FUS gene is dual-coding with both proteins united in molecular hallmarks of amyotrophic lateral sclerosis

Marie A. Brunet$^{1,2,*}$, Jean-Francois Jacques$^{1,2}$, Sonya Nassari$^{3}$, Giulia E. Tyzack$^{4,5}$, Philip McGoldrick$^{6}$, Lorne Zinman$^{7}$, Steve Jean$^{3}$, Janice Robertson$^{6}$, Rickie Patani$^{4,5}$, Xavier Roucou$^{1,2,*}$

1 Department of Biochemistry, Université de Sherbrooke, Sherbrooke, Québec, Canada;
2 PROTEO, Quebec Network for Research on Protein Function, Structure, and Engineering;
3 Anatomy and Cell Biology Department, Université de Sherbrooke, Sherbrooke, Québec, Canada;
4 The Francis Crick Institute, 1 Midland Road, London, UK;
5 Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, Queen Square, London, UK;
6 Tanz Centre for Research in Neurodegenerative Diseases, University of Toronto, Toronto, Canada;
7 Division of Neurology, Department of Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Canada.

*Corresponding authors:
Xavier Roucou, Department of Biochemistry, Université de Sherbrooke, 3201 Jean Mignault, Sherbrooke, Québec J1E 4K8, Canada, Tel. (819) 821-8000x72240; E-Mail: xavier.roucou@usherbrooke.ca
ABSTRACT:

Novel functional coding sequences (altORFs) are camouflaged within annotated ones (CDS) in a different reading frame. We discovered an altORF nested in the FUS CDS encoding a conserved protein, altFUS. We thus demonstrate the dual-coding nature of the Amyotrophic Lateral Sclerosis (ALS)-associated FUS gene. AltFUS is endogenously expressed in human tissues, notably in the motor cortex and motor neurons of healthy controls and ALS patients. AltFUS inhibits autophagy, a pathological hallmark presently and incorrectly attributed to the FUS protein. AltFUS is pivotal in two other pathological hallmarks: loss of mitochondrial membrane potential and accumulation of FUS/TDP-43 cytoplasmic aggregates. Suppression of altFUS expression in a FUS-ALS Drosophila model protects against neurodegeneration. Thus, wild-type altFUS is essential for ALS-like phenotypes arising from mutated FUS. Some mutations found in ALS patients are overlooked because of their synonymous effect on the FUS protein, yet we showed they exert a deleterious effect via their missense consequence on the overlapping altFUS protein. These findings suggest that both proteins, FUS and altFUS, are involved in the aetiology and pathological hallmarks of ALS.
Amyotrophic Lateral Sclerosis (ALS) is a fatal motor neuron degenerative disease characterized by a progressive loss of motor neurons in the brain, brainstem and spinal cord. Most cases are sporadic (sALS), although up to 10% are familial (fALS). The first gene associated with ALS (SOD1) was described in 1993. Since then, over 25 genes have been associated with fALS, sALS or both, with the four most common being C9ORF72, SOD1, FUS, and TARDBP. Notably, mutations in FUS are linked to both fALS and sALS, and present some of the earliest and more rapidly progressive forms of the disease. FUS is also involved in other neurodegenerative diseases, such as frontotemporal dementia (FUS-FTLD), all characterized by FUS cytoplasmic inclusions.

FUS is a nuclear RNA-binding protein, with a C-terminal nuclear localization signal (NLS). The protein is involved in RNA processing, DNA repair and cellular proliferation, although its functions are not precisely elucidated. Most of the mutations associated with neurodegenerative diseases alter FUS NLS. More recently, mutations in FUS 3’UTR were described in ALS patients and linked to an increased level of FUS mRNA and protein. Surprisingly, over-expression of wild-type FUS provokes an aggressive ALS phenotype in mice and fruit flies, in accordance with findings in yeast and mammalian cells. The mechanism of the wild-type or ALS-linked mutated FUS toxicity remains unclear.

With currently non-annotated proteins being increasingly reported, we hypothesized that the toxicity resulting from wild-type FUS over-expression may come from another, unseen, actor. These novel proteins are coded by alternative open reading frames (altORFs) that are located within “non-coding” RNAs (ncRNA), within the 5’ or 3’ “untranslated” regions (UTR) of mRNAs, or overlapping a known coding sequence (CDS) within a different frame of an mRNA. Serendipitous discoveries and ribosome profiling have recently highlighted the distribution of
altORFs throughout the human genome, and the consequences of their absence from current
databases\textsuperscript{26}. For example, mass spectrometry-based proteomics has become the gold standard
for protein identification and has been extensively used in ALS studies\textsuperscript{28,29}. However, if a protein
is not annotated, it is not included in the protein database (e.g. UniProtKB), and thus cannot be
detected by mass spectrometry. An estimated 50% of mass spectra from a proteomics
experiment are unmatched at the end of the analysis\textsuperscript{26,30}.

Genome annotations must avoid spurious ORF annotations. Thus unless functional
characterization has been published, they rely upon 2 arbitrary criteria: a minimum length of 100
codons and a single ORF per transcript. Several groups developed tools to challenge such criteria,
such as the sORFs repository\textsuperscript{31} and the OpenProt database\textsuperscript{27}, which offer a data-driven broader
view of eukaryotic proteomes. The OpenProt database is based on a polycistronic model of ORF
annotation\textsuperscript{27} and reports any ORF longer than 30 codons within any frame of an mRNA or ncRNA.
It contains currently annotated proteins (RefProts), novel Isoforms, and novel alternative proteins
(altProts). Here, we used the OpenProt database (www.openprot.org) to ask whether \textit{FUS} may
encode additional proteins that could explain the toxicity of the wild-type protein over-
expression, and bring novel insights concerning \textit{FUS}-linked ALS pathogenicity\textsuperscript{26,27}.
AltFUS is a novel 170 amino acid protein, endogenously expressed in cell lines and tissues.

We began by querying OpenProt\textsuperscript{27} predictions for FUS canonical mRNA (\textit{ENST00000254108} or \textit{NM_004960}), which led to 8 predicted altORFs, either overlapping the coding sequence (CDS) or within the 3'UTR (Extended Data Table 1). Amongst these, IP_243680 or altFUS, a 170 codon altORF, presents convincing experimental evidence of expression (OpenProt v1.3). AltFUS overlaps the FUS CDS in an open reading frame shifted by one nucleotide (Fig. 1a, Extended Data). FUS is a complex gene with 13 annotated transcripts resulting from alternative splicing. Based on the GTEx expression data in brain tissues and nerves, five transcripts are more abundant and represent 85% of all transcripts (Extended Data Fig. 1b). Three of them (\textit{FUS-206, FUS-211, FUS-203}) are non-coding according to Ensembl, while only two (\textit{FUS-211, FUS-203}) are non-coding according to OpenProt (Fig. 1b). Ensembl\textsuperscript{32} annotates two transcripts as coding (\textit{FUS-201 and FUS-202}), either for the 526 amino acid FUS protein or its 525 amino acid isoform. From OpenProt prediction, these two transcripts also encode altFUS (IP_243680), or its 169 amino acid isoform (IP_243691) respectively. Moreover, the second most abundant transcript in brain tissues and nerves (\textit{FUS-206}), representing about 20% of all transcripts, is non-coding according to Ensembl, but OpenProt predicts it contains the altFUS CDS (Fig. 1b, Extended Data Fig. 1c). Thus, of the five most abundant transcripts in brain tissues and nerves, two code for both FUS and altFUS proteins, one codes for altFUS alone, and the remaining two are non-coding.

Published Ribo-seq data in Human, retrieved from the Gwips portal\textsuperscript{33}, revealed an accumulation of initiating ribosomes around the altFUS initiating methionine, in association with an increase in the density of elongating ribosomes over altFUS CDS (Fig. 1c). These results suggest that altFUS is translated. Similar results were observed in Mouse (Extended Data Fig. 1d). Additionally, we retrieved nucleotide conservation scores (PhyloP) for FUS transcripts over 100 vertebrates. PhyloP scores range from -10 (highly variable) to 10 (highly conserved). PhyloP scores over the FUS CDS...
are under a constraint at the altFUS CDS locus (average score of 2.6 instead of 4 elsewhere on the FUS CDS), which is consistent with a selection pressure across 2 overlapping frames (Extended Data Fig. 1e)\(^34\). We then retrieved altFUS protein sequences over 84 species and observed a strong protein conservation across mammals, and primates notably (75 to 99.4 % of sequence identity - Extended Data Table 2, Extended Data Fig. 1f and Extended Data Alignment File). Thus, AltFUS is well conserved, with domains showing little to no sequence variations (Fig. 1d).

Based on the OpenProt database, AltFUS was identified in multiple proteomics datasets, with up to 7 confident unique peptides, representing a 41 % sequence coverage (Fig. 1e, Extended Data Table 1). To validate altFUS protein expression, we developed a custom antibody targeting two unique altFUS peptides (Extended Data Fig. 2a) and tested it using three constructs: FUS, altFUS and FUS\(^{(Ø)}\). The latter is a monocistronic FUS version, where all altFUS methionines are mutated for threonines in a manner synonymous for FUS (Extended Data Fig. 2b-d). Thus, the FUS protein sequence is unchanged, but the altFUS sequence does not contain any methionines. Transfection of HEK293 cells revealed expression of both proteins, FUS and altFUS, from the FUS nucleotide sequence (Fig. 1f). As expected, altFUS expression was lost with the monocistronic FUS\(^{(Ø)}\) construct. HEK293 cells transfected with a siRNA targeting FUS mRNA showed a significant knockdown of both proteins, FUS and altFUS; whereas altFUS endogenous expression was visible in scrambled control siRNA and mock transfected cells (Fig. 1g). These results validate the specificity of the custom antibody and demonstrate altFUS endogenous expression in HEK293 cultured cells.

