Supplementary figures and tables

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] & 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,

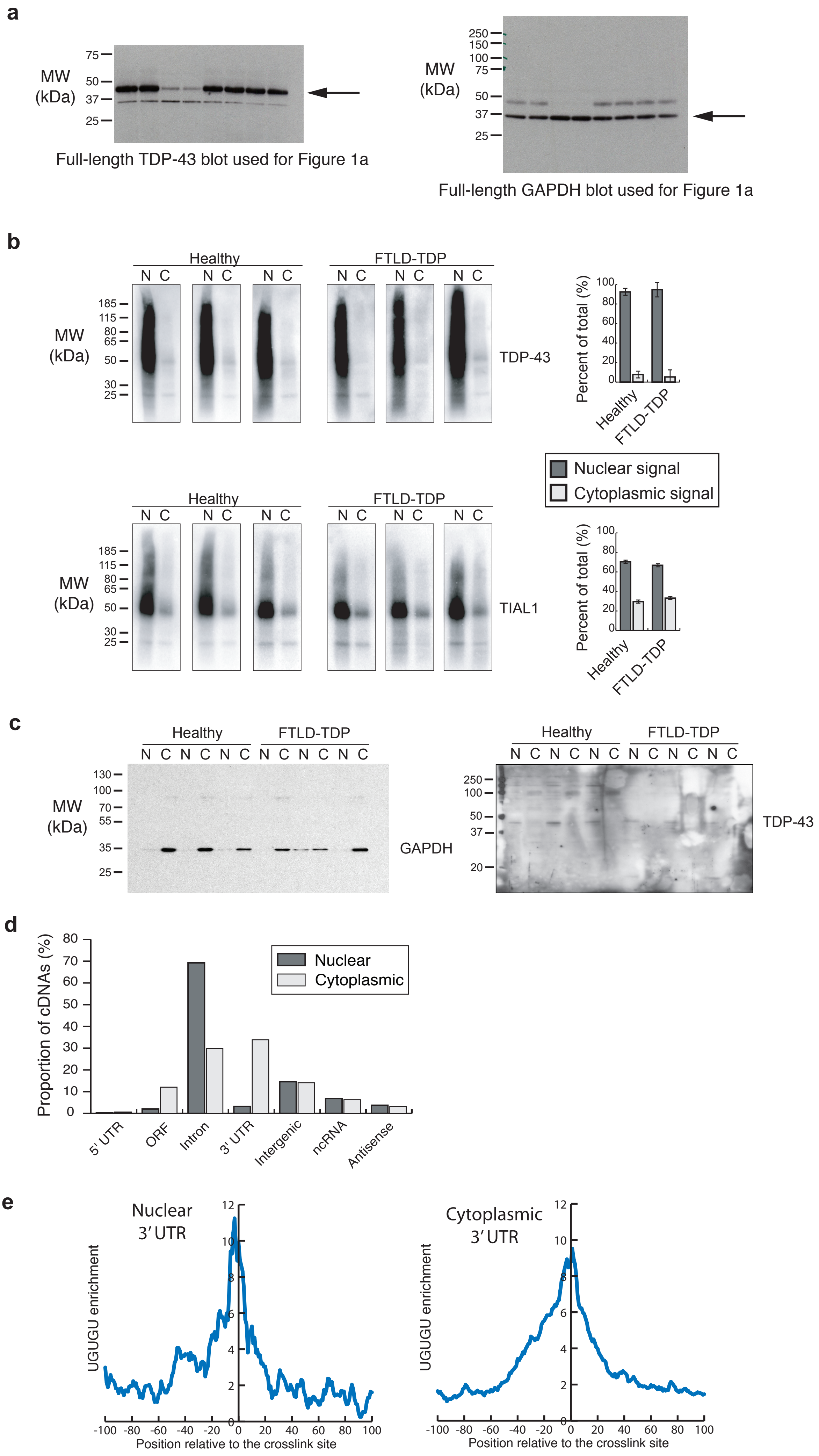
[4]Scientific Institute IRCCS E. Medea, Via don L. Monza 20, 23842 Bosisio Parini (LC), Italy

[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

[*]equal contribution

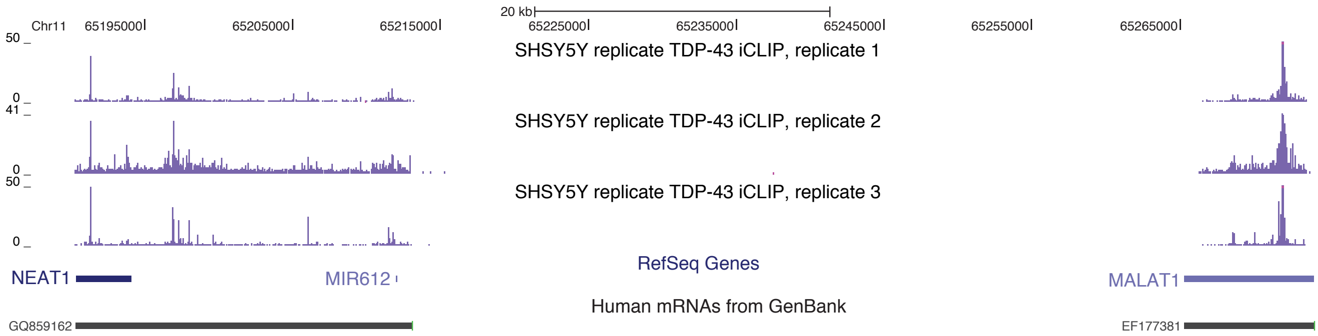
Supplementary Figure 1


Full-length TDP-43 blot used for Figure 1a

Full-length GAPDH blot used for Figure 1a

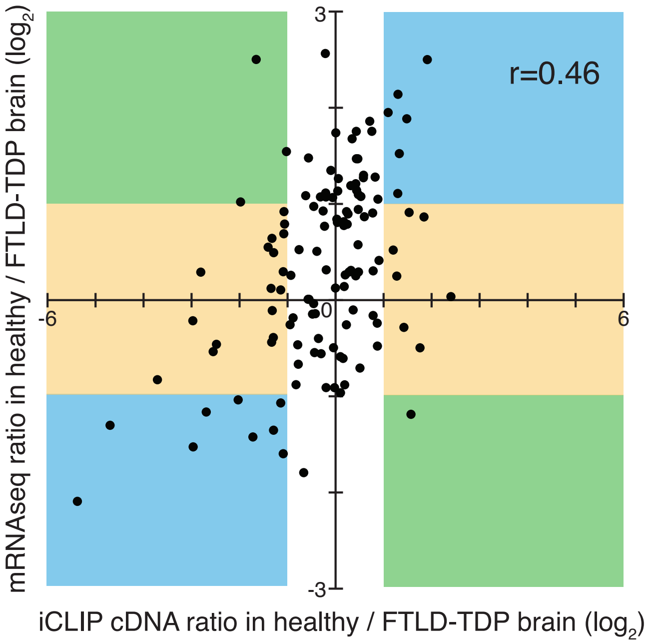
Supplementary Figure 1 cont.

