Appendix

FUS regulates splicing of minor introns: Implications for ALS

Stefan Reber, Jolanda Stettler, Giuseppe Filosa, Martino Colombo, Daniel Jutzi, Silvia C. Lenzken, Christoph Schweingruber, Rémy Bruggmann, Angela Bachi, Silvia M.L. Barabino, Oliver Mühlemann, Marc-David Ruepp

Contents:

Appendix Materials and Methods

Plasmids

pSUPuro-FUS (target sequence 5'-GGACAGCAGCAAAGCTATA-3') and pSUPuro-scrambled (scrambled target sequence 5'-ATTCTCCGAACGTGTCACG-3') are described elsewhere (Raczynska et al., 2015); (Buhler et al., 2006).

To create the C-terminally FLAG-tagged expression constructs, a parental construct was generated by inserting a double stranded oligonucleotide coding for a Gly-Ser-$(Gly)_{15}$ linker followed by the FLAG peptide (GSG15-FLAG) into the HindIII and XbaI sites of pcDNA3. The double stranded oligonucleotide contained XhoI and BamHI sites, separated by seven nucleotides, upstream of the GSG15-FLAG for subsequent insertion of PCR-amplified FUS and EBFP cDNAs. FUS was amplified from the full length ImaGenes Clone IRAUp969F059D.

To create the N-terminally MS2-tagged expression constructs, a fragment coding for the MS2 coat protein was PCR-amplified from pCMV-PABPN1-MS2-HA (Eberle et al., 2008) and cloned into the NheI, HindIII sites of pcDNA3.1(+) using PCR primers MS2CP NheI f and MS2CP HindIII r (see Appendix Table S2 for oligonucleotide sequences). Full length and truncated FUS coding sequences where PCR-amplified from the pcDNA3-FUS-GSG15-FLAG construct described above and cloned subsequently into the XhoI, ApaI sites of pcDNA3-MS2 or the HindIII, XbaI sites of pcDNA3.1(+)-MS2 respectively. pcDNA3.1(+)-MS2-FUS was cloned using FUS HindIII f and XbaIFUSwtR and cloned into the HindIII, XbaI sites of pcDNA3.1(+)-MS2. pcDNA3.1(+)-MS2-FUS P525L was created using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent) and the mutagenic primer QC P525L. pcDNA3.1(+)-MS2-FUS 285-526 was cloned by PCR amplification using FUS 285 fwd and XbaIFUSwtR and cloned into the HindIII, XbaI sites of pcDNA3.1(+)-MS2. To create pcDNA3.1(+)-MS2-FUS 1-371 NLS, FUS HindIII f and BamHI FUS dCT R were used to PCR amplify a fragment encoding amino acids 1-371 of FUS, which was then cloned into the HindIII, BamHI sites of pcDNA3.1(+)-MS2. The C-terminal nuclear import signal was added by inserting a double stranded oligo (annealed FUS import sig F and FUS import sig R, see Appendix Table S2) into the BamHI site to create pcDNA3.1(+)-MS2-FUS 1-137 NLS. To create pcDNA3.1(+)-MS2-FUS 1-165 NLS, the coding sequence for

amino acids 1-165, followed by the nuclear localization signal, was ordered as gBlocks Gene Fragment (IDT) and cloned into the HindIII, XbaI sites of pcDNA3.1(+)-MS2. pcDNA3.1(+)-MS2-FUS dFETbm 1-165 NLS using pcDNA3.1(+)-MS2-FUS 1-165 NLS as template and was generated by fusion PCR using i) MS2CP NheI and FUS dFET r, and ii) FUS dFET f and XbaIFUSwtR. The two resulting PCR products were purified, mixed and amplified using MS2CP NheI and XbaIFUSwtR. The fusion PCR product was purified and cloned into the NheI, XbaI sites of pcDNA3.1(+)-MS2. To create pcDNA3.1(+)-MS2-FUS P525L SV40 NLS, FUS P525L SV40 NLS was PCR-amplified from pcDNA3.1(+)-MS2-FUS using the forward primer MS2CP HindIII fw and the mutagenic reverse primer P525L NLS ApaI r, which adds the coding sequence of the SV40 nuclear localization signal, and subsequently cloned into the HindIII, ApaI sites of pcDNA3.1(+). To create pcDNA3.1(+)-MS2-FUS 1-165 SV40 NLS and pcDNA3.1(+)-MS2-FUS 1-165 PKIα NES, a PCR fragment obtained from pcDNA3.1(+)-MS2-FUS 1-165 NLS using the forward primer MS2CP HindIII fw and either of the two reverse primers FUS 165 rev NES and FUS 165 rev NLS. The PCR fragments were then digested using Hind III and XbaI and cloned into the same sites of pcDNA3.1(+)-MS2.

The pcDNA6F-meG, pcDNA6F-FUS WT and pcDNA6F-FUS-P525L plasmids were created as follows: A synthetic double-stranded oligonucleotide (GeneArt, Life Technologies) containing the chimeric intron from pCI-neo followed by a FLAG tagged monomeric EGFP residing between XbaI and BamHI cloning sites was cloned into the SacI, PmeI sites of pcDNA6/TR, thereby exchanging the Tet-Repressor with FLAG-mEGFP. The mEGFP was then excised with XbaI and BamHI and PCR amplified FUS or FUS-P525L was cloned into the linearized vector.

The mSCN8A minigene was created by correcting the pCMVTnT-SCN8A^{medJ} minigene (a kind gift of Miriam Meisler, University of Michigan, Ann Arbor) to pCMVTnT-SCN8A using the QuikChange Lightning Multi Site-Directed Mutagenesis Kit (Agilent). The human SCN8A minigene, consisting of the minor intron flanked by the two exons, was ordered as DNA String (GeneArt, Life Technologies) and cloned into pcDNA3.1.

The p120-MS2bs-ex7 minigene was created as follows: HindIII, EcoRV digested fragment containing exon 7 of p120 comprising two MS2 binding sites which was ordered as a DNA String (GeneArt, Life Technologies) was cloned into the HindIII, ExoRV sites of pRSV2-p120 (McNally et al., 2006) to generate pRSV2-p120-MS2bs-ex7.

Antibodies

The polyclonal rabbit anti-FUS antibody and the Y12 monoclonal antibody are described elsewhere (Lerner et al., 1981, Raczynska et al., 2015). The hybridoma cells to produce the monoclonal FUS antibody (19B2) raised against hexahistidine-tagged FUS were generated by Paratopes Ltd and kindly provided by Ramesh S. Pillai, EMBL, Grenoble. A directly labeled FUS antibody was made by labeling 1 mg of monoclonal FUS antibody with the 800CW high molecular weight protein labeling kit (LI-COR Biosciences). Additional antibodies that were used: Anti-Flag M2 Affinity gel (A2220, Sigma-Aldrich), anti-Flag M2 antibody (F1804, Sigma-Aldrich), rabbit anti-FLAG (ABIN99294, Antibodies Online), rabbit anti-FUS (A300-294A, Bethyl Laboratories), goat anti-hnRNP H antibody (sc-10042, Santa Cruz Biotechnology), mouse anti-tyrosine tubulin (T9028, Sigma-Aldrich), rabbit anti-hnRNP H antibody (a kind gift from Douglas Black, HHMI/UCLA), mouse anti-U1A (Santa Cruz, sc-101149), rabbit anti-U1C (Bethyl, A303-947A), goat anti-U1-70K (Santa Cruz, sc-9571), rabbit anti-U11/U12-20K (LS-C178592, LifeSpan BioSciences), rabbit anti U11-59K (ab131258, Abcam), rabbit anti-SmD3 (AP9280b, Abgent), rabbit anti-U11-31K (PA5-20772, Thermo Fisher Scientific), rabbit anti-MS2 coat protein (ABE76, Merck Millipore), donkey anti-mouse IRDye680LT (LI-COR Biosciences, 925-68022), donkey anti-rabbit IRDye680LT (LI-COR Biosciences, 926-68023), donkey anti-goat IRDye680LT (LI-COR Biosciences, 925-68024), donkey anti-mouse IRDye800CW (LI-COR Biosciences, 926-32212), donkey anti-rabbit IRDye800CW (LI-COR Biosciences, 926-32213), donkey anti-goat IRDye800CW (LI-COR Biosciences, 926-32214).