AltFUS endogenous expression was visible in control human tissues, HEK293 and HeLa cell lines (Fig. 1h). Since FUS gene is associated with ALS, which predominantly affects the motor cortex in addition to the ventral spinal cord, we retrieved motor cortex lysates from 3 ALS patients with a C9orf72 mutation (most common) and 3 sporadic ALS patients. AltFUS endogenous expression
was detected in all cases (Fig. 1i). Furthermore, as ALS is a motor neuron disease, we derived functional ventral spinal motor neurons from induced pluripotent stem cells (iPSCs)\textsuperscript{35,36} from healthy controls and ALS patients carrying valosin-containing protein mutations (3 lines per group). AltFUS endogenous expression was detected in all samples (Fig. 1j). We noticed that brain (Fig. 1h) and motor cortex lysates (Fig. 1i), as well as iPSCs-derived motor neurons (Fig. 1j), from healthy controls and ALS patients, presented a higher band detected with the custom altFUS antibody. This band is not present in cultured cell lines or other tissues. It could come from a non-specific signal or a post-translational modification on altFUS that is specific to the motor cortex and spinal cord motor neurons (Fig. 1h-j). We could also observe a lower band in the line 2 of controls motor neurons, which may correspond to a degradation product or an initiation at a downstream methionine in altFUS sequence. This band has never been observed in other samples so far. We demonstrate that the \textit{FUS} gene encodes two proteins, FUS and altFUS, both endogenously expressed in the motor cortex and spinal cord motor neurons, to the two archetypal regions predominantly affected in ALS.

AltFUS is a mitochondrial protein, involved in ALS-associated loss of mitochondrial membrane potential

FLAG-tagged altFUS (altFUS-FLAG) displayed a strong co-localization with a common mitochondrial marker, TOMM20 (Fig. 2a). Additionally, mitochondrial extracts showed an enrichment in altFUS-FLAG (Fig. 2b). Cellular fractionation of cells over-expressing untagged altFUS further validated altFUS mitochondrial localization (Fig. 2c). The endogenous altFUS protein was found in the mitochondrial fraction, although it displayed a weak cytoplasmic signal as well (Fig. 2d), consistent with the immunofluorescence data (Fig. 2a). Furthermore, cells over-expressing altFUS showed an altered mitochondrial network, with a significant increase in fragmented mitochondria.
Mitochondrial fragmentation is observed in FUS-linked ALS models, and here we reproduced a similar effect when over-expressing altFUS alone. Thus, we wondered whether altFUS played a role in other mitochondrial dysfunctions observed in ALS. To this end, we reproduced an ALS-linked FUS mutant: FUS-R495x. This mutant leads to a premature stop codon before FUS NLS and is linked to severe fALS and sALS cases. In this construct, altFUS is still present and not affected by the mutation. Similarly to FUS, we also generated the monocistronic construct FUS-R495x, which contains synonymous mutations for FUS-R495x (Extended Data Fig. 3b), but prevents altFUS expression. V5-FUS-FLAG and V5-FUS-FLAG-R495x did not express altFUS, but only the FUS protein, wild-type or ALS-linked mutant R495x respectively (Extended Data Fig. 3c).

We first investigated the effect of altFUS on the mitochondrial membrane potential using the potential sensitive dye TMRE (Fig. 2g, h, Extended Data Fig. 3d). As previously described, over-expression of bicistronic FUS or FUS-R495x led to a decrease in mitochondrial membrane potential. The mitochondrial membrane potential remained normal when over-expressing monocistronic FUS or FUS-R495x, underlining the role of altFUS. However, over-expression of altFUS alone did not alter the mitochondrial membrane potential, which suggests both proteins cooperate for this ALS pathological hallmark.

To further evaluate the possible implication of altFUS in ALS, we investigated its protein interactors. Using stimulated emission depletion microscopy (STED), we observed that altFUS localized in puncta following a cristae-like pattern inside the mitochondria, delimited using an outer-membrane mitochondrial marker, TOMM20 (Fig. 2i, j). We then used size-exclusion chromatography on mitochondrial extracts to isolate altFUS-FLAG macromolecular complexes (Fig. 2k). Following a FLAG affinity purification and mass spectrometry analysis, we identified

(globular) compared to mock cells that displayed more tubular structures (Fig. 2e, f, Extended Data Fig. 3a).
interacting proteins (Fig. 2l, Extended Data Table 3). A gene enrichment analysis to the Human mitochondrial proteome identified three significantly enriched biological processes: autophagy-related pathways, mitochondrial metabolism and cellular response to stress (Fig. 2m). Disruptions within these are pathological hallmarks of ALS\textsuperscript{3,21,39,40}.

**AltFUS inhibits autophagy and drives the accumulation of FUS- and TDP43-positive cytoplasmic aggregates**

Following on these results, we hypothesized that the inhibition of autophagy observed with ALS-linked FUS mutants may instead be attributed to altFUS. We used the mCherry-GFP-LC3 reporter to track the autophagic flux by confocal microscopy (Extended Data Fig. 4a). Under basal conditions, cells displayed red and yellow foci as expected (Fig. 3a). An accumulation of yellow foci was observed when cells were treated with bafilomycin, an inhibitor of autophagy. Similarly, cells over-expressing altFUS displayed a significant accumulation of yellow foci (Fig. 3a). Our results were consistent with previously published data\textsuperscript{11} as cells transfected with FUS or FUS-R495x displayed a decreased autophagic flux (Fig. 3a, Extended Data Fig. 4b). This accumulation of yellow foci was absent in cells that express monocistronic FUS constructs, thus lacking altFUS expression (FUS\textsuperscript{(Ø)} or FUS\textsuperscript{(Ø)-R495x}). We used bafilomycin followed by LC3 probing to further validate the impact of altFUS on autophagy (Fig. 3b, Extended Data Fig. 4c). Similarly, an inhibition of autophagy was observed only in cells over-expressing altFUS. Furthermore, in cells over-expressing monocistronic FUS\textsuperscript{(Ø)-R495x}, the inhibition of autophagy could be restored by co-transfecting altFUS (Fig. 3b). These results establish altFUS, rather than FUS, as the protein responsible of the inhibition of autophagy.

Furthermore, altFUS interactome analysis suggested a role in the cellular stress response, which is known to be altered in ALS with a TDP-43 cytoplasmic accumulation in 98 % of patients\textsuperscript{41,42}. In
FUS-linked ALS and some sALS cases, FUS cytoplasmic aggregates or mislocalization are also observed. We demonstrated that cells over-expressing FUS-R495x displayed cytoplasmic aggregates that were positive for both FUS-R495x and TDP-43 (Fig. 3c). In cells over-expressing the monocistronic FUS(Ø)-R495x construct, thus lacking altFUS expression, FUS-R495x displayed a more diffuse cytoplasmic localization, and TDP-43 remained in the nucleus (Fig. 3c). FUS cytoplasmic aggregates were significantly more numerous and larger when altFUS was co-expressed (Fig. 3d, e). Accumulation of FUS-R495x and TDP-43 in cytoplasmic aggregates could be reconstituted by co-transfecting altFUS and the monocistronic FUS(Ø)-R495x construct (Fig. 3c). These observations were repeated across all 7 ALS-linked FUS mutations tested (Extended Data Fig. 4d-f). The cytoplasmic aggregates were also TIAR-positive as observed with ALS-linked FUS mutants in previous work (Extended Data Fig. 5a, b-d). These results suggest a strong interplay between FUS and altFUS in ALS and that altFUS plays a pivotal role in the assembly of cytoplasmic FUS/TDP-43 aggregates.

**ALT FUS PROTECTS AGAINST NEURODEGENERATION IN FUS-ASSOCIATED DROSOPHILA MODELS**

In order to investigate the role of altFUS in ALS, we generated Drosophila models expressing either the bicistronic, FUS and FUS-R495x constructs, or the monocistronic, FUS(Ø) and FUS(Ø)-R495x, constructs. We used the Elav-GeneSwitch-GAL4 Driver strain, as previously described, as it allows for an inducible over-expression in motor neurons and avoids lethality at the larval stage from FUS over-expression in the central nervous system. First, we generated flies containing the sequences for UAS-t-altFUS, UAS-t-FUS, UAS-t-FUS(Ø), UAS-t-FUS-R495x or UAS-t-FUS(Ø)-R495x. These flies were then crossed with the Elav-GeneSwitch-GAL4 driver strain (Fig. 4a). UAS-t-mCherry flies were used as controls. Selected F1 individuals were then divided into 2 groups with equal proportions of males/females. The first group received standard food, while the other received RU-486 treated food. The treatment induces a conformational change in the Elav-
GeneSwitch driver, which allows activation of the UAS promoter and thus expression of the target protein. We retrieved flies at selected time points to validate protein expression in the RU-486 treated population through time, while the controls showed no expression (Fig. 4b).

The motor neuron degeneration linked to ALS provokes a progressive locomotion loss measurable with a well-described climbing assay\(^4^8\). The control populations did not show any significant locomotion loss at day 1, 10 nor 20 (Fig. 4c-e); similarly to the RU-486 treated control group (mCherry transgenic flies – Fig. 4c). AltFUS flies did not show any significant locomotion loss through time (Fig. 4c). This result is consistent with the in cellulo data showing altFUS alone is not sufficient to provoke pathological hallmarks. As previously shown with this model\(^1^9\), the bicistronic FUS flies, which express both FUS and altFUS proteins, displayed a significant locomotion loss (Fig. 4d). Bicistronic ALS-linked FUS-R495x flies showed an even greater motor neuron degeneration through time compared to FUS (Fig. 4e). Monocistronic FUS\(^0\) (Fig. 4d) and FUS\(^0\)-R495x (Fig. 4e) flies, which do not express altFUS, displayed both a delay in the onset of motor neuron degeneration and a reduced drop in climbing success at 20 days post-induction (40% vs. 60% for FUS\(^0\), 30% vs. 70% for FUS\(^0\)-R495x). These results in Drosophila confirm a role for altFUS in ALS-like neurodegeneration in vivo, are consistent with our in cellulo observations and highlight the pathological cooperation between FUS and altFUS.

**ALS-linked mutations, synonymous for FUS, alter AltFUS and lead to TDP-43 cytoplasmic aggregates**

As of today, over 50 mutations in the FUS gene have been associated with ALS\(^5\). However, most of these locate at the carboxyl end of the protein and as such have no effect on altFUS. We wondered whether mutations altering altFUS may have been overlooked as non-consequential in the FUS reading frame. We retrieved FUS synonymous mutations found in ALS patients, with an
allelic frequency below 0.01 %, from previous studies and the ALS Variant Server (http://als.umassmed.edu/ - **Extended Data Table 4**). The retrieved mutations clustered on the altFUS locus, with 60 % of FUS synonymous mutations found in sALS patients and 50 % of FUS synonymous mutations found in fALS patients, which is significantly higher than expected by chance (34 %) (**Extended Data Table 4**). We selected 4 mutations for further analysis based on the residue conservation: altFUS-P31L, altFUS-A38V, altFUS-A46V and altFUS-R64P. We generated them in GFP-FUS constructs: GFP-FUS(P31L-FLAG)-S44=; GFP-FUS(A38V-FLAG)-G51=; GFP-FUS(A46V-FLAG)-G59= and GFP-FUS(R64P-FLAG)-S77= (**Fig. 5a**). All altFUS mutants still localized to the mitochondria (**Fig. 5b**).