f



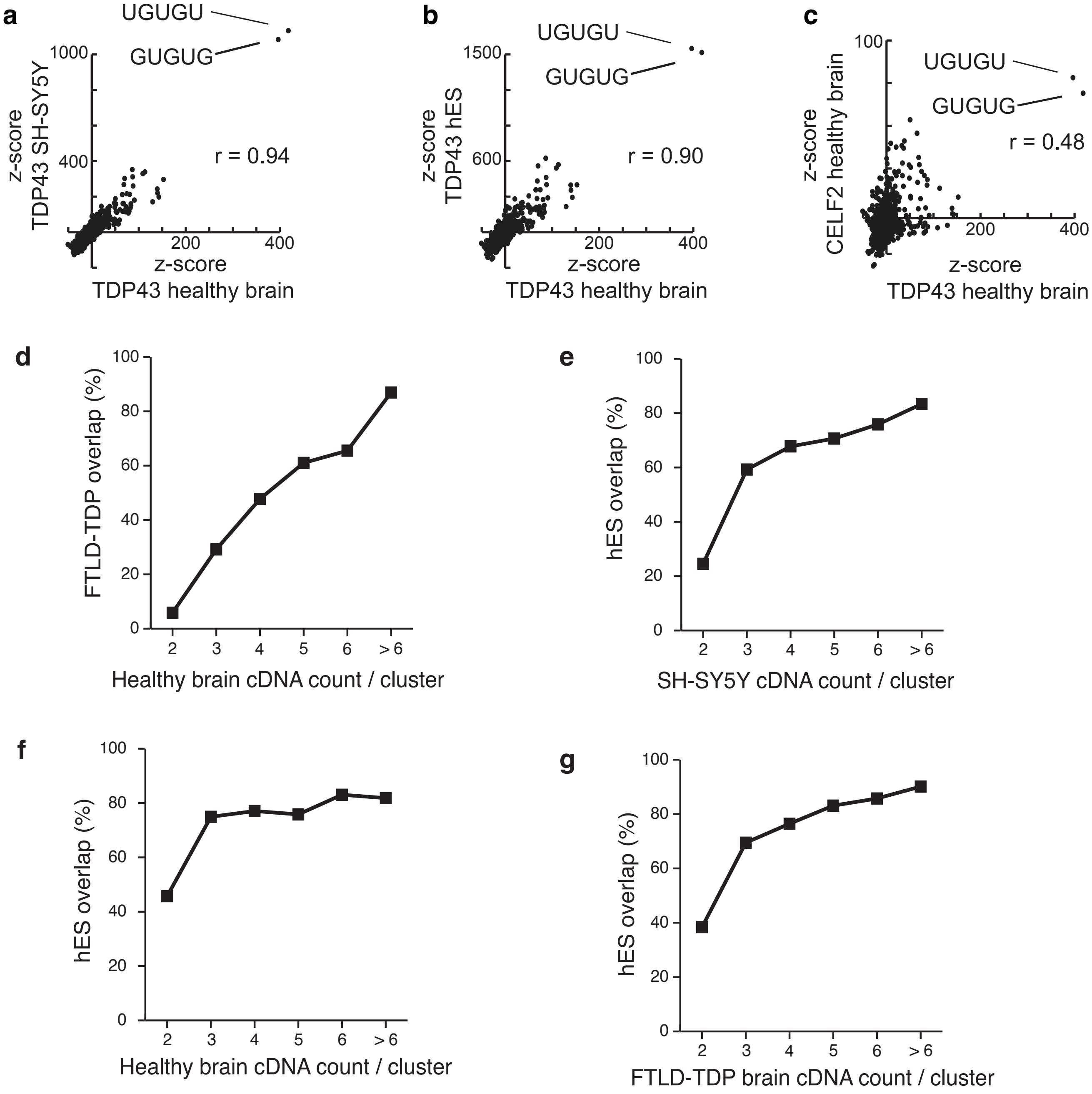
Supplementary Figure 1: Analysis of TDP-43 crosslinking to RNAs in nuclei and cytoplasm. (a) TDP-43 and GAPDH Western analysis of extracts used for CLIP as shown in Fig 1a. (b) To analyse the proportion of TDP-43 RNA complexes isolated by iCLIP from nuclei and cytoplasm, 32P-labelled RNA bound to TDP-43 gel was isolated from crosslinked healthy or FTLD-TDP samples and with high RNAse condition. Radioactivity was quantified using phosphoimager, and average and standard deviation are shown on the right. (c) Western blot analysis of GAPDH and TDP-43 was used as a control of successful fractionation. (d) The proportion of cDNAs (out of all cDNAs that mapped to human genome) from the TDP-43 iCLIP experiments in nuclear or cytoplasmic fraction of SH-SY5Y cells that mapped to different RNA regions. (e) Analysis of UGUGU enrichment positions compared to randomised data around TDP-43 crosslink sites from iCLIP experiments in nuclear or cytoplasmic fraction of SH-SY5Y cells. (f) iCLIP cDNA counts for TDP-43 crosslink positions in the region of chromosome 11 that contains the NEAT1 and MALAT1 genes in three biological replicate experiments from SH-SY5Y cell line. The blue bars represent sequences on the sense strand of the genome, and the height of the bars corresponds to the cDNA count at individual crosslink sites.

Supplementary Figure 2



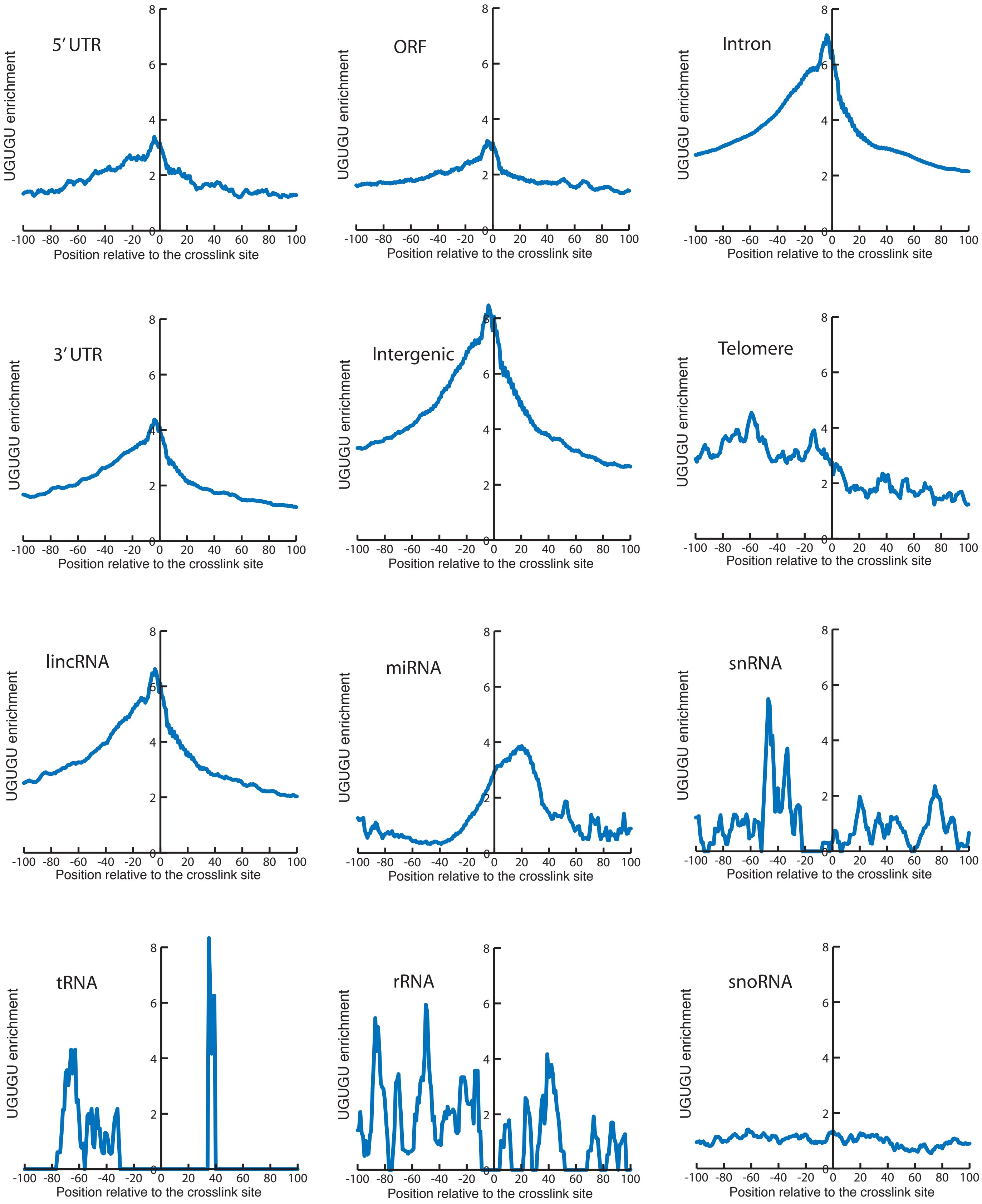
Supplementary Figure 2: Comparison of iCLIP binding with mRNA expression.
The log2 healthy/FTLD-TDP brain ratios of cDNA counts within the 3’UTR are shown for mRNAs with at least 10 cDNAs identified both in iCLIP and in mRNAseq in either healthy or FTLD-TDP brain. The plot areas corresponding to an at least two-fold change both in iCLIP and mRNAseq are shaded in blue, whereas those with opposite change are shaded in green.

