

Quantitative analysis of cryptic splicing associated with TDP-43 depletion

Jack Humphrey^{1,2}, Warren Emmett^{1,3,5}, Pietro Fratta⁴,
Adrian M. Isaacs² and Vincent Plagnol¹

September 19, 2016

¹ University College London Genetics Institute, Gower Street, London WC1E 6BT, UK

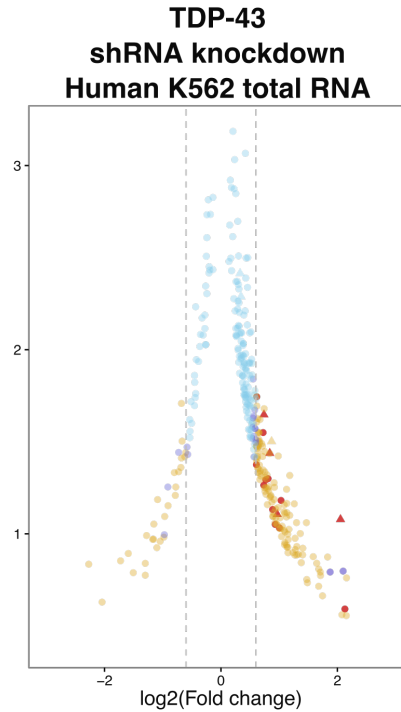
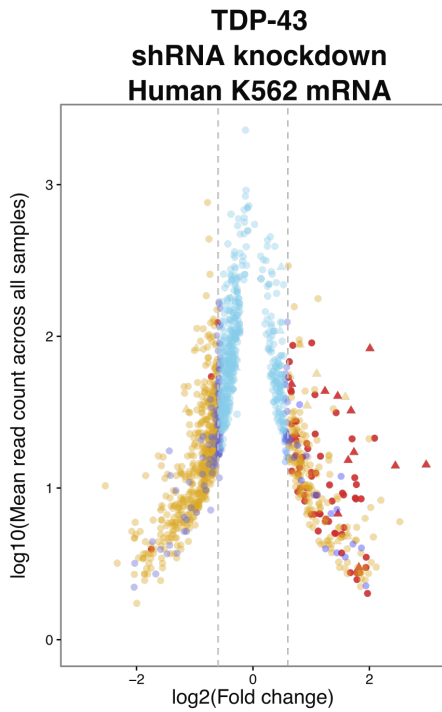
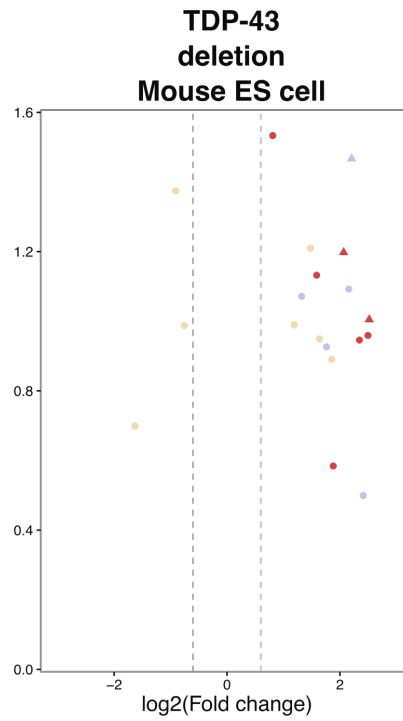
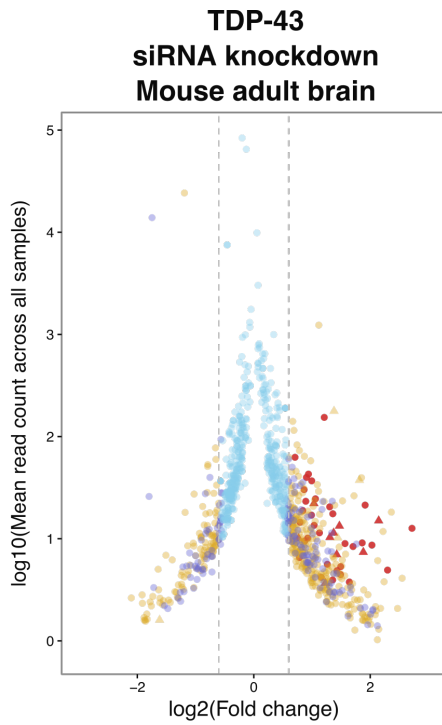
² Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

³ Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK

⁴ Sobell Department of Motor Neuroscience and Movement Disorders, UCL Institute of Neurology, Queen Square, London WC1N 3BG, UK

⁵ RNA Networks Laboratory, The Francis Crick Institute, 1 Midland Road, London NW1 1AT, UK

Supplementary Figures S1 and S4

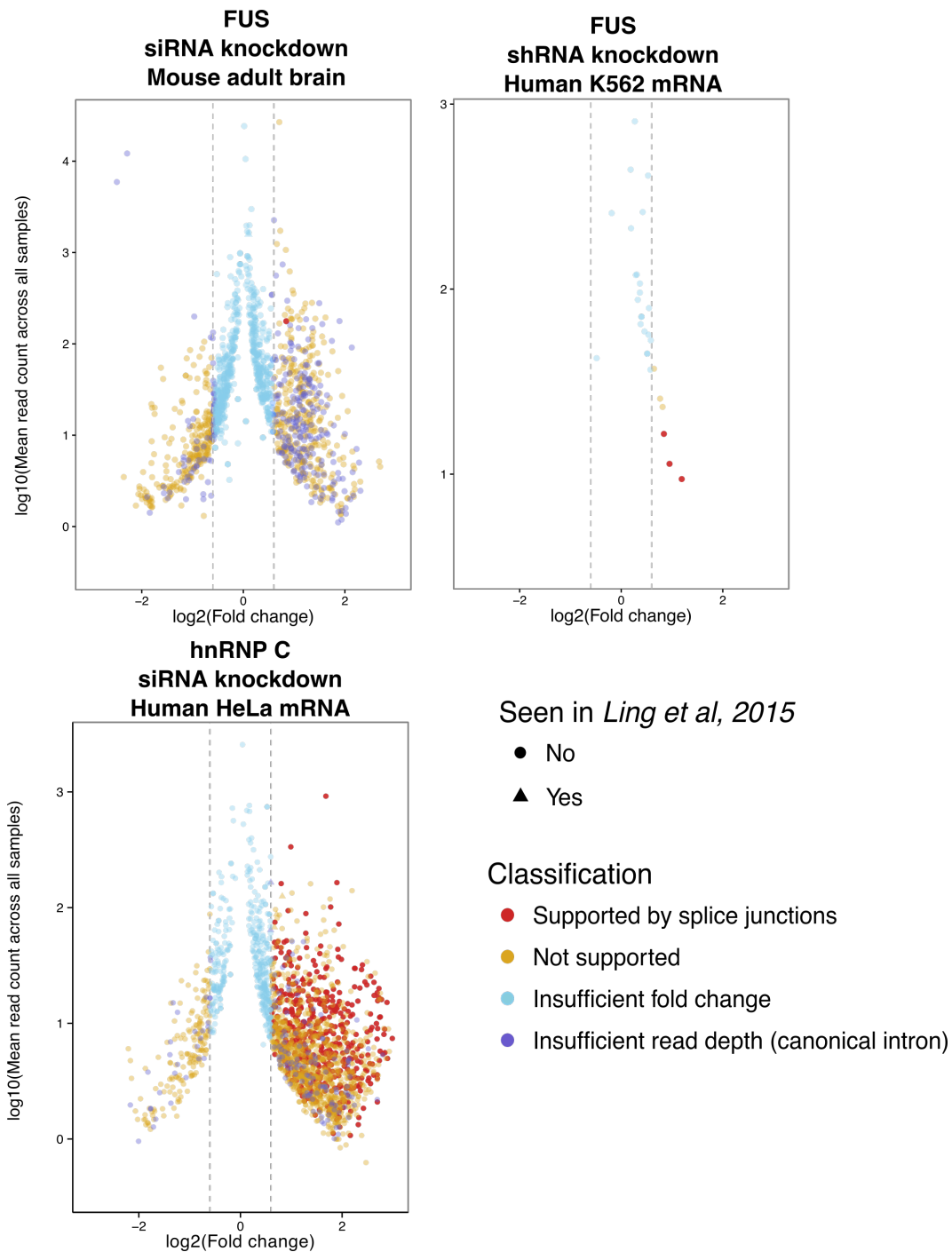


Seen in *Ling et al, 2015*

- No
- ▲ Yes

Classification

- Supported by splice junctions
- Not supported
- Insufficient fold change
- Insufficient read depth (canonical intron)



Supplementary Figure 1. Pre-classification output of *CryptEx* pipeline demonstrates a variety of novel splicing events in TDP-43, FUS and hnRNP C depletion data. Every novel splicing event plotted by mean depth of reads covering the novel event against log₂(fold change) between depletion and control samples. The cryptic exon classifier throws out any splicing event where the canonical intron in which the cryptic splicing event appears is represented by less than five spliced reads (purple) or where the |log₂(fold change)| < 0.6 (light blue). Splicing events are classified as cryptic exons if the spliced reads agree with the rest of the reads and have at least 1 spliced read per sample (red). Those that fail this step are coloured orange.

Human cryptic exons flanked by 100nt

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)