To investigate whether these mutations may provoke an ALS phenotype, we quantified the number of cells presenting TDP-43 aggregates. All 4 altFUS mutants showed a 1.8 to 2.4 fold increase compared to wild-type altFUS (**Fig. 5c, d**). This result indicates that altFUS mutations potentiate TDP-43 cytoplasmic aggregation, a pathological hallmark in 98 % of ALS cases. Hence, some *FUS* mutations, synonymous for the FUS protein, exert a deleterious effect through their missense consequence on the altFUS protein.

**CONCLUSIONS: FUS AND ALT FUS, ENCODED BY THE SAME GENE, COOPERATE IN ALS-LINKED TOXIC DYSFUNCTIONS**

Despite considerable advances in the field, current genome annotations still uphold arbitrary assumptions, such as the monocistronic nature of eukaryotic genes. Here, we demonstrate *FUS* is a bicistronic gene. We discovered FUS CDS contains a second protein-coding sequence in a shifted frame overlapping its prion-like intrinsically disordered domain, regions known to host dual-coding events. This novel protein, named altFUS, is not an isoform but an entirely new sequence of 170 amino acids. AltFUS is endogenously expressed in human tissues and cultured cell lines, as demonstrated by ribosome profiling, mass spectrometry and with a custom antibody. AltFUS is notably expressed in the motor cortex and iPSCs-derived motor neurons of healthy
controls and ALS patients. Because altFUS is embedded within the FUS CDS, this discovery is of crucial importance to the field. Indeed, over-expression studies on FUS actually implicate two proteins: FUS and altFUS. Similarly FUS knockdown or knockout studies actually inhibit expression of both proteins51. Moreover, previous work has shown that gene editing techniques targeting a specific CDS do not necessarily result in knockout of the gene in case of dual-coding gene52. Our discovery thus suggests that FUS edited cells or models, notably targeting its last exons53–55, might only impair FUS protein expression but not altFUS, thus not resulting in a true FUS knockout. Our work provides an unprecedented and much needed reconsidered view of FUS coding potential to better understand its role, notably in the context of ALS and other neurodegenerative diseases.

Following the discovery of altFUS, we developed tools in order to differentiate the specific roles and phenotypes of FUS and altFUS. Our study demonstrates that altFUS is necessary for three ALS pathological hallmarks: mitochondrial fragmentation and loss of membrane potential, inhibition of autophagy, and cytoplasmic aggregation of FUS/TDP-43. The inhibition of autophagy was observed when over-expressing altFUS alone, absent when over-expressing FUS alone (wild-type or ALS-associated R495x mutant) and reconstituted when co-expressing FUS and altFUS. This demonstrates that the inhibition of autophagy, previously described in FUS-linked ALS11,21, has been incorrectly associated to the FUS protein. AltFUS inhibits autophagy. AltFUS is necessary but not sufficient for the mitochondrial membrane potential loss and cytoplasmic aggregation of FUS and TDP-43. Although TDP-43 aggregates are not commonly seen with FUS mutations45, this study lends support for altFUS conspiring with FUS and TDP-43 to lead to this molecular hallmark of ALS. Our data suggests the stoichiometry between FUS and altFUS may be important for the development of cytoplasmic aggregates in ALS. Both proteins are required to observe the phenotype, highlighting a functional alliance between FUS and altFUS. This pathological synergy was also observed in the Drosophila model.
The physiological function of altFUS is still unclear, although our work provides evidence for its role in mitochondrial dynamics and the cellular response to stress. One mechanism put forward in ALS is that the disease originates from a sub-optimal resolution of cellular stresses, which can come from environmental sources or mutated proteins\(^40,56\). We have shown that altFUS, not FUS, inhibits autophagy, most likely via its interaction partners (Extended Data Table 3). Since an inhibition of autophagy inhibits the dissociation of stress granules\(^40\), we suggest altFUS potentiates stress granule accumulation under stress conditions, and that FUS phase separation properties then lead to the formation of solid and toxic aggregates\(^57\). Further work is needed to fully understand the role of altFUS in physiological and pathological conditions, yet our study shows this novel protein plays a crucial role in FUS-linked gain-of-toxic dysfunctions observed in ALS. Hence, FUS should be recognized and annotated as a polycistronic gene.

Recent studies have addressed the toxicity resulting from over-expression of wild-type FUS. Bogaert and colleagues used FUS domain truncation mutants to investigate wild-type FUS toxicity. A FUS mutant lacking its N-terminal intrinsically disordered domain, thus lacking altFUS, displayed reduced toxicity. This study corroborates our findings that the absence of altFUS reduces the toxicity. Furthermore, Bogaert and colleagues concluded that FUS N-terminal synergizes with the C-terminal domain to mediate toxicity in Drosophila\(^47\). Here, we showed that altFUS synergizes with FUS to mediate ALS pathological features in cultured cells and toxicity in Drosophila. Despite different laboratories and different techniques, our data are in agreement with theirs. Yet, we point to an alternative (not necessarily mutually exclusive) explanation whereby altFUS, not the FUS N-terminal domain, synergizes for toxicity. Additionally, this discussion shows that not being aware of overlapping CDSs, especially in deletion studies, precludes alternative interpretations.

Current genome annotations guide the interpretation of data and the design of studies, however they also affect the way we screen for pathological mutations\(^56\). Most of the ALS-linked FUS
mutations affect the carboxyl end of the protein, as shown with the example used in this study, the R495x mutation\textsuperscript{13}. These mutations do not alter the altFUS protein, which is embedded at the beginning of the FUS CDS and span on exons 3 to 6. How could altFUS be important for the disease if most pathological mutations do not alter it? To answer this question, one needs to grasp how much genome annotations shape today’s research. For example, when screening for pathological mutations, those that are synonymous for FUS are discarded early in analyses as insignificant\textsuperscript{58}. Yet, a synonymous mutation for FUS may not be for altFUS. Our work shows that FUS synonymous mutations found in patients cluster on altFUS genomic locus. We tested 4 of these mutations, synonymous for the FUS protein and missense for the altFUS protein, and we found that each of them potentiated TDP-43 cytoplasmic aggregation, a pathological hallmark of ALS\textsuperscript{42}. Thus, FUS is a double-edge genetic source of ALS, with both of its proteins involved in the aetiology of the disease.

This work demonstrates the importance of genome annotation in our quest to better understand physiological and pathological mechanisms and serves as a proof-of-concept to acknowledge the polycistronic coding potential of eukaryotic genes\textsuperscript{26}. For example, taking the 16 genes associated with ALS as reviewed in Taylor and colleagues, 56\% (9) of them display at least one currently non-annotated protein that has been detected by mass spectrometry (44\%) or ribosome profiling (31\%) according to the OpenProt database (\textbf{Extended Data Table 5})\textsuperscript{3,27}. Several studies have highlighted the biological function of currently non-annotated proteins, and intriguingly a functional cooperation between proteins from the same gene is often observed\textsuperscript{23,26,52,59–61}. Furthermore, a quantitative proteomic study looked at the relative abundance of two proteins from the bicistronic \textit{MIEF1} gene, MiD51 and altMiD51. The major product of \textit{MIEF1} gene is the novel protein, altMiD51 (now annotated as L0R8F8 in UniProtKB), as shown in cell lines and
healthy colon tissue. AltFUS is the first example of an alternative protein linked to a pathology but is unlikely to be the only one.
METHODS

FUS constructs

FUS and altFUS sequences were obtained from Bio Basic Gene Synthesis service. All FUS constructs were subcloned into pcDNA3.1- (Invitrogen) using Gibson assembly (New England Biolabs, E26115). FUS and altFUS wild-type sequences correspond to that of the Human FUS canonical transcript (ENST00000254108 or NM_004960). FUS and altFUS proteins were tagged with V5 (GKPIPNPLLGLDST) and 2 FLAG (DYKDDDDKDYKDDDDK) respectively. FUS was tagged on the N-terminal, altFUS was tagged on the C-terminal. For immunofluorescence assays, N-terminal GFP-tagged FUS was also cloned into pcDNA3.1- by Gibson assembly. The necessary gBlocks were purchased from IDT. The monocistronic constructs FUS(Ø) and FUS(Ø)-R495x were generated by mutating all altFUS methionines (ATG) to threonines (ACG). These mutations are synonymous in the FUS CDS (TAT > TAC, both coding for tyrosine). The altFUS mutated sequence was obtained from Bio Basic Gene Synthesis service, and then subcloned in FUS sequences in pcDNA3.1- using Gibson assembly. The bicistronic constructs are named as follows throughout the article: FUS, FUS-R495x, or FUS(FLAG) and FUS(FLAG)-R495x when altFUS is FLAG-tagged in the +2 reading frame. The monocistronic constructs are named as follows throughout the article: FUS(Ø) or FUS(Ø)-R495x to indicate altFUS absence.