Supplementary Figure 3



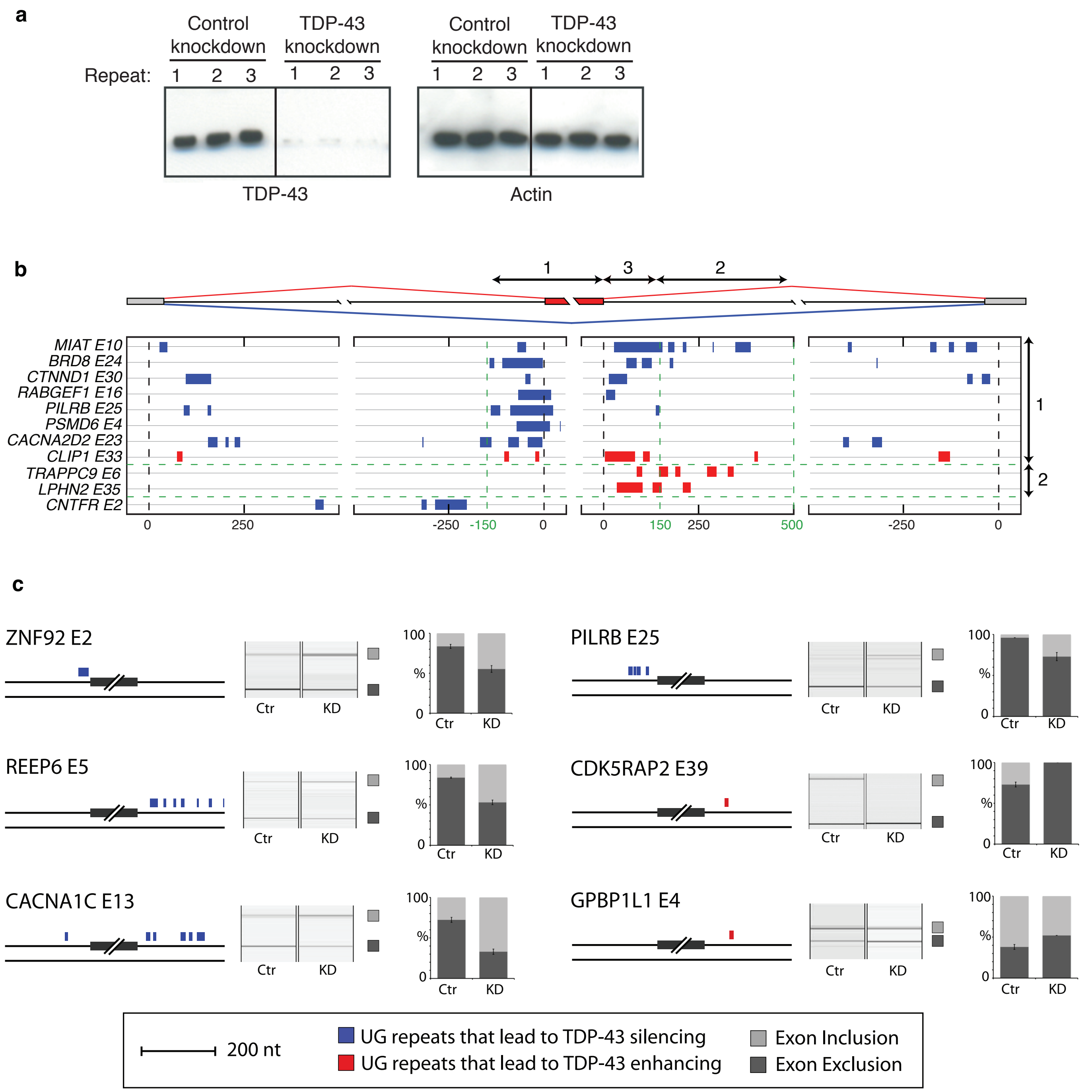
Supplementary Figure 3: Analysis of experimental overlap. (a-c) z-scores of pentamer occurrence withinthe 61 nt sequence surrounding crosslink sites (-30 nt to +30 nt) are shown for comparisons of iCLIP experiments from different samples. The sequences of the two most enriched pentamers and the Pearson correlation coefficient (r) between the two samples are given. (d) Overlap of binding to the TDP-43 crosslinking clusters between healthy and FTLD-TDP brain samples, depending on the number of cDNAs that mapped to the crosslinking clusters in healthy brain. (e) Overlap of binding to the TDP-43 crosslinking clusters between SH-SY5Y and hES cell lines, depending on the number of cDNAs that mapped to the crosslinking clusters in SH-SY5Y cells. (c) Overlap of binding to the TDP-43 crosslinking clusters between healthy brain and hES cell lines, depending on the number of cDNAs that mapped to the crosslinking clusters in SH-SY5Y cells. (f) Overlap of binding to the TDP-43 crosslinking clusters between SH-SY5Y and hES cell lines, depending on the number of cDNAs that mapped to the crosslinking clusters in SH-SY5Y cells.

Supplementary Figure 4



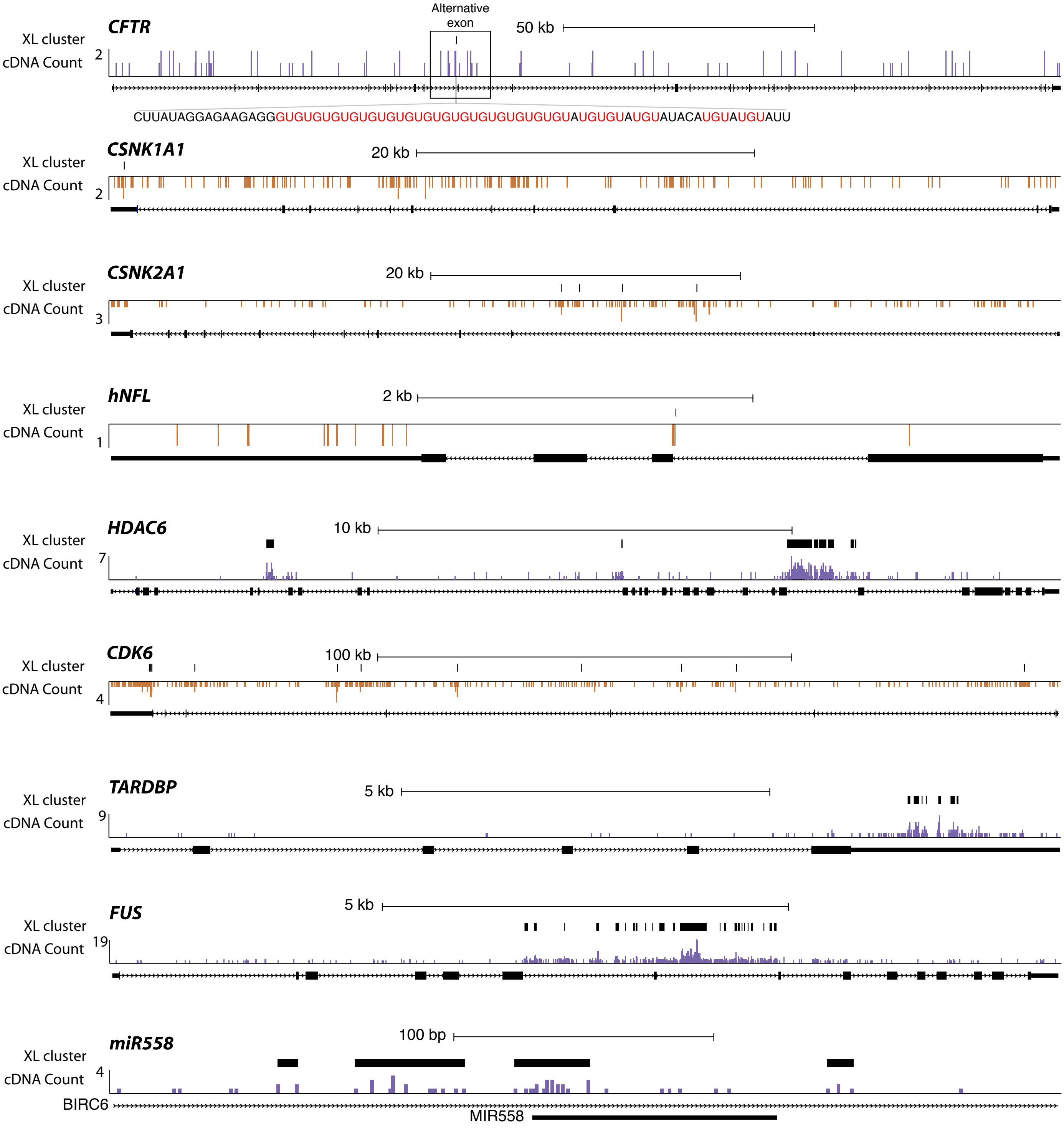
Supplementary Figure 4: UGUGU enrichment in different types of RNAs. Analysis of positions of UGUGU enrichment compared to randomised data in all TDP-43 iCLIP experiments (all data grouped) in different RNAs.