Immunoprecipitations for mass spectrometric analysis

293T cells transfected with pcDNA3-FUS-GSG15-FLAG or pcDNA3-EBFP-GSG15-FLAG were harvested by trypsinization, followed by centrifugation at 200 x g for 5 min at 4°C. The sedimented cells were washed once with PBS and suspended in ice cold hypotonic gentle lysis buffer (10 mM Tris-HCl pH 7.5, 10 mM NaCl, 2 mM EDTA, 0.1% Triton X-100, 1 x Halt Protease Inhibitor [Pierce]) to a final concentration of 1 x 10^7 cells / ml, either in the absence (RNase free) or in the presence of 0.2 mg/ml RNase A (RNase treated). The cells were lysed for 10 min on ice, followed by supplementation of NaCl to 150 mM final concentration and

further incubation on ice for 5 minutes. The lysate was then cleared from insoluble particles by centrifugation (16'100 x g for 15 min at 4 °C). Supernatant was recovered and incubated with anti-FLAG™ M2 Affinity Gel (20 µl/1x10^7 cells; Sigma) for 1.5 h head over tail at 4 °C. The solution was centrifuged at 1000 x g for 5 min at 4 °C. The affinity gel was suspended in 1 ml NET-2 (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.05% Triton-X-100) and washed five times by subsequent suspension and centrifugation steps. For the high salt interactors (750 mM), the precipitate from RNase A treated extracts were washed three times with NET-2 supplemented with NaCl to a final concentration of 750 mM. After the last wash, the slurry was transferred to a new tube with NET-2 and the buffer was completely removed with a syringe (0.4 x 19 mm, $27G^{3/4}$, BD Microlance). To elute the precipitated proteins from the affinity gel, the resin was incubated with 1 bed volume of elution buffer (NET-2, 0.5 x protease inhibitor, 1 mg/ml FLAG peptide) and incubated 30 min head over tail at 4°C. The slurry was transferred to a Bio-Rad Microspin column and centrifuged at 500 x g for 1 min at 4 °C. Eluates were mixed with SDS-Gel loading buffer, boiled for 5 min at 95 °C and loaded on an 8.25 % SDS-polyacrylamide gel. After electrophoresis, gels were stained with colloidal Coomassie (0.08% Coomassie Brilliant Blue G-250, 20% EtOH, 8% ammonium sulfate, 1.598% phosphoric acid) and destained with double distilled water.

RT-qPCR

HeLa cells from splicing assays were harvested by trypsinization. 2 x 10^5 cells where set aside for western blotting to monitor protein expression and/or knockdown efficiencies. The RNA of the remaining cells was isolated using TRIZOL. The RNA samples were DNase treated using the TURBO DNA-free™ Kit (Life Technologies). Reverse transcription of total RNA and RT-qPCR was performed as described in (Raczynska et al., 2015). Analysis was performed as described in (Metze et al., 2013). Primers used for qPCR are listed in Appendix Table S2. The statistical significance of qPCR results was determined by unequal variances *t*-test.

Proximity ligation assay

The Duolink® In Situ reagents from Olink® Bioscience were used. 2 x 10^4 HeLa cells per well were seeded in

200 µl of DMEM +/+ medium in an 8-well plate and incubated over night at 37 °C in a humidified 5 % CO_2 incubator. The following day, the cells were washed with 200 µl of PBS and fixed with 200 µl of 4% paraformaldehyde for 20 min at room temperature. The fixed cells were washed three times with 200 µl PBS for 5 min, permeabilized with 0.5 % Triton X-100 in PBS for 10 min at room temperature. Cells were then washed three times with 0.05 % Tween in TBS for 5 min and one drop of 1 x Duolink II blocking solution was added per well. After incubation for 1 h at 37 °C, the blocking solution was replaced by 100 µl of antibodies diluted 1:100 in 1 x Duolink II antibody diluent, incubated for 2 h at room temperature and then washed twice with 1 x Duolink II wash buffer A for 5 min. The host specific minus/plus PLA® probes were diluted 1:5 in 1 x antibody diluent. 100 µl of this mixture was added per well and incubated for 1 h at 37 °C in the dark. The slides were washed twice with 1 x wash buffer A for 5 min and 60 µl of ligase solution (1 x Duolink II ligation buffer, 1.5 U ligase) was added per well. The slides were incubated for 30 min at 37 °C and washed twice with wash buffer A for 2 min. The rolling circle amplification was carried out by adding 60 µl of amplification-polymerase solution (1 x Duolink II amplification buffer, 7.5 U polymerase) to each well and incubated for 100 min at 37 °C. Finally, the slides were washed twice with 1 x Duolink II wash buffer B, once with 0.1 x wash buffer B and then mounted with 40 µl of Duolink II mounting medium. The slides were stored at 4 °C overnight and analysed with fluorescence microscopy (Leica DMI6000 B with Leica DFC360 FX monochrome digital camera).

Immunofluorescence

Cells were seeded in 8 well chambers and fixed 24 h later with 0.2 % dithiobis(succinimidyl propionate) (DSP) in 1 x PBS supplemented with 2 mM $MgCl_2$ and 10 % glycerol for 30 min at 37 °C. The cells were washed five times with the crosslinking buffer without DSP and stored at 4 °C in 1 x PBS supplemented with 200 mM glycine and 0.1 % sodium azide to quench the crosslinking reagent. The cells were then washed once in 1 x TBS followed by permeabilization/blocking with 1 x TBS, 0.5 % Triton-X-100, 6 % BSA for 30 min at room temperature (RT). The primary antibodies were diluted in 1 x TBS, 0.1 % Triton-X-100, 6 % BSA and incubated for 1 h at 37 °C followed by another incubation at RT, followed by three 5 minute washes with 1 x TBS, 0.1 % Triton-X-100, 6 % BSA. The secondary antibodies were diluted in 1 x TBS, 0.1 % Triton-X-100, 6

% BSA and incubated on the slides for 1 h at 37 °C followed by 30 min at RT. Unbound antibodies were removed by two 5 min washes, followed by incubation of the cells with DAPI (100 ng/ml) in TBS for 10 min at RT. Slides were washed again three times 5 for min with TBS before mounting with Mowiol.

Fluorescence *in situ* hybridization (FISH) in combination with immunofluorescence

The 6-FAM-labeled U11 and U12 snRNA antisense probes were made by *in vitro* transcription using 5-Ethynyl-UTP and subsequent clicking of 6-FAM-azide to the modified UTP. As template, linearized vectors harbouring the U11 or U12 snRNA antisense sequence under the control of the T7 promoter (ordered from GenArt, Life Technologies) were used as templates. RNA was *in vitro* transcribed in 40 mM Tris-HCl pH 7.9, 6 mM $MgCl_2$, 10 mM DTT, 10 mM NaCl, 2 mM spermidine, 1 mM ATP, 1 mM CTP, 1 mM GTP, 1 mM 5-Ethynyl-UTP, 1 U/µl RNase inhibitor, 0.001 U/µl inorganic pyrophosphatase, 30 ng/µl DNA template, 1 U/µl T7 RNA polymerase and incubated for 45 min at 37 °C, followed by supplementation of T7 RNA polymerase to 2 U/µl and incubation for another 45 min at 37 °C. The 5-Ethynyl-labelled RNA was purified over the MEGAclear Transcription Clean-Up kit (Life technologies) according to the manufacture's protocol. By click chemistry, the RNA was labelled with 6-FAM-azide using the BaseClick RNA labelling kit (BaseClick GmBH). Unincorporated 6-FAM-azide was removed by purification of the reaction over the MEGAclear Transcription Clean-Up kit, followed by subsequent gel exclusion purification (Sephadex G50, GE Healthcare). This procedure yielded 6-FAM-labelled U11- and U12 antisense RNA probes with 2.9 and 3.2 dyes/100 nucleotides, respectively.