Cell culture, transfections, western blots and immunofluorescence

HEK293 and HeLa cells cultures tested negative for mycoplasma contamination (ATCC 30–1012K). Transfections, immunofluorescence, confocal analyses and western blots were carried out as previously described. For FUS knock-down, 150 000 HEK293 cells in a 6-well plate were transfected with 25 nM FUS SMARTpool: siGENOME siRNA (Dharmacon, Canada, L-009497-00-0005) or ON-TARGET plus Nontargeting pool siRNAs (Dharmacon, D-001810-10-05) with DharmaFECT one transfection reagent (Dharmacon, T-2001–02) according to the manufacturer’s
protocol. Cell media was changed every 24 hrs and cells were processed 72 hrs after transfection. For immunofluorescence, primary antibodies were diluted as follows: anti-Flag (Sigma, F1804) 1/1000, anti-TOMM20 (Abcam, ab186734) 1/500, anti-V5 (Cell Signalling Technologies, #13202) 1/1000, anti-TDP-43 (ProteinTech, 10782-2-AP) 1/500, and anti-TIAR (Cell Signalling Technologies, #8611) 1/1600. For western blots, primary antibodies were diluted as follows: anti-Flag (Sigma, F1804) 1/8000, anti-V5 (Sigma, V8012) 1/8000, anti-actin (Sigma, A5441) 1/40000, anti-FUS (Abcam, ab84078) 1/500, anti-altFUS (Abcam, custom antibody) 1/3000, anti-LC3 (Cell Signalling Technologies, #2775) 1/1000, anti-Hsp70 (Thermo Fisher Scientific, MA3-028) 1/1000, anti-Tubulin (Thermo Fisher Scientific, a11126) 1/2000 and anti-VDAC (Abcam, ab15895) 1/10000. The altFUS antibody was generated by injection two rabbits, each with 2 unique altFUS peptide. The purified antibody from rabbit 2 was used in this study at a 1/2000 dilution. Mitochondrial morphology was evaluated using the microP tool. A minimum of 100 cells per replicate were counted across 3 independent experiments (n = 3, i.e. minimum 300 cells for each experimental condition). Colocalization analyses were performed using the JACoP plugin (Just Another Colocalization Plugin) implemented in Image J software, as previously described. When specified, images obtained by confocal microscopy on the Leica TCS SP8 STED 3X were deconvolved using the Huygens software (Scientific Volume Imaging B.V., Hilversum, Netherlands). The software uses a signal reassignment algorithm for deconvolution, identical deconvolution parameters were applied to all images. The default parameters were used, including the Classic Maximum Likelihood Estimation (CMLE) algorithm, signal to noise ratio, and background estimation radius. The maximum iteration number was set at 30. Human tissue lysates for altFUS endogenous expression were purchased from Zyagen Laboratories (San Diego, California, USA).

**Ribo-seq data and conservation analyses**
Global aggregate reads for initiating ribosomes and elongating ribosomes footprints across all available studies were downloaded from the Gwips portal (https://gwips.ucc.ie/), for *Homo sapiens* and for *Mus musculus*. For altFUS protein conservation analysis, all FUS mRNAs with at least EST evidence were retrieved across all available species from NCBI RefSeq. We performed an *in silico* 3-frame translation and retrieved the best matching protein sequence per species that displayed a minimum of 20 % sequence identity with the Human altFUS sequence over 25 % of Human altFUS length. AltFUS homologous sequences were found in 83 species, and we manually added that of *Drosophila melanogaster* that displayed a 37.5 % sequence identity over 19 % of the Human altFUS length. All retrieved altFUS sequences were then aligned using Clustalω with default parameters.

**Human induced pluripotent stem cell differentiation into motor neurons**

Directed differentiation to human iPSC-motor neurons was performed as previously reported\(^{35}\). Briefly, iPSCs were maintained on Geltrex (Life Technologies) with Essential 8 Medium media (Life Technologies), and passaged using EDTA (Life Technologies, 0.5mM). All cell cultures were maintained at 37°C and 5% carbon dioxide. For motor neuron differentiation, iPSCs were differentiated to neuroepithelium by plating to 100% confluency in chemically defined medium consisting of DMEM/F12 Glutamax, Neurobasal, L-Glutamine, N2 supplement, non-essential amino acids, B27 supplement, β-mercaptoethanol (Life Technologies) and insulin (Sigma). Treatment with the following small molecules from day 0-7: 1µM Dorsomorphin (Millipore), 2µM SB431542 (Tocris Bioscience), and 3.3µM CHIR99021 (Miltenyi Biotec). At day 8, cells patterned for 7 days with 0.5µM retinoic acid and 1µM Purmorphamine. At day 14 spinal cord motor neuron precursors were treated with 0.1µM Purmorphamine for a further 4 days before being terminally differentiated for >10 days in 0.1 µM Compound E (Enzo Life Sciences) to promote cell cycle exit.
Throughout the neural conversion and patterning phase (D0-18) the neuroepithelial layer was enzymatically dissociated twice (at D4-5 and D10-12) using dispase (GIBCO, 1 mg ml-1).

**Preparation of tissue lysates of the motor cortex of ALS patients**
Approximately 100mg of motor cortex from 4 sporadic ALS and 4 C9orf72-ALS cases was lysed in 10x RIPA (50mM Tris HCl pH7.8, 150mM NaCl, 0.5% sodium deoxycholate, 1% NP40; supplemented with protease inhibitors and EDTA) volume using TissueLyzer equipment (Qiagen). Lysates were incubated on ice 20 minutes followed by centrifugation at 20,000xg for 20 minutes at 4°C. Supernatant was taken as ‘RIPA fraction’ and pellets resuspended in RIPA and SDS (final concentration of 2%). 3 sporadic ALS and 3 C9orf72-ALS samples were subsequently used as they were sufficiently concentrated to load 100 ug of proteins onto SDS-page gels.

**Mitochondrial extracts and cellular fractionation**
Mitochondrial extracts were prepared as previously described. Briefly, HEK293 cells grown up to 80% confluence, were rinsed twice with PBS and gathered using a cell scraper. Cells were pelleted by centrifugation at 500 x g for 10 mins at 4 °C. Supernatant was discarded and cells suspended in mitochondrial buffer (210 mM mannitol, 70 mM sucrose, 1 mM EDTA, 10 mM HEPES-NaOH, pH 7.5, 0.5 mM PMSF and EDTA-free protease inhibitor (Thermo Fisher Scientific)). Cells were disrupted by 15 consecutive passages through a 25G1 0.5 x 25 needle syringe on ice, followed by a 3 min centrifugation at 2 000 x g at 4 °C. Supernatant was collected and the pellet suspended in mitochondrial buffer. The cell disruption was repeated four times and all retrieved supernatants containing mitochondria were again passed through syringe needle in mitochondrial buffer and cleared by centrifugation for 3 mins at 2000 x g at 4 °C. Supernatants were pooled and centrifuged for 10 mins at 13 000 x g at 4 °C to pellet mitochondria. The pellet was suspended in 200 μL of mitochondrial buffer until further processing. Cellular fractionation was performed using the Cell Fractionation Kit (#9038S, Cell Signaling Technology). Briefly, HEK293 cells were
grown up to 80 % confluence, washed twice with PBS and gathered using a cell scraper. Cells were
spun at 350 x g for 5 mins at 4 °C and 2.5 x 10^6 cells were suspended in 500 μL of ice-cold PBS. An
aliquot of 100 μL was spun at 350 x g for 5 mins at 4 °C and resuspended in SDS buffer (4 % SDS,
Tris-HCl 100 mM pH 7.6) and kept as WCL (Whole Cell Lysate). The rest of the collected cells
(remaining 400 μL) were spun at 500 x g for 5 mins at 4 °C. Supernatant was discarded and pellet
resuspended in 500 μL of CIB (Cytoplasmic Isolation Buffer) from the kit, vortexed for 5 secs and
incubated on ice for 5 mins. After centrifugation at 500 x g for 5 mins at 4 °C, the supernatant was
collected as the cytosplasmic fraction. The pellet was resuspended in 500 μL of MIB buffer
(Membrane Isolation Buffer) from the kit, vortexed for 15 secs and incubated on ice for 5 mins.
After centrifugation at 8 000 x g for 5 mins at 4 °C, the supernatant was collected as the membrane
and organelles fraction. To each 100 μL of fraction was added 60 μL of loading buffer 1 x (from
ColdSpring Laemmli sample buffer: 50 mM Tris pH 6.8, 2 % SDS, 10 % Glycerol, 5 % β-mercaptoethanol) before processing for western blot.

Mitochondrial membrane potential measurements
Mitochondrial membrane potential was measured by flow cytometry in HEK293 cells using TMRE
(Tetramethylrhodamine, Ethyl ester, Abcam, ab113852). FCCP was used as a positive control to
validate each independent experiment. Cells were grown up to 80 % confluence and washed twice
with PBS. The cells were then incubated for 5 mins at 37 °C, 5 % CO₂ with PBS/A (0.2 % BSA in PBS)
solution (experimental) or 3 μM FCCP in PBS/A solution (positive control). Then, 100 nM of TMRE
was added and cells were incubated 15 mins at 37 °C, 5 % CO₂. After incubation, cells were
trypsinized and centrifuged at 800 x g for 5 mins at 4 °C and resuspended in 500 μL of PBS and
kept on ice. Cells were immediately analysed by flow cytometry. A gate for living cells was set, as
well as a second gate to filter out cell doublets. TMRE fluorescence (PE-A) was recorded over a
minimum of 50,000 gated cells for each experimental condition. The mean TMRE fluorescence intensity was measured over 3 independent experiments for each experimental condition.

### Stimulated Emission Depletion (STED) microscopy

Samples were prepared as described above for confocal microscopy. A Leica TCS SP8 STED 3X was used with a 100x objective lens and immersion oil for dual-color STED images. Images were obtained by sequential scanning of a given area. The combination of Alexa Fluor 488 (Thermo Fisher Scientific, A-11017) and Alexa Fluor 568 (Thermo Fisher Scientific, A-21069) dyes was chosen for STED imaging. Alexa Fluor 488 dye was excited with a white light laser (WLL) at 488 nm and was depleted using the 660 nm STED laser. Alexa Fluor 568 dye was excited with a WLL at 561 nm and was depleted using the 660 nm STED laser. The STED laser (660 nm) was applied at 80% of maximum power.