Supplementary Figure 5



Supplementary Figure 5: Analysis of predictive value of iCLIP crosslink clusters and UG repeats. (a) Western blots validation of TDP-43 knockdown in SH-SY5Y cells. Analysis of TDP-43 protein levels in mock-depleted and TDP-43-depleted SH-SY5Y cells with TDP-43 antibody. Actin antibody is used as a loading control. (b) The map of iCLIP crosslink clusters at positions within 500 nt of alternative exons and flanking exons that predicted TDP-43 regulated exons. The exons with 1 > ΔIrank ≥ 0.3 that were validated by RT-PCR are shown (red clusters in enhanced and blue clusters in silenced exons). Group 1 is identified by clusters in the region 150-0 nt upstream of the exon, group 2 by clusters in the region within 150nt upstream of the exon, but not upstream of the exon. (c) RT-PCR validated TDP-43 regulated alternative splicing events as predicted by UG-rich sequences within defined intronic and exonic regions surrounding alternative exons. At least 13 nucleotides had to be part of GUG or UGU trimers within a 30 nt sliding window in order to be considered as a UG-rich sequence. The UG-rich sequences are shown within 50 nt of exon and 200 nt of intron surrounding intron/exon boundaries in red if exon inclusion is enhanced by TDP-43, and in blue if it is silenced by TDP-43. Example PCR gels are shown for each exon, together with charts representing the levels of alternative exon inclusion and skipping in control (Ctr) and TDP-43 knockdown (KD) SH-SY5Y cells.

Supplementary Figure 6



Supplementary Figure 6: iCLIP binding sites in previously identified TDP-43 targets. hES iCLIP cDNA counts are shown for a variety of previously studied TDP-43 targets and other RNAs of interest. Crosslinking to the sense strand is represented by purple cDNA bars, and orange represents crosslinking to the antisense strand. Significant cDNA clusters (FDR < 0.05) are indicated by black bars above the cDNA counts. Exon/intron structure of the transcript is indicated below. In the first panel, the alternative exon 9 is marked by a box and the RNA sequence underlying the TDP-43 binding cluster is given below (UG and GU motifs in pink).

Supplementary table 1: Experimental annotation and barcoding.

Cells/tissue	**5' barcode**	**5' random barcode**	**3' barcode**	**Sequencing file**	**Sequence Length**
SH-SY5Y cells, replicate 1	CG,AAA	NN_0:1,NN_4:5	AT	iCLIP_TDP-43_SHSY5Y_High_hu_NNCGNNAAA_20091102_LUjt4_8.fq (high band on cDNA gel), iCLIP_TDP-43_SHSY5Y_Low_hu_NNCGNNAAA_20091102_LUjt5_8.fq (low band on cDNA gel)	54 nt
SH-SY5Y cells, replicate 2	CA	NNN_2:4	CC	iCLIP_TDP-43_SHSY5Y__hu_CANNN_20100222_LUjt3_3.fq	72 nt
SH-SY5Y cells, replicate 3	GGTT	NNN_0:2,NN_7:8	CC	iCLIP_TDP-43_SH-SY5Y_Total_Hs_NNNGGTTNN_20101125_LUjt8_1.fq	54 nt
SH-SY5Y cells, nuclear fraction	CAAT	NNN_0:2,NN_7:8	CC	iCLIP_TDP-43_SH-SY5Y_Nuclear_Hs_NNNTTGTNN_20101125_LUjt8_2.fq	54 nt
SH-SY5Y cells, cytoplasmic fraction	TTGT	NNN_0:2,NN_7:8	CC	iCLIP_TDP-43_SH-SY5Y_Cytoplasmic_Hs_NNNCAATNN_20101125_LUjt8_3.fq	54 nt

Embrionic stem cells	GA	NNN_2:4	AT	iCLIP_TDP-43_Embrionic-Stem-Cells__hu_GANNN_20100222_LUjt3_4.fq	72 nt

Healthy MTG Brain (C23, 20 year old female)	CG,CGA	NN_0:1,NN_4:5	GA	iCLIP_TDP-43_Brain_High_hu_NNCGNNCGA_20091102_LUjt4_2.fq (high band on cDNA gel), iCLIP_TDP-43_Brain_Low_hu_NNCGNNCGA_20091102_LUjt5_2.fq (low band on cDNA gel)	54 nt
Healthy MTG Brain (C30, 25 year old male)	TG	NNN_2:4	TG	iCLIP_TDP-43_Brain-C30__hu_TGNNN_20090724_LUjt1_1.fq	54 nt
Healthy MTG Brain (C25, 21 year old male)	TG	NNN_2:4	GA	iCLIP_TDP-43_Brain-C25__hu_TGNNN_20090724_LUe5_1.fq	54 nt

FTLD-TDP MTG Brain (F20, 51 year old male)	CA	NNN_2:4	CC	iCLIP_TDP-43_Brain-F20__hu_CANNN_20090724_LUjt1_3.fq	54 nt
FTLD-TDP MTG Brain (F21, 74 year old female)	GA	NNN_2:4	AT	iCLIP_TDP-43_Brain-F21__hu_GANNN_20090724_LUjt1_4.fq	54 nt
FTLD-TDP MTG Brain (F24, 48 year old male)	TG	NNN_2:4	TG	iCLIP_TDP-43_Brain-F24__hu_TGNNN_20090724_LUjt2_1.fq	54 nt

Supplementary table 2: Mapping information for iCLIP data.

	Sequence reads matching barcode	Reads mapping to genome	Unique cDNA reads after random barcode evaluation	Genomic positions of unique cross-link sites
SH-SY5Y cells, replicate 1	4211565	1973355	165027	143737
SH-SY5Y cells, replicate 2	5219220	3481515	792465	761229
SH-SY5Y cells, replicate 3	9145095	3633864	182011	165524
SH-SY5Y cells, nuclear fraction	1002293	564165	13785	13213
SH-SY5Y cells, cytoplasmic fraction	5650910	1169230	22820	19889
Human embrionic stem cells	9723824	6899033	2404697	2250869
Healthy brain (C23)	2001567	515205	103352	96879
Healthy brain (C30)	2375728	812621	75179	72279
Healthy brain (C25)	3495324	1998680	12274	11965
FTLD-TDP brain (F20)	2509880	473407	49209	47786
FTLD-TDP brain (F21)	1645633	609928	76136	75445
FTLD-TDP brain (F24)	1795393	296003	69605	68563

Supplementary table 3: RNA regions with a significant, greater than 0.025% change in proportion of cDNAs between control and FTLD-TDP brain (p value < 0.05 by Student's t-test, one-tailed, unequal variance).

