, Gitler AD. A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity. Proc Natl Acad Sci U S A. 2008; 105:6439–6444. [PubMed: 18434538]
9. Voigt A, et al. TDP-43-mediated neuron loss in vivo requires RNA-binding activity. PLoS ONE. 2010; 5:e12247. [PubMed: 20806063]
10. König J, et al. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol. 2010; 17:909–915. [PubMed: 20601959]
11. Clemson CM, et al. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. Mol Cell. 2009; 33:717–726. [PubMed: 19217333]
12. Cleveland DW, Rothstein JD. From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. Nat Rev Neurosci. 2001; 2:806–819. [PubMed: 11715057]
13. Ule J, et al. CLIP identifies Nova-regulated RNA networks in the brain. Science. 2003; 302:1212–1215. [PubMed: 14615540]
14. Ayala YM, et al. TDP-43 regulates its mRNA levels through a negative feedback loop. Embo J. 2011; 30:277–288. [PubMed: 21131904]
15. Polymenidou M, et al. Disrupted processing of long pre-mRNAs and widespread RNA missplicing are components of neuronal vulnerability from loss of nuclear TDP-43. Nat neurosci. 2011
16. Sephton CF, et al. Identification of neuronal RNA targets of TDP-43-containing Ribonucleoprotein complexes. J Biol Chem. 2010
17. Strong MJ, et al. TDP43 is a human low molecular weight neurofilament (hNFL) mRNA-binding protein. Mol Cell Neurosci. 2007; 35:320–327. [PubMed: 17481916]
18. Ayala YM, Misteli T, Baralle FE. TDP-43 regulates retinoblastoma protein phosphorylation through the repression of cyclin-dependent kinase 6 expression. Proc Natl Acad Sci U S A. 2008; 105:3785–3789. [PubMed: 18305152]

IDT Exhibit 1029.0010

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

19. Buratti E, et al. Nuclear factor TDP-43 can affect selected microRNA levels. FEBS J. 2010; 277:2268–2281. [PubMed: 20423455]
20. Kametani F, et al. Identification of casein kinase-1 phosphorylation sites on TDP-43. Biochem Biophys Res Commun. 2009; 382:405–409. [PubMed: 19285963]
21. Fiesel FC, et al. Knockdown of transactive response DNA-binding protein (TDP-43) downregulates histone deacetylase 6. EMBO J. 29:209–221. [PubMed: 19910924]
22. O'Connor L, et al. Bim: a novel member of the Bcl-2 family that promotes apoptosis. EMBO J. 1998; 17:384–395. [PubMed: 9430630]
23. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol. 2008; 18:290–298. [PubMed: 18515081]
24. Witten JT, Ule J. Understanding splicing regulation through RNA splicing maps. Trends Genet. 2011
25. Black DL. Mechanisms of alternative pre-messenger RNA splicing. Annu Rev Biochem. 2003; 72:291–336. [PubMed: 12626338]
26. Hung LH, et al. Diverse roles of hnRNP L in mammalian mRNA processing: a combined microarray and RNAi analysis. Rna. 2008; 14:284–296. [PubMed: 18073345]
27. Rockman MV, Wray GA. Abundant raw material for cis-regulatory evolution in humans. Mol Biol Evol. 2002; 19:1991–2004. [PubMed: 12411608]
28. Kashi Y, King DG. Simple sequence repeats as advantageous mutators in evolution. Trends Genet. 2006; 22:253–259. [PubMed: 16567018]
29. Legendre M, Pochet N, Pak T, Verstrepen KJ. Sequence-based estimation of minisatellite and microsatellite repeat variability. Genome Res. 2007; 17:1787–1796. [PubMed: 17978285]
30. Wu LS, et al. TDP-43, a neuro-pathosignature factor, is essential for early mouse embryogenesis. Genesis. 2010; 48:56–62. [PubMed: 20014337]
31. Matter C, Pribadi M, Liu X, Trachtenberg JT. Delta-catenin is required for the maintenance of neural structure and function in mature cortex in vivo. Neuron. 2009; 64:320–327. [PubMed: 19914181]
32. Ishiyama N, et al. Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion. Cell. 141:117–128. [PubMed: 20371349]
33. Kouchi Z, et al. p120 catenin recruits cadherins to gamma-secretase and inhibits production of Abeta peptide. J Biol Chem. 2009; 284:1954–1961. [PubMed: 19008223]
34. Yang Q, et al. Regulation of neuronal survival factor MEF2D by chaperone-mediated autophagy. Science. 2009; 323:124–127. [PubMed: 19119233]
35. Ness JM, et al. Selective involvement of BH3-only Bcl-2 family members Bim and Bad in neonatal hypoxia-ischemia. Brain Res. 2006; 1099:150–159. [PubMed: 16780816]
36. Becker EB, Bonni A. Pin1 in neuronal apoptosis. Cell Cycle. 2007; 6:1332–1335. [PubMed: 17568190]
37. Tripathi V, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Mol Cell. 2010; 39:925–938. [PubMed: 20797886]
38. Wang Z, et al. iCLIP predicts the dual splicing effects of TIA-RNA interactions. PLoS Biol. 2010; 8:e1000530. [PubMed: 21048981]

