ALS-associated FUS mutation reshapes the RNA and protein composition of stress granules

Davide Mariani [1,2,†], Adriano Setti[2,†], Francesco Castagnetti [1], Erika Vitiello [1],
Lorenzo Stufera Mecarelli[1,2], Gaia Di Timoteo[2], Andrea Giuliani[2], Angelo D'Angelo[2],
Tiziana Santini[2], Eleonora Perego[1], Sabrina Zappone[1], Nara Liessi[3], Andrea Armirotti[3],
Giuseppe Vicidomini[1] and Irene Bozzoni [1,2,4,*]

[1]Center for Human Technologies, Istituto Italiano di Tecnologia, Via Enrico Melen 83, 16153, Genoa, Italy
[2]Department of Biology and Biotechnologies "C. Darwin", Sapienza University of Rome, Piazzale Aldo Moro 5, 00185, Rome, Italy
[3]Analytical Chemistry Lab, Istituto Italiano di Tecnologia, Via Morego 30, 16163, Genoa, Italy
[4]Center for Life Nano-& Neuro-Science, Fondazione Istituto Italiano di Tecnologia, Viale Regina Elena 291, 00161, Rome, Italy

*To whom correspondence should be addressed. Tel: +39 0649912202; Email: irene.bozzoni@uniroma1.it
†The first two authors should be regarded as Joint First Authors.

Abstract

Stress granules (SG) are part of a cellular protection mechanism where untranslated messenger RNAs and RNA-binding proteins are stored upon conditions of cellular stress. Compositional variations due to qualitative or quantitative protein changes can disrupt their functionality and alter their structure. This is the case of different forms of amyotrophic lateral sclerosis (ALS) where a causative link has been proposed between the cytoplasmic de-localization of mutant proteins, such as FUS (Fused in Sarcoma), and the formation of cytotoxic inclusions. Here, we describe the SG transcriptome in neuroblastoma cells and define several features for RNA recruitment in these condensates. We demonstrate that SG dynamics and RNA content are strongly modified by the incorporation of mutant FUS, switching to a more unstructured, AU-rich SG transcriptome. Moreover, we show that mutant FUS, together with its protein interactors and their target RNAs, are responsible for the reshaping of the mutant SG transcriptome with alterations that can be linked to neurodegeneration. Our data describe the molecular differences between physiological and pathological SG in ALS-FUS conditions, showing how FUS mutations impact the RNA and protein composition of these condensates.

Graphical abstract
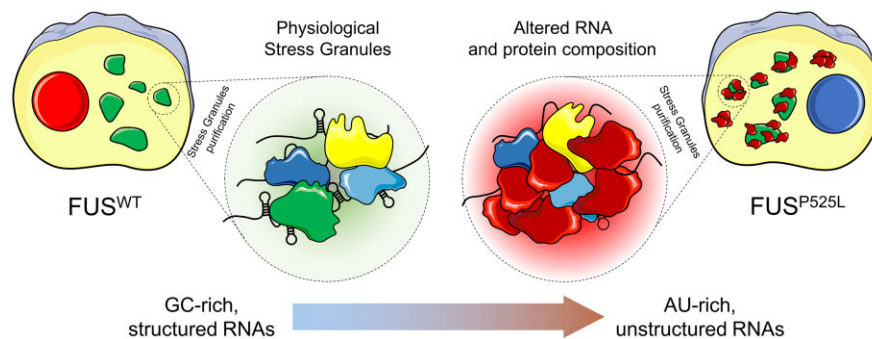
Introduction

A well-defined intracellular compartmentalization is vital for eukaryotic cells, with membrane-bound organelles physically separating specialized molecular processes from the rest of the cellular environment. Alongside these structures, a large number of membrane-less organelles (MLOs) has been identified in recent years as crucial in the regulation of gene expression and metabolic activity of cells (1–3). Lacking a lipid bilayer, these assemblies are separated from the neighboring cytoplasm or nucleoplasm by liquid–liquid demixing, conferring them stability, but also dynamic properties of exchange with the molecular surroundings (4–6). The liquid-like behavior of MLOs, mainly constituted by diverse RNAs and RNA-binding proteins (RBPs), results from the summation of several weak interactions involving its constituents (7).

RNA is a pivotal component of such condensates, since the intrinsic characteristics of individual transcripts, such as length, structure and nucleotide composition, are sufficient

to influence the number of RNA–RNA and RNA–protein interactions within a granule, thereby affecting its physical properties (8–11). Moreover, it has been observed that the presence of intrinsically disordered regions (IDRs) in RBPs is a molecular determinant of their tendency to phase-separate, both *in vitro* and in living cells (12); IDRs are often coupled to canonical RNA binding domains within the same protein, providing the foundation for establishing multivalent interactions with RNAs and other proteins, thus leading to self-assembling, separated granules (13,14).

The formation of cytoplasmic stress granules (SG) in response to variations in cell homeostasis stimuli is a primordial protection mechanism (15). SG result from the condensation of untranslated, stalled messenger RNAs (mRNAs) and several IDR-containing RBPs, functioning as a temporary reservoir and being readily disassembled as protein translation is reactivated after stress (16). SG assembly is a multi-step event, where initial phase separation of individual ribonucleoparticles leads to the coalescence of stable *cores*, surrounded by a liquid, quickly remodeled *shell* (17).

Post-mitotic cells like neurons are highly susceptible to dysregulation of SG assembly and disassembly (18–20). Mutations in numerous protein components of these condensates are causative of the formation of permanent, insoluble protein–RNA assemblies; these aggregates have been causatively linked to the pathogenesis of neurodegenerative disorders and constitute a hallmark of several motor-neuron diseases, such as amyotrophic lateral sclerosis (ALS) (21,22).

A paradigmatic example is represented by FUS/TLS (Fused in Sarcoma/Translocated in Liposarcoma), whose mutations account for 4% of familial ALS cases (23). The protein embeds a IDR which is responsible for self-assembly and spontaneous aggregation (24–26); ALS-linked mutations are clustered in the nuclear localization sequence (NLS), reducing its nuclear reimport and resulting in an aberrant cytoplasmic accumulation and subsequent recruitment in SG (27,28). As mutant FUS accumulates in the cytoplasm of motor-neurons, it triggers the activation of major stress pathways and impacts SG number and dynamics, finally leading to an increased vulnerability of these cells to harmful stimuli (29).

In this work, we focused on the RNA content of SG in a neuronal-like system, demonstrating that neuronal RNAs are specifically enriched in these condensates in response to oxidative stress. We show that the expression of the mislocalized FUS[P525L] is sufficient to impact the dynamics of SG and to strongly modify their transcriptome by enriching AU-rich RNAs, at the expense of GC-rich transcripts. The modification of the nucleotide composition caused by the presence of mutant FUS reduces the global structuration propensity of SG RNAs, by recruiting poorly structured transcripts. Moreover, we show that the reorganization of the SG transcriptome in FUS[P525L] condition affects specific classes of RNAs linked to neurodegenerative processes. Finally, we show that the presence of FUS[P525L] also impacts on the SG proteome; thus, while the direct binding of mutant FUS to RNA is not sufficient to explain the altered engagement of transcripts into SG, the differential recruitment of FUS-interacting RBPs contributes to the aberrant RNA composition of pathological granules.

Materials and methods

Cell culture, maintenance and treatment

SK-N-BE were cultured in adherence in RPMI-1640 supplemented with 10% Fetal Bovine Serum (FBS), 1 mM sodium pyruvate, $1\times$ GlutaMAX[TM], 100 U/ml penicillin and 100 µg/ml streptomycin (all reagents were purchased from Thermo Fisher Scientific). Cells were frozen in culture medium supplemented with 20% FBS and 10% Dimethyl Sulfoxide (DMSO, Sigma Aldrich), and quickly thawed when needed in a 37°C water bath. To induce FUS overexpression, cells were incubated with 50 ng/ml doxycycline hyclate (Sigma Aldrich) in culture medium for 24 h. To induce SG formation, cells were incubated with 0, 5 mM sodium arsenite (Sigma Aldrich) at 37°C for 1 h; the effective condensation of GFP-G3BP1 was inspected under a fluorescence microscope before cell collection.

Cloning and SK-N-BE stable cell line generation

The plasmid containing an N-terminal GFP-tagged form of human G3BP1 was a gentle gift from Roy Parker's Lab (University of Colorado Boulder). The GFP-G3BP1 sequence was subcloned in the ePB-BSD-EIF1a vector (30); briefly, GFP-G3BP1 was PCR-amplified with CloneAmp HiFi PCR Premix (TakaraBio) and cloned in the acceptor plasmid linearized with XbaI and NotI restriction enzymes (Thermo Fisher Scientific) using the InFusion Cloning kit following manufacturer protocol (TakaraBio); the NotI site was not maintained to reduce the Tm of the cloning primers. Primer sequences are listed in Supplementary Table S5.

Stable cell lines were generated taking advantage of the transposon-based PiggyBac technology. A PiggyBac HyPB transposase expression plasmid (31) was co-transfected with the ePB-EIF1a-GFP-G3BP1 vector with a 1:10 ratio; cells were selected in 10 µg/ml of blasticidin S (Thermo Fisher Scientific) for 7 days and then tested for consistent expression of the fluorescent protein. SK-N-BE cell lines expressing FUS[WT] and FUS[P525L] under a doxycycline-inducible promoter were already present in the lab (32).

RNA extraction and analysis

RNA extraction was performed with the Direct-zol Miniprep RNA Purification Kit (Zymo Research) with on-column DNase treatment, according to the manufacturer's instructions, eluting in 25 µl of RNase-free water (50 µl for total RNA Input samples). Total RNA was retro-transcribed with PrimeScript™ RT Reagent Kit (TakaraBio) according to the manufacturer's instructions. For quantitative PCR (qPCR) analysis, the PowerUp SYBR Green Master Mix (Thermo Fisher Scientific), following the manufacturer's protocol, in 15 µl reactions, on a QuantStudio5 qPCR system (Thermo Fisher Scientific) with standard cycling conditions. For SG purification experiments, data were expressed as % of input: briefly, the Ct of Input samples for each target were adjusted to account for the dilution by subtracting the log2 of the dilution factor. Then, the % of input was calculated as $2(\text{adjusted input Ct} - \text{IP Ct})$ for each target. The sequences of primers used for qRT-PCR are listed in Supplementary Table S5.

Protein extraction and western blot analysis

Protein extracts were obtained using RIPA buffer, supplied with $1\times$ Complete Protease Inhibitor Cocktail (Merck). Protein concentration was assessed using BCA protein assay (Thermo Fisher Scientific). Protein electrophoresis was performed using 4–15% Mini-PROTEAN TGX Precast acrylamide gel (Bio-Rad), and proteins were transferred to 0.45 µm nitrocellulose membrane, using the TransBlot Turbo

System (Bio-Rad), according to the manufacturer's instructions. Membranes were stained with Ponceau, blocked with 5% non-fat dry milk (Sigma Aldrich) for 1 h and incubated overnight at 4°C with the following primary antibodies: anti-GFP (Thermo Fisher Scientific, A-11122, 1:1000), anti-TIAR (BD Transduction, 610352, 1:1000), anti-FUS (SantaCruz Biotech, sc-47711, 1:2000), anti-GAPDH-HRP (SantaCruz Biotech, sc-47724, 1:1000), anti-FLAG (Merck, F3165, 1:1000), anti-ELAVL1 (SantaCruz Biotech, sc-5261, 1:1000), anti-DHX36 (Proteintech, 13159–1-AP, 1:500) and anti-βActin (SantaCruz Biotech, sc-69879, 1:1000). After secondary antibody incubation for 1 h at room temperature (RT), protein detection was carried out with Immobilon Crescendo Western HRP substrate (Merck) using ChemiDoc™ MP System and images were analyzed using Image Lab™ Software (Bio-Rad).

Puromycilation assay

FUS expression was induced in SK-N-BE cells as stated above. After 1 h of stress with 0, 5 mM sodium arsenite, the medium was removed and cells were washed twice in warm medium. Cells were then incubated in culture medium supplemented with 10 μg/ml puromycin (Merck, P4512) for 0, 3 or 4 h. After the incubation, cells were washed twice in phosphate-buffered saline (PBS), then lysed in RIPA buffer and sonicated to optimize protein solubilization. After protein quantification with BCA and normalization of the lysates, 15 μg of each sample were subjected to western blot. Puromycin signal was detected using an anti-puromycin antibody (Merck, MABE343, 1:2000). Total protein and puromycin signals were quantified using Image Lab™ Software (Bio-Rad).

Immunofluorescence and fluorescence correlation spectroscopy analyses

Cells were plated on glass coverslips and treated as previously described. Samples were fixed in 4% paraformaldehyde for 10 min at RT, followed by three washes in complete PBS. Permeabilization and blocking were performed by incubating the coverslips in PBS supplemented with 2% bovine serum albumin (BSA), 0.2% Triton X-100 for 30 min at RT. Samples were then incubated with primary antibody diluted in PBS with 1% BSA for 1 h at RT, followed by three washes for 5 min in PBS. The following antibodies were used: anti-G3BP1 (Cell Signaling, 17798, 1:100), anti-FLAG (Merck, F3165, 1:500), anti-ELAVL1 (SantaCruz Biotech, sc-5261, 1:150), anti-DHX36 (Proteintech, 13159–1-AP, 1:250), anti-TDP43 (Abcam, ab109535, 1:100), anti-TIAR (BD Transduction, 610352, 1:200). Fluorophore-conjugated secondary antibodies (Thermo Fisher Scientific) were diluted 1:300 in PBS with 1% goat/donkey serum and incubated for 45 min at RT. After three PBS washes, nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI, Sigma; diluted 1ug/ml in PBS) for 3 min at RT and coverslips were mounted with Prolong Diamond Mounting Media (Thermo Fisher Scientific, *P*-36961) on microscope slides. Fixed samples were imaged on a Nikon Instrument A1 Confocal Laser Microscope equipped with 60× and 100× objectives. Confocal images were collected with NIS-Elements AR software (Nikon): ND Acquisition module was used for multipoint acquisition of Z-stack images (250 nm Z-spacing). Images were analyzed with FIJI-ImageJ open-source software. Briefly, after Gaussian blur filtering and background removal, Pearson correlation coeffi-

cient was calculated launching the JaCoP plugin on the Maximun Intensity Projections of Z-planes. Fluorescence intensity was measured by creating a mask of the G3BP1 granules (thresholded according to the Otsu algorithm) and obtaining the mean gray value of the ELAVL1, DHX36 or TDP43 signal for each granule.

Fluorescence correlation spectroscopy (FCS) measurements were performed as described in (33,34). Cells were plated and FUS expression was induced as previously described; before starting the acquisition, culture medium was replaced with FluoroBrite™ DMEM (Thermo Fisher Scientific) supplemented with 10% FBS, and sodium arsenite was added at a final concentration of 0.5 mM for the 'stressed' samples. Acquisition of GFP/RFP fluorescence signals was obtained in a custom-built laser scanning microscope equipped with a single-photon array detector from the Molecular Microscopy and Spectroscopy division of the Italian Institute of Technology in Genoa. A variant of conventional FCS (spot-variation FCS) was employed to quantify the diffusion mode of GFP-G3BP1. In spot-variation FCS, multiple diffusion times are measured at different detection volume sizes. Thanks to the spatial information granted by the single-photon array detector (33,35,36), spot-variation FCS can be implemented within single measurements, by integrating off-line the signal coming from different subsets of the pixel detector, such as the central pixel only, the sum of the inner 3×3 pixels, and the sum of the entire 5×5 pixels. The diffusion modes are evaluated by fitting a linear function (diffusion law) to the diffusion times against the area of the detection volumes (37,38). When the intercept of the linear function is zero, the molecules freely diffuse in the detection area; when the value is higher than zero, the molecules are isolated in domains. Conversely, when it is lower than zero, the molecules are constrained inside a meshwork.