		healthy brain samples							FTLD-TDP brain samples								
		cDNA count per region			percentage of cDNAs per region			Average % of cDNAs per region	cDNA count per region			percentage of cDNAs per region			Average % of cDNAs per region	% change (Healthy / FTLD-TDP)	p-value (Healthy / FTLD-TDP)
Gene Name	Region	C25	C30	C23	C25	C30	C23		F20	F21	F24	F20	F21	F24			
SLC1A2	3UTR	35	289	300	0.285	0.384	0.290	0.320	35	105	29	0.071	0.138	0.042	0.084	0.236	0.003
NRXN3	intron	30	158	332	0.244	0.210	0.321	0.259	36	119	76	0.073	0.156	0.109	0.113	0.146	0.013
OPCML	intron	18	178	201	0.147	0.237	0.194	0.193	32	91	81	0.065	0.120	0.116	0.100	0.092	0.025
ACCN1	intron	17	135	168	0.139	0.180	0.163	0.160	19	98	56	0.039	0.129	0.080	0.083	0.078	0.039
STX1B	3UTR	6	66	75	0.049	0.088	0.073	0.070	0	17	9	0.000	0.022	0.013	0.012	0.058	0.009
KCNQ3	intron	10	86	139	0.081	0.114	0.134	0.110	13	68	32	0.026	0.089	0.046	0.054	0.056	0.041
SHANK2	intron	7	88	73	0.057	0.117	0.071	0.082	10	26	16	0.020	0.034	0.023	0.026	0.056	0.042
HS6ST3	intron	7	73	111	0.057	0.097	0.107	0.087	18	35	16	0.037	0.046	0.023	0.035	0.052	0.030
DPP6	intron	6	73	115	0.049	0.097	0.111	0.086	14	36	18	0.028	0.047	0.026	0.034	0.052	0.049
PLXNA4	intron	5	68	66	0.041	0.090	0.064	0.065	6	14	8	0.012	0.018	0.011	0.014	0.051	0.034
NPTXR	3UTR	9	40	56	0.073	0.053	0.054	0.060	4	12	5	0.008	0.016	0.007	0.010	0.050	0.004
NDRG4	3UTR	10	27	66	0.081	0.036	0.064	0.060	4	16	6	0.008	0.021	0.009	0.013	0.048	0.029
FXYD6	3UTR	6	62	44	0.049	0.082	0.043	0.058	2	8	14	0.004	0.011	0.020	0.012	0.046	0.025
ATP2B2	intron	6	54	78	0.049	0.072	0.075	0.065	1	25	17	0.002	0.033	0.024	0.020	0.046	0.011
CAMK1D	intron	8	74	98	0.065	0.098	0.095	0.086	13	39	31	0.026	0.051	0.045	0.041	0.045	0.015
CAMK2A	intron	6	67	81	0.049	0.089	0.078	0.072	10	26	19	0.020	0.034	0.027	0.027	0.045	0.027
GABBR2	intron	8	43	70	0.065	0.057	0.068	0.063	9	20	15	0.018	0.026	0.022	0.022	0.041	0.000
PTPRJ	intron	10	35	52	0.081	0.047	0.050	0.059	6	17	15	0.012	0.022	0.022	0.019	0.041	0.028
MSRA	intron	6	64	79	0.049	0.085	0.076	0.070	14	31	15	0.028	0.041	0.022	0.030	0.040	0.024
GRIN2B	intron	8	38	70	0.065	0.051	0.068	0.061	6	25	15	0.012	0.033	0.022	0.022	0.039	0.004
DAB1	intron	9	59	95	0.073	0.078	0.092	0.081	16	47	25	0.033	0.062	0.036	0.043	0.038	0.017
ENC1	3UTR	7	58	73	0.057	0.077	0.071	0.068	19	27	14	0.039	0.035	0.020	0.031	0.037	0.006
XYLT1	intron	4	48	77	0.033	0.064	0.075	0.057	7	22	14	0.014	0.029	0.020	0.021	0.036	0.046
C1orf61	intron	4	47	37	0.033	0.063	0.036	0.044	1	10	6	0.002	0.013	0.009	0.008	0.036	0.026
SLC1A2	ORF	5	39	23	0.041	0.052	0.022	0.038	4	0	3	0.008	0.000	0.004	0.004	0.034	0.025
DSCAML1	intron	6	67	65	0.049	0.089	0.063	0.067	14	23	32	0.028	0.030	0.046	0.035	0.032	0.048
CALM1	3UTR	10	43	68	0.081	0.057	0.066	0.068	12	28	33	0.024	0.037	0.047	0.036	0.032	0.015
NTRK3	intron	8	69	72	0.065	0.092	0.070	0.076	18	47	24	0.037	0.062	0.034	0.044	0.031	0.030
AL049766.1	intron	4	39	29	0.033	0.052	0.028	0.038	1	10	3	0.002	0.013	0.004	0.006	0.031	0.017
TMEM132D	intron	6	42	80	0.049	0.056	0.077	0.061	12	24	26	0.024	0.032	0.037	0.031	0.030	0.029

Supplementary table 3 cont: RNA regions with a significant, greater than 0.025% change in proportion of cDNAs between control and FTLD-TDP brain (p value < 0.05 by Student's t-test, one-tailed, unequal variance).

		healthy brain samples							FTLD-TDP brain samples								
		cDNA count per region			percentage of cDNAs per region			Average % of cDNAs per region	cDNA count per region			percentage of cDNAs per region			Average % of cDNAs per region	% change (Healthy / FTLD-TDP)	p-value (Healthy / FTLD-TDP)
Gene Name	Region	C25	C30	C23	C25	C30	C23		F20	F21	F24	F20	F21	F24			
JAKMIP1	intron	5	45	46	0.041	0.060	0.045	0.048	4	12	23	0.008	0.016	0.033	0.019	0.029	0.019
LRFN2	intron	3	37	52	0.024	0.049	0.050	0.041	3	12	10	0.006	0.016	0.014	0.012	0.029	0.030
AC019118.2	intron	6	32	49	0.049	0.043	0.047	0.046	5	22	9	0.010	0.029	0.013	0.017	0.029	0.014
KCNB1	intron	4	20	47	0.033	0.027	0.045	0.035	3	6	4	0.006	0.008	0.006	0.007	0.028	0.017
GABRG3	intron	9	40	65	0.073	0.053	0.063	0.063	12	36	25	0.024	0.047	0.036	0.036	0.027	0.019
TNR	intron	5	30	66	0.041	0.040	0.064	0.048	12	15	14	0.024	0.020	0.020	0.021	0.027	0.036
DLGAP3	intron	3	37	24	0.024	0.049	0.023	0.032	2	7	4	0.004	0.009	0.006	0.006	0.026	0.044
KCNIP1	intron	6	28	27	0.049	0.037	0.026	0.037	3	12	9	0.006	0.016	0.013	0.012	0.026	0.021
ODZ4	intron	8	36	63	0.065	0.048	0.061	0.058	20	23	18	0.041	0.030	0.026	0.032	0.026	0.010
EML6	intron	6	23	40	0.049	0.031	0.039	0.039	7	12	8	0.014	0.016	0.011	0.014	0.026	0.017
SCAPER	intron	1	6	13	0.008	0.008	0.013	0.010	21	29	16	0.043	0.038	0.023	0.035	-0.025	0.022
ARHGEF4	intron	2	22	14	0.016	0.029	0.014	0.020	20	39	30	0.041	0.051	0.043	0.045	-0.025	0.008
U1_121	ncRNA	2	4	2	0.016	0.005	0.002	0.008	20	21	25	0.041	0.028	0.036	0.035	-0.027	0.005
HIF3A	intron	1	19	6	0.008	0.025	0.006	0.013	14	36	32	0.028	0.047	0.046	0.041	-0.027	0.017
KDM6A	intron	0	3	17	0.000	0.004	0.016	0.007	15	27	26	0.030	0.035	0.037	0.034	-0.028	0.009
AC058791.2	intron	5	19	55	0.041	0.025	0.053	0.040	30	54	49	0.061	0.071	0.070	0.067	-0.028	0.030
NCOA1	intron	1	9	19	0.008	0.012	0.018	0.013	17	38	26	0.035	0.050	0.037	0.041	-0.028	0.006
ERBB2IP	intron	3	9	16	0.024	0.012	0.015	0.017	20	35	36	0.041	0.046	0.052	0.046	-0.029	0.002
GPR98	intron	3	16	17	0.024	0.021	0.016	0.021	34	34	28	0.069	0.045	0.040	0.051	-0.031	0.034
NCAM2	intron	8	38	86	0.065	0.051	0.083	0.066	43	82	67	0.087	0.108	0.096	0.097	-0.031	0.031
ADCY2	intron	1	5	34	0.008	0.007	0.033	0.016	25	38	30	0.051	0.050	0.043	0.048	-0.032	0.027
MAPK10	intron	5	14	59	0.041	0.019	0.057	0.039	38	64	36	0.077	0.084	0.052	0.071	-0.032	0.049
KIAA1267	intron	0	6	37	0.000	0.008	0.036	0.015	24	43	26	0.049	0.056	0.037	0.048	-0.033	0.037
NEBL	intron	2	30	44	0.016	0.040	0.043	0.033	34	65	34	0.069	0.085	0.049	0.068	-0.035	0.032
ZBTB16	intron	1	22	39	0.008	0.029	0.038	0.025	21	47	54	0.043	0.062	0.078	0.061	-0.036	0.029
PRKG1	intron	2	22	31	0.016	0.029	0.030	0.025	44	39	35	0.089	0.051	0.050	0.064	-0.038	0.042
NFIA	intron	1	19	27	0.008	0.025	0.026	0.020	30	45	46	0.061	0.059	0.066	0.062	-0.042	0.006
SNORA14_1	ncRNA	1	2	4	0.008	0.003	0.004	0.005	33	28	31	0.067	0.037	0.045	0.049	-0.045	0.018
MALAT1	ncRNA	10	94	196	0.081	0.125	0.190	0.132	134	198	243	0.272	0.260	0.349	0.294	-0.162	0.009
NEAT1	ncRNA	12	143	99	0.098	0.190	0.096	0.128	123	307	308	0.250	0.403	0.442	0.365	-0.237	0.018