IDT Exhibit 1029.0011

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

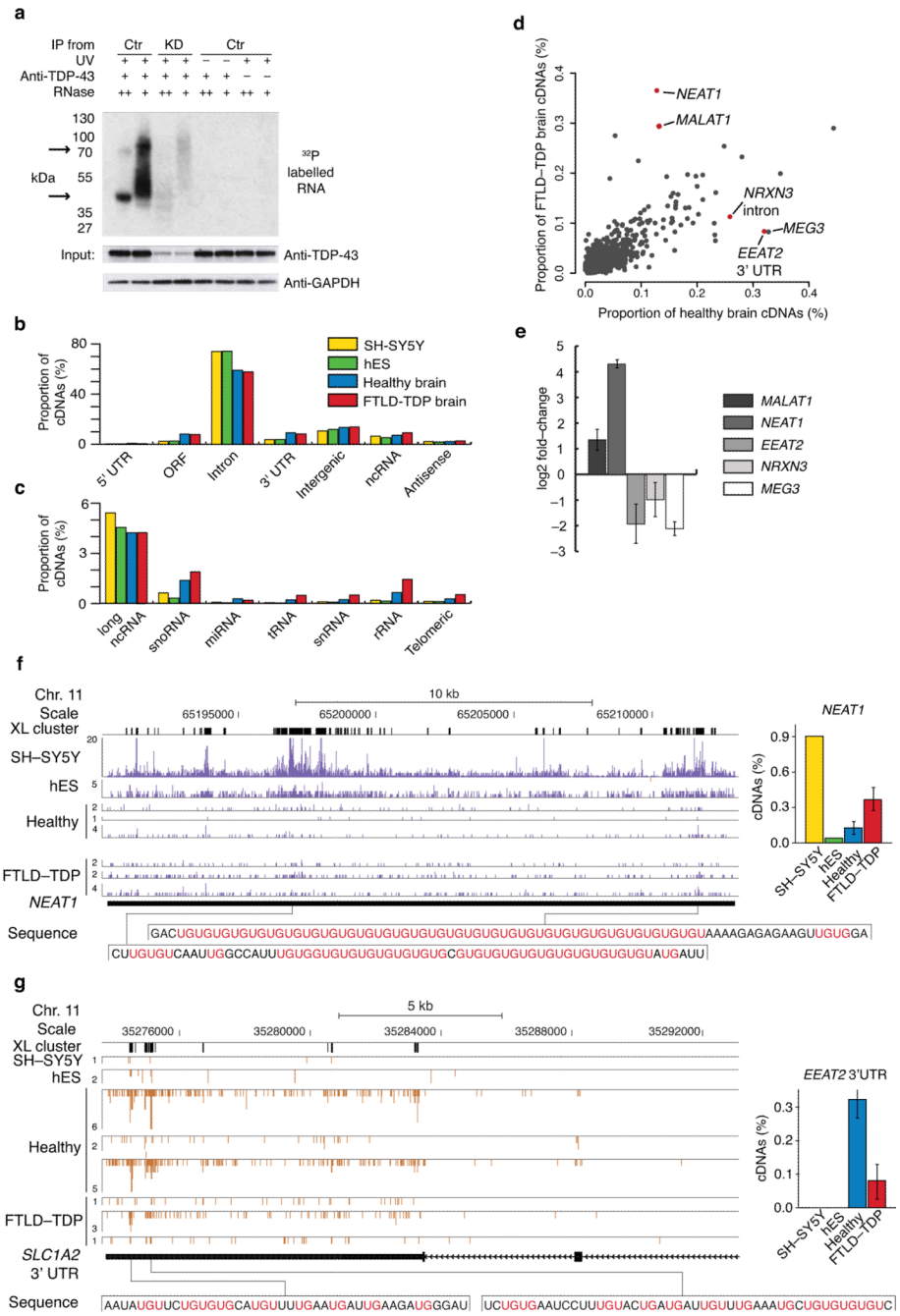

Figure 1. Comparison of TDP-43 RNA binding in healthy and FTLD-TDP brain

(a) To validate specificity of TDP-43 antibody for iCLIP, the ^{32}P-labelled RNA bound to TDP-43 gel was isolated from control (Ctr) or knockdown (KD) HeLa cells in the presence or absence of UV crosslinking or anti-TDP-43 antibody. High and low RNase concentrations were used to confirm the presence of RNA bound to TDP-43. The arrows mark the positions on the gel corresponding to the size of TDP-43 monomer or dimer. TDP-43 Western analysis of input extracts confirmed TDP-43 knockdown, and GAPDH was used as a loading control. The full image is shown in Supplementary Fig. 1a **(b)** The proportion of cDNAs (out of all cDNAs that mapped to human genome) from the TDP-43

IDT Exhibit 1029.0012

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

iCLIP experiments in the four types of samples that mapped to different RNA regions. **(c)** The proportion of cDNAs that mapped to different types of ncRNAs. **(d)** The proportion of cDNAs that mapped to individual RNAs with at least 10 cDNAs in any experiment. The RNAs with the largest significant change between control and FTLD-TDP brain (difference in proportion of cDNAs > 0.1% and p value < 0.05 by Student's t-test, one-tailed, unequal variance) are