Single-molecule fluorescence *in situ* hybridization (smFISH) analysis

FISH probes for NORAD, UNC5B, SMARCA5, CLSTN1 and MBNL1 transcripts (25 for each target) were designed by utilizing the Stellaris® RNA FISH Probe Designer (Biosearch Technologies) and labelled with ddUTP-ATTO565 (72464–5MG-F, Merck) following the protocol from Gaspar *et al.* (39). Otherwise, FISH probes for GAPDH transcript [labeled with Quasar 570 Dye (SMF-2026–1)] were purchased from LGC Biosearch Technologies. Hybridization steps were performed following the protocol from Santini *et al.* (40), followed by immunofluorescence for G3BP1 (Cell Signaling, 17798, 1:100). After DAPI staining, coverslips were permanently mounted using ProLong Diamond Antifade Mountant (*P*-36961, Thermo Fisher Scientific). Samples were imaged on a Nikon Instrument A1 Confocal Laser Microscope equipped with a 1.49 NA 100× objective (Apo TIRF 100× Oil, Nikon, Tokyo, Japan). Confocal images were collected with NIS-Elements AR software (Nikon): ND Acquisition module was used for multipoint acquisition of Z-stack images (250 nm Z-spacing). FIJI-ImageJ opensource software was used for image analysis. Transcript enrichments in G3BP granules were computed on Maximum Intensity Projection of Z-planes after subtracting the FISH signals colocalizing with the nucleus. FISH spots were identified and quantified using RS-FISH plugin (41) over the entire microphotograph field.

SG purification

SG purification for the RNA sequencing experiments was performed according to (42), with minor adjustments. After lysis, the protein concentration of samples was measured through BCA assay (Thermo Fisher Scientific); 3 mg of lysate were diluted in 1 ml of SG Lysis buffer and used for the purification, saving 5% as Total RNA input (stored in TRIZOL L/S, Thermo Fisher Scientific). SG were pelleted as in (42), while the immunoprecipitation step was performed by incubation with antibody-coupled beads. For each sample, 50 μl of DEPC-treated Protein G Dynabeads (Thermo Fisher Scientific) were incubated with 12.5 μg of anti-GFP antibody in rotation at 4°C for 1 h, then washed thrice in 1 ml of SG Lysis Buffer and finally resuspended in 150 μl of SG Lysis Buffer. Prepared beads were added to the precleared SG-enriched fraction and incubated in rotation at 4°C for 2 h. Washes and beads elution were performed as in (43). To isolate SG for qRT-PCR analysis, mass spectrometry and western blot validation, Dynabeads were replaced with 25 μl/sample of GFP-Trap beads (ChromoTek), which are covalently bound to a highly efficient alpaca-raised nanobody against the tag. For qRT-PCR, beads were washed and RNA was eluted as in (42). To isolate proteins for mass spectrometry and WB analysis, after capture the beads were washed twice with 1 ml Wash Buffer [10 mM TRIS-HCl (pH = 7.5), 150 mM NaCl, 0.05% Igepal, 0.5 mM ethylenediaminetetraacetic acid (EDTA)] in rotation for 10 min at 4°C, then washed twice with 1 ml High Salt Wash Buffer [10 mM TRIS-HCl (pH = 7.5), 500 mM NaCl, 0.05% Igepal, 0.5 mM EDTA] in rotation for 10 min at 4°C. Beads were resuspended in 40 μl of 1× Laemmli Sample Buffer (Bio-Rad) supplemented with 50 mM dithiothreitol (DTT), and heated to 85°C for 10 min; the supernatant was collected and used for western blot analysis, or further processed for mass spectrometry (see the 'SG proteome analysis' section).

SG RNA sequencing (RNA-seq) and bioinformatic analysis

Purified SG RNAs and relative inputs were retrieved in absence of doxy-induction (four replicates); in condition of doxycycline-induced overexpression of FUS^{P525L} (two replicates) and FUS^{WT} (two replicates). RNA libraries for all samples were produced using Stranded Total RNA Prep with Ribo-Zero Plus (Illumina). All samples were sequenced on an Illumina Novaseq 6000 Sequencing system. Trimmomatic (44) and Cutadapt (45) were used to remove adapter sequences and poor quality bases; minimum read length after trimming was set to 35. Reads aligning to ribosomal RNAs were filtered out; this first alignment was performed using Bowtie2 software (v2.4.2) (https://bowtie-bio.sourceforge.net/bowtie2/index.shtml). STAR software (46) was used to align reads to GRCh38 genome. PCR duplicates were removed from all samples using MarkDuplicates command from Picard suite (https://broadinstitute.github.io/picard/). Uniquely mapping fragments were counted for each annotated gene (Ensemble release 99) using htseq-counts software (47). edgeR software (48) was used to compare SG-enriched RNAs to relative input samples. RNAs with $\log_2$ fold change ($\log_2$FC) > 1 and false discovery rate (FDR) < 0.05 were defined *enriched* in SG core, while those with $\log_2$FC < −1 and FDR < 0.05 were defined *depleted*; all the others were labelled as *neither*. SG from SMMC-7721

(49) cell line were re-analyzed using the previously described pipeline.

In the RNA-seq analysis, the high variability of the library size of immunoprecipitation (IP) samples could affect the number of SG-enriched RNAs identified. This could introduce technical bias especially when we need to compare the enriched RNAs of IPs performed in different conditions. To overcome this bias, we performed a 'Resampling analysis'. We decided to use the observed number of fragments in FUS^{P525L} condition experiment as reference and then we randomly sampled fragments from the alignment files of the other conditions in order to mirror the library size composition of the reference. This approach allowed us to obtain samples with simulated library sizes and comparable levels of technical variability. The RNA fragments random sampling from BAM files was performed with Picard suite using FilterSamReads function and using random subsets of fragments identifiers list as input (https://broadinstitute.github.io/picard). Then, for each condition, RNA quantification and SG-enriched RNAs identification was performed as previously described.

To define the SG-enriched RNAs, we used the already specified $\log_2$FC threshold for each condition. However, the high variability of the immunoprecipitation-based experiments could result in high variability in the SG enrichment efficiency. Experiments with low enrichment efficiency could be penalized by a fixed threshold compared to others with a high efficiency, affecting the comparison of the defined enriched RNA sets. To overcome this technical limit, we performed an 'Enrichment convergence analysis'. For each analyzed condition, we first ranked the expressed RNAs by their SG enrichment ($\log_2$FC) and then we selected a fixed number of enriched RNAs equal for each dataset. We then compared these three equally sized RNA sets using Venn diagrams. We repeated this analysis selecting different fixed numbers of enriched RNAs, gradually reducing the inclusion of the noise from the background transcriptome (fixed numbers of top enriched RNAs: 5000, 4500, 4000, 3500, 3000, 2500, 2000, 1500, 1000, 500).

Differentially expressed genes between FUS^{WT} and FUS^{P525L} condition were identified comparing INPUT samples using egdeR software (FDR < 0.05). For this analysis, a counts matrix produced from not-deduplicated BAM files was used. The details of the main RNA-seq analyses are available in Supplementary Table S1.

RNA features analysis was performed using for each gene the longest functional isoform as reference. Only protein coding isoforms were taken in consideration for protein coding genes. RNA sequences and features were retrieved from Ensemble database (release 99) using biomaRt R package (https://bioconductor.org/packages/release/bioc/html/biomaRt.html).

Gene Ontology term enrichment analysis

Gene Ontology (GO) was performed using WebGestalt R tool (v0.4.4) (50) applying weighted set cover reduction.

Literature survey

Literature survey analysis was performed using Entrez.esearch function of Biopython python library (https://academic.oup.com/bioinformatics/article/25/11/1422/330687). Bibliographic results indicated by PUBMED IDs in Supplementary Table S2 were obtained by searching on PubMed database the combination of key words ('amyotrophic lateral sclerosis'

and 'neurodegeneration') with the genes belonging to COMMON, LOSS and GAIN groups linked by 'AND' boolean logic proposition. We considered ALS linked genes also ANKRD12 and VSP4B, although not resulting from the described literature investigation. Indeed, ANKRD12 was described as downregulated in murine spinal cord of TDP43 N390D/+ mutants (51), while VSP4B is transcriptionally regulated by TDP43 and its derepression upon TDP43 downregulation leads to dendritic loss (52,53). Results and references are listed in the Supplementary Table S2.

K-mers enrichment analysis

The frequency of every possible k-mer of k = 7 was calculated in the SG enriched or depleted RNAs in each analyzed condition (FUS^{P525L}, FUS^{WT} and no DOXY). Subsequently, Z-scores for each possible 7mer were calculated based on the 7mers frequencies distribution and the difference between 7mer Z-scores in the SG enriched and depleted groups was calculated. Then, for each condition, the 7mers were ranked by the defined Z-scores difference and the top 50 and the bottom 50 7mers were selected as SG enriched or depleted 7mers, respectively. The sequence logos of the top 50 SG enriched 7mers were generated using Logomaker Python package (54). After that, the average transcript frequency of each 7mer was calculated for SG-enriched RNAs and background distribution (neither group), and their ratio (fold-change, FC) was computed. Then, the log$_2$FC of the top enriched and depleted 7mers in all the analyzed conditions were used for Heatmap representation and the k-means clustering was performed using ComplexHeatmap R package (https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html).

RNA structure prediction and analysis

In RNA structure propensity analysis, three independent random samplings of 200 mRNAs from GAIN and LOSS groups were performed imposing the similarity of 5′UTR, CDS and 3′UTR regions between groups. The similarity of the distributions was successively confirmed using Mann–Whitney U test (P value > 0.05). RNA secondary structures were predicted using RNAfold algorithm (55), using standard parameters, and CROSS algorithm (http://service.tartaglialab.com/update_submission/736048/9393232b64), using human PARS data settings. For RNAfold predictions the minimum free energy (MFE) structure of each RNA was taken in consideration and the structuration score was calculated from the dot-bracket annotation by the number of structured nucleotides over the length of the region/transcript. For RNA accessibility analysis the previously described sampling procedure was applied to GAIN and LOSS groups, selecting in each sample 50 mRNAs that were also detected in the DMS-MapSeq experiment conducted in HEK293T cell line (56). DMS-MapSeq data were retrieved from RASP database (57).

G-quadruplex analysis

G-quadruplex (G4) structures were analyzed using data from both rG4-Seq (58) and rG4-Seq 2.0 experiments (59). These methodologies use RNA sequencing to detect reverse transcription (RT) stops induced by the G4 structure stabilization through potassium (K$^+$) or potassium and pyridostatin (KPDS). The liftover of G4 genomic coordinates (rG4-Seq) from hg19 to hg38 genome assembly has been performed

using liftOver tool from UCSC (https://genome.ucsc.edu/cgi-bin/hgLiftOver). To map G-quadruplex coordinates to the RNAs identified in our system, we used bedtools intersect from BEDtools suite version 2.29.1 (60) with '-s' parameter. In the analysis performed using the data from rG4-Seq, the structures identified were stratified in three equally sized groups according to their stability score defined as the fraction of reads with RT stops (58). The three groups were labelled as 'low-stability', 'medium-stability' and 'high-stability'. G-quadruplex common set between K$^+$ and KPDS experiments from rG4-Seq1 has been identified by intersecting their genomic coordinates with bedtools intersect using the parameter previously described. Consensus G-quadruplexes among rG4-Seq and rG4-Seq 2.0 has been provided by the author of rG4-Seq 2.0 paper in Additional file 5. For the enrichment analysis, the proportion of G-quadruplex structures was compared among the *depleted*, *enriched* and *neither* groups using a Fisher's exact test to assess statistical significance (P value < 0.05).

FUS^{P525L} HITS-CLIP analysis

RNA-sequencing dataset of FUS^{P525L} HITS-CLIP in SK-N-BE cells were retrieved from GEO database with GSE242771 accession number. Reads preprocessing, alignment and postprocessing were carried out as described in Di Timoteo *et al.* (61) GRCh38 genome was divided into 200 bp long non-overlapping bins using the makewindows function from BEDtools suite (60). Properly paired reads falling in each bin were counted using the BEDtools intersect tool considering the position of the first nucleotide in read 2 (R2); in this way, sample-specific RNA fragments count files were created. Then, CLIP-tags enriched bins, were identified using two complementary approaches. At first, for each CLIP replicate, we employed the Piranha peak caller (62) with the following parameters '-no_pval_correct -r -s -u 0 -p 0.05' and using the respective INPUT counts as covariate. We then selected Piranha resulting bins in at least one replicate, and, for each bin, we further evaluated the over-representation of CLIP-tags compared to the INPUT RNA fragments using Fisher's exact test. Bins with a log$_2$ odds ratio (log$_2$OR) ≥ 1 and a FDR corrected P value ≤ 0.05 were considered significantly enriched in CLIP samples. Finally, to evaluate the consistency of enriched regions across replicates, we applied an irreproducible discovery rate (IDRate) approach (63) retaining only bins with an IDRate of ≤ 0.05. We then merged the enriched and reproducible bins using BEDtools merge to define the peak regions. Peaks annotation was performed using BEDtools intersect tool with Ensembl (release 99) gtf (Supplementary Table S4). CLIP signal was calculated by subtracting the CPM values in INP sample to the respective CLIP sample. For bigwig files it was calculated using bigwigCompare function from deepTools (https://deeptools.readthedocs.io/en/develop/).

The 'Signal saturation analysis' of HITS-CLIP allows us to determine if the RNA-sequencing depth is sufficient to saturate the CLIP signal and thus identify the optimal number of protein binding regions. The analysis was performed by generating 5 sub-samplings of CLIP-tags from each CLIP replicate: using Samtools v1.3 (64) with the -s option, 90% of the CLIP-tags were randomly selected for the first sub-sampling, 80% for the second, 70% for the third, 60% for the fourth and, finally, 50% for the fifth. For each of these sub-samples, peak calling was performed as described previously against the

non-sub-sampled INPUT. Subsequently, BEDTools (60) was used to intersect the peak regions from each sub-sampling with those from the total CLIP-tags, calculating the signal saturation index as the percentage of original RNA interactors re-called in the sub-sample (Supplementary Figure S4).

The comparison between FUS^{P525L} HITS-CLIP in SK-N-BE (GSE242771) and endogenous or FLAG FUS^{P525L} PAR-CLIP in iPSC-derived MN (GSE118347) was performed through the stratification of the MN transcriptome into 10 equally sized groups based on the FUS binding score, defined by dividing the number of T to C conversions by the total number of T in the exons. The group of RNA with no transition sites was assigned the label '0'. Subsequently, the fraction of HITS-CLIP interactors present in each RNA group was calculated (Supplementary Figure S4).

Meta-region analysis of FUS^{P525L} HITS CLIP peak regions was conducted using a custom Python script. Peak summits were determined based on the average CLIP signal (CLIP-INPUT) from the two biological replicates. Following this, a region extending from −200 nt to +200 nt around the peak summit was analyzed using a sliding window of 10 nt, with steps of 5 nt. The windows were then ordered based on their relative positions within transcripts. Using BEDTools intersect and the exon annotation from Ensembl release 99, only those windows that included exonic regions were selected for further analysis. Then, the sequence for each of these windows was extracted using BEDTools getfasta with the -s parameter, from the GRCh38 genome. Finally, for each position relative to peak summit, the frequency of windows containing at least one 7mer of the AU-rich or GC-rich group was calculated.

Meta-transcript profiles were generated based on the lengths of the 5′UTR, CDS or 3′UTR regions of the specified group of transcripts. These information, together with the transcript sequences were retrieved from the Ensembl 99 database using the biomartR package (https://bioconductor.org/packages/release/bioc/html/biomaRt.html). Each mRNA region was divided into an equal number of bins, ensuring the proportions among the lengths of the transcripts regions were maintained. The bin ranges for each RNA were compiled into a BED file, and the sequences for each bin were extracted using BEDtools getfasta from the transcriptome FASTA file. For the analysis of the HITS-CLIP peaks frequency in the meta-transcript of FUS^{P525L} interactors, the transcript bin ranges were mapped to the GRCh38 genomic coordinates using a custom python script. The peaks were then assigned to the bins using BEDtools intersect (60). Graphical representations of meta-region and meta-transcript analyses were performed using ComplexHeatmap R package (https://bioconductor.org/packages/release/bioc/html/ComplexHeatmap.html).