Supplementary table 4: RT-PCR primers for validation of splicing events identified by the microarray, iCLIP crosslink clusters or UG repeats.

Validation of exons predicted by microarray

Gene Symbol	Exon	alternative exon (hg19)	flanking exons (hg19)	Strand	dI rank	Array dI (%)	PCR dI (%)	Cont. stdev	KD stdev	Forward Primer	Reverse Primer	PCR product	
AL080249.26	E8	chr20:37054584-37054696	chr20:37050889-37055062	-	-3.15	-24	-8	1	2	CTGACCTCTGTCCTGTGTGC	GCTTCCCCTACATCATCCAA	121	238
BCL2L11	E5	chr2:111881628-111881716	chr2:111881447-111907621	+	1.93	18	32	0	1	TCTGAGTGTGACCGAGAAGG	TCTTGGGCGATCCATATCTC	245	155
CDK5RAP2	E39	chr9:123169291-123169526	chr9:123166392-123170625	-	7.78	32	27	3	0	AAGGCGTATGAGAAGCTGGA	CCTGGGAGGAATCAAACAGA/ TTCTAGTTGCAAGCGCAGAG	135	103
CENPO	E3	chr2:25016722-25016834	chr2:25016371-25022544	+	-2.07	-18	-31	4	3	GGTTTTGAAGACGTGGATGG	TCTCGCAGACGCCTTAGTTT	291	177
DDR1	E12	chr6:30853403-30853457	chr6:30848903-30856465	+	-1.57	-14	-36	3	1	CCGGGAAGAGCGATGAGA	CCTTCATGTCAGCATCTCCA	187	131
FAM114A2	E3	chr5:153417949-153417994	chr5:153414528-153418417	-	-2.64	-20	-20	2	1	ACGCCGAGTCGGCTGCCGTG	ATTCTTGGCTGGCTCACAGT/ CAGGTCACCGCCTCCCTAAC	148	135
GPBP1L1	E4	chr1:46151249-46151292	chr1:46126898-46152084	-	1.09	10	14	3	1	CCATTTTGGGAAGAGCTTTG	CAATCATTTGAAAGGCCACA	196	150
HP1BP3	E11	chr1:21102162-21102498	chr1:21100104-21103090	-	-1.67	-21	-18	4	3	AAAGGGGGAACCTGAGAATG/ CCCTGGTCTTTTCCTCCTTC	GCCCAGGAAGGAATTGTTTT	140	123
KIF1B	E16	chr1:10339156-10339196	chr1:10338187-10342458	+	1.12	14	14	6	3	AAGGACCTTCTTCGTGCTCA	CGTCAAGCCCACCTGACTAC	160	108
KIF2A	E19	chr5:61668266-61668378	chr5:61662309-61669514	+	1.75	17	26	3	0	TCTCGTACCTGCATGATTGC	GGTGCATTATTGGACGAACA	256	142
MADD	E40	chr11:47348096-47348103	chr11:47346129-47348289		1.37	23	-13	2	1	CTGGAAGGGATCAACCTCAA	ACAAAGACGCCTCGAACTGT	101	92
MAST2	E11	chr1:46466632-46466651	chr1:46463472-46468477		1.38	13	29	5	1	CCCGGAATTTCTCTCCAAAT	CCTGAAGAGGGCAAAGAGG	121	100
MEF2D	E11	chr1:156446287-156446306	chr1:156445030-156446804	-	-2.04	-17	-29	7	6	AGGCAGGAAAGGGGTTAATG	CTGAGTAAACTCGGCGTTGC	136	115
NAV2	E23	chr11:20072836-20072879	chr11:20070742-20075626	+	2.12	18	18	4	3	CGACTCCTTGACTCCCTTTG	CTGCAGAATGGGACCGTAAC	191	125
no symbol	E2	chr19:58836642-58836727	chr19:58826403-58838189	+	3.37	26	79	3	0	TGGCGAACTGACAAACTCTG	CTGAGTGGGAGAGCAGCAG	550	147
OGT	E9	chrX:70775041-70775235	chrX:70774443-70775804	+	-2.22	-11	-31	4	1	AAGAGGCACGCATTTTTGAC	GCGAACTGCCTCTTCAATGT	338	142
PFKP	E20	chr10:3167290-3167441	chr10:3162237-3172011	+	-3.97	-24	-13	3	3	ATGGTTCCCGCTACTGTGTC/ ACAGGCTTTGCCTGACAAGT	TGGTCTCGATGATGAACACG	186	148

dI: change in the percentage of exon including isoform

dI rank: significance ranking of splicing changes as determined by ASPIRE3

Supplementary table 4 cont.: RT-PCR primers for validation of splicing events identified by the microarray, iCLIP crosslink clusters or UG repeats.