### Fast Protein Liquid Chromatography (FPLC) and affinity-purification mass spectrometry (AP-MS)

Mitochondrial extracts of HEK293 cells were centrifuged at 13,000 x g for 10 mins at 4 °C to remove the supernatant and were resuspended in FPLC buffer (50 mM Tris-HCl, 1 mM EDTA, 150 mM NaCl, 1% Triton X-100, pH 7.5, filtered with 0.2 μm filters) at 2 mg/mL for a total of 4 mg of mitochondrial proteins. Samples were incubated on ice for 15 mins and then centrifuged at 10,000 x g for 5 mins at 4 °C and the supernatant was loaded in the injection syringe without disrupting the pellet. The FPLC was performed on a HiLoad 16/60 Superdex 200 pg column (GE Healthcare, Chicago, USA) at 4 °C. The column was pre-equilibrated with the FPLC buffer for up to 0.2 CV (column volume) and the sample was applied at a flow rate of 0.5 mL/min with a pressure alarm set at 0.5 MPa. The elution was performed over 72 fractions of 1.5 mL for a maximum of 1.1 CV. For altFUS probing by western blot, proteins were precipitated from 150 μL of each 4 fractions in technical duplicates. First, 600 μL of methanol was added to each tube and mixed gently, before adding 150 μL of chloroform. Tubes were gently inverted 10 times before adding 450 μL of milliQ
H₂O and vortexing briefly. After centrifugation at 12 000 x g for 3 mins, the upper phase was discarded, and 400 μL of methanol was added. Tubes are centrifuged at 16 000 x g for 4 mins and the pellet was resuspended in loading buffer. For interactome analysis by mass spectrometry, fractions of interest (8 to 14) were pooled together and incubated at 4 °C overnight with magnetic FLAG beads (Sigma, M8823) pre-conditioned with FPLC buffer. The beads were then washed 3 times with 5 mL of FPLC buffer, and 5 times with 5 mL of 20 mM NH₄HCO₃ (ABC). Proteins were eluted and reduced from the beads using 10 mM DTT (15mins at 55 °C), and then treated with 20 mM IAA (1 hour at room temperature in the dark). Proteins were digested overnight by adding 1 μg of trypsin (Promega, Madison, Wisconsin) in 100 μL ABC at 37 °C overnight. Digestion was quenched using 1 % formic acid and supernatant was collected. Beads were washed once with acetonitrile/water/formic acid (1/1/0.01 v/v) and pooled with supernatant. Peptides were dried with a speedvac, desalted using a C18 Zip-Tip (Millipore Sigma, Etobicoke, Ontario, Canada) and resuspended into 30 μl of 1% formic acid in water prior to MS analysis.

**Mass-spectrometry analysis**

Peptides were separated in a PepMap C18 nano column (75 μm × 50 cm, Thermo Fisher Scientific). The setup used a 0–35% gradient (0–215 min) of 90% acetonitrile, 0.1% formic acid at a flow rate of 200 nL/min followed by acetonitrile wash and column re-equilibration for a total gradient duration of 4 h with a RSLC Ultimate 3000 (Thermo Fisher Scientific, Dionex). Peptides were sprayed using an EASYSpray source (Thermo Fisher Scientific) at 2 kV coupled to a quadrupole-Orbitrap (QExactive, Thermo Fisher Scientific) mass spectrometer. Full-MS spectra within a m/z 350–1600 mass range at 70,000 resolution were acquired with an automatic gain control (AGC) target of 1e6 and a maximum accumulation time (maximum IT) of 20 ms. Fragmentation (MS/MS) of the top ten ions detected in the Full-MS scan at 17,500 resolution, AGC target of 5e5, a maximum IT of 60 ms with a fixed first mass of 50 within a 3 m/z isolation window at a normalized
collision energy (NCE) of 25. Dynamic exclusion was set to 40 s. Mass spectrometry RAW files were searched with Andromeda search engine implemented in MaxQuant 1.5.5.1. The digestion mode was set at Trypsin/P with a maximum of two missed cleavages per peptides. Oxidation of methionine and acetylation of N-terminal were set as variable modifications, and carbamidomethylation of cysteine was set as fixed modification. Precursor and fragment tolerances were set at 4.5 and 20 ppm respectively. Files were searched using a target-decoy approach against UniprotKB (*Homo sapiens* 03/2017 release) with the addition of altFUS sequence for a total of 92,949 entries. The false discovery rate (FDR) was set at 1% for peptide-spectrum-match, peptide and protein levels. Protein interactions were then scored using the SAINT algorithm, with Mock cells as control and the magnetic FLAG beads in HEK293 cells. Proteins with a SAINT score above 0.99 were considered, as well as those presenting a SAINT score above 0.88 with a minimum of two unique peptides.

**Biological processes and cellular compartment enrichment analysis**

Proteins identified in altFUS interactome were screened for cellular compartment and biological processes enrichment using Gene Ontology (GO) enrichment. Proteins were queried against the whole Human Proteome for cellular compartment and against the Human mitochondrial proteome (MitoCarta 2.0) for biological processes. The statistical analysis used a Fisher’s Exact test with a FDR set at 1%.

**Autophagic flux measurements**

The mCherry-GFP-LC3 was used to evaluate the autophagic vesicles within HeLa cells by confocal microscopy. Before fusion with the lysosome, the LC3 molecules on the autophagosome display a yellow fluorescence (combined mCherry and GFP fluorescence). After fusion, the GFP fluorescence is quenched by the lysosomal pH, and as such the LC3 molecules display a red signal (mCherry alone). This allows a visual representation of the autophagic flux in a given cell. Cells
treated with 50 nM Bafilomycin for 4 hours were used as a positive control to validate each independent experiment. Observations were made across 2 technical duplicates for each biological condition, across 3 independent experiments (n=3). Alternatively, the autophagic flux was also evaluated by LC3 probing before and after bafilomycin treatment (50 nM for 4 hours). The quantification corresponds to the treated / untreated ratio of LC3-II abundance.

**Cytoplasmic aggregates measurements**

Images of HeLa cells were taken by confocal microscopy and then processed using the Image J 3D Objects Counter plugin. FUS cytoplasmic aggregates were then quantified in number and size ($\mu m^2$) for each cell. A total of 100 cells across two technical replicates were taken for each independent experiment (n=3, i.e. a minimum of 300 cells per biological conditions).

**Transgenic Drosophila and climbing assay**

The bicistronic constructs, FUS and FUS-R495x, and the monocistronic constructs, altFUS, FUS($^{\Omega}$) and FUS($^{\Omega}$)-R495x, were subcloned in the pUASTattB expression vector for site specific insertion into attP2 on chromosome 3. Transgenic flies were generated by Best Gene (Best Gene Inc., California, USA). The Elav-GeneSwitch-GAL4 driver (stock number: 43642, genotype: y[1] w[*]; P[w[/+mC]=elav-Switch.O]GSG301) and the UAS-mCherry flies (stock number: 35787, genotype: y[1] sc[*] v[1]; P[y[+t7.7] v[+t1.8]=UAS-mCherry.VALIUM10]attP2) was purchased from Bloomington (Bloomington Drosophila Stock Center, Indiana, USA). All stocks were in a w$^{1118}$ background and were cultured on standard medium at 25°C or room temperature. Transgenic flies were crossed with the Elav-GeneSwitch-GAL4 driver strain. The F1 was equally divided in two groups with equal proportion of males and females: one group will feed on standard food supplemented with ethanol (0.2 % - control flies), the other on standard food supplemented with RU-486 at 10 $\mu M$ diluted in ethanol (induced flies). The climbing assay was performed as previously described$^{48}$. Briefly, flies were transferred into an empty vial and tapped to the bottom.
After 18 s, the number of flies at the top of the tube were considered successful. The assay was done at day 1, 10 and 20 post-induction, across 4 independent F1. Five flies were taken at day 1, 10 and 20 post-induction to validate expression of the proteins of interest.

**Statistical analyses and representation**

Unless otherwise stated, the statistical analysis carried was a two-way ANOVA with Tukey’s multiple comparison correction. The box plots represent the mean with the 5 to 95 % percentile. The bar graphs represent the mean, and error bars correspond to the standard deviation. When using parametric tests, normality of data distribution was verified beforehand using the Shapiro-Wilk test.

**DATA AVAILABILITY STATEMENT**

The OpenProt database is available at [www.openprot.org](http://www.openprot.org). The GTEx portal is available via [www.gtexportal.org](http://www.gtexportal.org). The Gwips portal is available at [www.gwips.ucc.ie](http://www.gwips.ucc.ie). The proteomics data are available on the PRIDE repository with the accession PXD------ ([Fig 2l-m](#)). Any other relevant data are available from the corresponding authors upon reasonable request.

**ACKNOWLEDGEMENTS**

This research was supported by CIHR grants MOP-137056 and MOP-136962, by an ALS Canada Project Grant, and by a Canada Research Chair in Functional Proteomics and Discovery of Novel Proteins to X.R. X.R and S.J are members of the Fonds de Recherche du Québec Santé (FRQS)-supported Centre de Recherche du Centre Hospitalier Universitaire de Sherbrooke. P.M. is supported by the ALS Double Play Christopher Chiu Postdoctoral Fellowship. We thank J. Ule, M-J. Boucher and D. Hunting for helpful discussions. The autopsy program is supported by the James Hunter and Family ALS Initiative. The Genotype-Tissue Expression (GTEx) Project was supported
by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on April 2019.

**CONTRIBUTIONS**

MA.B and X.R designed and wrote the study. MA.B and JF.J did the experiments, and MA.B did the analyses and figures. S.N and S.J assisted with the Drosophila experiments. P.M, L.Z and J.R provided the motor cortex tissue lysates. GE.T and R.P provided the iPSCs-derived motor neurons. All authors proofread the manuscript.
1. Morgan, S. & Orrell, R. W. Pathogenesis of amyotrophic lateral sclerosis. *Br. Med. Bull.* **119**, 87–98 (2016).

2. Mathis, S., Goizet, C., Soulages, A., Vallat, J.-M. & Masson, G. L. Genetics of amyotrophic lateral sclerosis: A review. *J. Neurol. Sci.* **399**, 217–226 (2019).

3. Taylor, J. P., Brown Jr, R. H. & Cleveland, D. W. Decoding ALS: from genes to mechanism. *Nature* **539**, 197–206 (2016).

4. Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).

5. Nguyen, H. P., Broeckhoven, C. V. & Zee, J. van der. ALS Genes in the Genomic Era and their Implications for FTD. *Trends Genet.* **0**, (2018).

6. Zou, Z.-Y., Liu, M.-S., Li, X.-G. & Cui, L.-Y. Mutations in FUS are the most frequent genetic cause in juvenile sporadic ALS patients of Chinese origin. *Amyotroph. Lateral Scler. Front. Degener.* **17**, 249–252 (2016).

7. Millecamps, S. *et al.* Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. *J. Med. Genet.* **49**, 258–263 (2012).

8. Corcia, P. *et al.* A novel mutation of the C-terminal amino acid of FUS (Y526C) strengthens FUS gene as the most frequent genetic factor in aggressive juvenile ALS. *Amyotroph. Lateral Scler. Front. Degener.* 1–4 (2017) doi:10.1080/21678421.2016.1265564.

9. Deng, H., Gao, K. & Jankovic, J. The role of FUS gene variants in neurodegenerative diseases. *Nat. Rev. Neurol.* **10**, 337–348 (2014).