1		1e-12	-2.972e+01	12.63%	0.44%	166.5bp (201.8bp)
2 *		1e-11	-2.706e+01	11.58%	0.41%	74.6bp (106.1bp)
3 *		1e-11	-2.685e+01	6.32%	0.01%	89.4bp (0.0bp)
4 *		1e-11	-2.682e+01	23.16%	3.45%	105.3bp (216.9bp)
5 *		1e-11	-2.659e+01	7.37%	0.04%	81.5bp (66.3bp)
6 *		1e-10	-2.470e+01	20.00%	2.71%	75.4bp (216.6bp)
7 *		1e-10	-2.440e+01	24.21%	4.35%	94.6bp (182.4bp)
8 *		1e-10	-2.354e+01	14.74%	1.29%	73.1bp (104.5bp)
9 *		1e-10	-2.341e+01	11.58%	0.62%	94.0bp (108.6bp)
10 *		1e-10	-2.340e+01	10.53%	0.45%	73.9bp (85.8bp)

Mouse cryptic exons flanked by 100nt

Rank	Motif	P-value	log P-pvalue	% of Targets	% of Background	STD(Bg STD)
1 *		1e-11	-2.658e+01	19.23%	0.54%	24.4bp (82.6bp)
2 *		1e-10	-2.528e+01	21.15%	0.92%	65.3bp (66.0bp)
3 *		1e-9	-2.276e+01	15.38%	0.36%	58.9bp (40.8bp)
4 *		1e-9	-2.245e+01	17.31%	0.60%	81.4bp (73.2bp)
5 *		1e-9	-2.156e+01	11.54%	0.09%	78.6bp (9.1bp)
6 *		1e-9	-2.103e+01	13.46%	0.24%	72.1bp (74.4bp)
7 *		1e-8	-2.067e+01	9.62%	0.03%	102.8bp (24.8bp)
8 *		1e-8	-2.002e+01	15.38%	0.53%	75.7bp (119.5bp)
9 *		1e-8	-1.890e+01	9.62%	0.05%	47.7bp (61.9bp)
10 *		1e-7	-1.832e+01	15.38%	0.70%	124.3bp (143.7bp)

Supplementary Figure 4. Motif finding with *HOMER*. The top ten motifs reported by the algorithm when comparing flanked cryptic exons with adjacent intronic sequence. The red asterisk indicates that the motif is potentially a false positive result.