One day before fixation, 6 x 10^6 HeLa cells, previously transfected as described above with pcDNA3-MS2-FUS or pcDNA3-MS2-FUS P525L, were seeded into a well of an 8-Chamber Slide [Semadeni. 7647]. Cells were rinsed with PBS and fixed with 4 % paraformaldehyde in PBS for 15 min at RT. Fixed cells where washed twice with 70 % EtOH and permeabilized for at least 24 h at 4 °C in 70 % EtOH. After permeabilization, cells were washed three times for 5 min with PBS at RT and blocked three times for 10 min at RT with blocking solution (1 % BSA, 2 mM Ribonucleoside Vanadyl Complex (RVC) [Sigma Aldrich, R3380] in PBS). Thereafter, cells were incubated with primary antibody (rabbit anti-MS2 [Merck Milipore, ABE76]) diluted 1:500 in antibody diluent (1 % BSA. 0.4 U/µl NxGen RNAse Inhibitor [Lucigen, 30281**]** in

PBS) for 1 h at 37 °C and 1 h at RT. Afterwards, cells were washed three times for 10 min at RT with blocking solution and subsequently incubated with secondary antibody (goat α-rabbit IgG conjugated to DyLight 594 [Bethyl. A120-201D4]) diluted 1:250 in antibody diluent for 45 min at RT. After immunostaining, cells were washed three times with PBS and post-fixed for 5 min at RT with 4 % paraformaldehyde in PBS. Subsequently, cells were rinsed two times with 2 x SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0). Before the probe was added to the cells, they were pre-hybridized for 10 min at RT with pre-hybridization buffer (15 % formamide, 10 mM sodium phosphate, 2 mM RVC in 2x SSC, pH 7.0). 50 ng probe was dissolved in 100 µl hybridization buffer (15 % formamide, 10 mM sodium phosphate, 10 % dextran sulfate [Merck Milipore, S4030], 0.2 % BSA, 0.5 g/l *E.coli* tRNA, 0.5 g/l sheared salmon sperm DNA, 2 mM RVC in 2x SSC). Hybridization was performed at 42 °C for 12-16 h. To prevent evaporation of the hybridization buffer, the chamber was sealed with parafilm. The next day, the cells were subsequently washed (all wash steps at 42 °C) two times for 30 min with pre-hybridization buffer, two times for 15 min with 0.1 % Triton in 2 x SSC, two times for 15 min with 0.1 % Triton in 1 x SSC, and finally three times for 10 min in high stringency wash solution (20 % formamide, 2 mM RVC in 0.05 x SSC). After washing, cells were rinsed twice with 2 x SSC. The slides were mounted with Vectashield mounting medium containing 1.5 µg/µL DAPI [Vector Laboratories, H-1200], covered with a coverslip and sealed with nail polish. Images were acquired with a DFC360 FX monochrome camera [Leica Biosystems] mounted on a Leica DMI6000 B microscope equipped with a modified light source to visualize the infrared spectrum. Analysis of images was performed with LAS AF software [Leica Biosystems].

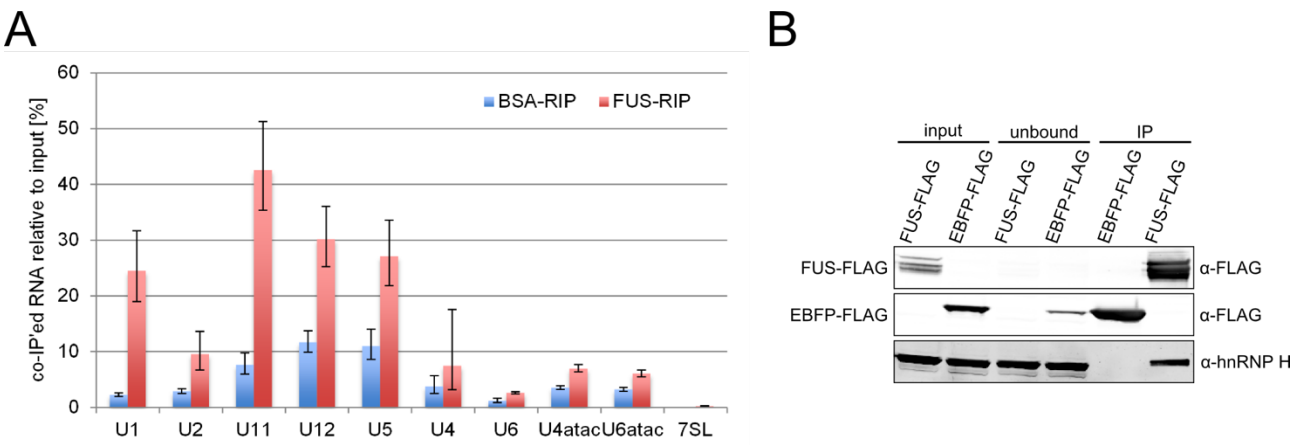
Immunoblotting

If not stated otherwise, whole-cell extract of 1 x 10^5 – 2 x 10^5 cell equivalents per lane were loaded. The proteins were run on a 6 %, 8 % or 10 % SDS-PAGE or on a NuPAGE 4-12 % Bis-Tris Midi Gel (Life Technologies). Proteins from SDS-PAGE were transferred on a nitrocellulose membrane (Optitran BA-S 85, Whatman) using a TE77 ECL Semi-Dry Transfer Unit (Amersham Biosciences). To transfer proteins from a NuPAGE gel, the iBlot Gel Transfer Device (Life Technologies) was used according to the manufacturer's manual. Membranes were blocked with 5 % non-fat dry milk in 0.1 % Tween in TBS and subsequently

incubated at 4 °C overnight with the primary antibodies. Thereafter, the membranes were incubated with the fluorescence-labeled secondary antibodies for 1.5 h at room temperature. The dried membranes were analysed with the Odyssey Infrared Imaging System (Li-Cor).

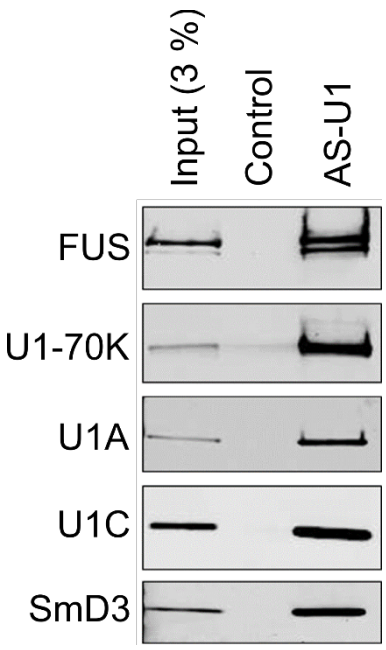
Generation of FUS KO SH-SY5Y cells and high-throughput sequencing

Three wells of a six well plate containing 80 % confluent SH-SY5Y neuroblastoma cells were each transfected with 2.7 µg FUS knockout ZeoR matrix (GeneArt, Life Technologies) and 1.3 µg of plasmid coding for the Cas9 endonuclease and the guide RNAs which target the first intron of FUS (pU6gDNA-Cas9-GFP; target sequence: 5'-TGGATGTCCACCAAGACCTTGG-3' or 5'-TCCAATGGTTAAGGCTTCTGGG-3' (Sigma Aldrich)) using Lipofectamine 3000 (Thermo Scientific) according to the manufacturers manual. One day later, the transfections were repeated with each 6.2 µg FUS KO ZeoR matrix and 0.8 µg pU6gDNA-Cas9-GFP. Two days later, the cells were passaged and pooled on one 15 cm plate. The day after, Zeocin selection was started (250 µg/ml) and maintained until cells were analysed for expression of the FUS gene. Colonies originating from single cells were isolated and expanded. Total RNA was isolated, reverse transcribed and FUS mRNA levels were measured by RT-qPCR as described above. Positive clones were further expanded and analysed by western blot. From two such clonal cell lines and from wild-type SH-SY5Y cells, we extracted total RNA and performed mRNA-seq. Four biological replicates of both FUS KO clones and of the wild-type cells were prepared and total RNA was subsequently poly(A) selected to enrich for mRNAs. From these poly(A)-enriched RNA samples, cDNA libraries were generated that were then sequenced on a Illumina HiSeq3000 machine.