SG proteome analysis

After immunoprecipitation and washes (see the 'SG purification' section), beads were washed in 1 ml of Digestion buffer (50 mM ammonium bicarbonate) and resuspended in 120 μl of Digestion buffer. Disulfide bonds were reduced on-beads by adding DTT to a final concentration of 10 mM and incubating at 56°C for 30 min. Proteins were alkylated by adding Iodoacetamide to a final concentration of 20 mM and incubating beads at RT for 20 min. On-beads trypsin digestion was then performed by resuspending the beads in 200 μl of Digestion Buffer supplemented with 2 μl of 0.5 μg/μl Sequencing Grade Trypsin (Merck). Samples were incubated overnight at 37°C

with constant agitation at 750 rpm, then each sample (∼8 μg of peptides) was dried using a Speed-Vac and resuspended in 16 μl acetonitrile, 3% added with 0.1% formic acid. Tryptic peptides were desalted on a trapping column then loaded on a Aurora C18 column (75 mm × 250 mm, 1.6 μm particle size – Ion Opticks, Fitzroy, Australia) and separated using a Dionex Ultimate 3000 nano-LC system (Thermo Fisher Scientific). Eluents were A: water + 0.1% formic acid and B: water:acetonitrile = 20:80 (v/v) + 0.1% formic acid. Injection volume was 2 μl, flow was 0.300 μl/min, the column temperature was 40°C, samples were eluted with a gradient program: 0.0–1.0 min, 3% B; 1.0–5.0 min, 3–5% B; 5.0–35.0 min, 5–31% B; 35.0–40.0 min, 31–44% B; 40.0–41.0 min, 44.0–95% B; 41.0–46.0 min, 95% B; and 46.0–47.0 min, back to 3% B. The column was then reconditioned for 13 min. The total run time was 60.0 min. Separated peptides were analyzed using an Orbitrap Exploris 480 Mass Spectrometer (Thermo Fisher Scientific), in positive electrospray ionization mode. The capillary voltage was set to 2.0 kV, the RF lens was set to 40% and the normalized AGC target % was set to 300, maximum injection time 50 ms. The acquisition was performed in Data Dependent mode (DDA) with a scan range set from 350 to 1500 m/z, resolution 120 000. The intensity threshold defined for a data dependent scan was set to 10000, MS/MS spectra were acquired in High Energy Collisional Dissociation mode with a collision energy of 30%. The raw data were processed using Thermo Proteome Discoverer software (65), 1% FDR threshold and at least two peptides were used for protein identification. Proteins from Mass Spectrometry data were annotated using biomaRt (ensemble 109 release).

Analysis of protein–RNA and protein–protein interactions

RBPs *in vitro* binding propensity to SG 7mers signature was performed using *Homo sapiens* RNAcompete (66) dataset from CISBP-RNA database.

Protein–RNA interactions *in silico* predictions performed by catRAPID software were retrieved from RNAct database (67). For each couple of interacting protein and RNA, the GENCODE basic isoform with the highest interaction score was selected as representative. To compare the RBP binding propensity between GAIN and LOSS RNAs avoiding biases due to different RNA lengths between the two groups we used RNA samples already described in the 'RNA structure prediction and analysis' section. The predicted preferential binding with GAIN over LOSS group was performed calculating for each SG protein the median normalized score with GAIN RNAs and with LOSS RNAs. Then, Z-scores were calculated from the ratios of median normalized scores (GAIN/LOSS) of the whole set of SG proteins.

Protein–RNA interactions analysis of cells-derived datasets was performed using data from POSTAR3 database (CLIPdb) (68) and the already described RNA samplings from GAIN and LOSS groups. To define RBP targets, we used BEDtools suite version 2.29.1 (60) intersecting the RBP peak regions with the Ensembl (release 99) exon annotation relative to the representative isoforms of the selected genes. The RBP binding propensity with GAIN over LOSS group was then defined as the log$_2$OR between GAIN RNAs identified as interactors of the RBP and LOSS RNAs interactors. Additionally, the statistical significance of the enrichment analysis was evaluated using Fisher's exact test, with P values adjusted for multiple

comparisons using the Benjamini–Hochberg FDR correction (Supplementary Table S4).

Protein–protein physical Interactions were retrieved from STRING database (v11.5).

Quantification and statistical analysis

The distribution and deviation of data shown in the figures of this work, the statistical tests used to calculate significant differences, and the exact value of n (e.g. number of biological replicates of the experiments) are denoted in figure legends. In main figure legends 'SD' stands for 'standard deviation' and 'SEM' stands for 'standard error mean'. Significance values were depicted in the figures using the following key legend: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

Results

Characterization of the RNA composition of neuroblastoma SG

To investigate the RNA composition of neuronal SG, we adapted the purification protocol described by Khong *et al.* (42) (Figure 1A) to a SK-N-BE neuroblastoma cell line expressing the SG core protein G3BP1 fused to a GFP tag (SK-GFP-G3BP1). Fluorescence microscopy showed that upon oxidative stress these cells display GFP-G3BP1 cytoplasmic foci (Supplementary Figure S1A). We performed an additional immunofluorescence control for the SG component TIAR, finding that 80.7% of TIAR-positive granules contain GFP-G3BP1. Interestingly, the concomitant staining with a G3BP1 antibody, thus recognizing also the endogenous protein, showed a stronger colocalization with TIAR (93.8%), demonstrating that the TIAR-positive granules which are negative for GFP signal are indeed SG. (Supplementary Figure S1B and C). Western blot analysis confirmed the specific enrichment of the endogenous SG marker TIAR in the IP fraction of SK-GFP-G3BP1 only in stress conditions compared to unstressed controls, suggesting that the SG are enriched as expected (Supplementary Figure S1D). After RNA extraction, qRT-PCR was performed on targets previously identified as enriched or depleted in granules of U2OS cells (43). AHNAK, CENPF, HUWE1 and TRIO were selected as positive controls and displayed a strong enrichment in the IP fraction under stress induction, while the negative controls ATP5O, COX8A and LGALS1 were not precipitated (Supplementary Figure S1E). These controls allow to conclude that the particles that we purified are *bona fide* SG.

After these controls, we generated four replicates of purified SG and corresponding total RNA inputs and performed total RNA sequencing. Multidimensional scaling plot of leading fold change (Supplementary Figure S1F) and Pearson's correlation matrix (Supplementary Figure S1G) confirmed the expected clusterization of the samples. We proceeded to RNA-sequencing analysis comparing SG-enriched (SG-enr) versus Input samples (INP). Among the 12 318 RNAs detected in the system (>1 FPKM in INP or SG-enr) we identified 1816 *enriched* ($\log_2 FC > 1$ and FDR < 0.05) and 2939 *depleted* ($\log_2 FC < -1$ and FDR < 0.05) transcripts in SG (Figure 1B and Supplementary Table S1). Most of the RNAs that we identified as enriched in SG were mRNAs (85% of the enriched fraction) and among the non-coding fraction (15%) the most represented category was long non-coding RNAs (9%) (Figure 1C).

In order to validate in an independent manner the SG engagement of the different categories of RNAs we selected candidate transcripts from the *enriched* and *depleted* categories and performed smFISH. GAPDH was selected as representative for the '*depleted*' group, while NORAD for the '*enriched*' species. The co-localization analysis of smFISH signals with endogenous G3BP1 foci in stress conditions confirmed the differential association to SG of the two classes of transcripts (Figure 1D and E).

Comparison of SG transcriptome datasets

In order to identify the discriminating features of the neuroblastoma SG transcriptome, we compared it to the ones of osteoblastoma (U2OS) (43) and endocervical adenocarcinoma (SMMC-7721 cell line, see the 'Materials and methods' section) (49) (Supplementary Table S1). We observed a strong correlation of the enrichment scores ($\log_2 FC$) among all the three datasets (Supplementary Figure S2A), indicating that the methodology is highly reproducible even when applied in different systems. When comparing the SG transcriptomes, we identified a common *set* consisting of 502 RNAs (Figure 2A). These transcripts are mostly mRNAs (Supplementary Figure S2B) and are enriched in GO categories related to cytoskeletal and structural components, including neuronal-specific factors (Supplementary Figure S2C and Supplementary Table S2). Noteworthy, neuronal transcripts, even if less represented in the transcriptomes of the other analyzed cell lines (Supplementary Figure S2D), show a significant enrichment in the SG common *set* (Supplementary Figure S2C). Considering that length was confirmed as a relevant feature for SG engagement in any cell type analyzed (Figure 2B), their enrichment may be due to their tendency to be longer than other gene categories (e.g. metabolic process, mitochondrion, endoplasmic reticulum and translation, Figure 2C).

However, the SG RNA content shows strong divergences in the different systems, with 37.8%, 40.4% and 52.7% of the SG transcriptomes differing in SK-N-BE, U2OS and SMMC-7721 granules, respectively. Interestingly, the SG transcriptome in neuroblastoma cells is enriched for neuronal categories such as 'neuron to neuron synapse', 'neuronal cell body' and 'axon part', which are absent in the other systems (Figure 2D, Supplementary Figure S2E, F and G, and Supplementary Table S2). These data indicate that, in our neural-like system, transcripts with neuronal functions have propensity to be included in SG, and this behavior correlates with their length compared to transcripts belonging to other categories (Figure 2C). Furthermore, we observed that the length of the CDS and 3′UTR, and less of 5′UTR (Supplementary Figure S2H), are crucial features for SG engagement.

However, length and RNA biotype are not the only discriminating features for recruitment into granules. In fact, when we analyzed the nucleotide composition, we observed that the SK-N-BE and SMMC-7721 datasets have a higher GC content than the U2OS set (Figure 2B), suggesting that the inclusion of specific transcripts leads to different overall nucleotide composition and thus could be associated with different physical properties of SG. Noteworthy, these evidences are not dependent on the thresholds selected to define enriched RNAs, since we observed the same results by stratifying the dataset according to the enrichment score
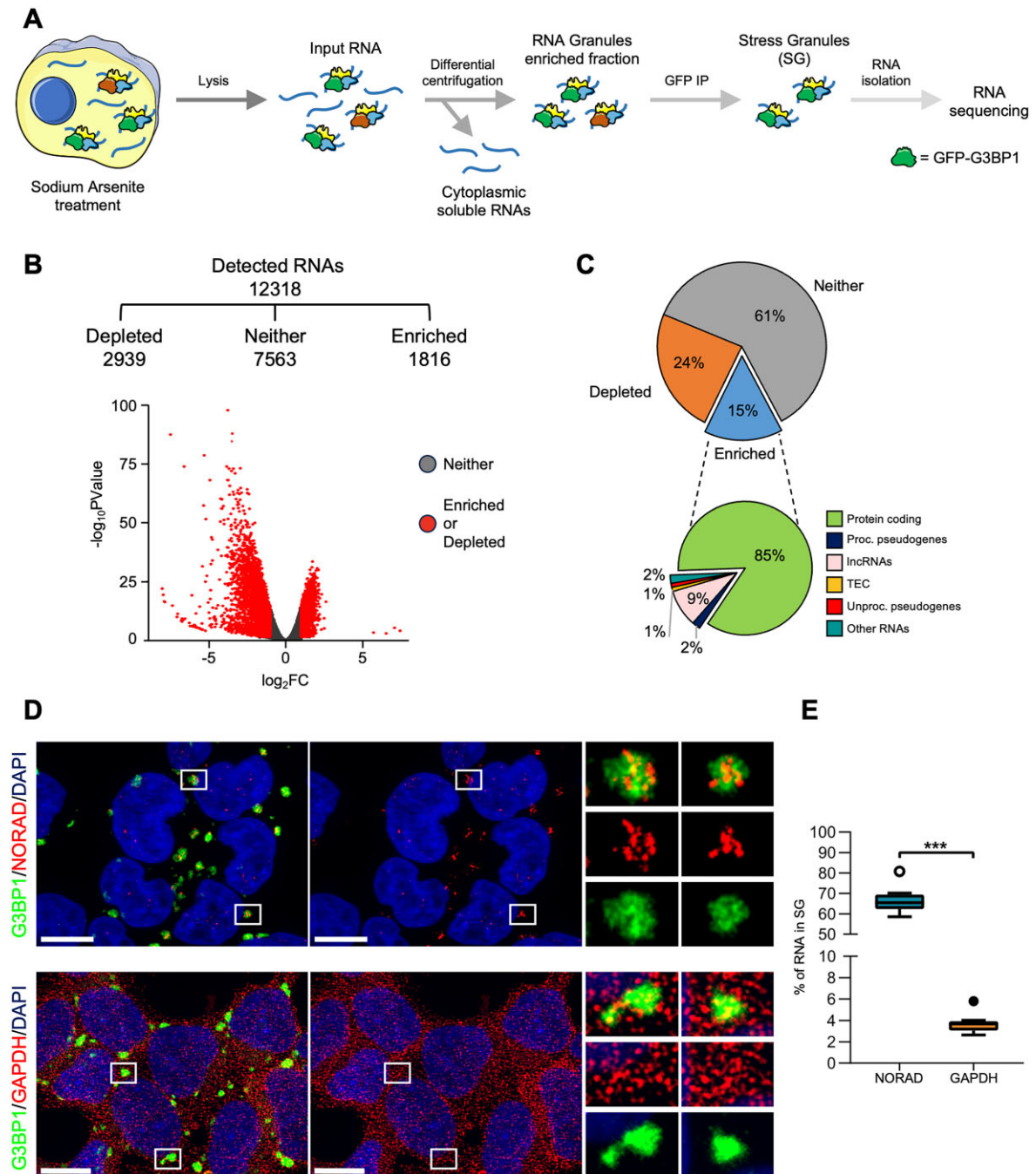
Figure 1. (**A**) Schematic representation of the SG purification protocol performed to isolate and analyze SG RNA composition. (**B**) Upper panel. Summary of the number of RNAs detected in RNAseq of SG purification experiment, and their partition among *enriched*, *neither* and *depleted* groups. Lower panel. Volcano plot showing the $\log_2$FC SG/INP and the $-\log_{10}P$ value of SG enrichment for each RNA detected. Significantly *enriched* and *depleted* RNAs (FDR < 0.05, $|\log_2$FC$| > 1$) or *neither* RNAs are indicated by red or gray dots, respectively. (**C**) Upper panel. Pie chart depicting the proportions of SG-*enriched*, *depleted* and *neither* RNAs. Lower panel. Pie chart depicting the proportions of RNA biotypes among the SG-enriched RNAs. (**D**) Representative confocal images of single-molecule fluorescence *in situ* hybridization (smFISH) for NORAD (upper panel) and GAPDH (lower panel) transcripts (red), combined with immunofluorescence for G3BP1 (green) in SK-N-BE cells subjected to sodium arsenite stress. Nuclei were counterstained with DAPI. Scale bar = 10 μm. Right panels show digital magnifications of the G3BP1 granules highlighted by squares. (**E**) Boxplot depicting the distribution of the percentage of NORAD and GAPDH RNA localization within G3BP1-marked SG, derived from confocal images analysis. A total of 691 and 714 cells were counted for NORAD and GAPDH, respectively, from three independent biological replicates. Statistical significance was assessed with two-tailed, unpaired T-test.
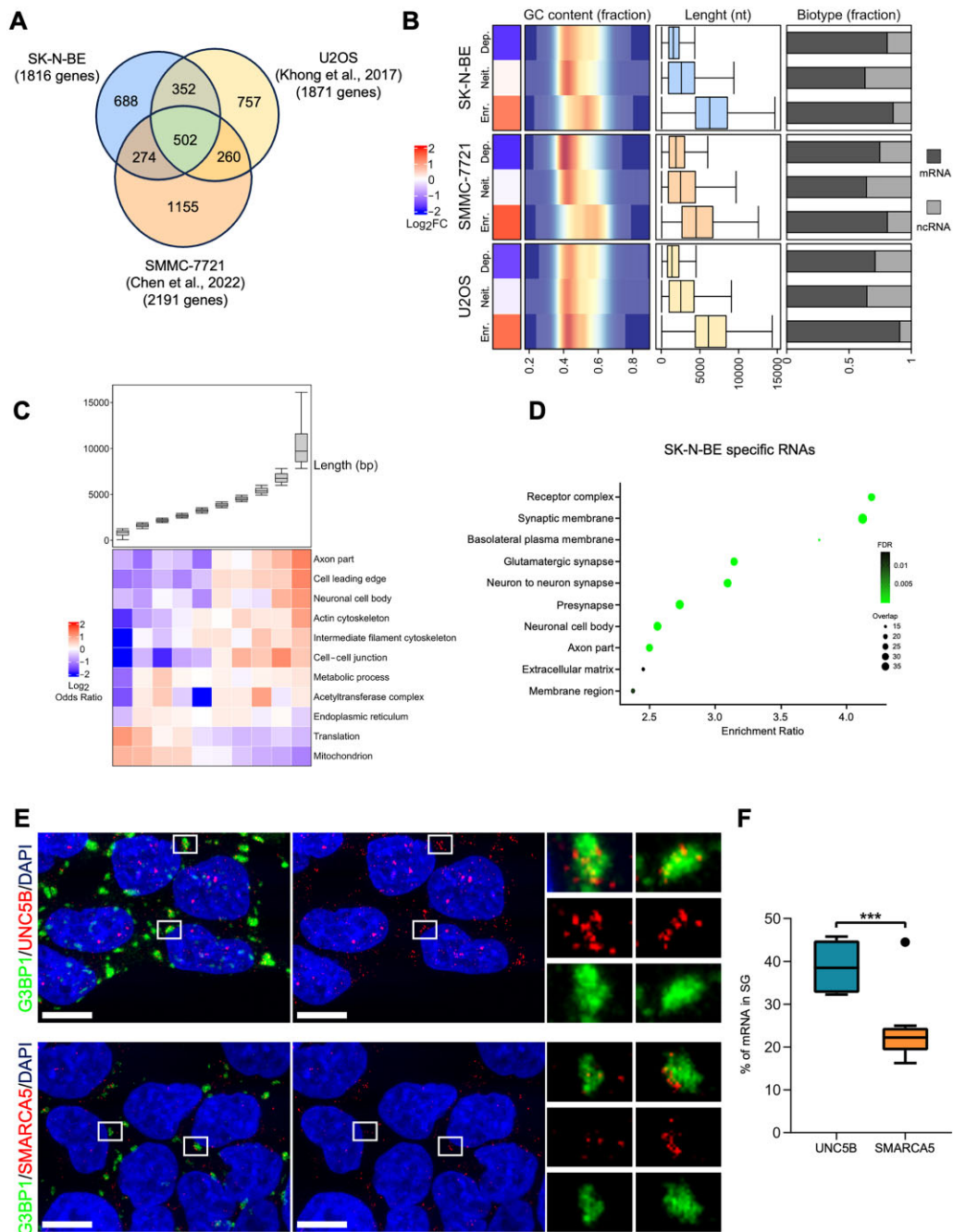
Figure 2. (**A**) Venn diagram depicting the overlap between SG-enriched RNAs in neuroblastoma (SK-N-BE, this study), osteoblastoma (U2OS) and endocervical adenocarcinoma (SMMC-7721) cell lines. (**B**) Heatmap depicting for SK-N-BE, SMMC-7721, U2OS cell lines the median SG enrichment score (log_2FC SG/INP) of the *depleted*, *neither* and *enriched* transcripts. Red color scale reports positive SG enrichment scores while blue color scale describes negative values. Heatmap density plot annotation reports the distribution of RNA GC content (ratio of GC nucleotides over the transcript length) for each transcripts group. Red color scale describes high-density of values while blue color scale describes low-density. Heatmap boxplot annotation reports global transcripts length (nt) for the specified RNA groups. Blue/orange/yellow boxplot color refers to the SK-N-BE, SMMC-7721 or U2OS system, respectively. Heatmap barplot annotation reports transcript biotypes fractions in each specified RNA group. Dark-gray color describe the fraction of mRNAs, while light-gray color the fraction of ncRNAs. (**C**) Heatmap (lower panel) depicting the relationship between RNA length and GO categories enriched in 'SG common *set*'. The 10 transcripts groups represented were defined by length-based stratification of the transcriptome. The log_2OR depicted in the heatmap shows the depletion (blue) or the enrichment (red) of each GO category in the defined group compared to the remaining fraction of the transcriptome. Boxplot (upper panel) represents the distributions of RNAs length for the 10 transcripts groups. (**D**) Dotplot depicting the cellular component GO over-represented categories of the SK-N-BE SG specific RNAs. Only significant categories (FDR < 0.05) were depicted. X-axis represents category enrichment score, while Y-axis reports the GO category description. Dot size represents the amount of RNAs in the analyzed group that overlap the category, while green colors report the significativity of the enrichment for CC categories. (**E**) Representative confocal images of smFISH for UNC5B (upper panel) and SMARCA5 (lower panel) transcripts (red), combined with immunofluorescence for G3BP1 (green) in SK-N-BE cells subjected to sodium arsenite stress. Nuclei were counterstained with DAPI. Scale bar = 10 μm. Right panels show digital magnifications of the G3BP1 granules highlighted by squares. (**F**) Boxplot depicting the distribution of the percentage of UNC5B and SMARCA5 RNA localization within G3BP1-marked SG, derived from confocal images analysis. 581 and 626 cells were counted for UNC5B and SMARCA5, respectively, from three independent biological replicates. Statistical significance was assessed with two-tailed, unpaired T-test.

