

Supplementary material accompanying the manuscript “Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain”

Boris Rogelj^{1, 6, §}, Laura E. Easton^{2, §}, Gireesh Bogu³, Lawrence W. Stanton³, Gregor Rot⁴, Tomaž Curk⁴, Blaž Zupan⁴, Yoichiro Sugimoto², Miha Modic², Nejc Haberman², James Tollervey^{2, 7}, Ritsuko Fujii⁵, Toru Takumi⁵, Christopher E. Shaw^{1, §, *} & Jernej Ule^{2, §, *}

Legends:

Supplementary Figure S1. Comparison of RNA regions with enriched binding of FUS, TDP-43 and U2AF65.

a) To validate the specificity of iCLIP procedure for isolating FUS-bound RNA, we separated the ³²P-labeled RNA crosslinked to FUS using the SDS-PAGE gel. FUS was isolated from mouse embryonic day 18 brain from wild-type (WT) or FUS^{-/-} (KO) mice in the presence or absence of antibody to FUS (anti-FUS). We used high and low RNase concentrations to confirm the presence of RNA bound to FUS.

b) Regionally-normalised iCLIP cDNA density was determined in 0.5 kb regions within genes longer than 50 kb that are expressed in brain. Regions with greater than 3-fold enrichment compared to the average cDNA density within the gene were identified for each protein, and the overlap between these is shown.

Supplementary Figure S2. Qiaxcel validation images.

Analysis of alternative splicing of selected genes in E18 brains from wildtype and *FUS*^{-/-} mice. Qiaxcel electropherograms are on the left, and quantification charts on the right show the average percentage of the mRNA isoform that excludes (dark grey) or includes (light grey) the regulated exon, standard deviation and the stars mark the significance of splicing change (* <0.05, ** < 0.01, *** <0.001, t-test, two-tailed, unequal variance).

Supplementary Table S1. Annotation and barcoding of iCLIP experiments.

Supplementary Table S2. Mapping information for iCLIP data.

Supplementary Table S3. Genes with transcript level changes in the *FUS*^{-/-} mouse brains.

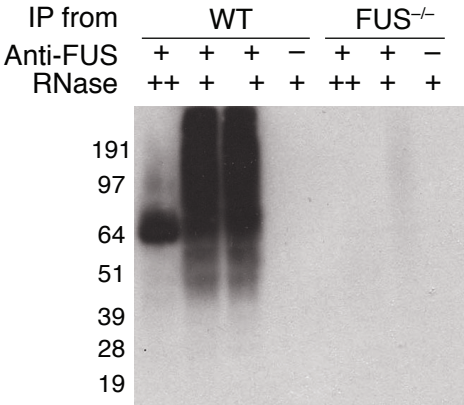
Supplementary Table S4. Validated exons and associated primers.

Significance of changes observed by PCR was determined using the two-tailed, unequal variance t-test.

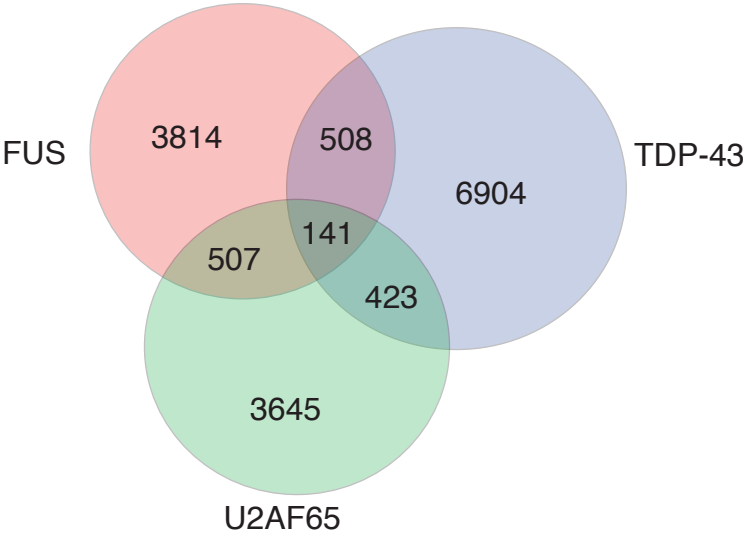
Supplementary Table S5. Significant GO terms associated with the exons regulated by FUS.