10. Vance, C. *et al.* Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science* **323**, 1208–1211 (2009).
11. Marrone, L. et al. FUS pathology in ALS is linked to alterations in multiple ALS-associated proteins and rescued by drugs stimulating autophagy. *Acta Neuropathol. (Berl.)* (2019) doi:10.1007/s00401-019-01998-x.

12. Dormann, D. et al. ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. *EMBO J.* 29, 2841–2857 (2010).

13. Naumann, M. et al. Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nat. Commun.* 9, 335 (2018).

14. Zou, Z.-Y. et al. Screening of the FUS gene in familial and sporadic amyotrophic lateral sclerosis patients of Chinese origin. *Eur. J. Neurol.* 19, 977–983 (2012).

15. Dini Modigliani, S., Morlando, M., Errichelli, L., Sabatelli, M. & Bozzoni, I. An ALS-associated mutation in the FUS 3′-UTR disrupts a microRNA–FUS regulatory circuitry. *Nat. Commun.* 5, 4335 (2014).

16. Sabatelli, M. et al. Mutations in the 3′ untranslated region of FUS causing FUS overexpression are associated with amyotrophic lateral sclerosis. *Hum. Mol. Genet.* 22, 4748–4755 (2013).

17. Ajmone-Cat, M. A. et al. Increased FUS levels in astrocytes leads to astrocyte and microglia activation and neuronal death. *Sci. Rep.* 9, 4572 (2019).

18. Ju, S. et al. A Yeast Model of FUS/TLS-Dependent Cytotoxicity. *PLoS Biol.* 9, (2011).

19. Chen, Y. et al. Expression of human FUS protein in Drosophila leads to progressive neurodegeneration. *Protein Cell* 2, 477–486 (2011).

20. Miguel, L. et al. Accumulation of insoluble forms of FUS protein correlates with toxicity in Drosophila. *Neurobiol. Aging* 33, 1008.e1–15 (2012).

21. Ling, S.-C. et al. Overriding FUS autoregulation in mice triggers gain-of-toxic dysfunctions in RNA metabolism and autophagy-lysosome axis. *eLife* 8, e40811 (2019).
22. Nolan, M., Talbot, K. & Ansorge, O. Pathogenesis of FUS-associated ALS and FTD: insights from rodent models. *Acta Neuropathol. Commun.* **4**, (2016).

23. Samandi, S. *et al.* Deep transcriptome annotation enables the discovery and functional characterization of cryptic small proteins. *eLife* **6**, e27860 (2017).

24. Saghatelian, A. & Couso, J. P. Discovery and characterization of smORF-encoded bioactive polypeptides. *Nat. Chem. Biol.* **11**, 909–916 (2015).

25. Delcourt, V., Staskevicius, A., Salzet, M., Fournier, I. & Roucou, X. Small Proteins Encoded by Unannotated ORFs are Rising Stars of the Proteome, Confirming Shortcomings in Genome Annotations and Current Vision of an mRNA. *Proteomics* (2017) doi:10.1002/pmic.201700058.

26. Brunet, M. A., Levesque, S. A., Hunting, D. J., Cohen, A. A. & Roucou, X. Recognition of the polycistronic nature of human genes is critical to understanding the genotype-phenotype relationship. *Genome Res.* (2018) doi:10.1101/gr.230938.117.

27. Brunet, M. A. *et al.* OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes. *Nucleic Acids Res.* **47**, D403–D410 (2019).

28. Umoh, M. E. *et al.* A proteomic network approach across the ALS-FTD disease spectrum resolves clinical phenotypes and genetic vulnerability in human brain. *EMBO Mol. Med.* e201708202 (2017) doi:10.15252/emmm.201708202.

29. Collins, M. A., An, J., Hood, B. L., Conrads, T. P. & Bowser, R. P. Label-Free LC–MS/MS Proteomic Analysis of Cerebrospinal Fluid Identifies Protein/Pathway Alterations and Candidate Biomarkers for Amyotrophic Lateral Sclerosis. *J. Proteome Res.* **14**, 4486–4501 (2015).
Chick, J. M. et al. A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. Nat. Biotechnol. 33, 743–749 (2015).

Olexiouk, V., Van Criekinge, W. & Menschaert, G. An update on sORFs.org: a repository of small ORFs identified by ribosome profiling. Nucleic Acids Res. 46, D497–D502 (2018).

Zerbino, D. R. et al. Ensembl 2018. Nucleic Acids Res. 46, D754–D761 (2018).

Michel, A. M., Kiniry, S. J., O’Connor, P. B. F., Mullan, J. P. & Baranov, P. V. GWIPS-viz: 2018 update. Nucleic Acids Res. 46, D823–D830 (2018).

Pavesi, A. Asymmetric evolution in viral overlapping genes is a source of selective protein adaptation. Virology 532, 39–47 (2019).

Hall, C. E. et al. Progressive Motor Neuron Pathology and the Role of Astrocytes in a Human Stem Cell Model of VCP-Related ALS. Cell Rep. 19, 1739–1749 (2017).

Luisier, R. et al. Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS. Nat. Commun. 9, 1–15 (2018).

Carri, M. T., D’Ambrosi, N. & Cozzolino, M. Pathways to mitochondrial dysfunction in ALS pathogenesis. Biochem. Biophys. Res. Commun. 483, 1187–1193 (2017).

Deng, J. et al. FUS Interacts with HSP60 to Promote Mitochondrial Damage. PLOS Genet. 11, e1005357 (2015).

Nakaya, T. & Maragkakis, M. Amyotrophic Lateral Sclerosis associated FUS mutation shortens mitochondria and induces neurotoxicity. Sci. Rep. 8, 15575 (2018).

Monahan, Z., Shewmaker, F. & Pandey, U. B. Stress granules at the intersection of autophagy and ALS. Brain Res. 1649, 189–200 (2016).

Aulas, A. & Vande Velde, C. Alterations in stress granule dynamics driven by TDP-43 and FUS: a link to pathological inclusions in ALS? Front. Cell. Neurosci. 9, (2015).
42. Hergesheimer, R. C. et al. The debated toxic role of aggregated TDP-43 in amyotrophic lateral sclerosis: a resolution in sight? *Brain* **142**, 1176–1194 (2019).

43. Deng, H.-X. et al. FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. *Ann. Neurol.* **67**, 739–748 (2010).

44. Tyzack, G. E. et al. Widespread FUS mislocalization is a molecular hallmark of amyotrophic lateral sclerosis. *Brain J. Neurol.* (2019) doi:10.1093/brain/awz217.

45. Farrawell, N. E. et al. Distinct partitioning of ALS associated TDP-43, FUS and SOD1 mutants into cellular inclusions. *Sci. Rep.* **5**, 13416 (2015).

46. Lanson, N. A. et al. A Drosophila model of FUS-related neurodegeneration reveals genetic interaction between FUS and TDP-43. *Hum. Mol. Genet.* **20**, 2510–2523 (2011).

47. Bogaert, E. et al. Molecular Dissection of FUS Points at Synergistic Effect of Low-Complexity Domains in Toxicity. *Cell Rep.* **24**, 529-537.e4 (2018).

48. Chambers, R. P. et al. Nicotine increases lifespan and rescues olfactory and motor deficits in a Drosophila model of Parkinson’s disease. *Behav. Brain Res.* **253**, 95–102 (2013).

49. Kovacs, E., Tompa, P., Liliom, K. & Kalmar, L. Dual coding in alternative reading frames correlates with intrinsic protein disorder. *Proc. Natl. Acad. Sci.* **107**, 5429–5434 (2010).

50. Pancsa, R. & Tompa, P. Coding Regions of Intrinsic Disorder Accommodate Parallel Functions. *Trends Biochem. Sci.* **41**, 898–906 (2016).

51. Scekic-Zahirovic, J. et al. Toxic gain of function from mutant FUS protein is crucial to trigger cell autonomous motor neuron loss. *EMBO J.* **35**, 1077–1097 (2016).

52. Delcourt, V. et al. The Protein Coded by a Short Open Reading Frame, Not the Annotated Coding Sequence, Is the Main Gene Product of the Dual-Coding Gene MIEF1. *Mol. Cell. Proteomics MCP* **17**, 2402–2411 (2018).
53. Kino, Y. et al. FUS/TLS deficiency causes behavioral and pathological abnormalities distinct from amyotrophic lateral sclerosis. Acta Neuropathol. Commun. 3, (2015).

54. Hicks, G. G. et al. Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. Nat. Genet. 24, 175–179 (2000).

55. An, H. et al. ALS-linked FUS mutations confer loss and gain of function in the nucleus by promoting excessive formation of dysfunctional paraspeckles. Acta Neuropathol. Commun. 7, 7 (2019).

56. Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat. Rev. Neurol. 9, nrneurol.2013.203 (2013).

57. Qamar, S. et al. FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-π Interactions. Cell 173, 720-734.e15 (2018).

58. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. Off. J. Am. Coll. Med. Genet. 17, 405–424 (2015).

59. Bergeron, D. et al. An out-of-frame overlapping reading frame in the ataxin-1 coding sequence encodes a novel ataxin-1 interacting protein. J. Biol. Chem. 288, 21824–21835 (2013).

60. Vanderperre, B. et al. An overlapping reading frame in the PRNP gene encodes a novel polypeptide distinct from the prion protein. FASEB J. 25, 2373–2386 (2011).

61. Rathore, A. et al. MIEF1 Microprotein Regulates Mitochondrial Translation. Biochemistry 57, 5564–5575 (2018).
62. Peng, J.-Y. et al. Automatic Morphological Subtyping Reveals New Roles of Caspases in Mitochondrial Dynamics. *PLOS Comput. Biol.* 7, e1002212 (2011).

63. Choi, H. et al. SAINT: Probabilistic Scoring of Affinity Purification - Mass Spectrometry Data. *Nat. Methods* 8, 70–73 (2011).
**FIGURES' LEGEND**

**Figure 1: FUS is a bicistronic gene**

**a.** FUS gene bicistronic annotation, with the canonical FUS CDS (in blue, +1 frame) and altFUS CDS (in green, +2 frame) represented on the FUS canonical transcript (ENST00000254108 or NM_004960). Sequence length proportions are respected, the scale bar corresponds to 300 nucleotides.