($\log_2$FC SG/INP) in SG and comparing groups of equal numerosity (Supplementary Figure S2I).

In order to prove that indeed RNA recruitment in SG could vary in different systems we performed smFISH of two RNAs: UNC5B that resulted enriched in SG only in SK-N-BE cells, and SMARCA5 that resulted enriched in U2OS and SMMC-7721, but not in SK-N-BE cells. Despite their similar length (6841 nt and 7684 nt respectively), UNC5B has a high-GC content (60.2%) and has neuronal functions (69), while SMARCA5 has a lower GC content (38.4%) and it is not associated to any neuronal activity. The co-localization analysis of smFISH signals with endogenous G3BP1 foci indicated the specific SG enrichment only for UNC5B (Figure 2E and F).

The ALS-associated FUS^{P525L} mutation affects SG properties and dynamics

Mis-localization of mutant FUS in the cytoplasm and its accumulation in inclusions is known to be a central hallmark of ALS-associated mutations. Moreover, it is known that the mis-localized FUS co-localizes with other RBPs in SG as a consequence of oxidative stress (28). Despite this evidence, few efforts have been made to characterize the implications of mutant FUS localization on the physical-chemical features of SG and their RNA composition. To elucidate this aspect, we generated two stable SK-N-BE cell lines overexpressing under a doxycycline-inducible promoter either the wild-type FUS protein (FUS^{WT}) or a version carrying the P525L mutation on the NLS, which is associated to severe and juvenile ALS (FUS^{P525L}) (70). We induced FUS overexpression with 50 ng/ml doxycycline obtaining expression levels of ectopic FUS comparable to the endogenous levels. Indeed, the amount of sequenced RNA fragments that have the P525L point mutation are slightly more abundant than the wild type endogenous fragments (Supplementary Figure S3A); similarly, the protein levels of induced FUS are comparable to the endogenous ones (Supplementary Figure S3B). We showed that the overexpressed FUS^{P525L} is physically recruited into SG, as demonstrated by biochemical purification (Supplementary Figure S3C). We also verified the correct FUS localization and observed that, while overexpressed FUS^{WT} is completely localized in the nucleus, FUS^{P525L} is cytoplasmic and co-localizes upon stress with the SG core protein G3BP1 in specific foci (Supplementary Figure S3D and E). Moreover, we compared the number of SG in both conditions. As reported in other systems (71), we found an increased number of SG per cell in the presence of FUS^{P525L} (Supplementary Figure S3F).

We then analyzed GFP-G3BP1 dynamics in FUS^{WT} and FUS^{P525L} cells, using fluorescence fluctuation spectroscopy with a SPAD array detector, to evaluate changes in the dynamics of condensation due to FUS mutation (33,34). In agreement with previous observations, we found that in both cell lines G3BP1 movement decreases upon stress induction (Figure 3A; *D*app measures the speed of movement of individual molecules). The decrease of the diffusion coefficient upon stress induction reflects the internalization of G3BP1 inside SG. We then analyzed the diffusion times at different detection areas (integrating the fluorescence signal at different pixels of the SPAD array detector) to perform the so-called diffusion law. When the intercept of the linear regression of the diffusion times at different detection areas is bigger than zero, the sample has a domain-confined type of movement; when it is lower

than zero, the sample has a meshwork-confined type of movement. The results of Figure 3B show a more confined type of movement in presence of FUS^{P525L} compared to FUS^{WT}, suggesting a less dynamic conformation for SG when the mutant FUS is included.

Evidence that more solid granules form in the presence of mutant FUS was recently provided by the discovery that the inclusion of FUS^{P525L} in SG causes their delayed dissolution after the release of oxidative stress (61). Since SG are generated after the block of cap-dependent translation (72), we wondered whether the persistence of aberrant granules could be associated with a reduced translation recovery following stress release. To address this question, we performed a protein puromycilation assay on both FUS^{WT} and FUS^{P525L} cells after 3 and 4 h of stress recovery, and measured puromycin incorporation by western blot. Interestingly, we observed that the protein levels were significantly lower in FUS^{P525L} compared to FUS^{WT} at 4h after stress release (Figure 3C and Supplementary Figure S3G), indicating that the persistence of aberrant granules is associated to reduced translation.

Altogether, these data indicate that SG physiology is strongly affected by mutant FUS; in particular, the presence of FUS^{P525L} in SG induces an increased number of granules that are globally less fluid and more difficult to dissolve. Moreover, the presence of persistent aberrant granules and the reduced translation in mutant FUS conditions, which could be mediated by other mechanisms together with SG endurance (73,74), indicate that cells could have greater difficulty recovering from a stress state.

ALS-related FUS mutation P525L affects SG RNA composition

Since we observed variations in SG physical properties, we wondered whether the RNA content of these granules is altered in presence of FUS^{P525L}. Therefore, we generated duplicated samples of purified SG and performed total RNA sequencing with matched total RNA inputs in condition of doxycycline induction of FUS^{WT} and FUS^{P525L}. Leading Fold change analysis (Supplementary Figure S3H) and dendrogram of relative correlation distances (Supplementary Figure S3I) confirmed that samples show the expected clusterization in INP and SG-enr conditions and in FUS^{WT} or FUS^{P525L} backgrounds. Therefore, for each condition we compared the SG-enriched samples with matched input: in FUS^{WT} conditions we identified 2178 *enriched* ($\log_2$FC > 1 and FDR < 0.05) and 2990 *depleted* RNAs ($\log_2$FC < -1 and FDR < 0.05) ; using the same thresholds, we observed 1510 SG-*enriched* and 2757 *depleted* RNAs in FUS^{P525L} (Figure 3D and Supplementary Table S1). Besides, we compared the SG transcriptome in the three datasets we generated: FUS no DOXY (without FUS induction), FUS^{WT} and FUS^{P525L}. First, we observed that 87.5% of no DOXY SG RNAs are present also in FUS^{WT} SG, as expected from the mainly nuclear localization of the overexpressed wild type (WT) protein. Conversely, investigating the impact of mutant FUS, we observed that 822 RNAs enriched in SG in both no DOXY and FUS^{WT} conditions, are excluded from FUS^{P525L} SG (thus, we renamed this group of transcripts 'LOSS'). Moreover, we also identified a consistent group of 556 RNAs that are specifically recruited in SG in presence of the mutation, but not in the other two conditions (group renamed as 'GAIN'). Despite the strong differences observed, a
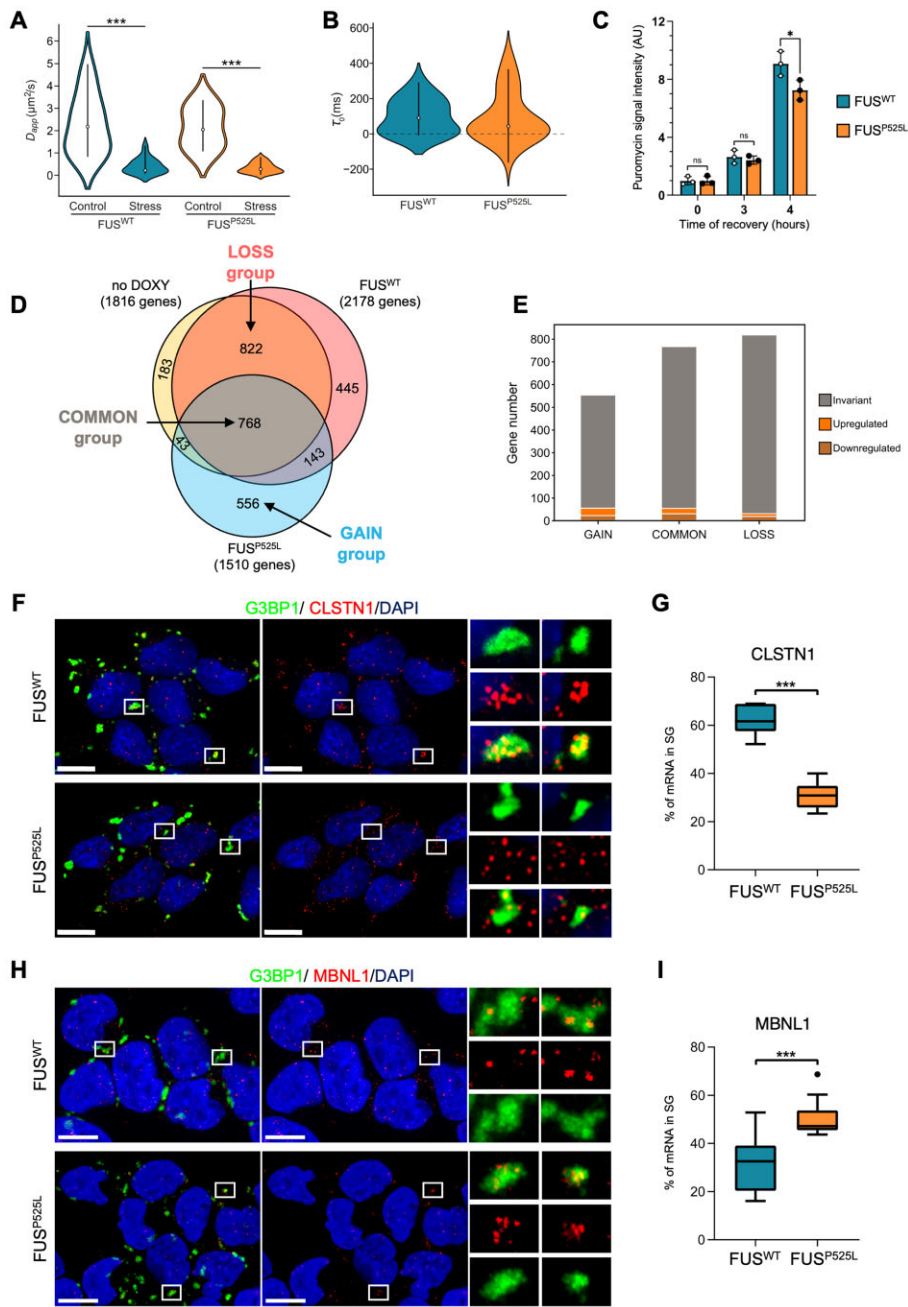
Figure 3. (**A**) Violin plot showing spot-variation FCS measurements of the diffusion coefficient (Dapp) of GFP-G3BP1 in live SK-N-BE cells overexpressing FUS^{WT} and FUS^{P525L}, both in control and stress conditions. Dapp measures the speed of movement of individual molecules. (**B**) Violin plot comparing the diffusion times (t_0) of individual GFP-G3BP1 molecules within formed SG in FUS^{WT} and FUS^{P525L} conditions. The diffusion time t_0 is obtained by integrating the fluorescence signal at different pixels of the SPAD array detector and defines the confinement of movement of the analyzed molecules. (**C**) Barplot depicting the quantification of puromycin signal from the puromycilation assay after 0, 3 and 4 h of recovery after stress. Puromycin signal is normalized on the total protein signal from Ponceau staining. $n = 3$. Statistical significance was assessed with two-tailed, unpaired T-test. (**D**) Venn diagrams showing the overlap between SG-enriched RNAs in no DOXY, FUS^{WT} and FUS^{P525L} conditions. The amount of RNAs resulted enriched in SG of each condition was indicated. Arrows indicate the fraction of RNAs defined as FUS^{P525L} GAIN, COMMON and FUS^{P525L} LOSS. (**E**) Stacked Barplot depicting the amount of deregulated RNAs in FUS^{P525L} versus FUS^{WT} INPUT comparison among the GAIN, COMMON and LOSS groups. (**F**) Representative confocal images of smFISH for CLSTN1 transcript (red), combined with immunofluorescence for G3BP1 (green) in FUS^{WT} and FUS^{P525L}-expressing SK-N-BE cells subjected to sodium arsenite stress. Nuclei were counterstained with DAPI. Scale bar = 10 μm. Right panels show digital magnifications of the G3BP1 granules highlighted by squares. (**G**) Boxplot depicting the distribution of the percentage of CLSTN1 RNA localization within G3BP1-marked SG in FUS^{WT} and FUS^{P525L}-expressing SK-N-BE cells, derived from confocal images analysis. 536 and 405 cells were counted in FUS^{WT} and FUS^{P525L} cells, respectively, from three independent biological replicates. Statistical significance was assessed with two-tailed, unpaired T-test. (**H**) Representative confocal images of smFISH for MBNL1 transcript (red), combined with immunofluorescence for G3BP1 (green) in FUS^{WT} and FUS^{P525L}-expressing SK-N-BE cells subjected to sodium arsenite stress. Nuclei were counterstained with DAPI. Scale bar = 10 μm. Right panels show digital magnifications of the G3BP1 granules highlighted by squares. (**I**) Boxplot depicting the distribution of the percentage of MBNL1 RNA localization within G3BP1-marked SG in FUS^{WT} and FUS^{P525L}-expressing SK-N-BE cells, derived from confocal images analysis. 591 and 618 cells were counted in FUS^{WT} and FUS^{P525L} cells, respectively, from three independent biological replicates. Statistical significance was assessed with two-tailed, unpaired T-test.