Supplementary Figure 1

A

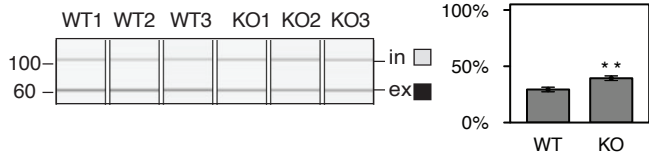


B



Supplementary Figure 2

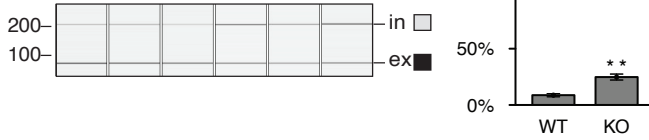
Bat2d; E46



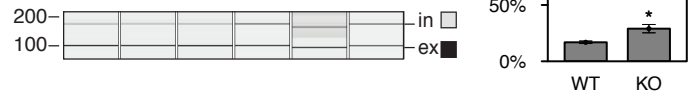
Enah; E15



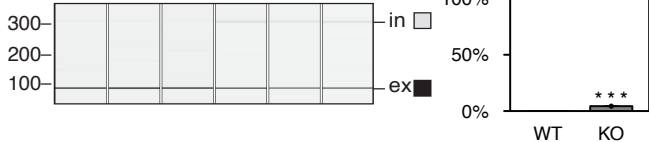
Ntng1; E11



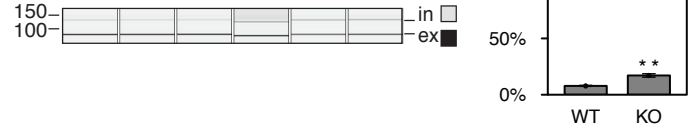
Slitrk4; E2



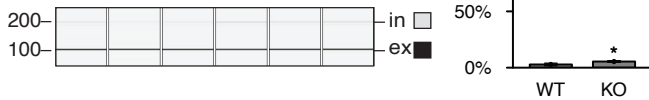
Dtna; E9



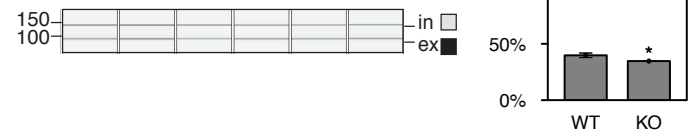
Npnt; E3



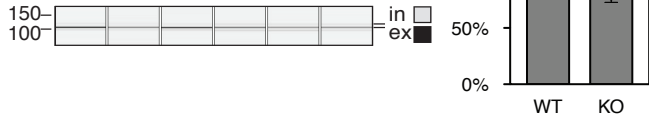
Mapt; E15



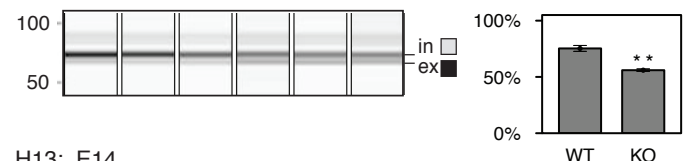
Lims1; E16



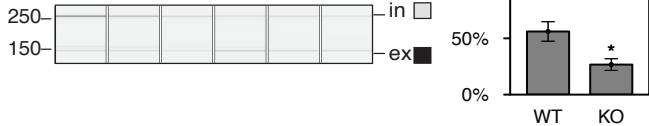
Enah; E5