marked in red. The long ncRNA MEG3 (maternally expressed 3) with the largest decrease in TDP-43 binding in FTLD-TDP is also marked. **(e)** Real-time PCR analysis of transcripts with largest TDP-43 iCLIP changes in total RNA isolated from control and FTLD-TDP brain samples. **(f)** iCLIP cDNA counts for TDP-43 crosslink positions in the *NEAT1* gene in experiments from SH-SY5Y and hES cell lines, and healthy and FTLD-TDP tissue from different individuals. The blue bars represent sequences on the sense strand of the genome, and the height of the bars corresponds to the cDNA count. The positions of significant crosslink clusters (XL cluster) are shown on top, and the RNA sequence underlying the two man clusters is shown below, with UG repeats in pink. The graph shows the average proportion of cDNAs that map to *NEAT1* transcript in each experiment (out of all cDNAs maping to the human genome), as well as standard deviation between experiments from different individuals. **(g)** iCLIP cDNA counts for TDP-43 crosslink positions in the *SLC1A2* gene. The orange bars represent sequences on the antisense strand of the genome. The graph shows the average proportion of cDNAs that map to *SLC1A2* 3′ UTR in different experiments, as well as standard deviation between experiments from different individuals. The RNA sequence underlying the peak crosslinking sites is shown below, with UG repeats in pink.