set of 768 transcripts is maintained in all the three conditions and always engaged in SG ('COMMON' group) (Figure 3D).

In order to check whether such alterations in SG composition could be due to diversities in the FUS^{P525L} and FUS^{WT} transcriptomes, we performed differential expression analysis in INP samples of FUS^{P525L} and FUS^{WT}. The results indicated that the two conditions differ for only a small fraction of transcripts (Supplementary Figure S3J, 244 among the downregulated transcripts and 274 in the upregulated ones) and that the contribution of such RNAs to the LOSS, COMMON and GAIN groups is negligible (Figure 3E).

Moreover, using 'Resampling analysis' (see the 'Materials and methods' section), we confirmed that the observed alterations of SG transcriptome in presence of FUS^{P525L} are not dependent on the technical variability of RNA-seq library sizes (Supplementary Figure S3K), and 'Enrichment convergence analysis' (see the 'Materials and methods' section) confirmed that it is does not depend on the enrichment threshold (log$_2$FC SG/INP) used to define the SG transcriptomes (Supplementary Figure S3L).

To understand whether specific RNA categories are associated to the observed reorganization of SG transcriptome, we performed GO terms enrichment analysis (Supplementary Table S2). RNAs that are enriched in all conditions (COMMON) include functions related to neural development, cell-cell adhesion and transcription regulators (Supplementary Figure S3M). Instead, in the LOSS group, we found RNAs related to cell signaling pathways, and cytoskeletal and neuronal components (Supplementary Figure S3N). Finally, in the GAIN group we found fewer categories mainly related to transcriptional regulation (Supplementary Figure S3O).

To biochemically validate the differential recruitment of individual RNAs into SG in the different conditions analyzed, we selected several transcripts enriched in the LOSS and GAIN groups, which were described to be linked to neurodegeneration or ALS (COMMON = 311 genes, LOSS = 321 genes and GAIN = 172 genes; see Supplementary Table S2, tab 'Literature survey' for references). Through qRT-PCR amplification on independent SG purification samples, we confirmed the strong differential enrichment of the selected RNAs in the mutant versus the WT conditions (Supplementary Figure S3P and Q).

Furthermore, in order to validate the altered SG localization of RNAs in the indicated groups, we performed smFISH selecting transcripts representative of the LOSS (CLSTN1) and GAIN (MBNL1) categories. Imaging confirmed the biochemical data, showing that CLSTN1 is preferentially excluded from mutant SG (Figure 3F and G), while MBNL1 is strongly enriched (Figure 3H and I).

The inclusion or exclusion of RNAs associated to neurodegeneration from FUS^{P525L} SG is particularly intriguing. Indeed, in both cases, these transcripts are vulnerable in the presence of mutant FUS: the GAIN group suffers from the inability of FUS^{P525L} SG to dissolve promptly upon stress release (61,75), while the LOSS group is not protected from stress insults since they are excluded from SG.

Altogether, these data suggest that the ALS-associated FUS^{P525L} mutation has a strong effect on SG dynamics and induces a dramatic reorganization of the SG transcriptome which includes components previously associated to neuronal degeneration.

FUS^{P525L} affects SG transcriptome nucleotide composition

We then decided to investigate whether FUS mutation also affects RNA length, UTRs and CDS size and GC content of SG-enriched RNAs. Thus, we considered how SG enrichment (log$_2$FC SG/INP) relates with those characterizing features.

As we observed for the no DOXY condition, length is a key feature to determine SG entry for both FUS^{WT} and FUS^{P525L} samples, with longer RNAs being more efficiently recruited (Figure 4A). Moreover, we noticed that the length of CDS and 3′UTR of mRNAs is relevant for SG engagement in all the three conditions, while the contribution of the 5′UTR in this enrichment is less obvious. However, despite the length of RNAs and mRNAs regions being a common criterion for SG engagement, we could detect some relevant differences: RNAs in the FUS^{P525L} SG are globally longer (Supplementary Figure S4A), and this increased size is mainly due to the length of the 3′UTRs. Indeed, conversely to what it was observed for the 3′UTR, both 5′UTR and CDS are slightly shorter in the FUS^{P525L} SG transcriptome (Supplementary Figure S4B).

Furthermore, we also analyzed the fraction of protein coding transcripts and we observed that most of the RNAs identified as enriched were mRNAs, with comparable percentages in all the three groups (80% in FUS^{WT} to 85% FUS^{P525L} and 85% no DOXY) (Figure 4A).

However, we observed the strongest difference in the GC content of the compared transcriptomes, with a consistent reduction of the GC content in the FUS^{P525L} condition compared to no DOXY and FUS^{WT} conditions (Figure 4A and Supplementary Figure S4C). Indeed, analyzing the GC content in no DOXY and FUS^{WT} conditions, we observed a low correlation degree with SG enrichment ($\rho = 0.065$ no DOXY and $\rho = 0.04$ FUS^{WT}, in Supplementary Figure S4D), but when we focused on the FUS^{P525L} condition we strikingly observed a strong anti-correlation ($\rho = -0.3$, Supplementary Figure S4D). This evidence not only highlights that SG-enriched RNAs in FUS^{P525L} condition have a greater AU content (Figure 4A) compared to the others, but also that this enrichment is a feature associated with RNA entry in SG only in the presence of mutant FUS.

Noteworthy, this evidence is not dependent on the thresholds selected to define enriched RNAs. Indeed, we got the same results by stratifying the dataset according to the enrichments score in SG (Supplementary Figure S4E).

To further investigate the relationship between sequence composition and SG recruitment, we asked whether specific subsequences are particularly represented in SG enriched transcripts. To highlight this aspect, we performed a k-mer (k = 7) enrichment analysis comparing enriched and depleted RNAs. Analyzing the top 50 7mers over-represented in SG RNAs, we found that transcripts enriched in both FUS^{WT} and no DOXY conditions are characterized by a higher content of GC-rich 7mers, while the mutant FUS condition preferentially recruits RNAs containing AU-rich 7mers (Figure 4B). Indeed, K-means clustering analysis of subsequence enrichment in SG identified two defined clusters of 7mers (cluster 1 and 5), that have divergent patterns of enrichment: the cluster 1, specifically enriched in FUS^{P525L}, displayed AU-rich sequences, while the cluster 5 that is specifically enriched in no DOXY and FUS^{WT} conditions displayed GC-rich ones (Figure 4C and Supplementary Table S2).
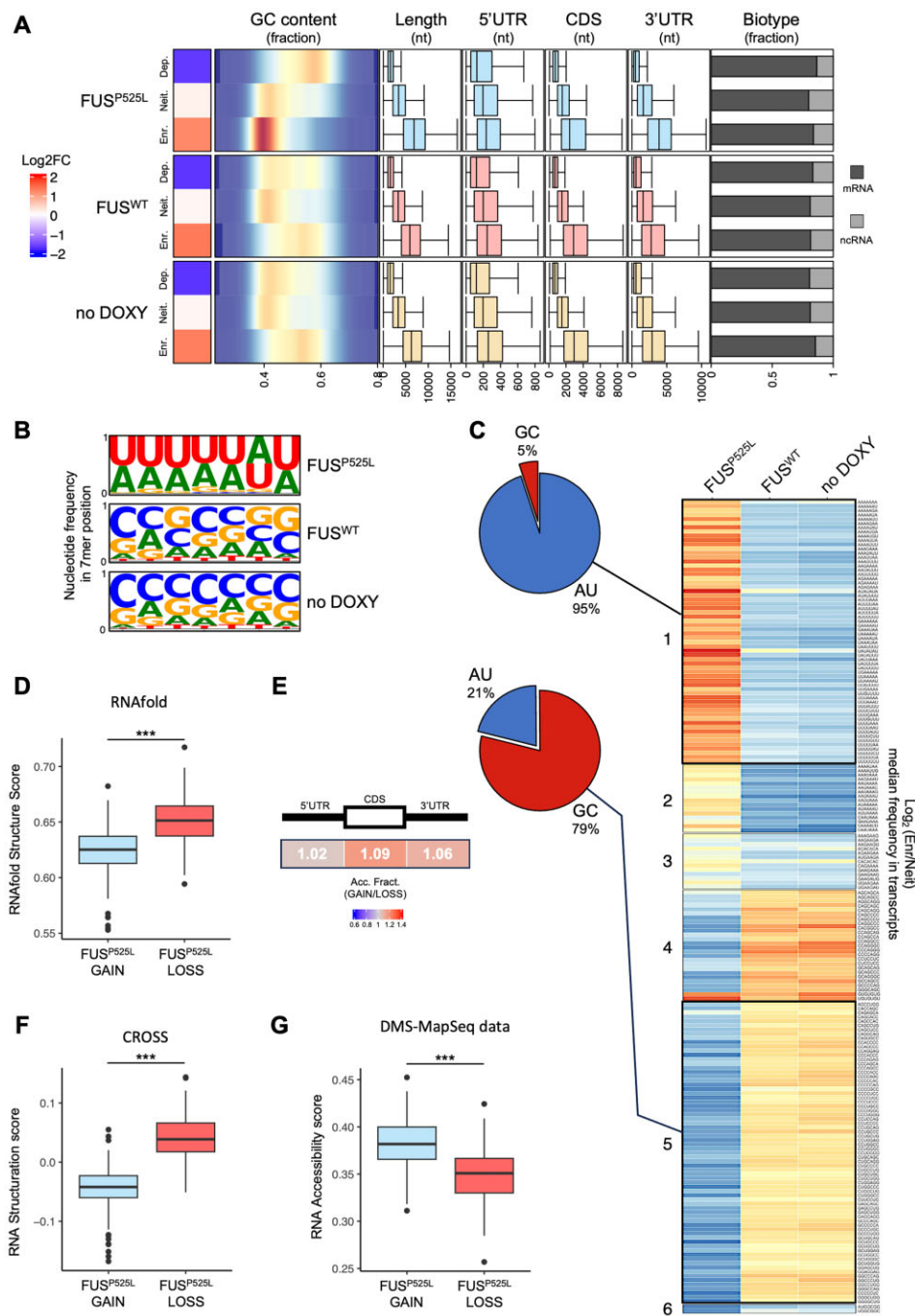
Figure 4. (**A**) Heatmap depicting for FUS[P525L], FUS[WT] and no DOXY conditions the median SG enrichment score (log$_2$FC SG/INP) of the *depleted*, *neither* and *enriched* transcripts. Red color scale reports positive SG enrichment scores while blue color scale describes negative values. Heatmap density plot annotation reports the distribution of RNA GC content (ratio of GC nucleotides over the transcript length) for each transcript group. Red color scale describes high-density of values while blue color scale describes low-density. Heatmap boxplot annotation reports global transcripts length (nt), 5′ UTR, CDS and 3′UTR length (nt) for the specified RNA groups. Light-blue/light-red/yellow boxplot color refers to the FUS[P525L], FUS[WT] and no DOXY conditions, respectively. Heatmap barplot annotation reports transcript biotypes fractions in each specified RNA group. Dark-gray color describe the fraction of mRNAs, while light-gray color the fraction of ncRNAs. (**B**) Sequence logo of the top 50 7mers over-represented in SG-enriched RNAs compared to SG depleted in no DOXY, FUS[WT] and FUS[P525L] conditions. (**C**) Heatmap showing the log$_2$FC of 7mers frequency in SG *enriched* over *neither* RNAs in no DOXY, FUS[WT] and FUS[P525L] conditions. Only the union of the top 50 over-represented and 50 under-represented 7mers in each condition is depicted. Cluster number is reported on the left. Pie charts represent the percentage of AU or GC nucleotides in the 7mers enriched (cluster 1) and depleted (cluster 5) in FUS[P525L] SG. (**D**) Boxplot depicting the distribution of RNA structuration degree predicted using RNAfold algorithm between FUS[P525L] GAIN and LOSS groups. The differences between groups were calculated using Mann–Whitney U test. The first sampling was reported. Size of sampling = 200 RNAs per group. (**E**) Representative heatmap showing the fold between average accessibility of FUS[P525L] GAIN over LOSS groups. The differences between groups were calculated using Mann–Whitney U test. Only significant differences were depicted (*P* value < 0.05). The first sampling was reported. Size of sampling = 200 RNAs per group. (**F, G**) Boxplot depicting the distribution of RNA structuration score predicted by CROSS algorithm (**F**) and RNA average accessibility calculated using DMS-MapSeq data (**G**) between FUS[P525L] GAIN and LOSS groups. The differences between groups were calculated using Mann–Whitney U test. The first sampling was reported. Size of sampling = 200 RNAs per group for CROSS (**F**) and 50 RNAs per group for DMS-MapSeq (**G**).

Moreover, we observed that AU-rich sequences are predominantly enriched in the CDS of GAIN transcripts, while GC-rich sequences of LOSS transcripts are more frequent in their 3′UTR. This finding highlights that the mRNA regions that may serve as primary drivers for granule recruitment are different between the WT SG and the aberrant ones (Supplementary Figure S4F and G).

To better understand the relationship between different RNA composition and the phenotype of reduced SG dynamics observed in FUS[P525L], we decided to deeper investigate the properties of the LOSS and GAIN transcripts, which represent the strong changes in transcriptome architecture. To do that, we compared the RNA secondary structures propensity of mRNAs from the two groups. In order to have a more reliable comparison, we sampled an equal number of RNAs from the two groups; moreover, since mRNA functional regions (5′UTR, CDS and 3′UTR) have different structural and interaction properties, we compared mRNAs with similar 5′UTR, CDS and 3′UTR length (Mann–Whitney U test, P value > 0.05; Supplementary Figure S4H, I and J). We then applied the RNAfold algorithm (55) to predict the minimum free energy (MFE) RNA secondary structure of each transcript, calculating the structuration score as the number of paired nucleotides over the length of the full transcripts or regions of them. Interestingly, we found that RNAs in the GAIN group were globally less structured than the LOSS group (Figure 4D and Supplementary Figure S4K). Moreover, comparing accessibility (defined as the fraction of un-paired nucleotides over the length of the full transcripts or regions of them) in each mRNA region, we found that CDS displays the stronger contribution to this feature, with 9% increased accessibility in FUS[P525L] GAIN compared to the FUS[P525L] LOSS group (Figure 4E).