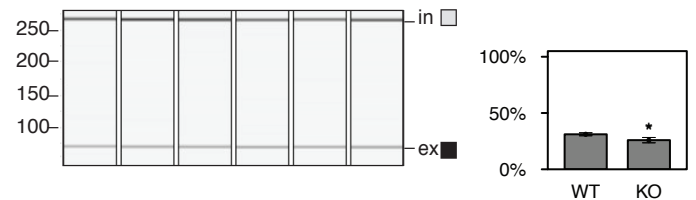
Ablim; E23



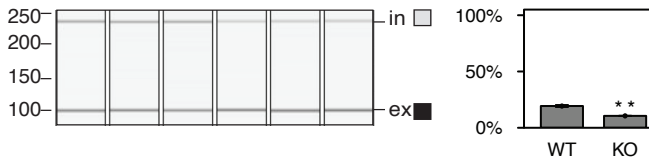
Adnp; E16



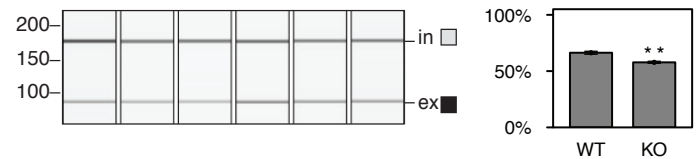
H13; E14



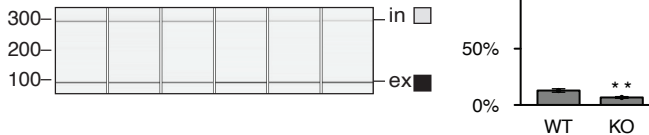
Smg7; E15



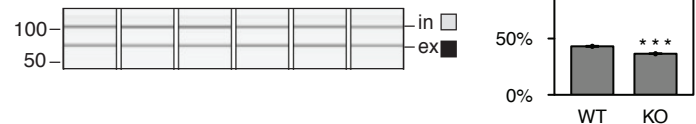
Sort1; E21



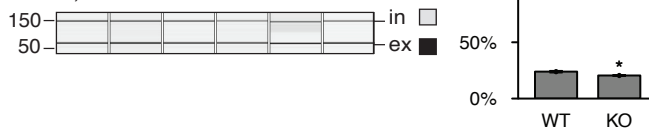
Mtf2; E4



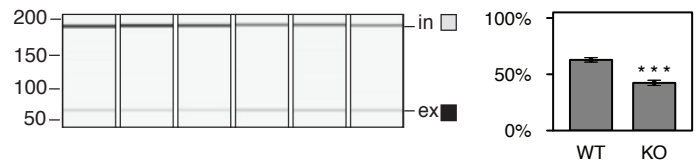
Tsc1; E2



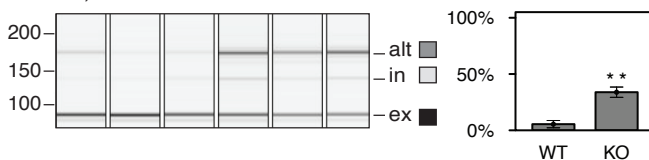
Ewsr1; E21



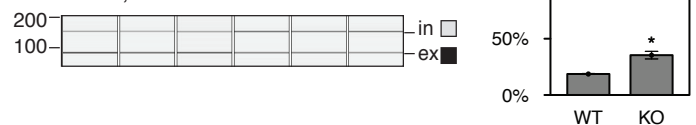
Tmem209; E7



Rmnd1; E5



AK134610; E2



Supplementary table 1: Experimental annotation and barcoding.