IDT Exhibit 1029.0013

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

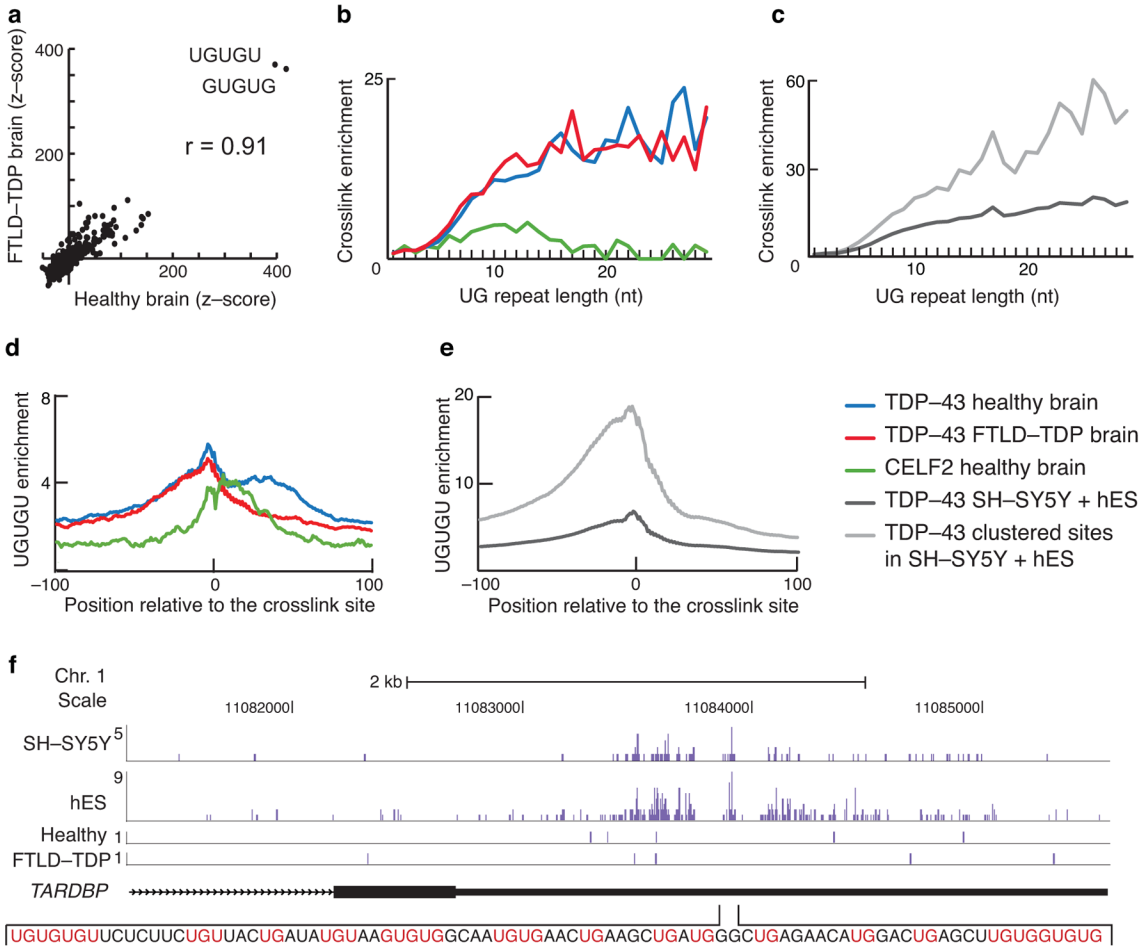

Figure 2. TDP-43 binding motif analysis

(a) z-scores of penamer occurrence within the 61 nt sequence surrounding all crosslink sites (−30 nt to +30 nt) are shown for healthy and FTLD-TDP brain iCLIP. The sequences of the two most enriched pentamers and the Pearson correlation coefficient (r) between the two samples are given. **(b)** Enrichment of crosslinking compared to randomised data in UG repeats of different lengths in TDP-43 iCLIP experiments from healthy brains (blue) and FTLD-TDP brains (red), and CELF2 experiments in healthy brain (green). **(d)** Enrichment of crosslinking compared to randomised data on UG repeats of different lengths in TDP-43 iCLIP experiments from SH-SY5Y and hES cells, where we either included all crosslink sites (dark grey), or only those mapping to crosslink clusters (light grey). **(e)** Analysis of positions of UGUGU enrichment compared to randomised data around TDP-43 crosslink sites in healthy brains (blue) and FTLD-TDP brains (red), and CELF2 experiments in healthy brain (green). **(e)** Analysis of positions of UGUGU enrichment compared to randomised data in TDP-43 iCLIP experiments from SH-SY5Y and hES cells, where we either included all crosslink sites (dark grey), or only those mapping to crosslink clusters (light grey). **(f)** TDP-43 crosslinking in its own transcript (*TARDBP).* The replicate experiments are summed and shown in a single track. The RNA sequence underlying the peak crosslinking site is shown, with UG and GU dinucleotides in pink.