**b.** FUS gene model, inspired from the GTEx portal, where exon combinations for the five most abundant transcripts in the brain are represented below (sorted by decreasing abundance). FUS CDS is coloured in blue, altFUS CDS is coloured in green, transcripts are identified with Ensembl names (FUS-202: ENST00000380244; FUS-206: ENST00000487509; FUS-211: ENST00000568901; FUS-203: ENST00000568901; FUS-201: ENST00000254108).

**c.** Ribo-seq data over the FUS gene (from the Gwips portal) with initiating ribosomes indicated in blue bars and elongating ribosome footprints indicated with the blue curve.

**d.** AltFUS protein conservation across Human (*Homo sapiens*), Chimpanzee (*Pan troglodytes*), Rat (*Rattus norvegicus*), Mouse (*Mus musculus*) and Dog (*Canis lupus familiaris*). Residues are coloured based on their degree of conservation from white (none) to red (in all species).

**e.** Mass spectrometry-based evidence for altFUS protein with detected peptides indicated in grey and numbered according to Extended Data Table 1. The peptide corresponding to the displayed MS/MS spectra below is indicated in red (from re-analysis of PMID 26186194 by OpenProt).

**f.** Expression of both FUS and altFUS from the FUS construct in HEK293 cells by western blot, and expression of FUS with the monocistronic construct FUS(o) (representative image from n=3). The western blot shows a low exposition (endogenous FUS and altFUS non visible).

**g.** AltFUS endogenous expression in HEK293 cells using a siRNA targeting FUS mRNA as negative control and over-expression of altFUS CDS as positive control (representative image from n=3). The asterisk indicates a protein species detected with the anti-altFUS antibody specifically in the brain.

**h.** AltFUS (arrow) endogenous expression in Human tissues (spinal cord, brain, muscles and kidney – 100 ug), in HEK293 and HeLa cultured cells (100 ug) and using the over-expression of altFUS CDS in HEK293 cells (50 ug) as positive control (representative image from n=3). The asterisk indicates a protein species detected with the anti-altFUS antibody specifically in the brain.

**i, j.** AltFUS (arrow) endogenous expression in the motor cortex of three C9orf72 and three sporadic ALS patients (i) or in iPSC-derived motor neurons of three lines from controls and from ALS patients (j) (representative image from n=3). The asterisk indicates a protein species detected with the anti-altFUS antibody specifically in the brain.
Figure 2: AltFUS is a mitochondrial protein involved in ALS-associated mitochondrial dysfunction

a, Images by confocal microscopy of altFUS-FLAG (green) in HeLa cells, using TOMM20 (red) as a mitochondrial marker (representative image from n=3, Pearson’s correlation r=0.92). The white scale bar corresponds to 10μm.  
b, AltFUS-FLAG enrichment in mitochondrial extracts from transfected HEK293 cells (representative image from n=3) with Hsp70 used as a mitochondrial marker (WCL = Whole Cell Lysate, Mito = Mitochondria).  
c, AltFUS mitochondrial expression in transfected HEK293 cells following fractionation (representative image from n=3), with Tubulin as a cytosolic fraction marker and VDAC as a mitochondrial fraction marker (WCL = Whole Cell Lysate, Cyto = Cytosol, Mito = mitochondria).  
d, Endogenous altFUS mitochondrial expression in HEK293 cells following fractionation (representative image from n=3), with Tubulin as a cytosolic fraction marker and VDAC as a mitochondrial fraction marker, and with siFUS transfected cells as a negative control and altFUS transfected cells as a positive control for altFUS expression (WCL = Whole Cell Lysate, Cyto = Cytosol, Mito = mitochondria).  
e, Representative images of the mitochondrial networks (TOMM20 in red) in mock and altFUS-FLAG (green) transfected HeLa cells (n=3). The white scale bar corresponds to 10μm.  
f, Proportion of tubules and globules in the mitochondrial network of mock and altFUS-FLAG transfected HeLa cells over a minimum of 100 cells across a technical duplicate per independent experiments (n=3, i.e. a minimum of 300 cells per biological conditions, p value < 0.001, Mann-Whitney U test).  
g, Representative traces of TMRE fluorescence measured by flow cytometry in mock cells and cells overexpressing the bicistronic FUS-R495x or the monocistronic FUS(Ø)-R495x constructs (n=4, minimum of 50 000 live cells per independent replicates). Mean fluorescence intensity of mock cells treated with the decoupling agent, FCCP, is indicated with a grey dotted line.  
h, Mean TMRE fluorescence intensity measures in mock cells, cells over-expressing altFUS, FUS, FUS(Ø), FUS-R495x or FUS(Ø)-R495x, or mock cells treated with FCCP across 3 independent experiments (n=4). Statistical significance is relative to the mock condition unless otherwise indicated (*** = p value < 0.001, **** = p value < 0.0001, ns = non-significant).  
i, Representative image by stimulated emission depletion microscopy (STED) of altFUS-FLAG (green) localization within mitochondria (TOMM20 marker in red). The white bar across the mitochondria represents the region of interest quantified in panel j. The white scale bar corresponds to 3μm.  
j, Relative fluorescence histogram for altFUS-FLAG and TOMM20 across the region of interest highlighted by a white line on panel i.  
k, AltFUS expression in fractions following size exclusion chromatography on HEK293 cells mitochondrial extracts (representative image from n=2).  
l, Predicted localisations of proteins identified in altFUS-FLAG affinity purification-mass spectrometry (AP-MS) from fractions 8 to 14 after size exclusion chromatography (k) on HEK293 cells mitochondrial extracts (n=2, confident identifications with a SAINT score > 0.85 across 2 independent experiments).  
m, Enrichment of biological processes in altFUS interacting proteins compared to the Human mitochondrial proteome (n=2, Fisher’s Exact test with FDR < 0.1 %).
Figure 3: AltFUS is necessary for ALS-associated inhibition of autophagy and accumulation of FUS/TDP-43 cytoplasmic aggregates

a, Images by confocal microscopy of mCherry-GFP-LC3 signal in HeLa cells across biological conditions: untreated mock, bafilomycin treated mock, altFUS, FUS, FUS(Ø), FUS-R495x and FUS(Ø)-R495x (representative images of n=3). The white scale bar corresponds to 10μm and the zoomed in region (right panel) is delimited as a white box. b, LC3-II accumulation after bafilomycin treatment from mock, altFUS, FUS, FUS(Ø), FUS-R495x, FUS(Ø)-R495x transfected cells and FUS(Ø)-R495x and altFUS co-transfected cells across 3 independent experiments (n=3). The quantification corresponds to the treated/untreated ratio of LC3-II abundance. Statistical significance is relative to the mock condition unless otherwise indicated (**** = p value < 0.0001, *** = p value < 0.001, ** = p value < 0.01, ns = non-significant). c, Images by confocal microscopy of altFUS (FLAG tagged-white), FUS (GFP tagged - green) and TDP-43 (red) signals in HeLa cells transfected with the bicistronic GFP-FUS(FLAG)-R495x or the monocistronic GFP-FUS(Ø-FLAG)-R495x constructs, or co-transfected with the monocistronic GFP-FUS(Ø-FLAG)-R495x and altFUS-FLAG constructs (representative images from n=3). The white scale bar corresponds to 10μm. d-e, Quantification of FUS cytoplasmic granules, number (d) and area (μm²) (e) in cells over-expressing the bicistronic (+) or monocistronic (-) construct for FUS, FUS-G156E, FUS-R495x, FUS-K510E, FUS-Q519x, FUS-Q519I-fs527x, FUS-R521C, and FUS-P525L. Statistical comparisons are made between bicistronic and monocistronic versions of each construct (**** = p value < 0.0001, *** = p value < 0.001, ** = p value < 0.01, * = p value < 0.05, ns = non-significant, one-way ANOVA test with Sidak’s multiple comparison).
Figure 4: AltFUS expression is necessary for the full FUS-linked ALS phenotype in *Drosophila*

a. Cross-breeding strategy for *Drosophila* generation using the Elav-GeneSwitch-GAL driver as an inducible expression system specific to the motoneurons. 
b. FUS and altFUS expression in mCherry (control), altFUS, FUS, FUS(Ø), FUS-R495x or FUS(Ø)-R495x expressing *Drosophila* from the control F1 and the RU-486 treated F1 (a) at 1, 10 or 20 days post-induction (representative image from n=2). 
c-e. Locomotion assay using percentage of climbing success in control and RU-486-treated transgenic *Drosophila* expressing mCherry or altFUS (c), the bicistronic FUS or the monocistronic FUS(Ø) (d), and the bicistronic FUS-R495x or the monocistronic FUS(Ø)-R495x (e) at day 1, 10 and 20 post-induction. Statistical comparison were made between each population (n=4). Indicated significance are between the monocistronic and the bicistronic transgenic flies of the RU-486-treated population (ns = non-significant, * = p value < 0.05, *** = p value < 0.001).
**Figure 5:** FUS mutations, synonymous for FUS but missense for altFUS, potentiate TDP-43 cytoplasmic aggregation

**a,** Graphical representation of FUS mutations found in ALS patients, that are synonymous (dark blue) for FUS protein (blue) but missense (red) for altFUS protein (green), on the canonical FUS mRNA (ENST00000254108 or NM_004960). **b,** Images by confocal microscopy of altFUS (FLAG tagged - green) and mitochondria (TOMM20 marker, red) in HeLa cells over-expressing altFUS-FLAG, altFUS-P31L-FLAG, altFUS-A38V-FLAG, altFUS-A46V-FLAG or altFUS-R64P-FLAG constructs (representative images from n=3). Deconvolution over a maximum of 30 iterations on the green and red channels was performed for the zoomed in pictures (the zoomed in regions are highlighted in white). The white scale bar corresponds to 10μm. **c,** Images by confocal microscopy of TDP-43 (yellow), FUS (GFP tagged - green) and altFUS (FLAG tagged - red) in HeLa cells over-expressing GFP-FUS(FLAG), GFP-FUS(P31L-FLAG)-S44=, GFP-FUS(A38V-FLAG)-G51=, GFP-FUS(A46V-FLAG)-G59= or GFP-FUS(R64P-FLAG)-S77= constructs (representative images from n=3). The white scale bar corresponds to 10μm. **d,** Quantification of cells with TDP-43 aggregates in HeLa cells from (c), where the data are represented as the fold-change compared to the GFP-FUS(FLAG) expressing cells. Statistical significance is indicated above the bars (n=3, **** = p value < 0.0001).
Extended Data - Legend

For: FUS gene is dual-coding with both proteins united in molecular hallmarks of amyotrophic lateral sclerosis, Brunet MA, et al.