| Cells/tissue | Protein | Antibody | 5' barcode | 5' random barcode | 3' barcode | Sequencing file | Sequence Length |
|------------------------------|---------|------------|------------|-------------------|------------|--|-----------------|
| Mouse brain E18, replicate 1 | FUS | NB-100-565 | TTGT | NNN_0:2, NN_7:8 | | iCLIP_FUS_Msbrain_1vs1000RNase_Mm_NNNTTGTNN_20120623_LUm01_5.fq | 70 nt |
| Mouse brain E18, replicate 2 | FUS | NB-100-565 | CAAT | NNN_0:2, NN_7:8 | | iCLIP_FUS_Msbrain_1vs2000RNase_Mm_NNNTCAATNN_20120623_LUm01_6.fq | 70 nt |
| Mouse brain E18, replicate 3 | FUS | NB-100-565 | AATA | NNN_0:2, NN_7:8 | | iCLIP_FUS_E18-brain_WT_Mm_NNNAATANN_20110428_LUI2_1.fq | 50 nt |
| Mouse brain E18, replicate 1 | None | NB-100-565 | TTAG | NNN_0:2, NN_7:8 | | iCLIP_FUS_Msbrain_Noantibody_Mm_NNNTTAGNN_20120623_LUm01_7.fq | 70 nt |
| Mouse brain E18, replicate 1 | TDP-43 | 10782-2-AP | TG | NNN_2:4 | TG | iCLIP_TDP-43_E18-brain_mm_TGNNN_20100222_LUjt3_1.fq | 50 nt |
| Mouse brain E18, replicate 2 | TDP-43 | 10782-2-AP | TGGGC | NN_0:1,NN_4:5 | TG | iCLIP_TDP-43_Brain_High_mm_NNTGNNGGC_20091102_LUjt4_1.fq | 50 nt |
| Mouse brain E18, replicate 3 | TDP-43 | 10782-2-AP | TGGGC | NN_0:1,NN_4:5 | TG | iCLIP_TDP-43_Brain_Low_mm_NNTGNNNGGC_20091102_LUjt5_1.fq | 50 nt |
| Mouse brain E18, replicate 1 | U2AF65 | U4758 | GGG | NNNN_3:6 | AT | iCLIP_U2af65_E18-brain_mm_GGGNNNN_20090826_LUd10_14.fq | 50 nt |
| Mouse brain E18, replicate 2 | U2AF65 | U4758 | GAA | NNNN_3:6 | TG | iCLIP_U2af65_E18-brain_mm_GAANNNN_20090826_LUd10_15.fq | 50 nt |

Supplementary table 2: Mapping information for iCLIP data.

| | Sequencing reads matching barcode | Reads mapping to genome | Unique cDNA reads after random barcode evaluation | Genomic positions of unique cross-link sites |
|---|-----------------------------------|-------------------------|---|--|
| Mouse brain E18, FUS, replicate 1 | 3384693 | 2588134 | 1808558 | 1739887 |
| Mouse brain E18, FUS, replicate 2 | 1063646 | 852980 | 597234 | 584818 |
| Mouse brain E18, FUS, replicate 3 | 4594375 | 3780081 | 1081221 | 1047143 |
| | | | | |
| Mouse brain E18, No antibody, replicate 1 | 13571 | 9134 | 6546 | 6394 |
| | | | | |
| Mouse brain E18, TDP-43, replicate 1 | 8231124 | 6782417 | 2566070 | 2338000 |
| Mouse brain E18, TDP-43, replicate 2 | 3934104 | 1803340 | 210296 | 177632 |
| Mouse brain E18, TDP-43, replicate 3 | 6357350 | 3871455 | 1642263 | 1368372 |
| | | | | |
| Mouse brain E18, U2AF65, replicate 1 | 4975623 | 4040201 | 3037287 | 2799821 |
| Mouse brain E18, U2AF65, replicate 2 | 2056872 | 1398745 | 978386 | 930940 |

Supplementary Table 3. Genes with transcript level changes in the FUS^{-/-} mouse brains.

| Gene Symbol | Position | Description | Fold change | T-test |
|--------------------|--------------------------|----------------------------------|--------------------|---------------|
| AC158301.4 | chr5:147077259-147079338 | | 2.12 | 1.9E-05 |
| Rtkn | chr6:83161299-83178013 | Rhotekin | 2.64 | 9.38E-05 |
| Fus | chr7:124064047-124078926 | Fused in sarcoma | 2.56 | 0.000134 |
| Olfr523 | chr7:136589891-136596282 | Olfactory receptor 523 | 0.49 | 0.009364 |
| BX088531.8 | chrX:120756178-120757964 | n/a | 0.42 | 0.010079 |
| 9230112D13Rik | chr14:31944972-31956154 | RIKEN cDNA 9230112D13 gene | 9.33 | 0.010257 |
| n/a | chr4:52840093-52856360 | n/a | 0.43 | 0.022047 |
| Gypa | chr8:79032919-79049965 | Glycophorin-A | 0.45 | 0.022227 |
| Xlr3a | chrX:68472944-68652039 | X-linked lymphocyte-regulated 3A | 0.48 | 0.024638 |
| Olfr393 | chr11:73928118-73929691 | Olfactory receptor 393 | 0.35 | 0.026061 |