IDT Exhibit 1029.0014

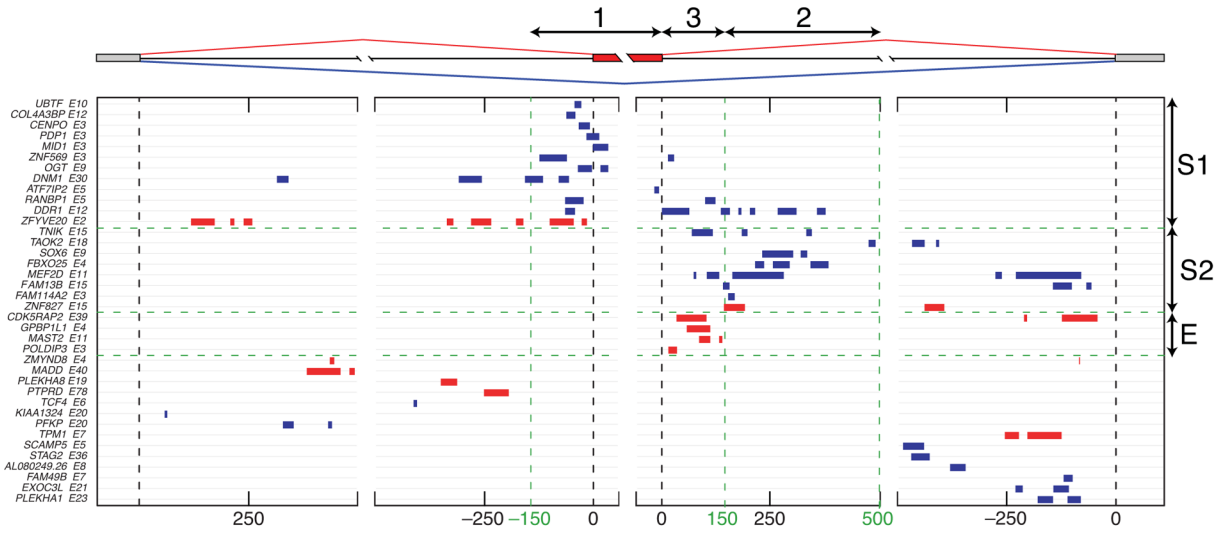

Figure 3. The RNA splicing map of TDP-43

The map of crosslink clusters at positions within 500 nt of alternative exons and flanking exons. The cassette exons with ΔIrank > 1 (enhanced exons, red clusters) or ΔIrank < −1 (silenced exons, blue clusters) with at least one crosslink cluster in these regions are shown. The exons are grouped by sequential analysis of crosslink cluster positions in three regions: group S1 is identified by clusters in region 1, 150nt upstream till the exon and within the exon; group S2 by clusters in region 2, 150–500nt downstream of the exon; and group E by clusters present 0–150nt downstream of the exon.