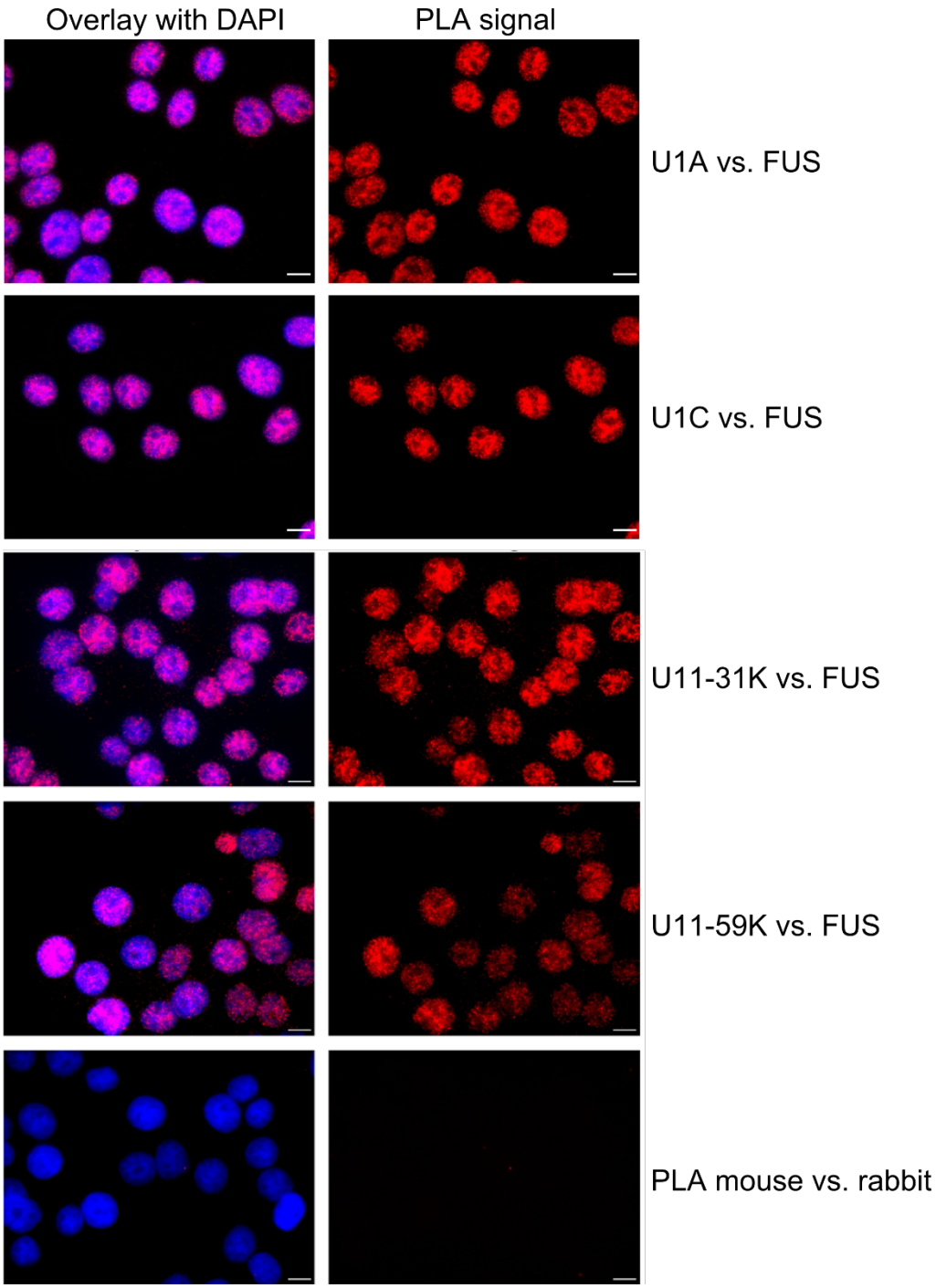
Appendix Figure S1. FUS interacts with the U11 snRNP and hnRNP H (related to Figs. 1 and 2)

A) HeLa nuclear extracts were subjected to immunoprecipitation with anti-FUS antibodies or BSA antibodies (negative control) and the levels of co-precipitated U snRNAs and 7SL RNA were quantitated by RT-qPCR. B) RNase treated extracts from 293T cells expressing FUS-FLAG or EBFP-FLAG were subjected to anti-FLAG immunoprecipitation. Inputs and unbound fractions (0.5 %) and 50 % of the immunoprecipitates were loaded on a 4-12% NuPage gel and subjected to western blotting using anti-hnRNP H and anti-FLAG antibodies.



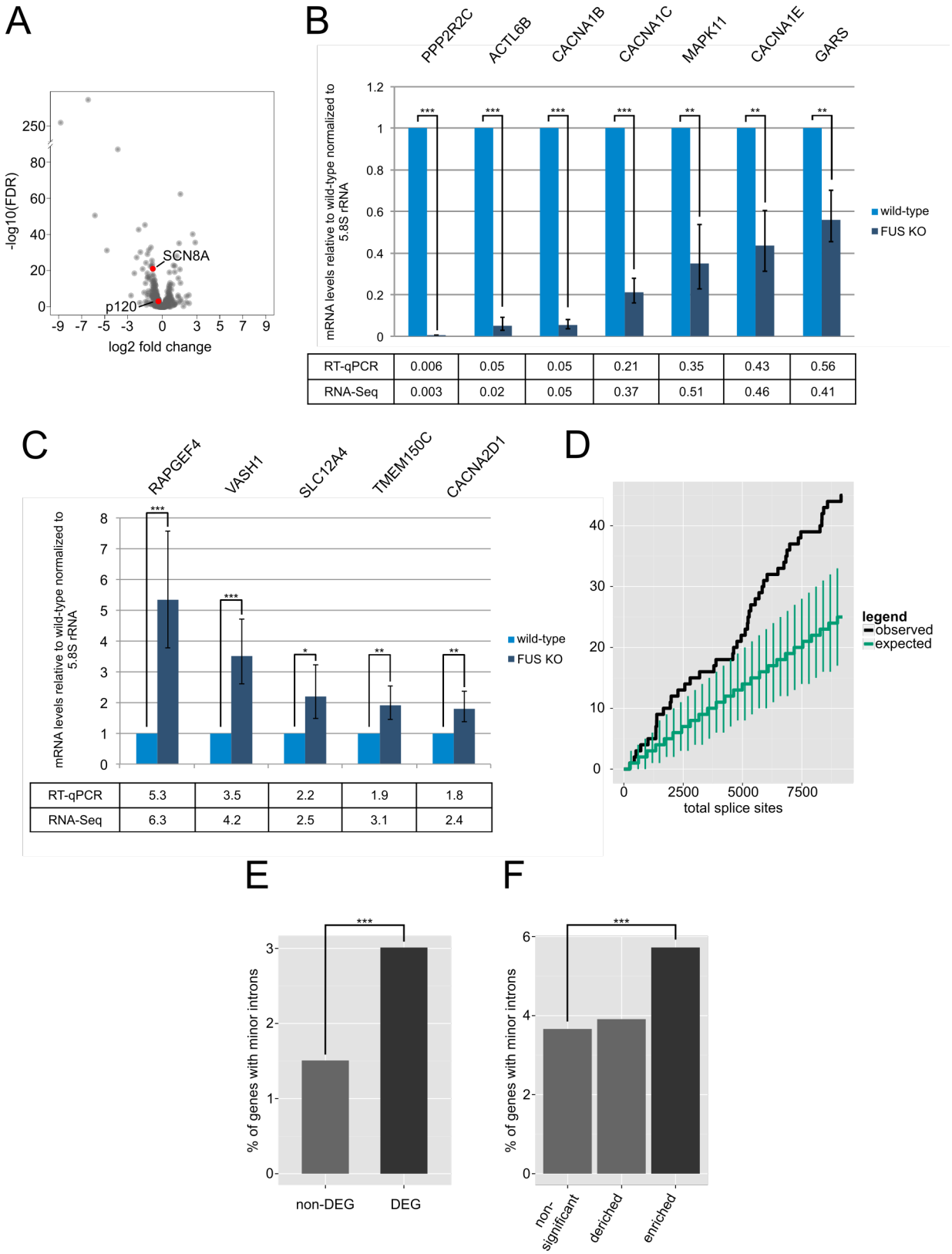
Appendix Figure S2. FUS interacts with the U1 snRNP (related to Figure 2)

Enrichment of FUS after U1 snRNP affinity purification with a biotinylated antisense oligonucleotide (AS-U1) complementary to U1 snRNA from HeLa nuclear extracts. As control, incubation of the magnetic strepatvidin beads with AS-U1 was omitted. After biotinylated antisense oligonucleotide pulldown, the purified complexes were eluted from the beads and subjected to 4-12 % NuPAGE gels. The blots were incubated with mouse anti-FUS-IRDye800CW, mouse anti-U1A, goat anti U1-70K and rabbit anti-SmD3, followed by detection with species specific IRDye680LT or IRDye800CW labelled secondary antibodies to confirm the presence of FUS, U1C, U1A, U1-70K and SmD3 in the AS-U1 enriched fraction. Input: 3 % of the used material.