Supplementary Table 4. RT-PCR primers for validation of splicing events identified by the microarray

| Gene symbol | Exon | Splice type* | Alternative exon (mm9) | Flanking exons (mm9) | Str. | dI rank | Array dI (%) | PCR dI (%) | p-value | Forward Primer | Reverse Primer | Excl. PCR prod. | Incl. PCR prod. |
|-------------|------|--------------|---------------------------|---------------------------|------|---------|--------------|------------|---------|-----------------------|-----------------------|-----------------|-----------------|
| Adnp | E16 | CE | chr2:168014783-168014894 | chr2:168013160-168032074 | - | 1.48 | 12 | 29 | 0.0123 | ACCCGCCGCCGTGCTCTA | CGTAAGAGAAGGGTCCACACA | 142 | 255 |
| AK134610 | E2 | CE | chrX:102585210-102585282 | chrX:102583276-102586815 | - | -3.81 | -20 | -17 | 0.0119 | TCCAGGAAGATGGCTGAATA | AGCTCTTGGAGCCTTGGTTT | 78 | 151 |
| Dtna | E9 | CE | chr18:23655984-23656204 | chr18:23633744-23693663 | + | -2.07 | -36 | -4 | 0.0008 | GGTTCTGGAAGTCCATGCT | CTGCCATGGTGTTCCTCTT | 81 | 302 |
| Enah | E5 | CE | chr1:183861306-183861316 | chr1:183860894-183865791 | - | 1.26 | 9.9 | 15 | 0.0403 | TGGGCCATTTTGAACCTTGAG | GCTACGCAGGACAGCACTAA | 96 | 108 |
| Enah | E15 | CE | chr1:183841729-183841790 | chr1:183838508-183843438 | - | -3.07 | -17 | -18 | 0.0065 | CGGCAGTAAGTACCTGTCA | GTCTGGACTCCATTGGCACT | 75 | 137 |
| H13 | E14 | CE | chr2:152529732-152529930 | chr2:152526005-152530358 | + | 2.04 | 15 | 5 | 0.0415 | GAGAAGTGGCCGAGATGTTT | TCTCGGCTGCTGGATCTTTA | 60 | 199 |
| Lba1 | E5 | CE | chr9:111247335-111247360 | chr9:111245750-111247899 | + | 1.82 | 26 | 5 | 0.0139 | TGGGATCTTCACAACGATGA | TCAAAGCAGGGCAAGAACT | 63 | 90 |
| Mapt | E15 | CE | chr11:104179472-104179563 | chr11:104171849-104182665 | + | -1.16 | -11 | -3 | 0.0351 | TCAGGTCGAAGATTGGCTCT | GGAGGTCACCTTGTCTCAGGT | 101 | 194 |
| Mtf2 | E4 | CE | chr5:108498182-108498382 | chr5:108494985-108502326 | + | 2.56 | 35 | 6 | 0.0078 | TTCTGTATGAAGCGGTTGG | CCAGTACCCTTGCCACTACTC | 89 | 310 |
| Npnt | E3 | CE | chr3:132594547-132594596 | chr3:132580518-132611011 | - | -1.34 | -12 | -9 | 0.0041 | GCAGGATTGACTGCTGCTG | TGTTTGACTGTGGTTGACA | 79 | 130 |
| Ntng1 | E11 | CE | chr3:109630546-109630679 | chr3:109585830-109635386 | - | -2.81 | -50 | -16 | 0.0023 | TGGACGATGAGAATGTGTGC | GTCCCTCCATTCTGGCAGT | 71 | 205 |
| Slitrk4 | E2 | CE | chrX:61527195-61527265 | chrX:61526285-61529975 | - | -2.59 | -18 | -12 | 0.0222 | GAGCGAGTCGAAAGCAGT | TCTCTTGCAATCAGCAAACA | 100 | 172 |
| Sort1 | E21 | CE | chr3:108158391-108158488 | chr3:108157160-108159149 | + | 2.26 | 12 | 9 | 0.0010 | CTTCTGAACCCACAAAGC | ACAAGCATCAGTCCCACGAT | 78 | 176 |
| Tmem209 | E7 | CE | chr6:30456787-30456917 | chr6:30455975-30458454 | - | 10.29 | 32 | 20 | 0.0003 | TGTGACATACTGGCCCTCT | CAAATCACGGGGAGGTGTAG | 59 | 190 |
| Nudt13 | E12 | CE | chr14:21129845-21129955 | chr14:21129020-21130693 | + | 1.93 | 26 | 4 | 0.1318 | CGAGTGTGCCCTCCAGTAA | CACCTCTCCGCAACTTCTC | 87 | 198 |