Validation of exons predicted by microarray

Gene Symbol	Exon	alternative exon (hg19)	flanking exons (hg19)	Strand	dI rank	Array dI (%)	PCR dI (%)	Cont. stdev	KD stdev	Forward Primer	Reverse Primer	PCR product	
PHF14	E3	chr7:11022000-11022786	chr7:11014573-11030330	+	-3.24	-20	-7	1	2	TCTCTGCTGGAAGCTCTTGA	CTCCCAGACAAACACAGCAA/GGAACCTTCTCCACTGTCCTT	126	149
POLDIP3	E3	chr22:42997977-42998062	chr22:42995800-42998776	-	15.88	60	67	1	4	TGCTCTGAAGCTCACCAAAA	GGAACGGAAGCTATACCATCAT	161	75
RANBP1	E5	chr22:20110104-20110220	chr22:20109945-20112831	+	-1.59	-14	-4	0	0	CTGAAGATCTGTGCCAACCA	TGCATTCTCAGCATTCAGGA	277	159
RAP1GAP	E20	chr1:21939530-21939639	chr1:21938603-21939669	-	-3.08	-24	-7	4	2	GATGGCAAAGTTGGTGTGTG	TGGCACACAGAAGGTGCTTA/GACATGCTCGTCAAAGGTGA	148	100
RGL1	E2	chr1:183666539-183666647	chr1:183605637-183711261	+	-2.69	-23	-28	2	3	GACGGAGTCGAAAGGTCAAA	CCACAGGTTTCACCTCCATC	239	129
RWDD1	E2	chr6:116894022-116894148	chr6:116892819-116895221	+	-10.98	-42	-44	2	1	GGGCCACGATGACAGATTAC	AGGAGGGTTGGATCAATGTG	237	109
SLC9A8	E7	chr20:48467348-48467381	chr20:48466218-48471975	+	-1.63	-15	-27	3	3	GGCAATCTCCGCTTTTGTAG	AAAGACCAGCATGTTGAGCA	184	149
TNIK	E15	chr3:170857260-170857345	chr3:170856169-170858188	-	-2.59	-22	-21	2	3	CGGCACTATGAGGAGCAGAT	CTTCCAATTGTTTGCGCTTA	172	86
UBTF	E10	chr17:42289713-42289822	chr17:42289375-42290187	-	-4.68	-27	-13	1	1	CACCCACGAGAAGAAGGTGT	ATGTTCAGCTCTGGGTGCTT/GTCCGAGAGCTGAGACCACT	195	84
ZFYVE20	E2	chr3:15139627-15139781	chr3:15138218-15140542	-	8.97	56	65	3	0	GAGTGACGGACGGTGGAG	TAAAATCTGGGAGGGCACAG/CCTTGATTGTTTATTCAACCACAT	146	130
ZNF569	E3	chr19:37957642-37957706	chr19:37956376-37957986	-	-10.99	-44	-9	1	2	GGCAGCGTTAAGTGTGGAAT	GCACCTAGGATTGTGCCTGT	194	128
ZNF827	E15	chr4:146684243-146684274	chr4:146682751-146686140	-	2.64	19	18	3	2	CCGGATGTTTGAGTGTGATG	GTCTTGAGGTGGCACTCCAT	192	159
CNOT1	E23	chr16:58590752-58590766	chr16:58589813-58590766	-	-4.1	-24	No change			CTCCCTGCAGTGAATAACGA	AAAGTGCTGGTTTGCCTCTG	135	120
TGOLN2	E6	chr2:85549865-85549881	chr2:85549865-85552038	-	-2.68	-27	No change			AAGGCCAGTGACTACCAACG	TCAAAGCTGAAACGAGAGCA/ACTGCTGCTTCATGGCAAAT	130	110
UNC13A	E37	chr19:17734352-17734356	chr19:17732680-17735473	-	1.09	27	No change			CCATGGAGAAAACCATCGTC	CTTGGCTGCATTGAAGATCA	149	143
ANXA1	E4	chr9:75769378-75769497	chr9:75769259-75773438	+	2.22	32	One PCR product only			CCAGTGTGAAATCTTCAGAGAA	GCCTGCTTGAGGAATTCTGATA	229	108
ATF7IP2	E5	chr16:10508432-10508469	chr16:10479980-10513438	+	-3	-29	One PCR product only			AAGTGGCTGTGGGTATTGCT	AGTGAGACCGTGTCTCAAAG	429	110

dI: change in the percentage of exon including isoform

dI rank: significance ranking of splicing changes as determined by ASPIRE3

Supplementary table 4 cont.: RT-PCR primers for validation of splicing events identified by the microarray, iCLIP crosslink clusters or UG repeats.

Validation of exons predicted by microarray

Gene Symbol	Exon	alternative exon (hg19)	flanking exons (hg19)	Strand	dI rank	Array dI (%)	PCR dI (%)	Cont. stdev	KD stdev	Forward Primer	Reverse Primer	PCR product	
CYP2C18	E5	chr10:96466542-96466717	chr10:96454835-96480153	+	2.41	26	One PCR product only			TCAGGATTCTGAGCTCTCCA	GGGTACTTCAGCAGGAGCAG	337	160
KIAA1324	E11	chr1:109729172-109730292	chr1:109727756-109730799	+	17.05	69	One PCR product only			CCAGCTTGCACAGACAAAGA	ACATGGGAAGAATGGAGCTG/ CCTCAAGGTCCTCGCTACAG	115	173
MAPK8	E11	chr10:49632558-49632628	chr10:49632203-49633931	+	-1.05	-12	One PCR product only			ACTACAGAGCACCCGAGGTC	TCCCTTCCTGGAAAGAGGAT/ CGTAAGTCCTTACTGTTGGTTGC	118	138
DNM1	E30	chr9:131015381-131015416	chr9:131013220-131016933	+	-1.02	-19	PCR didn't produce correct products			GCTTCCCCTGACCCTTTC	TCCGAGAAGAGGAGGGATCT	194	157
DOK3	E10	chr5:176930753-176931661	chr5:176930394-176931744	-	-1.34	-36	PCR didn't produce correct products			TACAGCTGGCCCTACCACTT	GGGTGCTGAAGGCAAAGAG/ CGCAGCAAACTCTCATCAAG	140	112
FAM153B	E23	chr5:175536111-175536140	chr5:175533594-175540344	+	1.19	41	PCR didn't produce correct products			CTGGAAGGGATCAACCTCAA	GTCCAGCTGCTCCAGTTCTC/ ATTCCTCCAGGACAAAGACG	103	171

Validation of exons predicted by iCLIP

Gene Symbol	Exon	alternative exon (hg19)	flanking exons (hg19)	Strand	dI rank	Array dI (%)	PCR dI (%)	Cont. stdev	KD stdev	Forward Primer	Reverse Primer	PCR product	
BRD8	E24	chr5:137495245-137495288	chr5:137492957-137495758	-	-0.86	-16	-12	1	1	GATTCTACCCGCAAACAGGA	TTAGGTTCGGCTCAAAGAGG	172	127
CACNA2D2	E23	chr3:50410497-50410516	chr3:50407812-50412168	-	-0.25	-7	-8	1	1	CAAGCCAATCTCAGTGACCA	TGTTGTCTGAGGCATTCAGG	154	133
CLIP1	E33	chr12:122760987-122760991	chr12:122758710-122762701	-	0.66	17	6	3	2	ACACTCAGGCAGACGAGGAT	CTGCTTCTGACATCATCTCCA	132	126
CNTFR	E2	chr9:34581094-34581203	chr9:34568980-34589553	-	0	-1	-21	0	2	ATCCGGCGGGCGGTGCTA	ATGGCAGTGTCACGTCAGAG	347	235
CTNND1	E30	chr11:57573934-57573950	chr11:57573508-57574387	+	-0.69	-17	-18	5	7	ACAATACTGGGCCACATGCT	TCATAGCCTCGAGCTGGACT	147	125
LPHN2	E35	chr1:82455342-82455379	chr1:82451040-82456075	+	0.36	8	18	3	2	GCACTTCAACACTTAATCAAGGA/ GCCTGAGCAGACAGACATGA	TGCAGCGAGTAGCTGTTGTT	120	106
MIAT	E10	chr22:27063149-27063273	chr22:27062901-27064121	+	-0.03	-3	-43	1	1	CTGTGTGTCGGCTGGATG	ACCCGAGGTCCAAAGAGAAG	213	88
PILRB	E25	chr7:99951519-99951635	chr7:99951107-99952766	+	-0.62	-14	-23	1	5	CTCTCACTGCCTCCCTTCAC	CTGCTCTCCACGTCCCTTAT	257	139
PSMD6	E4	chr3:64008320-64008477	chr3:64008200-64008951	-	-0.39	-7	-6	0	0	AAGAACCCCGACTTGCGTAT	GGATTTGCACAAGGCTTCAT	308	147

Supplementary table 4 cont.: RT-PCR primers for validation of splicing events identified by the microarray, iCLIP crosslink clusters or UG repeats.