Appendix Figure S3. Proximity ligation assay (PLA) reveals co-localization of FUS with U1 snRNP and U11 snRNP specific factors in the nucleus (related to Figure 2)

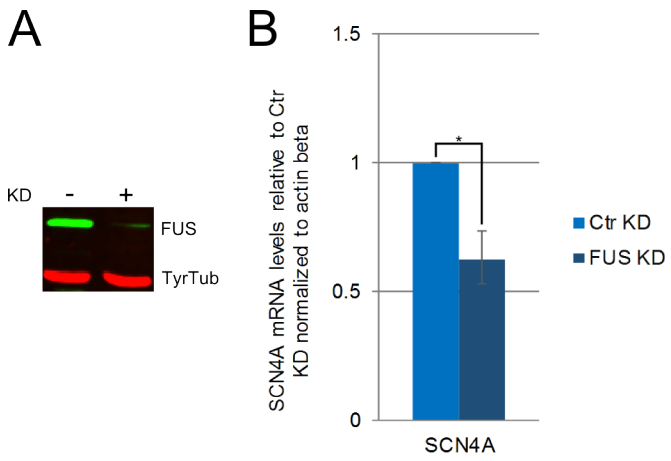
Shown are the overlays of the PLA signal with DAPI (left) as well as the PLA signals only (right) for U1A versus FUS (first row), U1C versus FUS (second row), U11-31K versus FUS (third row) and U11-59K versus FUS (fourth row). As a specificity control, cells were incubated only with the PLA+/- probes (omitting the primary antibodies) to estimate the background of the proximity ligation assay (PLA mouse vs. rabbit). Scale bar = 10 μm.



Appendix Figure S4. FUS binds preferably minor intron-containing mRNAs and regulates their splicing (related to Figure 4)

A) Volcano plot showing minor intron-containing genes that are differentially expressed in the FUS KO SH-

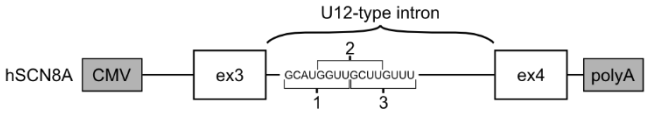
SY5Y cells. The log2 fold change between wt and FUS KO cells for each gene is legible from the x-axis. The statistical significance of each fold change is depicted on the y-axis (log10 fold change of the false discovery rate (FDR)). The p120 and the SCN8A genes are highlighted (red dots) since they were further investigated in this study. B) RT-qPCR validation of selected candidates among the top 30 downregulated genes. The relative mRNA levels compared to wild type samples are shown in the bar plot and the respective fold changes are depicted in the row "RT-qPCR". The corresponding values obtained from the mRNA sequencing experiment are indicated in the row "RNA-Seq" for comparison. Average fold changes and standard deviations of five biological replicas are shown. Single, double, and triple asterisks indicate *p*-values of <0.05, <0.01, and <0.001, respectively. C) As in B) but showing selected candidates selected among the top 30 upregulated genes. D) The plot shows the abundance of minor splice sites among the most differentially expressed splice sites in the RNA-seq data from mouse brains depleted for FUS (Lagier-Tourenne et al., 2012). The differentially expressed splice sites were ranked and are depicted on the x-axis, with the most differentially expressed sites positioned at the left end of the x-axis. The y-axis shows the number of minor splice sites among them. The black line represents the distribution observed in the RNA-seq data. This distribution was compared to what would be expected if minor and major splice sites were equally affected by FUS knockdown. This hypothetical distribution and its 5 % confidence interval (shown in green) were computed with a hypergeometric function. This analysis reveals a significant trend of enrichment for minor splice sites, indicating that the usage of minor splice sites is more often altered upon FUS KD relative to the usage of major splice sites. E) Analysis of RNA-seq results from (Lagier-Tourenne et al., 2012) focusing on genes containing at least one minor intron. The relative abundance of such genes over the total is depicted for the list of genes that can be detected but are not differentially expressed upon FUS KD (non-DEG) and for the genes that are differentially expressed (DEG). Compared to the non-DEG, the % of minor intron containing genes is significantly enriched in the DEG. Effect size: 2.00, *p*-value: 4e-10. F) Analysis of our FUS RIP results focusing on genes containing at least one minor intron. The relative abundance of such genes is depicted for the genes that can be detected but are not significantly changing upon FUS IP (non-significant), for the genes that are underrepresented (deriched) and for the genes that are overrepresented in the FUS RIP (enriched), respectively. The percentage of genes containing minor introns is significantly higher among the genes enriched in the FUS RIP compared to the other two groups (effect size: 1.56, *p*-value: 2e-11).



Appendix Figure S5. Endogenous SCN4A mRNA level is reduced by FUS depletion (related to Figure 5)

A) Western blot analysis of FUS protein levels after siRNA-mediated knockdown (KD) using control (-) or FUS targeting siRNAs (+). Total extracts from transfected human RH-30 cells were subjected to SDS-PAGE and western blotting with anti-FUS (green) and anti-tyrosine tubulin antibodies (loading control, red). B) RT-qPCR results showing the relative SCN4A mRNA levels in the RH-30 cells with the control (Ctr KD) or the FUS knockdown (FUS KD). Average values and standard deviations of four biological replicates are shown. Single asterisks indicate a *p*-value of <0.05.

A

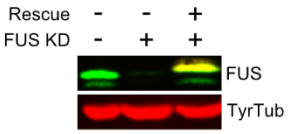


B

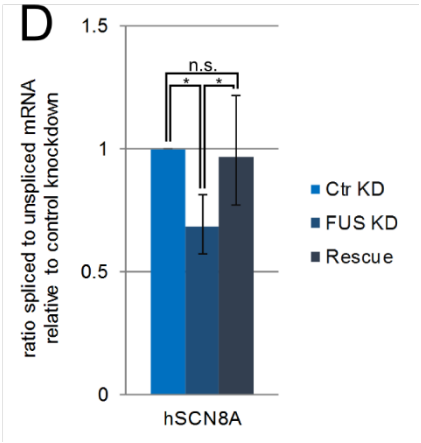
predicted hnRNP M binding sites (motif GGUUGGUU)