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts
IDT Exhibit 1029.0015

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

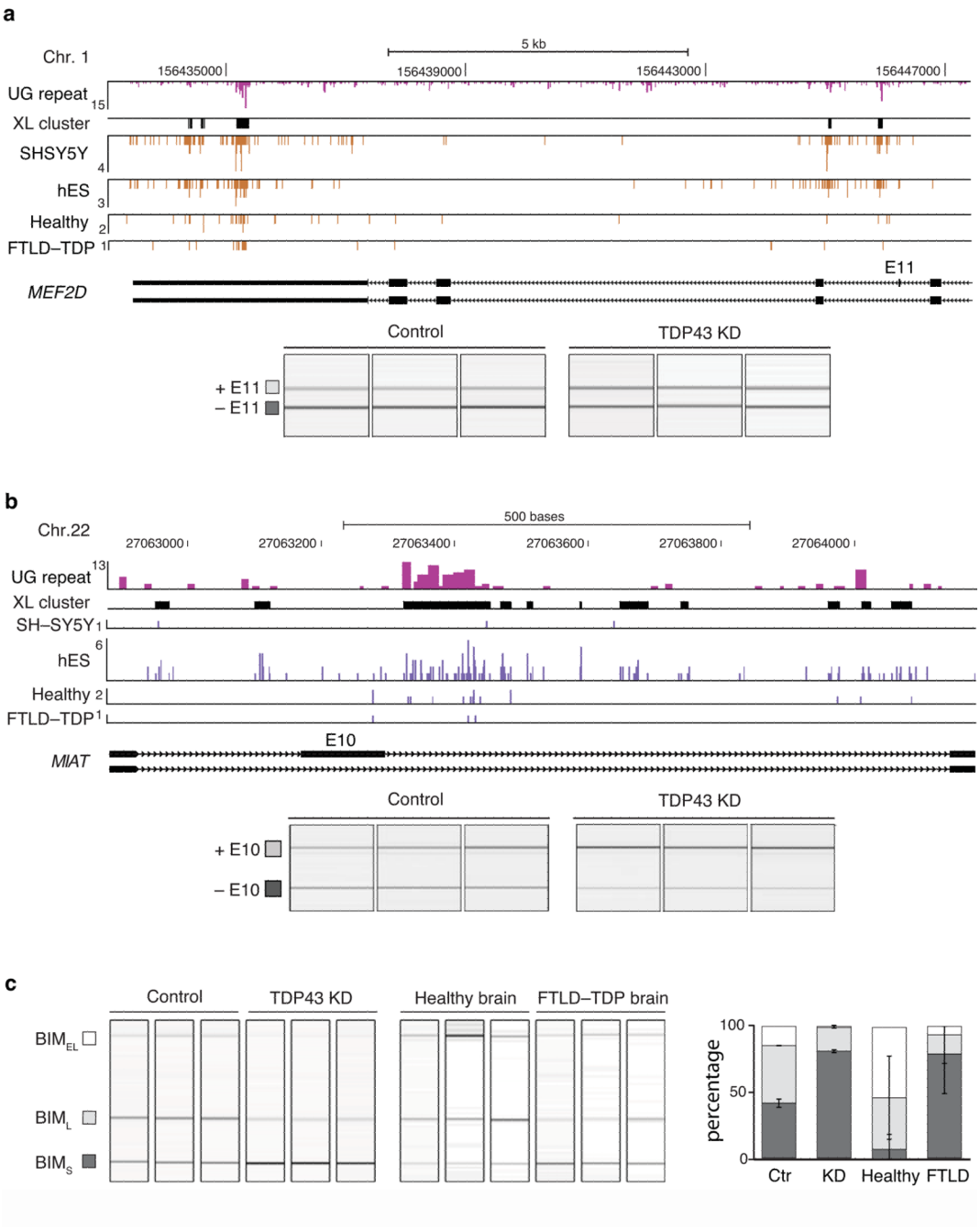

Figure 4. TDP-43 regulates splicing of non-coding and protein-coding RNAs

(a) TDP-43 crosslinking in *MIAT* ncRNA. UG repeat lengths are shown in pink bars and crosslink clusters are shown as black bars. Analysis of *MIAT* exon 10 inclusion (gel electropherogram, left, and quantification, right) in control and TDP-43 knockdown SH-SY5Y cells is shown below. **(b)** TDP-43 crosslinking in MEF2D protein-coding transcript. UG repeat lengths are shown in pink bars and crosslink clusters are shown as black bars. Analysis of *MEF2D* exon 11 inclusion (gel electropherogram, left, and quantification, right) in control and TDP-43 knockdown SH-SY5Y cells is shown below. **(c)** Analysis of *BIM* exon 3 splicing (gel electropherogram, left, and quantification, right) in control and TDP-43

IDT Exhibit 1029.0016

knockdown SH-SY5Y cells, and in healthy (C23, C25, C30) and FTLD-TDP (F19, F20, F21) brain samples.