| | | | | | | | | | | | | | |
|---------|-----|---------|--------------------------|--------------------------|---|--------|-------|------------|--------|---|--|----|-------------|
| Rasgrf1 | E2 | CE | chr9:89806287-89806430 | chr9:89805112-89818899 | + | -2.64 | -20 | -7 | 0.1714 | AGGAGTCCGGCGAGAAAC | TCCAAGGTCTTCTGGCTGTC / CACCAGCCTGAATGTTTTG | 70 | 98 |
| Sorbs1 | E27 | CE | chr19:40415176-40415240 | chr19:40414637-40418846 | - | -1.15 | -26 | -12 | 0.1961 | TTTTCGGAATTGGAGTTTGG | CTGCTTGGCAGAAGCTGAG | 64 | 130 |
| Kcnip2 | E2 | CE | chr19:45871582-45871676 | chr19:45870195-45890221 | - | -1 | -22 | Many bands | - | CGGCTCCTATGACCAGCTTA | GCACACCGTGGATAGTTCAA | 61 | 157 |
| Spp1 | E8 | CE | chr5:104867246-104867286 | chr5:104866755-104868284 | + | -1.49 | -17 | One band | - | TGACCCATCTCAGAAGCAGA | TGTGGTCATGGCTTTCATTG | 74 | 116 |
| Rims2 | E33 | CE | chr15:39441611-39441687 | chr15:39417309-39447815 | + | 1.35 | 36 | One band | - | GATGGCAGCATGAACAGCTA | TCACTGAACTGGCTGTCAGA | 83 | 161 |
| Nrcam | E33 | CE | chr12:45685823-45685974 | chr12:45677777-45691095 | + | 1.68 | 29 | Many bands | - | CGATAACGACTGTGGACGAA | TGGGTCCGATATCCACT | 64 | 217 |
| Bat2d | E46 | A5SS/CE | chr1:164604231-164604234 | chr1:164603788-164604234 | - | -3.22 | -33 | -10 | 0.0041 | GCGTTGAAGGCTGAACAAG | GAACCTCTGCTCTGCTTTG | 49 | 100 |
| Lims1 | E16 | A5SS | chr10:57881535-57881583 | chr10:57881535-57884202 | + | 1.1 | 28 | 5 | 0.0472 | CAAACGGGAGAGAGAAGCAA | TGGCTTCATGTCAAATTCCA | 87 | 137 |
| Smg7 | E15 | A5SS | chr1:154695727-154695864 | chr1:154693394-154695864 | - | 2.14 | 15 | 9 | 0.0027 | CCACGAGAAGTGAACCAAGG | TTTCTGGCTTCAGACT | 98 | 236 |
| Tsc1 | E2 | A5SS | chr2:28496811-28496839 | chr2:28496811-28507689 | + | 2.67 | 12 | 7 | 0.0003 | CTGTAGGCTGGAGGGACTGT | TGGTACATCAGTTCCAGTGC | 60 | 99 |
| Ablim1 | E23 | A3SS | chr19:57121556-57121561 | chr19:57121556-57123952 | - | 1.35 | 29 | 19 | 0.0024 | TTTACCGAAAACCACCCATC | CTTCTGAAGCCTTGCTCTGG | 72 | 120 |
| Rmnd1 | E5 | A3SS | chr10:5918133-5918172 | chr10:5917465-5918850 | + | -14.76 | -33 | -5/-24 | 0.0015 | CCTGTGGAACATGTAGGATCG | TGTAACCTGCCAGGGTGTG | 77 | 134/ 174 |
| Ewsr1 | I9 | RI | chr11:4978956-4979488 | chr11:4978956-4979488 | - | 8.28 | 53.29 | 3 | 0.0126 | AGAGCGAGGTGGCTTCAATA / ATGAGTCCCATCAATGGT | CCTAGATCAAGATCTGGTCCTC | 63 | 144 |
| Ewsr1 | E8* | TE | chr11:4980081-4982255 | chr11:4979522-4982255 | - | | | wt/kd =2.4 | 0.012 | TTGCGGTCAAGCTATCTCCT | CCAGCCCAACAATGGTTACT | | |