site #	Z-score	P-value
1	2.750	2.98e-03
2	3.123	8.89e-04
3	2.781	2.71e-03

C

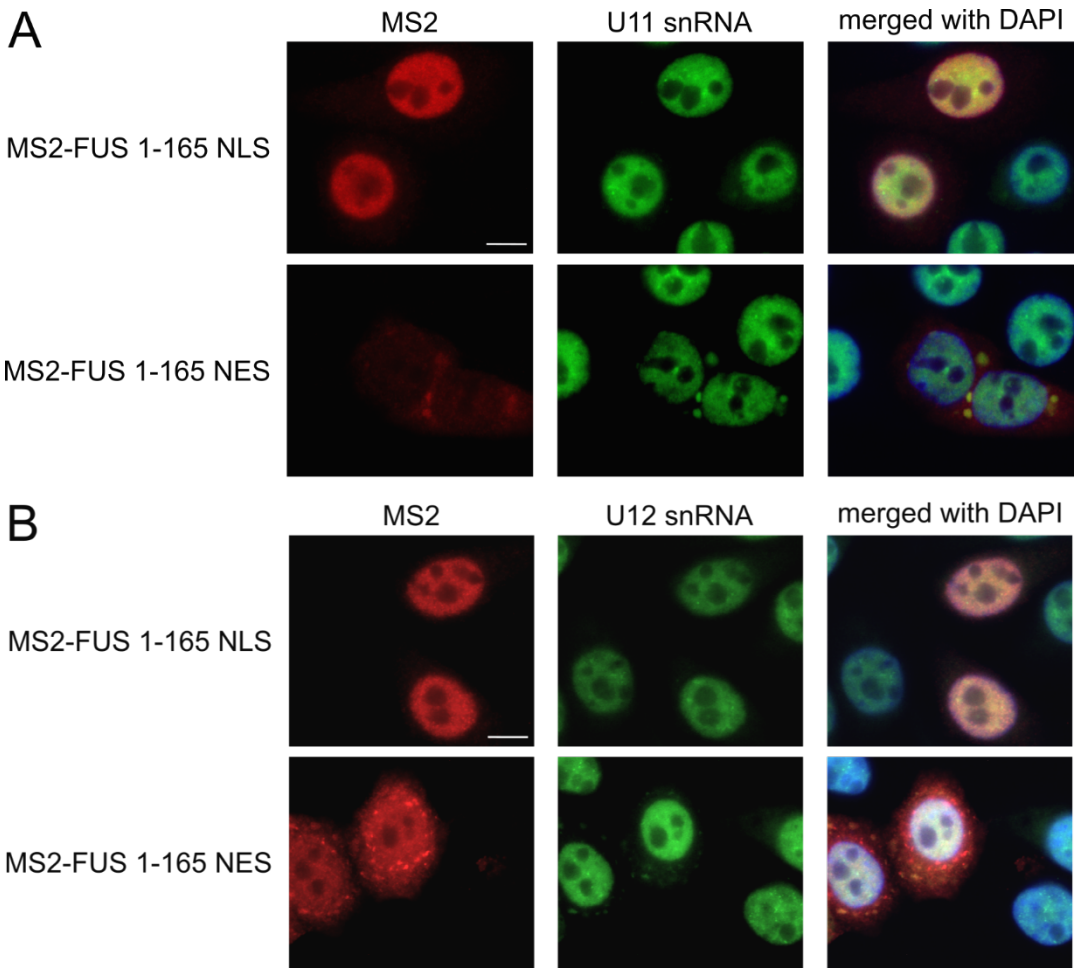


D



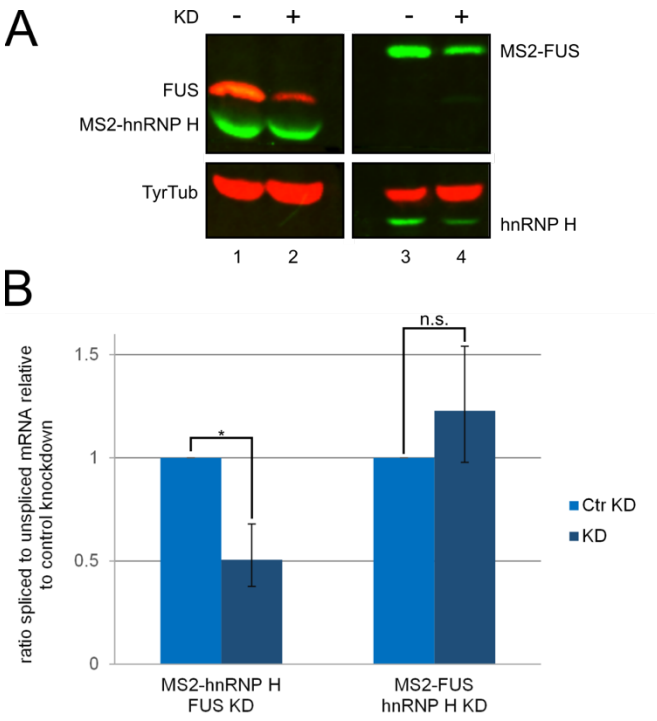
Appendix Figure S6. Human SCN8A minor intron splicing is affected by FUS depletion (related to Figure 5)

A) Schematic representation of the hSCN8A minor intron-containing minigene. Brackets 1, 2 and 3 in the intron designate putative hnRNP M binding sites predicted with RBPmap (Paz et al., 2014). B) Z-Scores and *p*-values of the predicted hnRNP M binding sites shown above (using RBPmap). C) Western blot analysis of FUS levels under control knockdown, FUS knockdown, and FUS rescue conditions. Total extracts from transfected HeLa cells were subjected to SDS-PAGE and western blotting with anti-FUS (green), anti-FLAG (red, upper part) and anti-tyrosine tubulin antibodies (loading control, red, lower part). The yellow signal in the upper part is due to the overlay of the red and the green channels. D) RT-qPCR results indicating the ratio of spliced to unspliced hSCN8A RNA under control knockdown (Ctr KD), FUS knockdown (FUS KD), and FUS knockdown with rescue by co-transfection of an RNAi-resistant FUS expression plasmid (Rescue). Average and standard deviations of five biological replicates are shown. Single asterisks indicate a *p*-value of <0.05.



Appendix Figure S7. The N-terminus of FUS is sufficient to mislocalize U11/12 snRNA (related to Figure 7)

A) Combined FISH and immunofluorescence on HeLa cells transiently expressing MS2-FUS 1-165 NLS (SV40 NLS, upper panel) or MS2-FUS FUS 1-165 NES (NES from human protein kinase α, lower panel). MS2 fusion proteins were visualized with anti-MS2 antibodies (red), U11 snRNA with 6-FAM azide labeled RNA complementary to the full length snRNA (green), and nuclei were stained with DAPI. Scale bar = 10 µm. B) As in A) but with a probe detecting U12 snRNA (green).



Appendix Figure S8. Tethered FUS acts independently of hnRNP H on p120-MS2 RNA

A) Total extract were subjected to SDS-PAGE and western blotting to assess the FUS and hnRNP H reduction after the respective knockdowns and the expression of the MS2 fusion proteins. The upper membrane piece of lanes 1-2 was incubated with antibodies against FUS (red) and MS2-hnRNP H (green), the lower piece with antibody against tyrosine tubulin (red, loading control). The upper membrane piece of lanes 3-4 was incubated anti-MS2 antibody, the lower piece with antibodies against hnRNP H (green) and tyrosine tubulin (red, loading control). B) RT-qPCR results indicating the ratio of spliced to unspliced p120-MS2 RNA when MS2-hnRNP H was tethered in the background of a control (Ctr KD) or a FUS knockdown (KD) is shown on the left, and on the right the same is shown when MS2-FUS was tethered in the background of a Ctr KD or a hnRNP H KD. Data are normalized to the total amount of the minigene mRNA. Average values and standard deviations of four biological replicates are shown. The asterisk indicates a *p*-value of <0.05.

Protein names	Gene names
Uncharacterized protein C11orf84	C11orf84
Calcium homeostasis endoplasmic reticulum protein	CHERP
ATP-dependent RNA helicase DDX1	DDX1
Probable ATP-dependent RNA helicase DDX17	DDX17
ATP-dependent RNA helicase DDX3X;ATP-dependent RNA helicase DDX3Y	DDX3X;DDX3Y
Probable ATP-dependent RNA helicase DDX5	DDX5
Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	DHX15
Eukaryotic translation initiation factor 4B	EIF4B
RNA-binding protein FUS	FUS
Heterogeneous nuclear ribonucleoprotein A3	HNRNPA3
Heterogeneous nuclear ribonucleoprotein D0	HNRNPD
Heterogeneous nuclear ribonucleoprotein F;Heterogeneous nuclear ribonucleoprotein F, N-terminally processed	HNRNPF
Heterogeneous nuclear ribonucleoprotein H;Heterogeneous nuclear ribonucleoprotein H, N-terminally processed	HNRNPH1
Heterogeneous nuclear ribonucleoprotein H2;Heterogeneous nuclear ribonucleoprotein H2, N-terminally processed	HNRNPH2
Heterogeneous nuclear ribonucleoprotein K	HNRNPK
Heterogeneous nuclear ribonucleoprotein M, isoform 2	HNRNPM
Heterogeneous nuclear ribonucleoprotein M	HNRNPM
Heterogeneous nuclear ribonucleoprotein U	HNRNPU
Heterogeneous nuclear ribonucleoprotein U-like protein 1	HNRNPUL1
Tyrosine-protein kinase JAK1	JAK1
S-adenosylmethionine synthase isoform type-2	MAT2A
Matrin-3	MATR3
6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 3;6-phosphofructo-2-kinase;Fructose-2,6-bisphosphatase	PFKFB3
Protein phosphatase 1B	PPM1B
Protein phosphatase 1B, isoform 2	PPM1B
Protein arginine N-methyltransferase 5;Protein arginine N-methyltransferase 5, N-terminally processed	PRMT5
U4/U6 small nuclear ribonucleoprotein Prp31	PRPF31
RNA-binding protein 10	RBM10
Splicing factor 45	RBM17
RNA-binding protein 6	RBM6
Serine/threonine-protein kinase RIO1	RIOK1
tRNA-splicing ligase RtcB homolog	RTCB
Scaffold attachment factor B1	SAFB
Splicing factor 3B subunit 1	SF3B1
Splicing factor 3B subunit 3	SF3B3
Serine/threonine-protein kinase 38	STK38
Serine/threonine-protein kinase 38-like	STK38L
Lamina-associated polypeptide 2, isoforms beta/gamma;Thymopoietin;Thymopentin	TMPO
U2 snRNP-associated SURP motif-containing protein	U2SURP
Methylosome protein 50	WDR77

Appendix Table S1. Conserved FUS interactors, related to Figure 1

The list contains the 40 proteins co-immunoprecpitated with FUS but not with EBFP in all replicates under all conditions (RNase treated, RNase free, RNase treated and washed with 750 mM NaCl).