Moreover, to corroborate this evidence with other approaches, we used CROSS algorithm (76), that is an RNA secondary structure predictor based on a machine learning approach, with parameters trained on RNA structures derived by PARS datasets. We observed a higher structuration score for LOSS RNAs compared to the GAIN group (Figure 4F and Supplementary Figure S4L), highlighting that the RNAs that lose their enrichment in SG in presence of mutant FUS are more structured. Finally, in accordance with the predictions, also RNA accessibility data derived from DMS-MapSeq methodology (56) confirmed less single stranded regions in the RNAs of LOSS group compared to the GAIN ones (Figure 4G and Supplementary Figure S4M).

Furthermore, we wondered whether this effect also reflects on non-canonical RNA structures such as G-quadruplexes. To answer this question, we used data from rG4-Seq1 (58) and rG4-Seq2 (59). These techniques are based on the stabilization of these structures by K+ or K+PDS (pyridostatin), which cause stops in the reverse transcription (RT) of RNA, subsequently detected through RNA sequencing. First, we asked whether the stable G-quadruplex structures were enriched in RNAs localized in SG. Thus, we analyzed the no DOXY control condition comparing '*depleted*', '*enriched*' and all the others RNAs (*neither*). Interestingly, we observed that the more stable G4 structures, which show a higher fraction of reads with RT stops, are enriched in RNAs engaged in SG compared to the other groups (Supplementary Figure S4N and O). This finding was also confirmed in rG4-Seq2, where the more stable G4s (the canonical and long loops), were more represented in the SG enriched group (Supplementary Figure S4P).

Notably, we confirmed the same results also using a set of high-confidence G-quadruplexes consensus in the two experiments (Supplementary Figure S4Q). Moreover, to understand if FUS[P525L] SG are less prone to contain G-quadruplex structures, we compared the types of G-quadruplexes between GAIN and LOSS RNAs in FUS[P525L] condition and found that the more stable types, the canonical ones, were more represented in the LOSS category, in accordance with the higher GC-content of this group (Supplementary Figure S4R).

Altogether, these data suggest that FUS[P525L]-mediated alteration of SG has a strong impact on the global GC content of the granules, and that the RNAs specifically lost in this condition are more structured, while the recruited ones are less prone to generate secondary structures.

Role of FUS[P525L] direct RNA binding in the reshaping of the SG transcriptome

In order to investigate the causal link between the presence of mutant FUS and the observed alterations in SG transcriptome, we focused on its RNA interactome. Indeed, since mutant FUS is almost completely localized in SG and directly binds RNA, we expected an enrichment of FUS interacting transcripts in FUS[P525L] SG compared to the WT condition.

We investigated this possibility by taking advantage of FUS[P525L] HITS-CLIP performed in SK-N-BE cells (61). Using Piranha software (62) to identify RNA regions enriched in CLIP tags compared to INPUT samples (peak calling per replicate FDR < 0.05 and $\log_2 OR$ > 1, reproducible peaks selected using IDR < 0.05) we found 9963 binding regions in 2652 interactors (Supplementary Table S4). As expected, we found that FUS[P525L] interacts mainly with mRNAs (Supplementary Figure S5A) and mainly in their 3′UTR (Supplementary Figure S5B).

Notably, a 'Signal saturation analysis' was performed by comparing the efficiency of peak calling using different levels of CLIP tags downsampling. This analysis allowed us to conclude that the sequencing depth sufficiently saturates the signal, enabling us to identify an optimal number of RNA interactors (Supplementary Figure S5C). Moreover, our FUS[P525L] HITS-CLIP signals well correlated with signals of mutant FUS PAR-CLIP (77) obtained with both the endogenous and the over-expressed protein (Supplementary Figure S5D). After these quality checks, we decided to investigate the contribution of FUS[P525L] direct interactors to the SG transcriptome. We found that even though mutant FUS direct interactors are enriched in the aberrant granules, they represent no more than the ∼ 55% of their components, suggesting that FUS direct binding alone is not sufficient to explain the differential recruitment of transcripts observed in mutant SG (Figure 5A). This evidence was further supported by the analysis of FUS[P525L] binding sites on FUS interactors. Indeed, we found that they are depleted for the AU-rich 7mers signature of the aberrant granule (Figure 5B) further supporting that FUS[P525L] alone is not sufficient to explain the transcriptome re-shape of the SG.

FUS[P525L] mutation alters the SG proteome

We then searched for other possible contributions to the remodeling of SG-associated RNAs, hypothesizing that the presence of mutant FUS could alter the proteome, thus indirectly impacting on the SG transcriptome. For this reason, aiming to characterize the protein content of granules in FUS[WT]

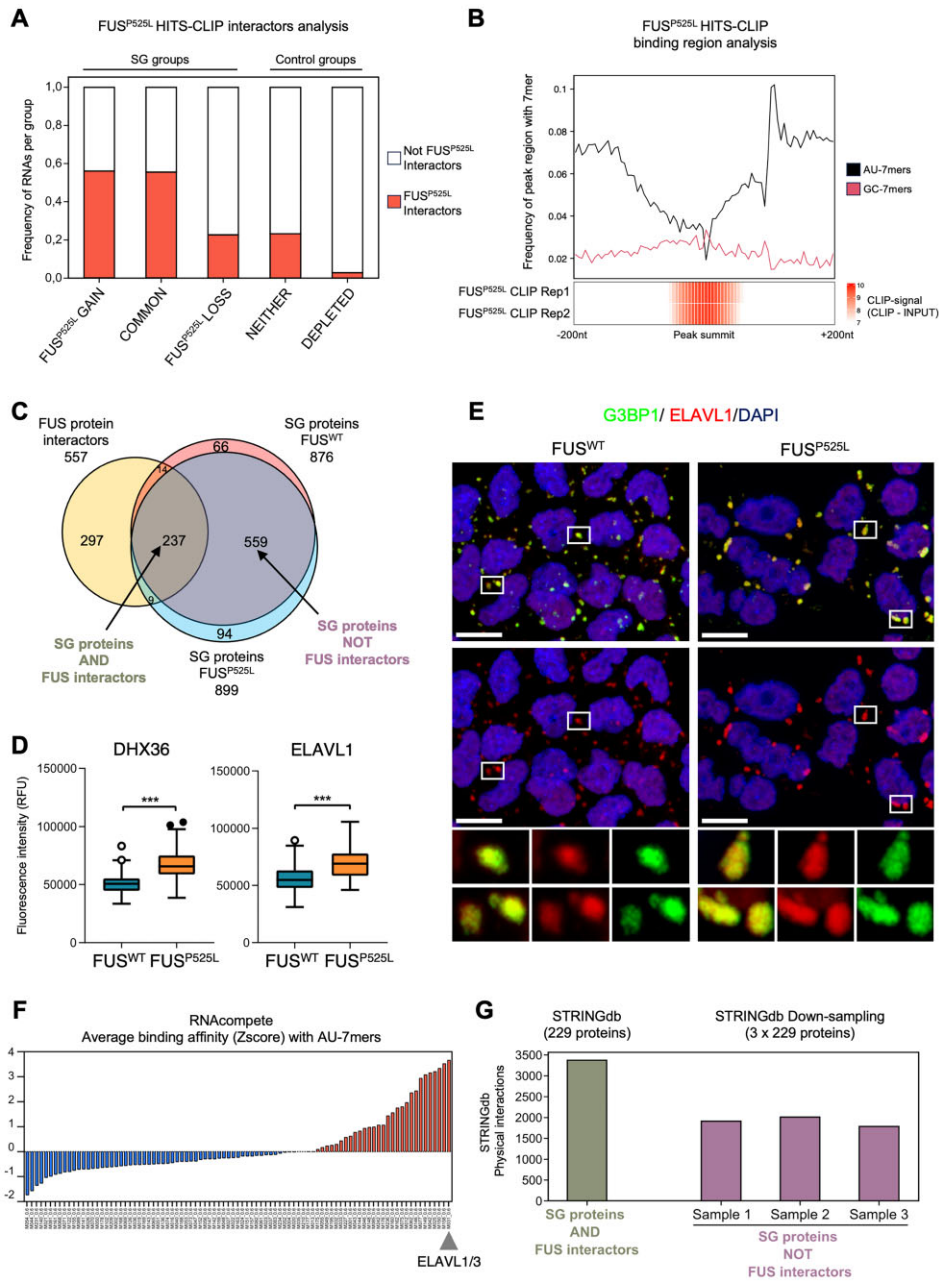**Figure 5.** (**A**) Barplot depicting the fraction of FUS^P525L HITS-CLIP interactors in the FUS^P525L GAIN, COMMON, FUS^P525L LOSS, NEITHER or DEPLETED transcript groups. X-axis reports the analyzed groups, while Y-axis reports the frequency of RNAs per group. Red color bars describe the FUS^P525L interactors, while white ones the not FUS^P525L interactors. For NEITHER and DEPLETED groups a common set of RNAs defined as *neither* or *depleted* in all the three datasets (no DOXY, FUS^P525L or FUS^WT) was taken in consideration for this analysis. (**B**) Meta-region profiles of FUS^P525L HITS-CLIP binding sites. Line plot in the top describes the frequency of windows in each position that contains AU-rich 7mers (black) or GC-rich 7mers (red). X-axis represents each analyzed position relative to the peak summit from -200nt to +200nt, while Y-axis reports the frequency of windows with 7mers. Heatmap in the bottom reports for each position the CLIP normalized signal of each biological replicate. Red color scale reports the intensity of the normalized signal. (**C**) Venn diagram depicting the overlap among SG proteomes in FUS^WT and FUS^P525L conditions and FUS protein interactors. Arrows indicate the fraction of proteins defined as "SG proteins AND FUS interactors" or "SG proteins NOT FUS interactors". (**D**) Boxplot depicting the distribution of the fluorescence intensity of DHX36 (left panel) and ELAVL1 (right panel) within TIAR and G3BP1-marked SG, derived from immunofluorescence analysis. For DHX36, 741 and 668 cells were counted for FUS^WT and FUS^P525L, respectively, from three independent biological replicates. For ELAVL1, 650 and 607 were counted for FUS^WT and FUS^P525L, respectively, from three independent biological replicates. Statistical significance was assessed with two-tailed, unpaired T-test. (**E**) Representative immunofluorescence for G3BP1 (green) and ELAVL1 (red) performed in SK-N-BE cells expressing FUS^WT and FUS^P525L under doxycycline control, subjected to 1 h of 0.5 mM sodium arsenite stress. Nuclei were counterstained with DAPI. Scale bar = 10 μm. Bottom panels show digital magnifications of the G3BP1 granules highlighted by squares. (**F**) Barplot depicting the average binding affinity (Z-score) for AU-rich 7mers in each RNAcompete experiment retrieved from CISBP-RNA database (*Homo sapiens*). X-axis reports the ID referred to each RNA compete experiment, while Y-axis reports the average Z-score for AU-7mers in the experiment. Red bars refer to experiments with positive average Z-scores, while blue bars refer to experiments with negative ones. A gray arrow indicates the RNAcompete experiment with the highest average Z-score for AU-rich 7mers, and the associated RBPs (ELAVL1/3). (**G**) Barplot depicting the enrichment of protein–protein physical interactions (PPI) in the "SG proteins AND FUS interactors" compared to the three samplings performed from the "SG protein NOT FUS interactors" group.

and FUS^{P525L} conditions, we performed SG purification followed by mass spectrometry, detecting 876 and 899 proteins in the two conditions, respectively (Supplementary Table S3). In order to assess the proper purification and detection of the SG proteome, we performed a GO enrichment analysis on the proteins identified in FUS^{WT} condition; since the over-expressed FUS^{WT} is almost exclusively nuclear, these samples could be used as reference control. As expected from already described SG proteomes (17,78), we found among the most enriched categories those related to translation, RNA processing, RNA binding and ribonucleoprotein granules (Supplementary Figure S5E and Supplementary Table S3).

Moreover, we compared the SG proteome in FUS^{WT} condition with other published SG proteome datasets from HEK293T cells (79) and U2OS cells (17), observing that a consistent portion of the SG localized proteins (77 common *set* proteins) are in common with all the three systems (Supplementary Figure S5F). Among the common *set* proteins, we found known SG components such as G3BP1, G3BP2 and CAPRIN1, helicases such as DHX36, RNA processing factors like XRN1 and CNOT1, and also RBPs such as STAU2, PUM1 and EWSR1 (Supplementary Figure S5G). Despite the common *set*, we also found proteins identified only in our dataset and not in the others (Supplementary Figure S5F), supporting the idea that the common SG *set* is accompanied by another group of proteins specific for the system under analysis (80).

Therefore, we studied the impact of FUS^{P525L} on the SG proteome by comparing the identified proteins from mass spectrometry data in FUS^{WT} or FUS^{P525L} conditions. Conversely to what we observed for SG transcriptome, we found strong similarities between the proteomes in both conditions. Indeed, among the 876 proteins detected in FUS^{WT} SG and the 899 proteins in FUS^{P525L} SG, 796 are shared by the two conditions (90.8% of FUS^{WT} and 89% of FUS^{P525L} respectively, Figure 5C), suggesting that the presence of mutant FUS does not cause a dramatic inclusion of new protein species in the SG. However, despite the qualitative similarity in SG protein contents, the increase of mutant FUS due to its cytoplasmic de-localization itself could have an impact on the quantitative recruitment of SG-associated proteins. Thus, in order to better understand the relationship between FUS protein interactors and the SG proteome, we took advantage of 5 published datasets of FUS protein interactors related to its soluble (81–83), condensed (84) or aggregated form (79). The latter dataset referred to a mis-localized and aggregated form of FUS carrying the R522G mutation, and displayed the highest amount of FUS protein interactors (488 proteins). Interestingly, across all datasets we found that a large fraction of FUS interactors is present in SG in both WT and mutant conditions, starting from a minimum of 40% up to 85% of overlap across the five interactomes analyzed (Figure 5C and Supplementary Figure S5H), meaning that a high portion of FUS protein interactors are already SG constituents. This evidence supports the hypothesis that mutant FUS is able to interact with a consistent set of proteins engaged in SG.

Interestingly, known FUS interactors, such as DHX36 (also known as RNA Helicase associated to AU-rich elements, RHAU) and ELAVL1, widely studied for their ability to bind AU-rich sequences (85,86), were found to be more abundant in SG in the presence of FUS^{P525L}, as demonstrated by western blot in comparison with another unperturbed SG marker (TIAR) (Supplementary Figure S5I). Since the expression of

these factors is comparable in the Input fractions of control and mutant conditions (Supplementary Figure S5J), their higher levels in FUS^{P525L} SG suggest that they can cooperate with FUS for the strong remodeling of SG transcriptome. To further prove their recruitment, we performed IF analysis confirming that both ELAVL1 and DHX36 are more abundant in FUS^{P525L} SG (Figure 5D and E, and Supplementary Figure S5K and L).

Making this result even more interesting is the fact that among the RBPs evaluated in *in vitro* RNA competition experiment (66), ELAVL1 displayed the highest affinity *in vitro* for the set of AU-rich 7mers of the aberrant granules (Figure 5F and Supplementary Table S4). Moreover, the comparison of binding affinity with the GC-rich 7mers demonstrated a high specificity of binding for the 7mer signature of FUS^{P525L} SG (Supplementary Figure S5M), suggesting that its increased recruitment may contribute to the alteration of the GC content in the granule.