CE = Cassette Exon, A5SS = alternative 5' splice site, A3SS = alternative 3' splice site, RI = retained intron, TE = terminal exon.

* Abundance of the terminal exon 9 was evaluated using real-time PCR with the shown primers, and normalised against the following primers in the Ewsr1 intron 8: GTCCAGGAGAAAGGGTGTCA, GTCGAATGAACCTGAGGAA.

Supplementary Table 5: Gene ontologies enriched in genes with splicing changes in the FUS knockout brain.

gene_exon records with splicing changes (108)
control; all evaluated gene_exon records (15975)

aspect: molecular process
number of tests: 1441

| p-value | term | redundance | reference genes | matching cluster genes | matching cluster genes |
|---------|--|------------|-----------------|------------------------|---|
| 0.00004 | positive regulation of cell-substrate adhesion | 0 | 54 | 5 | <i>Cd36_E4, Npnt_E3, Spp1_E2, Spp1_E8, Tsc1_E2</i> |
| 0.00094 | cellular protein complex assembly | 12 | 289 | 8 | <i>Actr3b_E3, Apc_E5, Kif23_E9, Mapt_E15, Rap1gds1_E7, Taz_E14, Taz_E15, Trim9_E10</i> |
| 0.00430 | negative regulation of apoptotic process | 33 | 453 | 9 | <i>Agap2_E15, Agt_E2, Apc_E5, Rtkn_E15, Rtkn_E6, Sort1_E21, Spp1_E2, Spp1_E8, Xiap_E5</i> |
| 0.00681 | nervous system development | 35 | 1281 | 17 | <i>Agt_E2, Ank3_E40, Apc_E5, Enah_E15, Gigyf2_E4, Grik1_E9, Hdac7_E20, Mapt_E15, Ndr2_E4, Ntn1_E11, Rasgrf1_E2, Sema6c_E6, Slitrk4_E2, Traf6_E2, Tsc1_E2, Ttc3_E12, Ttc3_E5</i> |
| 0.00834 | Rho protein signal transduction | 40 | 179 | 5 | <i>Rap1gds1_E7, Rasgrf1_E2, Rtkn_E15, Rtkn_E6, Tsc1_E2</i> |
| 0.01230 | axonogenesis | 100 | 357 | 7 | <i>Ank3_E40, Apc_E5, Enah_E15, Mapt_E15, Ntn1_E11, Sema6c_E6, Slitrk4_E2</i> |
| 0.01610 | protein catabolic process | 43 | 377 | 7 | <i>Agap2_E15, Agap3_E11, Apc_E5, Ddb1_E8, Trim9_E10, Ttc3_E12, Ttc3_E5</i> |