Name	5‘-Sequence-3‘	Purpose
Universal RT primer	GCTGTCAACGATACGCTACGTAACGGCATGACAGTGTTTTTTTTTTTTTTTTVN	RT
Universal reverse primer	GCTGTCAACGATACGCTACGTAAC	qPCR
U1 snRNA qPCR f	TTCGCGCTTTCCCCTGAA	qPCR
U2 snRNA qPCR f	CAGGAACGGTGCACCCAA	qPCR
U4 snRNA qPCR f	GACAGTCTCTACGGAGACTGAA	qPCR
U5 snRNA qPCR f	CCTTGCTTTGGCAAGGCTAAA	qPCR
U6 snRNA qPCR f	TCGTGAAGCGTTCCATATTTTTAA	qPCR
U11 snRNA qPCR f	GTCCCGGCGCCCTTTAA	qPCR
U12 snRNA qPCR f	GAGTTGCGATCTGCCCGAAA	qPCR
U4atac snRNA qPCR f	TGCTTTATTTTGGTGCAATTTTTGGAAAAATAA	qPCR
U6atac snRNA qPCR f	CGTGGCATCTAACCATCGTTTTTAA	qPCR
7SL RNA qPCR f	CCTGGGCAACATAGCGAGAC	qPCR
SNORD25 qPCR f	CTCCGTGAGGATAAATAACTCTG	qPCR
sybr p120spl r	AGCCTGGGCATCCTGCTC	qPCR
sybr p120un r	CCTGTTAAGGAACTGTGTCATGATG	qPCR
sypr p120 f	CACATTGGTGTGCACCTCCAAG	qPCR
sybr SCN4A f	ACAAGGGCAAGGCCATCTTC	qPCR
sybr SCN4Aspl r	CATGCTGAACAGCGCATGG	qPCR
sybr SCN4Aun r	GGTCAAGGAAAGTGAGGAAGCAG	qPCR
mSCN8A unspl f	GAGCCGGGTTTGTTTCCTAAGC	qPCR
mSCN8A unspl r2	TTGGACCATTCGGGAGGGTTAC	qPCR
mSCN8A spliced f	ATTTAGCGCCACTCCTGCCTTG	qPCR
mSCN8A spliced r	TCTTGGACCATTCGGGAGGGTTAC	qPCR
hSCN8A spliced f	ATTTAGTGCCACGCCTGCCTTG	qPCR
hSCN8A spliced r	TCTTCGACCAGTCAGGAGGGTTAC	qPCR
hSCN8A UNspl 2f	CCCTTGACTCTTCTCTACAG	qPCR
hSCN8A UNspl 2r	CATTCTTCGACCAGTCAG	qPCR
beta act fwd	TCCATCATGAAGTGTGACGT	qPCR
beta actin rev	TACTCCTGCTTGCTGATCCAC	qPCR
hSCN4A fwd	AAGGCCATCTTCCGCTTCTC	qPCR
hSCN4A rev	AGAAGCCTCGGGCCAGTATC	qPCR
ACTL6B fwd	GATGGAGCGGAGGTCATGTC	qPCR
ACTL6B rev	GAGCACTGGGTGCAGGTTTG	qPCR
GARS fwd	AATGCGCCTGGGAGATGCTG	qPCR
GARS rev	TCTCCATGTGCTGCCGGAAG	qPCR
PPP2R2C fwd	GAGCCCTCGGAGGATCTTTG	qPCR
PPP2R2C rev	CCGGCTTGATGTCCACGATG	qPCR
RAPGEF4 fwd	GGGAGGGCCTGATCATAGTC	qPCR
RAPGEF4 rev	GAATTGCTCTCGCGGGCAAG	qPCR
SLC12A4 fwd	CCCAACCTTACGACCGACTC	qPCR
SLC12A4 rev	TCTCCACGATGTCACCCTTC	qPCR
TMEM150C fwd	TGATCCTCCTGCAAGCTGTG	qPCR
TMEM150C rev	TGAAGCGCAGAACAGCTACC	qPCR
VASH1 fwd	GCGATGACTTCCGCAAGGAG	qPCR
VASH1 rev	ATCCTTCTTCCGGTCCTTGG	qPCR

CACNA2D1 fwd	GGCATTGGAAGCGGAGAAAG	qPCR
CACNA2D1 rev	GACTGCTGCGTGCTGATAAG	qPCR
CACNA1B fwd	GGGTCCTCTACAAGCAATCG	qPCR
CACNA1B rev	TGGTGGCCAGGATCATATAC	qPCR
CACNA1C fwd	GATTCCAACGCCACCAATTC	qPCR
CACNA1C rev	AGGAGTCCATAGGCGATTAC	qPCR
CACNA1E fwd	TGATGGGATCACCCAGTTTG	qPCR
CACNA1E rev	AATTCCAGGTGGCTCCTAAG	qPCR
MAPK11 fwd	ATGAGAGCGTTGAGGCCAAG	qPCR
MAPK11 rev	CAGCACCTCACTGCTCAATC	qPCR
5.8S rRNA fwd	GGTGGATCACTCGGCTCGT	qPCR
5.8S rRNA rev	GCAAGTGCGTTCGAAGTGTC	qPCR
MS2CP HindIII fw	TTTTAAGCTTGCCACCATGGCTTCTAACTTTACTCAGTTCGTTCTC	cloning
MS2CP NheI f	TTTTTGCTAGCGCCACCATGGCTTCTAACTTTAC	cloning
MS2CP HindIII r	TTTTTAAGCTTCTCATCCGCGTAGATGC	cloning
FUS HindIII f	AAAAAAGCTTGCCACCATGGCCTCAAAC	cloning
XbaIFUSwtR	AAATCTAGATTAATACGGCCTCTCCCTGCGATCCTG	cloning
QC P525L	ATCGCAGGGAGAGGCTGTATTGATAAGGCGG	mutagenesis
P525L NLS ApaI r	GATTGGGCCCTTCACTTGTCCTCCACTTTGCGTTTCTTTTTGGGATACAGCCTCTCCCTGCGATCC	cloning
QC SCN8A medJ to WT	CGAGAAACAAAGGAAGTGTTACTTACTCCACATTCTTGGAC	mutagenesis
FUS 285 fwd	ATTAAAGCTTGCCGCCACCATGAACACCATCTTTGTGCAAGGCCTGGGTG	cloning
BamHI FUS dCT R	CCTAGGAAAGCCACCACCCCGATTAAAGTCTGC	cloning
FUS import sig F	CTAGAAAAAAAGATCTAGGGGTGAGCACAGACAGGATCGCAGGGAGAGGCCGTATG	cloning
FUS import sig R	GATCCATACGGCCTCTCCCTGCGATCCTGTCTGTGCTCACCCCTAGATCTTTTTTT	cloning
FUS dFET f	TATGGAACTCAGTCAACTCCCCAGGGATATG	cloning
FUS dFET r	GGAGTTGACTGAGTTCCATAACTGCTCTGCTGGGAATAG	cloning
FUS 165 rev NES	AATATCTAGATCAGGATCCGATGTCCAGGCCGGCCAGCTTCAGGGCCAGAGATCCACTGCTGCT GTTGTACTG	cloning
FUS 165 rev NLS	AATATCTAGATCAGGCCTGCTTGTCCTCCACCTTTCTCTTCTTCTTGGGAGATCCACTGCTGCTGTT GTACTG	cloning
FUS siRNA	GGACAGCAGCAAAGCUAUAdTdT	mRNA knockdown
control siRNA	AGGUAGUGUAAUCGCCUUGdTdT	mRNA knockdown
hnRNP H siRNA	GGUAAAAGCAGUUGAAUUAdTdT	mRNA knockdown

Appendix Table S2. Sequences of oligonucleotides and siRNAs, related to experimental procedures

The table lists all oligonucleotides used in this study.