We also focused on the behavior of another RBP relevant in the ALS pathology, TDP43. Indeed, in FUS positive protein aggregates observed in ALS patients, TDP43 was not detected (87). This evidence prompted us to investigate whether the aberrant granule could recapitulate this behavior of pathological aggregates found in ALS patients. Indeed, performing IF analyses, we observed that TDP43 is quite completely nuclear and excluded from the FUS^{P525L} SG (Supplementary Figure S5N and O). Furthermore, analysis of TDP43 direct RNA interactors, identified in the seCLIP performed in MNs (88), revealed that they are depleted in the GAIN RNA category, compared to the LOSS one. Notably, this finding was also confirmed by comparing RNA groups with similar lengths sampled from GAIN and LOSS RNAs (Supplementary Figure S5P and Q).

To check whether RNAs characterizing the FUS^{P525L} GAIN group (those recruited in SG specifically upon FUS mutation) are preferentially bound by FUS protein interactors, we took advantage of RNAact database of protein–RNA interaction predictions based on catRAPID algorithm (67). To do that, we ranked the SG proteome for the preferential binding with the GAIN RNA group compared to the LOSS one (Supplementary Figure S5R). Interestingly, among the SG proteins that are more prone to interact with the GAIN RNAs, we found an enrichment of proteins that interact with FUS (Supplementary Figure S5R and Supplementary Table S4). We noticed a similar enrichment of FUS interactors as preferential binders of the GAIN RNAs over the LOSS ones even when we performed the same analysis using experimental data of RBP–RNA interactions from POSTAR3 (CLIPdb) database (68) (Supplementary Figure S5S and Supplementary Table S4). This evidence suggests that FUS^{P525L} mis-localization and increased cytoplasmic concentration favors the SG recruitment of both its RNA and protein interactors; moreover, the increased engagement of FUS protein interactors into SG could be responsible for the alteration of SG transcriptome by driving the inclusion of RNAs of the FUS^{P525L} GAIN group.

Finally, to more deeply investigate the SG solidification phenomenon, we analyzed the impact of FUS^{P525L} inclusion in SG on the interaction propensity of the SG proteome. Thus, we compared the protein–protein physical interactions (PPI, STRING database) that are expected to occur in the group of SG FUS interactors (named 'SG proteins AND FUS interactors') and in the remaining fraction of SG proteins (named 'SG proteins NOT FUS interactors'). Since these two

groups have different amounts of proteins (237 and 559, respectively; Figure 5C), we performed three downsampling of 'SG proteins NOT FUS interactors' group and compared the PPI of 'SG proteins AND FUS interactors' with those of the downsampled groups to obtain a more reliable comparison (Supplementary Figure S5T and U). Interestingly, the SG proteins that don't interact with FUS show around 40% less PPIs than FUS interactors (Figure 5G and Supplementary Figure S5V). This evidence indicates that mutant FUS localized in SG could affect the hub of protein–protein interactions of a huge set of SG components.

Taken together, our data suggest that the recruitment in SG of mutant FUS produces alteration of the transcriptome both through engagement of FUS direct RNA interactors, as well as by the recruitment of FUS protein interactors which in turn could convey in SG their bound RNAs; moreover, we suggest that the inclusion of FUS^{P525L} protein partners could reshape the protein–protein interaction landscape inside the granules.

Discussion

In this study, we addressed how mutations in the ALS-associated FUS protein affect SG molecular composition. Combining granule purification with high-throughput RNA sequencing, we characterized the composition of SG RNAs in human neuroblastoma cells (SK-N-BE), highlighting the recruitment of neuronal-specific RNAs. Also in our conditions, length is the main feature of the SG-enriched RNAs, in agreement with the description of SG RNA composition in other biological systems (43,49). Since longer RNA molecules usually harbor an increased number of possible interaction sites with both other RNAs and proteins, this feature emphasizes the role of multiple non-specific RNA interactions in the phase separation process (8). Moreover, we described the GC content as a diversifying feature of SG RNA in different cell types: the high GC content of neuronal SG could give them peculiar properties in term of RNA–protein interaction networks (89).

Moreover, using SK-N-BE cells expressing FUS^{WT} and FUS^{P525L}, we were able to reveal important physical alterations of SG condensates in which mutant FUS is embedded. Indeed, recruitment of FUS^{P525L} to SG increases their number and alters their molecular dynamics, leading to more aggregated SG with reduced physical mobility.

Notably, these more solid pathological granules also resulted to be long lasting after stress removal, with their persistence in the cytoplasm well correlating with the inability of the cell to recover physiological levels of translation. Other studies have pointed out the effect of mutant FUS on global translation even in absence of stress stimuli, acting through alteration of nonsense-mediated decay (73) or deregulation of FMRP (74). The alteration of translation associated with a longer retention of RNAs into SG could be an important link between the presence of mutant FUS and aberrant proteostasis observed in ALS pathology (90).

Furthermore, RNA-seq analyses revealed that FUS^{P525L} altered the RNA composition of SG, resulting in the inclusion of a new set of transcripts (GAIN group) and the exclusion of others (LOSS group), if compared to FUS^{WT}. Among the re-localized RNAs we found several transcripts described to be linked to neurodegenerative phenotypes and ALS. This is of interest since GAIN and LOSS transcripts could both be affected by the presence of mutant FUS in the cytoplasm. On one hand, the GAIN RNAs are entrapped in granules that are more resistant to dissolution, and would suffer for the inability of FUS^{P525L} SG to be promptly dissolved upon stress release (61,75) and to rapidly recover translation. On the other hand, the LOSS RNAs would not be protected from stress-related damage since excluded from SG (91).

FUS mis-localization was also shown to strongly alter the GC-content of recruited RNAs, shifting towards AU-rich sequences; the observed alteration in nucleotide composition implies a different structural propensity of the SG transcriptome. Indeed, lost RNAs result more likely to form secondary structures compared to the gained ones. This GC divergence is important to take into consideration in pathological contexts, since the RNA nucleotide composition is a crucial property for interaction with other RNAs and proteins (92) and its alteration could be at the base of the susceptibility to the irreversible aggregation process of these granules.

Finally, analyzing the FUS^{P525L} RNA interactome (61), we observed an enrichment of its direct RNA interactors, suggesting that a portion of transcripts is already associated with the mis-localized protein in the absence of stress, and that the formation of SG causes their engagement in the condensates. However, the extent of this enrichment cannot fully explain the transcriptome alteration observed in the SG, which in turn could be justified by the engagement of other RBPs, partners of FUS, that could indirectly contribute to the observed RNA compositional changes. Despite the fact that we did not observe a strong qualitative difference in the proteome components of SG in presence of mutant FUS, we found that some protein interactors of FUS are more engaged into SG and that they could convey their RNA targets into these granules. Therefore, the modification of the SG transcriptome could result from a double contribution of mutant FUS, which directly binds a portion of the SG recruited RNAs, but also drives the engagement of a subset of protein partners and, indirectly, of their interacting RNAs.

Furthermore, we have also observed that the FUS-interacting protein components of SG exhibit a significantly higher potential to form protein–protein physical interactions; this feature has been previously documented for the aggregated state of FUS^{R522G} interactors (79). Moreover, the physical interactions considered for this analysis include those involved in the formation of protein complexes (STRING database) and extend beyond contacts simply involving the low-complexity domains (LCDs) of proteins. This observation further highlights the importance of protein–protein interactions in the formation of aggregates and less dynamic SG.

While it has been suggested that RNA itself possesses aggregating potential (93), and undoubtedly contributes to the phenomena of phase separation, recent evidence supports the idea that RNA can act as a buffer in neurotoxic aggregation processes. This has been observed for TDP43 and FUS, where the role of the RNA-binding domain and its interaction with RNA is vital in buffering LCD-driven aggregation (94).

From our observations, we suggest that the altered dynamics of SG in FUS^{P525L} conditions and the increased condensate aggregation state are associated with interesting molecular features: (i) re-localization of relevant neuronal transcripts in SG, (ii) reduced architectural capacity of SG RNAs, involving a decreased propensity to form secondary structures, and (iii) increase in protein–protein interactions. In conclusion, our results shed new light on the importance of RNA features in the aggregation process involving SG. This not only suggests that specific and distinct RNA features are linked to patholog-

ical granules and aggregation processes, but also introduces intriguing possibilities for understanding how these phenomena can impact neural physiology and be causatively linked to neurodegeneration.

Data availability

All software, links to websites or tools used for this work are referred to in the 'Materials and methods' section or in the figure legends. The high-throughput sequencing data generated in this study have been deposited in the GEO database under accession code: GSE244751. The raw data concerning the protein mass spectrometry are available at the PRIDE Database with the following identifier: PXD049234. Additional dedicated scripts developed for this work are available upon request.

Supplementary data

Supplementary Data are available at NAR Online.

Acknowledgements

Funding

Conflict of interest statement

G.V. has a personal financial interest (co-founder) in Genoa Instruments, Italy.

References

1. Banani,S.F., Lee,H.O., Hyman,A.A. and Rosen,M.K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.*, **18**, 285–298.
2. Das,S., Vera,M., Gandin,V., Singer,R.H. and Tutucci,E. (2021) Intracellular mRNA transport and localized translation. *Nat. Rev. Mol. Cell Biol.*, **22**, 483–504.
3. Hirose,T., Ninomiya,K., Nakagawa,S. and Yamazaki,T. (2023) A guide to membraneless organelles and their various roles in gene regulation. *Nat. Rev. Mol. Cell Biol.*, **24**, 288–304.
4. Brangwynne,C.P., Eckmann,C.R., Courson,D.S., Rybarska,A., Hoege,C., Gharakhani,J., Jülicher,F. and Hyman,A.A. (2009) Germline P granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, **324**, 1729–1732.
5. Hyman,A.A., Weber,C.A. and Jülicher,F. (2014) Liquid–liquid phase separation in biology. *Annu. Rev. Cell Dev. Biol.*, **30**, 39–58.
6. Mitrea,D.M. and Kriwacki,R.W. (2016) Phase separation in biology; functional organization of a higher order. *Cell Commun. Signal.*, **14**, 1.
7. Gomes,E. and Shorter,J. (2019) The molecular language of membraneless organelles. *J. Biol. Chem.*, **294**, 7115–7127.
8. Van Treeck,B. and Parker,R. (2018) Emerging roles for intermolecular RNA–RNA interactions in RNP assemblies. *Cell*, **174**, 791–802.
9. Garcia-Jove Navarro,M., Kashida,S., Chouaib,R., Souquere,S., Pierron,G., Weil,D. and Gueroui,Z. (2019) RNA is a critical element for the sizing and the composition of phase-separated RNA–protein condensates. *Nat. Commun.*, **10**, 3230.
10. Roden,C. and Gladfelter,A.S. (2021) RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.*, **22**, 183–195.
11. Bevilacqua,P.C., Williams,A.M., Chou,H.-L. and Assmann,S.M. (2022) RNA multimerization as an organizing force for liquid–liquid phase separation. *RNA*, **28**, 16–26.
12. Boeynaems,S., Alberti,S., Fawzi,N.L., Mittag,T., Polymenidou,M., Rousseau,F., Schymkowitz,J., Shorter,J., Wolozin,B., Van Den Bosch,L., *et al.* (2018) Protein phase separation: a new phase in cell biology. *Trends Cell Biol.*, **28**, 420–435.
13. Molliex,A., Temirov,J., Lee,J., Coughlin,M., Kanagaraj,A.P., Kim,H.J., Mittag,T. and Taylor,J.P. (2015) Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, **163**, 123–133.
14. Wiedner,H.J. and Giudice,J. (2021) It's not just a phase: function and characteristics of RNA-binding proteins in phase separation. *Nat. Struct. Mol. Biol.*, **28**, 465–473.
15. Glauninger,H., Wong Hickernell,C.J., Bard,J.A.M. and Drummond,D.A. (2022) Stressful steps: progress and challenges in understanding stress-induced mRNA condensation and accumulation in stress granules. *Mol. Cell*, **82**, 2544–2556.
16. Protter,D.S.W. and Parker,R. (2016) Principles and properties of stress granules. *Trends Cell Biol.*, **26**, 668–679.
17. Jain,S., Wheeler,J.R., Walters,R.W., Agrawal,A., Barsic,A. and Parker,R. (2016) ATPase-modulated stress granules contain a diverse proteome and substructure. *Cell*, **164**, 487–498.
18. Shelkovnikova,T.A., Dimasi,P., Kukharsky,M.S., An,H., Quintiero,A., Schirmer,C., Buée,L., Galas,M.-C. and Buchman,V.L. (2017) Chronically stressed or stress-preconditioned neurons fail to maintain stress granule assembly. *Cell Death. Dis.*, **8**, e2788–e2788.
19. Advani,V.M. and Ivanov,P. (2020) Stress granule subtypes: an emerging link to neurodegeneration. *Cell. Mol. Life Sci.*, **77**, 4827–4845.
20. Jeon,P. and Lee,J.A. (2021) Dr. Jekyll and Mr. Hyde? Physiology and pathology of neuronal stress granules. *Front. Cell Dev. Biol.*, **9**, 609698.
21. Wolozin,B. and Ivanov,P. (2019) Stress granules and neurodegeneration. *Nat. Rev. Neurosci.*, **20**, 649–666.

22. Nedelsky,N.B. and Taylor,J.P. (2019) Bridging biophysics and neurology: aberrant phase transitions in neurodegenerative disease. *Nat. Rev. Neurol.*, **15**, 272–286.

23. Volk,A.E., Weishaupt,J.H., Andersen,P.M., Ludolph,A.C. and Kubisch,C. (2018) Current knowledge and recent insights into the genetic basis of amyotrophic lateral sclerosis. *Med. Gen.*, **30**, 252–258.

24. Murakami,T., Qamar,S., Lin,J.Q., Schierle,G.S.K., Rees,E., Miyashita,A., Costa,A.R., Dodd,R.B., Chan,F.T.S., Michel,C.H., *et al.* (2015) ALS/FTD mutation-induced phase transition of FUS liquid droplets and reversible hydrogels into irreversible hydrogels impairs RNP granule function. *Neuron*, **88**, 678–690.

25. Murray,D.T., Kato,M., Lin,Y., Thurber,K.R., Hung,I., McKnight,S.L. and Tycko,R. (2017) Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. *Cell*, **171**, 615–627.

26. Kato,M. and McKnight,S.L. (2021) The low-complexity domain of the FUS RNA-binding protein self-assembles via the mutually exclusive use of two distinct cross-β cores. *Proc. Natl Acad. Sci. U.S.A.*, **118**, e2114412118.

27. Bosco,D.A., Lemay,N., Ko,H.K., Zhou,H., Burke,C., Kwiatkowski,T.J., Sapp,P., McKenna-Yasek,D., Brown,R.H. and Hayward,L.J. (2010) Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. *Hum. Mol. Genet.*, **19**, 4160–4175.

28. Lenzi,J., De Santis,R., de Turris,V., Morlando,M., Laneve,P., Calvo,A., Caliendo,V., Chiò,A., Rosa,A. and Bozzoni,I. (2015) ALS mutant FUS proteins are recruited into stress granules in induced pluripotent stem cells (iPSCs) derived motoneurons. *Dis. Model. Mech.*, **8**, 755–66.

29. Szewczyk,B., Günther,R., Japtok,J., Frech,M.J., Naumann,M., Lee,H.O. and Hermann,A. (2023) FUS ALS neurons activate major stress pathways and reduce translation as an early protective mechanism against neurodegeneration. *Cell Rep.*, **42**, 112025.

30. Martone,J., Mariani,D., Santini,T., Setti,A., Shamloo,S., Colantoni,A., Capparelli,F., Paiardini,A., Dimartino,D., Morlando,M., *et al.* (2020) SMaRT lncRNA controls translation of a G-quadruplex-containing mRNA antagonizing the DHX36 helicase. *EMBO Rep.*, **21**, e49942.

31. Yusa,K., Zhou,L., Li,M.A., Bradley,A. and Craig,N.L. (2011) A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl Acad. Sci.*, **108**, 1531–1536.

32. Morlando,M., Dini Modigliani,S., Torrelli,G., Rosa,A., Di Carlo,V., Caffarelli,E. and Bozzoni,I. (2012) FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment. *EMBO J.*, **31**, 4502–4510.