Extended Data Table 1

OpenProt search results for FUS transcripts ENST00000254108 and NM_004960.3. The OpenProt version 1.3 was used. Each protein prediction is identified with a unique protein accession (starting with II_ for novel predicted isoform, and IP_ for predicted alternative proteins). All predicted alternative proteins are coloured in light blue. The protein localization is relative to the canonical FUS coding sequence (CDS). The furthered studied protein, named altFUS (IP_243680) is highlighted in dark blue. The workbook also contains the list of peptides identified in re-analysis of mass spectrometry experiments using the OpenProt database (version 1.3).

Extended Data Table 2

AltFUS conservation across vertebrates. The workbook contains the list of FUS transcripts retrieved from NCBI RefSeq database (transcript accession and nucleotidic sequences); the list of transcripts containing an altFUS sequence (transcript accession, species, and altFUS protein sequence); and the matrix of altFUS protein sequence identity between each species.

Extended Data Alignment File

Alignment of altFUS protein sequences across all 82 species. The alignment was done using Clustalω, and coloured based on conservation from white (highly variable residue) to dark blue (highly conserved residue).

Extended Data Table 3

Proteins identified by mass spectrometry following size exclusion chromatography (FPLC) and FLAG affinity purification on mitochondrial extracts from HEK293 cells over-expressing altFUS-FLAG. The proteins are listed by their UniProtKB accession alongside their annotated function, the number of unique peptides supporting the detection and the number of peptide spectrum matches (PSMs) within each experimental condition. Confidence of the interaction was scored using the SAINT algorithm. The SAINT score is also reported.

Extended Data Table 4

FUS mutations, synonymous for FUS but missense for altFUS. The first sheet contains mutations identified from a manually curated literature review. The second sheet retrieves variants in sALS cases from the ALS Variant Server (http://als.umassmed.edu/). The third sheet retrieves variants in fALS cases from the ALS Variant Server.

Extended Data Table 5

16 ALS-associated genes with currently non-annotated proteins detected in OpenProt database. The OpenProt version 1.3 was used. The workbook contains the list of novel predicted proteins from these genes, which have already be detected by mass spectrometry or ribosome profiling.
Extended Data Figure 1

a, AltFUS is encoded within an alternative open reading frame (ORF) overlapping the FUS canonical CDS. When read in the +1 frame, the FUS mRNA (here NM_004960.3) codes for a 526 amino acid protein (highlighted in blue), named FUS. In the +2 frame, the FUS mRNA contains a second ORF (highlighted in green) that codes for a novel 170 amino acid protein, named altFUS. The two proteins are not isoforms. b, GTEx portal data on FUS mRNA expression in ALS-related tissues are shown in blue colored scale. Transcripts are identified with Ensembl accessions (the number after the dot corresponds to the version used in the analysis) and are quantified across 14 tissues. Five transcripts share 85 % of all mRNA expression level (framed in red). c, Table compiling protein information relayed by Ensembl and OpenProt databases for the aforementioned transcripts. The FUS protein is highlighted in blue. The altFUS protein is highlighted in green. d, Ribo-seq data from the Gwips portal in Mus musculus shows an accumulation of initiating ribosomes at the FUS methionine and the altFUS methionine (green bars). Elongating ribosomes footprints density increase in the bicistronic region (light green curve). Data corresponds to global aggregates across the studies supported by Gwips. e, PhyloP nucleotidic conservation scores are represented in grey across the FUS mRNA (FUS-201), and in blue after noise reduction (Fast Fourier Transformation). The average PhyloP score on the bicistronic and the monocistronic region are represented as a dotted red line. The FUS CDS is represented by a blue box, the altFUS CDS is represented by the green box. f, Heatmap of altFUS protein sequences identity across 84 species. Primates, Rodents and Mammals in general display a strong protein conservation. The sequence identity is colored from blue (0 %) to red (100 %).
Extended Data Figure 2

a, AltFUS custom antibody strategy, with epitopes highlighted in bold on altFUS sequence (green), and ELISA test results on both bleeds from each rabbits. b, Nucleotidic mutations on FUS mRNA (NM_004960) are highlighted in red and mutate all altFUS methionines to threonines (altFUSØ).

c-d, Protein alignment of altFUS (c) and FUS (d) proteins from FUS bicistronic construct and the monocistronic (FUSØ) construct, residues are coloured based on sequence identity (dark blue for identical residues).
Mutated sequence
ACGGGCTTACCCCAACCCCGGCCAGCCAGGCTATTTCCAGCACAGCAGTCAGCCCTTACGAGCAGTTACAGGTGTTATAGCG
CAGTTCCACCGGACCTTTCGATACCGGCGAACAGCAAGGCTTCTTTAGCAGCCAGCGCAGACACACAGCTGACAACTCAGTCAAC
TCCCGAGGGACAGCTGCAGCTGAGGCTACGCAGCTACGGATACGGCATGAGCCAGACAGCCTTCCCATCAGGGGACACCGCTTCTTT
CAGGCAGCCACGCACTCCTCAACGACAACTCGGCAGCGACACCGCTCTCTTCAGTCTGCAGCGACAGGCACAGTCTAGCGACAGCCAGCAC
AGCTACACGCAAGCAGCTAGCTACGGGCAACAGCAACAGTCAATTCCCTACGAGGCGCTGACAGCTAGCTACGGACACAGGACCAAGTTGA
GAGTAGTGGTGAGGTGAGGTGAGTGAGTACAGCGCCAGCATACAGTCAAC

altFUS

1. MGYPTPSGPIAISRVAAPTSDVRVTVIASPRTLQAMARAAILLLMARARTQAMELSQLPR
   60
2. TGYPTPSGPIAISRVAAPTSDVRVTVIASPRTLQAMARAAILLLMARARTQAMELSQLPR
   60
3. DMRZAAMVARAPNRTGSSPTLTAMASSQLPAAPREVTVAVLRAAMGRSPVRGATASSS
   120
4. DTARLATVARAPNRTGSSPTLTATASSQLPAAPREVTVAVLRAAMGRSPVRGATASSS
   120
5. LAYDSSKAMDSSKAILPRAMDSTRTSTAATVVEEVEEVETTAKINPP
   170
6. LATDSSKATDSSKAILPRAMDSTRTSTAATVVEEVEEVETTAKINPP
   170

FUS

1. MASNTDYTQATQSYGAYPTQPGGYSQSQSQPYGQQSYGYSQSTDSYGGQSYSSYGG
   60
2. MASNTDYTQATQSYGAYPTQPGGYSQSQSQPYGQQSYGYSQSTDSYGGQSYSSYGG
   60
3. SQNTGYQTSTPQQYGGTSQGGSSSQQSSYQQYGGQAPPSSTSGSYGSSQQSQSQSQ
   120
4. SQNTGYQTSTPQQYGGTSQGGSSSQQSSYQQYGGQAPPSSTSGSYGSSQQSQSQSQ
   120
5. SYGGPQGQSSYSQQPYSYGGQQQSYQYGGQQQYNPPQYGQONQNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS
**Extended Data Figure 3**

a, Object image from MicroP processing of images of mock and altFUS transfected HeLa cells. Tubules are coloured in purple, globules are coloured in green. b, Graphical representation of bicistronic constructs FUS and FUS-R495x (mutation in red), and their monocistronic equivalent FUS(Ø) and FUS(Ø)-R495x. AltFUS CDS is highlighted in green when present, FUS CDS is highlighted in blue. c, V5-FUS, V5-FUS-R495x and altFUS-FLAG expression from bicistronic or monocistronic constructs in HEK293 cells (representative image from n=3). d, Representative traces of TMRE fluorescence measured by flow cytometry in mock cells and cells overexpressing the bicistronic constructs FUS or FUS-R495x (in red) in the top panels, and the monocistronic constructs, FUS(Ø) or FUS(Ø)-R495x (in blue) in the bottom panels (n=3, minimum of 50,000 live cells per independent replicates).
Extended Data Figure 4

a, Graphical representation of the mCherry-GFP-LC3 construct to study the autophagic flux. b, Images by confocal microscopy of mCherry-GFP-LC3 and altFUS (FLAG, in blue) signal in HeLa cells across biological conditions: untreated mock, bafilomycin treated mock, altFUS, FUS, FUS\(^{(\beta, \text{FLA})}\), FUS-R495x and FUS\(^{(\beta, \text{FLA})}.R495x\) (representative images of n=3). The white scale bar corresponds to 10μm and the zoomed in region (right panel) is delimited as a white box. c, LC3-II expression from untreated (-) or bafilomycin treated (+) cells transfected with mock, altFUS, FUS, FUS\(^{(\beta)}\), FUS-R495x, FUS\(^{(\beta, \text{FLA})}.R495x\), or co-transfected with FUS\(^{(\beta, \text{FLA})}.R495x\) and altFUS. d-f, Images by confocal microscopy of FLAG (white), GFP (green) and TDP-43 (red) signals in HeLa cells transfected with the bicistronic constructs (d), monocistronic constructs (e) or co-transfected with altFUS-FLAG and the monocistronic constructs (f) of 6 ALS-associated mutants: FUS-G156E, FUS-K510E, FUS-Q519x, FUS-Q519I-fs527x, FUS-R521C, and FUS-P525L (representative images from n=3). The white scale bar corresponds to 10μm.
Extended Data Figure 5

a, Images by confocal microscopy of FLAG (white), GFP (green) and TIA-1 (red) signals in HeLa cells transfected with the bicistronic GFP-FUS(FLAG)-R495x or the monocistronic GFP-FUS(Ø-FLAG)-R495x constructs, or co-transfected with the monocistronic GFP-FUS(Ø-FLAG)-R495x and altFUS-FLAG constructs (representative images from n=3). The white scale bar corresponds to 10μm. b-d, Images by confocal microscopy of FLAG (white), GFP (green) and TIA-1 (red) signals in HeLa cells transfected with the bicistronic constructs (b), monocistronic constructs (c) or co-transfected with altFUS-FLAG and the monocistronic constructs (d) of 6 ALS-associated mutants: FUS-G156E, FUS-K510E, FUS-Q519x, FUS-Q519I-fs527x, FUS-R521C, and FUS-P525L (representative images from n=3). The white scale bar corresponds to 10μm.