Validation of exons predicted by iCLIP

Gene Symbol	Exon	alternative exon (hg19)	flanking exons (hg19)	Strand	dI rank	Array dI (%)	PCR dI (%)	Cont. stdev	KD stdev	Forward Primer	Reverse Primer	PCR product	
RABGEF1	E16	chr7:66235311-66235454	chr7:66234013-66236870	+	-0.3	-11	-32	3	2	CTGCAGACAAAGCACACGAT	GCACAGGAGATCCGATTGAT	274	128
TRAPPC9	E6	chr8:141436715-141436740	chr8:141415798-141445211	-	0.53	9	20	5	1	GAACTGGTGGGAAGAGTGGA	TCTTAGCACGTCCGATCTCA	165	138
CLASP1	E22	chr2:122202516-122202538	chr2:122187754-122203025	-	0.44	21	No Change			TGCCAAATCCAAAGTCTCCT	CCCCGGTTATCAGGTGTAGA	180	156
ZBTB8OS	E4	chr1:33100370-33100393	chr1:33099712-33116034	-	1.03	22	No Change			GAGCCCCTCCAAACAGTAGA	AGGGTGCTTGGACAATGAAA	204	150
C19orf2	E3	chr19:30462101-30462134	chr19:30433572-30476130	+	-0.01	-1	No Change			GCGCGAGGAGCAGGAAAA	CCAGTTGTCCCCCAGTAAAA	179	214
EIF4G2	E20	chr11:10823209-10823321	chr11:10822635-10823596	-	0.07	3	No Change			ATCGCAGTTTGGAGAGATGG	CTGTCCCAGAGGTGGTGTTT	169	283
MDM4	E11	chr1:204506588-204506625	chr1:204501375-204507337	+	0.02	2	No Change			CTGGGACGTCAGAGCTTCTC	GCAGTGTGGGGATATCGTCT	155	223
WDR42A	E7	chr1:160231076-160231115	chr1:160213825-160232066	-	0	0	No Change			CATCGCTGGGAGTACTGGTT	GATGTCTCCCTCCCCTCTTC	166	240
BRD9	E15	chr5:883448-883522	chr5:881298-884053	-	-0.17	-6	One PCR product only			TGCCAAGAAATCCAAAAAGC	TCCCCGTTCCTCTTCAGATA	198	298
CLDN12	E4	chr7:90041322-90041363	chr7:90034923-90041958	+	-0.09	-4	One PCR product only			GTGTGAGAAGCAGGCTCAGA	CAGGAAGGAAAGGACTGTGG	179	224/278
EXOSC10	E14	chr1:11140572-11140606	chr1:11140572-11140821	-	0.03	2	One PCR product only			TCGCAGGGAAGATGAAAGTT	GCCTGCTGGATTAAAAGGTG	114	165
FUS	E11	chr16:31198124-31198157	chr16:31196501-31199646	+	0.02	0	One PCR product only			TGGCTATGAACCCAGAGGTC	GCTGTCCCGTTTTCTTGTTT	214	248
IMMT	E18	chr2:86380684-86380697	chr2:86378644-86385715	-	0	0	One PCR product only			CCAAGCTCGGGATGACTTTA	TTCTGCCAGCTCTCTGTTCA	160	175
INVS	E7	chr9:102894153-102894236	chr9:102892027-102977207	+	0	-1	One PCR product only			TTCTGAAGGCAGGAGCAGAT	TTCCAGATCCTTTTGCATCC	140	225
MEIS2	E2	chr15:37391657-37391800	chr15:37390401-37393093	-	-0.02	-2	One PCR product only			CGCGGAATTAAAAGGAACAA	GAGGGTCTCCGTACATGGAA	168	313
NAP1L1	E5	chr12:76462690-76462764	chr12:76461254-76467983	-	0	0	One PCR product only			CTGGCTCCCCATACTAGTCG	CTTGAAGGGCTGCAAGAATC	117	203
NAV2	E8	chr11:19905841-19905908	chr11:19901674-19913980	+	0.09	3	One PCR product only			CAGCAGTCAAAAGCACAAGC	TGCGTCGGTTGTTAGCAGTA	121	190
PRSS12	E9	chr4:119227101-119227158	chr4:119220094-119229591	-	-0.01	-1	One PCR product only			CAGTGGGGAACAATCTGTGA	AATCCTACCGCAGTGCATTC/CTCCTGCATCTTCACTGTGG	113	172/232
SEPT	E12	chrX:118759299-118759342	chrX:118750706-118763281	-	-0.2	-11	One PCR product only			AACTGCACCAGGACGAGAAG	GTGCAGCTGCTGAAGACAAG	194	239
SETD5	E27	chr3:9510204-9510259	chr3:9506357-9512143	+	-0.6	-9	One PCR product only			CCAGGAGAAGAGGAGTGTCG	TGCCAAATCCAGGTCTTTTC	129	187

Supplementary table 4 cont.: RT-PCR primers for validation of splicing events identified by the microarray, iCLIP crosslink clusters or UG repeats.

Validation of exons predicted by UG-rich motifs

Gene Symbol	Exon	alternative exon (hg19)	flanking exons (hg19)	Strand	dI rank	Array dI (%)	PCR dI (%)	Cont. stdev	KD stdev	Forward Primer	Reverse Primer	PCR product	
CACNA1C	E13	chr12:2656617-2656690	chr12:2622151-2659109	+	-0.6	-14	-40	2	3	CCTGAGAATGAGGACGAAGG	TTTTCTCCCTCGATGTCACC	192	118
PILRB	E25	chr7:99951519-99951635	chr7:99951107-99952766	+	-0.62	-14	-23	1	5	CTCTCACTGCCTCCCTTCAC	CTGCTCTCCACGTCCCTTAT	257	139
REEP6	E5	chr19:1496591-1496670	chr19:1496453-1497173	+	-0.56	-13	-32	1	3	GGGGCTCTCATGCTGTATCA	GGCTGCTTCACTTGTCCTTC	253	173
UNC84A	E16	chr7:889629-889670	chr7:889241-891021	+	0.65	11	5	1	1	GGGAGAATGAATGTGAGAGAGG	CAGAATCTGCAGCAAGAAGTAA	201	90
ZNF92	E2	chr7:64852816-64852941	chr7:64838914-64863254	-	-0.37	-19	-28	2	3	TGCGTCCTGTGCTGATAAAG	TTGGGCAAAATGAGAACACA	252	126
CNOT10	E2	chr3:32737066-32737319	chr3:32726975-32745360	+	-0.77	-12	One PCR product only			GAAGAACCCGAGTCGAAGC	TGCTGTGGGATTATGTGCAG	134	117
PDLIM3	E8	chr4:186429454-186429716	chr4:186427807-186435424	-	-0.42	-7	One PCR product only			AACAGCCCTAGGGGAAACAC/ GCCAATTGGGCTCTATTCAA	CCTGGAGCACTCTGAAGGAG	140	194
PTGFR	E8	chr1:78997899-78998038	chr1:78959227-79002091	+	-0.66	-14	One PCR product only			GCCCTTGGTGTTTCATTGTT	ATTCCAATGTTGGCCATTGT	330	190
SCUBE1	E7	chr22:43651495-43651583	chr22:43634961-43654225	-	-0.73	-23	One PCR product only			CAGAAGTACGCCCTCCACTC	ACACAGCTCCACATGCTGAC	267	178
SETD5	E27	chr3:9510204-9510259	chr3:9506357-9512143	+	-0.6	-9	One PCR product only			CCAGGAGAAGAGGAGTGTCG	TGCCAAATCCAGGTCTTTTC	186	129
KLRC4	E10	chr12:10542543-10542626	chr12:10541475-10544434	-	-0.66	-49	PCR didn't produce correct products			TCAAGATCTTCCCTCTCTGAGC/ TGGCTGTAAGTTGCCATTTG	TTTTGCCATCGTGTTGAAAA	292	208
PDGFB	E2	chr22:39636861-39636919	chr22:39631880-39639906	-	-0.32	-12	PCR didn't produce correct products			GCGCTCTTCCTGTCTCTCTG	TGACTGAGCAGGAATGGTGA/ ATCAAAGGAGCGGATCGAGT	102	111

dI: change in the percentage of exon including isoform

dI rank: significance ranking of splicing changes as determined by ASPIRE3