33. Slenders,E., Castello,M., Buttafava,M., Villa,F., Tosi,A., Lanzanò,L., Koho,S.V. and Vicidomini,G. (2021) Confocal-based fluorescence fluctuation spectroscopy with a SPAD array detector. *Light Sci. Appl.*, **10**, 31.

34. Perego,E., Zappone,S., Castagnetti,F., Mariani,D., Vitiello,E., Rupert,J., Zacco,E., Tartaglia,G.G., Bozzoni,I., Slenders,E., *et al.* (2023) Single-photon microscopy to study biomolecular condensates. *Nat. Commun.*, **14**, 8224.

35. Castello,M., Tortarolo,G., Buttafava,M., Deguchi,T., Villa,F., Koho,S., Pesce,L., Oneto,M., Pelicci,S., Lanzanó,L., *et al.* (2019) A robust and versatile platform for image scanning microscopy enabling super-resolution FLIM. *Nat. Methods*, **16**, 175–178.

36. Rossetta,A., Slenders,E., Donato,M., Zappone,S., Fersini,F., Bruno,M., Diotalevi,F., Lanzanò,L., Koho,S., Tortarolo,G., *et al.* (2022) The BrightEyes-TTM as an open-source time-tagging module for democratising single-photon microscopy. *Nat. Commun.*, **13**, 7406.

37. Wawrezinieck,L., Rigneault,H., Marguet,D. and Lenne,P.-F. (2005) Fluorescence correlation spectroscopy diffusion laws to probe the submicron cell membrane organization. *Biophys. J.*, **89**, 4029–4042.

38. Scipioni,L., Lanzanó,L., Diaspro,A. and Gratton,E. (2018) Comprehensive correlation analysis for super-resolution dynamic

39. Gaspar,I., Wippich,F. and Ephrussi,A. (2017) Enzymatic production of single-molecule FISH and RNA capture probes. *RNA*, **23**, 1582–1591.

40. Santini,T., Martone,J. and Ballarino,M. (2021) Visualization of Nuclear and Cytoplasmic Long Noncoding RNAs at Single-Cell Level by RNA-FISH. In: Bodega,B. and Lanzuolo,C. (eds.) *Capturing Chromosome Conformation: Methods and Protocols*. Springer US, NY, pp. 251–280.

41. Bahry,E., Breimann,L., Zouinkhi,M., Epstein,L., Kolyvanov,K., Mamrak,N., King,B., Long,X., Harrington,K.I.S., Lionnet,T., *et al.* (2022) RS-FISH: precise, interactive, fast, and scalable FISH spot detection. *Nat. Methods*, **19**, 1563–1567.

42. Khong,A., Jain,S., Matheny,T., Wheeler,J.R. and Parker,R. (2018) Isolation of mammalian stress granule cores for RNA-Seq analysis. *Methods*, **137**, 49–54.

43. Khong,A., Matheny,T., Jain,S., Mitchell,S.F., Wheeler,J.R. and Parker,R. (2017) The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Mol. Cell*, **68**, 808–820.

44. Bolger,A.M., Lohse,M. and Usadel,B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

45. Martin,M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.*, **17**, 10–12.

46. Dobin,A., Davis,C.A., Schlesinger,F., Drenkow,J., Zaleski,C., Jha,S., Batut,P., Chaisson,M. and Gingeras,T.R. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.

47. Anders,S., Pyl,P.T. and Huber,W. (2015) HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**, 166–169.

48. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.

49. Chen,S., Zhang,J. and Zhao,F. (2022) Screening linear and circular RNA transcripts from stress granules. *Genomics Proteomics Bioinformatics*, **21**, 886–893.

50. Liao,Y., Wang,J., Jaehnig,E.J., Shi,Z. and Zhang,B. (2019) WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.*, **47**, W199–W205.

51. Huang,S.-L., Wu,L.-S., Lee,M., Chang,C.-W., Cheng,W.-C., Fang,Y.-S., Chen,Y.-R., Cheng,P.-L. and Shen,C.-K.J. (2020) A robust TDP-43 knock-in mouse model of ALS. *Acta Neuropathol. Commun.*, **8**, 3.

52. Schwenk,B.M., Hartmann,H., Serdaroglu,A., Schludi,M.H., Hornburg,D., Meissner,F., Orozco,D., Colombo,A., Tahirovic,S., Michaelsen,M., *et al.* (2016) TDP-43 loss of function inhibits endosomal trafficking and alters trophic signaling in neurons. *EMBO J.*, **35**, 2350–2370.

53. Gimenez,J., Spalloni,A., Cappelli,S., Ciaiola,F., Orlando,V., Buratti,E. and Longone,P. (2023) TDP-43 epigenetic facets and their neurodegenerative implications. *Int. J. Mol. Sci.*, **24**, 13807.

54. Tareen,A. and Kinney,J.B. (2020) Logomaker: beautiful sequence logos in python. *Bioinformatics*, **36**, 2272–2274.

55. Lorenz,R., Bernhart,S.H., Höner zu Siederdissen,C., Tafer,H., Flamm,C., Stadler,P.F. and Hofacker,I.L. (2011) ViennaRNA package 2.0. *Algorithms Mol. Biol.*, **6**, 26.

56. Rouskin,S., Zubradt,M., Washietl,S., Kellis,M. and Weissman,J.S. (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature*, **505**, 701–705.

57. Li,P., Zhou,X., Xu,K. and Zhang,Q.C. (2021) RASP: an atlas of transcriptome-wide RNA secondary structure probing data. *Nucleic Acids Res.*, **49**, D183–D191.

58. Kwok,C.K., Marsico,G., Sahakyan,A.B., Chambers,V.S. and Balasubramanian,S. (2016) rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat. Methods*, **13**, 841–844.

fingerprinting of cellular compartments using the Zeiss Airyscan detector. *Nat. Commun.*, **9**, 5120.

59. Zhao,J., Chow,E.Y.-C., Yeung,P.Y., Zhang,Q.C., Chan,T.-F. and Kwok,C.K. (2022) Enhanced transcriptome-wide RNA G-quadruplex sequencing for low RNA input samples with rG4-seq 2.0. *BMC Biol.*, **20**, 257.

60. Quinlan,A.R. and Hall,I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.

61. Di Timoteo,G., Giuliani,A., Setti,A., Biagi,M.C., Lisi,M., Santini,T., Grandioso,A., Mariani,D., Castagnetti,F., Perego,E., *et al.* (2024) M6A reduction relieves FUS-associated ALS granules. *Nat. Commun.*, **15**, 5033.

62. Uren,P.J., Bahrami-Samani,E., Burns,S.C., Qiao,M., Karginov,F.V., Hodges,E., Hannon,G.J., Sanford,J.R., Penalva,L.O.F. and Smith,A.D. (2012) Site identification in high-throughput RNA–protein interaction data. *Bioinformatics*, **28**, 3013–3020.

63. Li,Q., Brown,J.B., Huang,H. and Bickel,P.J. (2011) Measuring reproducibility of high-throughput experiments. *Ann. Appl. Stat.*, **5**, 1752–1779.

64. Li,H., Handsaker,B., Wysoker,A., Fennell,T., Ruan,J., Homer,N., Marth,G., Abecasis,G., Durbin,R. and 1000 Genome Project Data Processing Subgroup1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

65. Orsburn,B.C. (2021) Proteome discoverer—a community enhanced data processing suite for protein informatics. *Proteomes*, **9**, 15.

66. Ray,D., Kazan,H., Cook,K.B., Weirauch,M.T., Najafabadi,H.S., Li,X., Gueroussov,S., Albu,M., Zheng,H., Yang,A., *et al.* (2013) A compendium of RNA-binding motifs for decoding gene regulation. *Nature*, **499**, 172–177.

67. Lang,B., Armaos,A. and Tartaglia,G.G. (2019) RNAct: protein–RNA interaction predictions for model organisms with supporting experimental data. *Nucleic Acids Res.*, **47**, D601–D606.

68. Zhao,W., Zhang,S., Zhu,Y., Xi,X., Bao,P., Ma,Z., Kapral,T.H., Chen,S., Zagrovic,B., Yang,Y.T., *et al.* (2022) POSTAR3: an updated platform for exploring post-transcriptional regulation coordinated by RNA-binding proteins. *Nucleic Acids Res.*, **50**, D287–D294.

69. Kim,S.H., Nichols,K.D., Anderson,E.N., Liu,Y., Ramesh,N., Jia,W., Kuerbis,C.J., Scalf,M., Smith,L.M., Pandey,U.B., *et al.* (2023) Axon guidance genes modulate neurotoxicity of ALS-associated UBQLN2. *eLife*, **12**, e84382.

70. Chiò,A., Restagno,G., Brunetti,M., Ossola,I., Calvo,A., Mora,G., Sabatelli,M., Monsurrò,M.R., Battistini,S., Mandrioli,J., *et al.* (2009) Two Italian kindreds with familial amyotrophic lateral sclerosis due to FUS mutation. *Neurobiol. Aging*, **30**, 1272–1275.

71. Lo Bello,M., Di Fini,F., Notaro,A., Spataro,R., Conforti,F.L. and La Bella,V. (2017) ALS-related mutant FUS protein is mislocalized to cytoplasm and is recruited into stress granules of fibroblasts from asymptomatic FUS P525L mutation carriers. *Neurodegener. Dis.*, **17**, 292–303.

72. Kedersha,N.L., Gupta,M., Li,W., Miller,I. and Anderson,P. (1999) RNA-binding proteins Tia-1 and Tiar link the phosphorylation of Eif-2α to the assembly of mammalian stress granules. *J. Cell Biol.*, **147**, 1431–1442.

73. Kamelgarn,M., Chen,J., Kuang,L., Jin,H., Kasarskis,E.J. and Zhu,H. (2018) ALS mutations of FUS suppress protein translation and disrupt the regulation of nonsense-mediated decay. *Proc. Natl Acad. Sci. U.S.A.*, **115**, E11904–E11913.

74. Birsa,N., Ule,A.M., Garone,M.G., Tsang,B., Mattedi,F., Chong,P.A., Humphrey,J., Jarvis,S., Pisiren,M., Wilkins,O.G., *et al.* (2021) FUS-ALS mutants alter FMRP phase separation equilibrium and impair protein translation. *Sci. Adv.*, **7**, eabf8660.

75. D'Ambra,E., Santini,T., Vitiello,E., D'Uva,S., Silenzi,V., Morlando,M. and Bozzoni,I. (2021) Circ-Hdgfrp3 shuttles along neurites and is trapped in aggregates formed by ALS-associated mutant FUS. *iScience*, **24**, 103504.

76. Delli Ponti,R., Marti,S., Armaos,A. and Tartaglia,G.G. (2017) A high-throughput approach to profile RNA structure. *Nucleic Acids Res.*, **45**, e35.

77. De Santis,R., Alfano,V., de Turris,V., Colantoni,A., Santini,L., Garone,M.G., Antonacci,G., Peruzzi,G., Sudria-Lopez,E., Wyler,E., *et al.* (2019) Mutant FUS and ELAVL4 (HuD) aberrant crosstalk in amyotrophic lateral sclerosis. *Cell Rep.*, **27**, 3818–3831.

78. Asadi,M.R., Sadat Moslehian,M., Sabaie,H., Jalaiei,A., Ghafouri-Fard,S., Taheri,M. and Rezazadeh,M. (2021) Stress granules and neurodegenerative disorders: a scoping review. *Front. Aging Neurosci.*, **13**, 650740.

79. An,H., Litscher,G., Watanabe,N., Wei,W., Hashimoto,T., Iwatsubo,T., Buchman,V.L. and Shelkovnikova,T.A. (2022) ALS-linked cytoplasmic FUS assemblies are compositionally different from physiological stress granules and sequester hnRNPA3, a novel modifier of FUS toxicity. *Neurobiol. Dis.*, **162**, 105585.

80. Markmiller,S., Soltanieh,S., Server,K.L., Mak,R., Jin,W., Fang,M.Y., Luo,E.-C., Krach,F., Yang,D., Sen,A., *et al.* (2018) Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell*, **172**, 590–604.

81. Groen,E.J.N., Fumoto,K., Blokhuis,A.M., Engelen-Lee,J., Zhou,Y., van den Heuvel,D.M.A., Koppers,M., van Diggelen,F., van Heest,J., Demmers,J.A.A., *et al.* (2013) ALS-associated mutations in FUS disrupt the axonal distribution and function of SMN. *Hum. Mol. Genet.*, **22**, 3690–3704.

82. Blokhuis,A.M., Koppers,M., Groen,E.J.N., van den Heuvel,D.M.A., Dini Modigliani,S., Anink,J.J., Fumoto,K., van Diggelen,F., Snelting,A., Sodaar,P., *et al.* (2016) Comparative interactomics analysis of different ALS-associated proteins identifies converging molecular pathways. *Acta Neuropathol.*, **132**, 175–196.

83. Kamelgarn,M., Chen,J., Kuang,L., Arenas,A., Zhai,J., Zhu,H. and Gal,J. (2016) Proteomic analysis of FUS interacting proteins provides insights into FUS function and its role in ALS. *Biochim. Biophys. Acta*, **1862**, 2004–2014.

84. Reber,S., Jutzi,D., Lindsay,H., Devoy,A., Mechtersheimer,J., Levone,B.R., Domanski,M., Bentmann,E., Dormann,D., Mühlemann,O., *et al.* (2021) The phase separation-dependent FUS interactome reveals nuclear and cytoplasmic function of liquid–liquid phase separation. *Nucleic Acids Res.*, **49**, 7713–7731.

85. Tran,H., Schilling,M., Wirbelauer,C., Hess,D. and Nagamine,Y. (2004) Facilitation of mRNA deadenylation and decay by the exosome-bound, DExH protein RHAU. *Mol. Cell*, **13**, 101–111.

86. Gallouzi,I.-E., Brennan,C.M., Stenberg,M.G., Swanson,M.S., Eversole,A., Maizels,N. and Steitz,J.A. (2000) HuR binding to cytoplasmic mRNA is perturbed by heat shock. *Proc. Natl Acad. Sci.*, **97**, 3073–3078.

87. Ling,S.-C., Polymenidou,M. and Cleveland,D.W. (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron*, **79**, 416–438.

88. Markmiller,S., Sathe,S., Server,K.L., Nguyen,T.B., Fulzele,A., Cody,N., Javaherian,A., Broski,S., Finkbeiner,S., Bennett,E.J., *et al.* (2021) Persistent mRNA localization defects and cell death in ALS neurons caused by transient cellular stress. *Cell Rep.*, **36**, 109685.

89. Sanchez de Groot,N., Armaos,A., Graña-Montes,R., Alriquet,M., Calloni,G., Vabulas,R.M. and Tartaglia,G.G. (2019) RNA structure drives interaction with proteins. *Nat. Commun.*, **10**, 3246.

90. McAlary,L., Chew,Y.L., Lum,J.S., Geraghty,N.J., Yerbury,J.J. and Cashman,N.R. (2020) Amyotrophic lateral sclerosis: proteins, proteostasis, prions, and promises. *Front. Cell. Neurosci.*, **14**, 581907.

91. Kong,Q. and Lin,C.G. (2010) Oxidative damage to RNA: mechanisms, consequences, and diseases. *Cell. Mol. Life Sci.*, **67**, 1817–1829.

92. Adinolfi,M., Pietrosanto,M., Parca,L., Ausiello,G., Ferrè,F. and Helmer-Citterich,M. (2019) Discovering sequence and structure

landscapes in RNA interaction motifs. *Nucleic Acids Res.*, **47**, 4958–4969.

93. Ripin,N. and Parker,R. (2022) Are stress granules the RNA analogs of misfolded protein aggregates? *RNA*, **28**, 67–75.

94. Mann,J.R., Gleixner,A.M., Mauna,J.C., Gomes,E., DeChellis-Marks,M.R., Needham,P.G., Copley,K.E., Hurtle,B., Portz,B., Pyles,N.J., *et al.* (2019) RNA binding antagonizes neurotoxic phase transitions of TDP-43. *Neuron*, **102**, 321–338.