Rank	Gene ID	Gene Name	Avg_log2FC	A4_log2FC	A5_log2FC
1	ENSG00000074211	PPP2R2C	-8.83E+00	-8.564058358	-9.089890376
2	ENSG00000077080	ACTL6B	-6.83E+00	-5.754254297	-7.909814379
3	ENSG00000145087	STXBP5L	-6.44E+00	-4.981741823	-7.907589885
4	ENSG00000148408	CACNA1B	-5.85E+00	-4.453404723	-7.254384054
5	ENSG00000162949	CAPN13	-4.81E+00	-4.535295327	-5.085049842
6	ENSG00000116254	CHD5	-3.86E+00	-3.874019421	-3.852484654
7	ENSG00000175426	PCSK1	-2.71E+00	-2.426762626	-2.996124924
8	ENSG00000151067	CACNA1C	-2.42E+00	-1.451833786	-3.396397885
9	ENSG00000104044	OCA2	-2.22E+00	-1.862467401	-2.579857088
10	ENSG00000172575	RASGRP1	-2.05E+00	-1.852168267	-2.247140465
11	ENSG00000198216	CACNA1E	-1.95E+00	-1.121822983	-2.78121317
12	ENSG00000117707	PROX1	-1.90E+00	-1.454397839	-2.348439949
13	ENSG00000138356	AOX1	-1.70E+00	-1.700848265	-1.694798908
14	ENSG00000151062	CACNA2D4	-1.68E+00	-1.308942537	-2.055596835
15	ENSG00000185386	MAPK11	-1.57E+00	-0.984075469	-2.14788285
16	ENSG00000006047	YBX2	-1.50E+00	-1.43415716	-1.567737968
17	ENSG00000155530	LRGUK	-1.46E+00	-0.509700074	-2.418452223
18	ENSG00000101311	FERMT1	-1.45E+00	-1.673820101	-1.223012334
19	ENSG00000007968	E2F2	-1.41E+00	-1.475427125	-1.347079278
20	ENSG00000007402	CACNA2D2	-1.37E+00	-0.344784721	-2.397300585
21	ENSG00000135773	CAPN9	-1.35E+00	-0.8757682	-1.822660511
22	ENSG00000130475	FCHO1	-1.29E+00	-1.281140613	-1.30831381
23	ENSG00000114805	PLCH1	-1.29E+00	-1.504785494	-1.08089809
24	ENSG00000128059	PPAT	-1.28E+00	-1.230055708	-1.320482151
25	ENSG00000155761	SPAG17	-1.27E+00	-1.424688033	-1.116547036
26	ENSG00000006283	CACNA1G	-1.27E+00	-1.563125232	-0.971483669
27	ENSG00000100206	DMC1	-1.24E+00	-1.137896452	-1.349089126
28	ENSG00000118473	SGIP1	-1.24E+00	-1.529189537	-0.945064578
29	ENSG00000145555	MYO10	-1.21E+00	-0.700000882	-1.711481116
30	**ENSG00000106105**	**GARS**	**-1.17E+00**	**-1.272977855**	**-1.065276696**

Appendix Table S3. Top 30 down-regulated genes, related to Figure 4.

The list contains the top 30 down-regulated genes in the SH-SY5Y FUS knockout cells; The genes are ranked according to their average log2 fold change (Avg_log2FC), which was calculated from the log2 fold changes of the clonal FUS knockout cell lines A4 and A5 compared to wild type SH-SY5Y cells. Genes highlighted in blue were validated by RT-qPCR (see Appendix Figure S4B).

Rank	Gene ID	Gene Name	Avg_log2FC	A4_log2FC	A5_log2FC
1	ENSG00000091428	RAPGEF4	2.87E+00	2.647081543	3.10046862
2	ENSG00000136048	DRAM1	2.65E+00	2.276529665	3.017300659
3	ENSG00000117228	GBP1	2.58E+00	0.029542987	5.132066996
4	ENSG00000155754	ALS2CR11	2.34E+00	1.158412536	3.524376411
5	ENSG00000136546	SCN7A	2.32E+00	1.738224769	2.909867791
6	ENSG00000156475	PPP2R2B	2.10E+00	1.43067182	2.767237651
7	ENSG00000169994	MYO7B	2.08E+00	1.295484625	2.865616975
8	ENSG00000117226	GBP3	1.85E+00	0.01550196	3.680588545
9	ENSG00000197558	SSPO	1.64E+00	1.886866636	1.393261312
10	ENSG00000092529	CAPN3	1.61E+00	1.679048351	1.535464819
11	ENSG00000071246	VASH1	1.59E+00	2.059672552	1.122530027
12	ENSG00000249242	TMEM150C	1.58E+00	1.633696966	1.534464552
13	ENSG00000132561	MATN2	1.54E+00	1.449939334	1.636624416
14	ENSG00000124067	SLC12A4	1.47E+00	1.327316046	1.603871495
15	ENSG00000147459	DOCK5	1.40E+00	1.121187869	1.683679152
16	ENSG00000168671	UGT3A2	1.40E+00	1.674775432	1.12252261
17	ENSG00000183091	NEB	1.39E+00	1.627511558	1.156660448
18	ENSG00000157445	CACNA2D3	1.38E+00	1.560042707	1.19013307
19	ENSG00000162976	PQLC3	1.26E+00	1.23379149	1.286949872
20	ENSG00000184343	SRPK3	1.17E+00	0.555249492	1.779772938
21	ENSG00000124159	MATN4	1.10E+00	0.932256443	1.275345944
22	ENSG00000153956	CACNA2D1	1.10E+00	1.278329808	0.926003303
23	ENSG00000168890	TMEM150A	1.09E+00	1.193516283	0.989972051
24	ENSG00000088538	DOCK3	1.07E+00	0.823163078	1.311812249
25	ENSG00000134516	DOCK2	1.05E+00	0.683823041	1.413831004
26	ENSG00000124721	DNAH8	1.00E+00	1.280071744	0.719891216
27	ENSG00000153214	TMEM87B	9.33E-01	0.70162314	1.164085695
28	ENSG00000198520	C1orf228	9.31E-01	1.580840437	0.28062978
29	ENSG00000102158	MAGT1	8.92E-01	1.072726002	0.710835272
30	ENSG00000204219	TCEA3	8.84E-01	0.957813386	0.810618665

Appendix table S4. Top 30 up-regulated genes, related to Figure 4.

The list contains the top 30 up-regulated genes in the SH-SY5Y FUS knockout clones; The genes are ranked according to their average log2 fold change (Avg_log2FC) calculated from the log2 fold changes of the clonal FUS knockout cell lines A4 and A5 compared to wild type SH-SY5Y cells. Genes highlighted in blue were validated by RT-qPCR (see Appendix Figure S4C).

Appendix references

Buhler M, Steiner S, Mohn F, Paillusson A, Muhlemann O (2006) EJC-independent degradation of nonsense immunoglobulin-mu mRNA depends on 3' UTR length. Nat Struct Mol Biol 13: 462-4

Eberle AB, Stalder L, Mathys H, Orozco RZ, Muhlemann O (2008) Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. PLoS Biol 6: e92

Lagier-Tourenne C, Polymenidou M, Hutt KR, Vu AQ, Baughn M, Huelga SC, Clutario KM, Ling SC, Liang TY, Mazur C, Wancewicz E, Kim AS, Watt A, Freier S, Hicks GG, Donohue JP, Shiue L, Bennett CF, Ravits J, Cleveland DW et al. (2012) Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. Nat Neurosci 15: 1488-97

Lerner EA, Lerner MR, Janeway CA, Jr., Steitz JA (1981) Monoclonal antibodies to nucleic acid-containing cellular constituents: probes for molecular biology and autoimmune disease. Proc Natl Acad Sci U S A 78: 2737-41

McNally LM, Yee L, McNally MT (2006) Heterogeneous nuclear ribonucleoprotein H is required for optimal U11 small nuclear ribonucleoprotein binding to a retroviral RNA-processing control element: implications for U12-dependent RNA splicing. The Journal of biological chemistry 281: 2478-88

Metze S, Herzog VA, Ruepp MD, Muhlemann O (2013) Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways. Rna 19: 1432-48

Paz I, Kosti I, Ares M, Jr., Cline M, Mandel-Gutfreund Y (2014) RBPmap: a web server for mapping binding sites of RNA-binding proteins. Nucleic Acids Res 42: W361-7

Raczynska KD, Ruepp MD, Brzek A, Reber S, Romeo V, Rindlisbacher B, Heller M, Szweykowska-Kulinska Z, Jarmolowski A, Schumperli D (2015) FUS/TLS contributes to replication-dependent histone gene expression by interaction with U7 snRNPs and histone-specific transcription factors. Nucleic Acids Res 43: 9711-28