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

IDT Exhibit 1029.0017

Europe PMC Funders Author Manuscripts
Europe PMC Funders Author Manuscripts

Table 1

Significant GO terms of the exons regulated by TDP-43.

GO term	p-value	reference exons	TDP-43 regulated exons	
organ morphogenesis	<0.001	811	13	BCL2L11_E5, COL4A3BP_E12, CTNND1_E30, DLC1_E2, FGFRL1_E4, FZD6_E2, MACF1_E68, MEF2D_E11, PLEKHA1_E21, PLEKHA1_E23, TFAP2A_E2, TLE1_E10, TPM1_E7
neural tube closure	0.001	45	3	DLC1_E2, FZD6_E2, TFAP2A_E2
homeostasis of number of cells	0.01	187	4	BCL2L11_E5, RPS24_E11, RPS24_E9, SOX6_E9
mitotic cell cycle	0.016	856	9	CLIP1_E33, HORMAD1_E4, KIF2A_E19, MYO16_E8, PMF1_E2, PSMD6_E4, RBM38_E3, SIRT7_E9, STAG2_E36
cell surface receptor linked signaling pathway	0.018	2005	16	ANXA1_E2, BRD8_E24, CNTFR_E2, CTNND1_E30, FZD6_E2, GRM4_E7, HOMER2_E5, IRAK3_E2, MACF1_E68, MADD_E40, PILRB_E25, PLEKHA1_E21, PLEKHA1_E23, PTPRD_E78, TLE1_E10, TNIK_E15
chordate embryonic development	0.021	475	6	BCL2L11_E5, COL4A3BP_E12, DLC1_E2, FZD6_E2, SOX6_E9, TFAP2A_E2
lipid transport	0.037	276	4	ANXA1_E2, ATP11C_E36, COL4A3BP_E12, STARD4_E5

IDT Exhibit 1029.0018

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

Table 2

Neuronal functions of proteins encoded by the alternative mRNA isoforms regulated by TDP-43.

Gene symbol	Gene name	neurologic disease association	neuronal function	Pubmed PMID
AP2 (TFAP2A)	transcription factor AP-2 alpha		cranial neural-tube closure	11137286
BIM (BCL2L11)	bcl-2 interacting mediator of cell death	ischemia-induced neuronal death	apoptotic signalling cascade	16780816 17568190
CNTFR	ciliary neurotrophic factor receptor	Amyotropic lateral sclerosis	neuronal survival	19386761
MADD	MAP-kinase activating death domain	Alzheimer's disease	neuronal survival under stress conditions	9482930
MEF2D	myocyte enhancer factor 2D	Parkinson's disease	neuronal survival	19119233
CDK5RAP2	CDK5 regulatory subunit associated protein 2	Primary autosomal-recessive microcephaly	maintainance of the neural progenitor pool	20471352
CTNND1	delta catenin	mental retardation syndrome Cri-du-Chat	maintenance of dendritic spines	19914181
DLC1	deleted in liver cancer 1		neural tube development	15710412
FZ3 (FZD3)	frizzled homolog 3		axonal guidance, neural tube closure	16495441
KIF2A	kinesin family member 2A		axonal branching	12887924
KIF1B	kinesin family member 1B	Charcot-Marie-Tooth neuropathy	axonal vesicle transport	11389829
SOX6	transcription factor SOX-6		cortical interneuron development	19657336
TLE1	transducin-like enhancer of split 1		cortical neuron differentiation	16314515
TNIK	TRAF2 and NCK interacting kinase		dendrite formation	20159449
UNC5C	unc-5 homolog C (netrin receptor)		spinal accessory motor neuron development	17543537

IDT Exhibit 1029.0019