Nucleocytoplasmic Proteomic Analysis Uncovers ETF1 and Nonsense Mediated Decay as Modifiers of ALS *C9orf72* Toxicity

Juan A. Ortega[1,*], Elizabeth L. Daley[1,*], Sukhleen Kour[2], Marisa Samani[1], Liana Tellez[1], Elizabeth A. Hall[1], Taylan Y. Esengul[1], Yung-Hsu Tsai[1], Tania F. Gendron[3], Christopher J. Donnelly[4,5], Teepu Siddique[1], Jeffrey N. Savas[1], Udai B. Pandey[2] and Evangelos Kiskinis[1,6,7,#]

[1]The Ken & Ruth Davee Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, 60611, USA

[2]Department of Pediatrics, Children's Hospital of Pittsburgh, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15224, USA

[3]Department of Neuroscience, Mayo Clinic, Jacksonville, FL, 32224, USA

[4]Department of Neurobiology, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15224, USA

[5]LiveLikeLou Center for ALS Research, University of Pittsburgh School of Medicine, Pittsburgh, PA, 15224, USA

[6]Department of Physiology, Northwestern University Feinberg School of Medicine, Chicago, IL, 60611, USA

[7]Lead Contact

*Equal Contribution

#Correspondence: evangelos.kiskinis@northwestern.edu

SUMMARY

The most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) is a hexanucleotide repeat expansion in *C9orf72* (C9-HRE). While RNA and dipeptide repeats produced by the C9-HRE disrupt nucleocytoplasmic transport, the proteins that become redistributed remain unknown. Here, we utilized subcellular fractionation coupled with tandem mass spectrometry and identified 126 proteins, enriched for protein translation and RNA metabolism pathways, which collectively drive a shift towards a more cytosolic proteome in C9-HRE cells. Amongst these was ETF1, which regulates translation termination and nonsense-mediated decay (NMD). ETF1 accumulates within elaborate nuclear envelope invaginations in patient iPSC-neurons and postmortem tissue and mediates a protective shift from protein translation to NMD-dependent mRNA degradation. Overexpression of ETF1 and the NMD-driver UPF1 ameliorate C9-HRE toxicity *in vivo*. Our findings provide a resource for proteome-wide nucleocytoplasmic alterations across neurodegeneration-associated repeat expansion mutations and highlight ETF1 and NMD as therapeutic targets in *C9orf72*-associated ALS/FTD.

KEYWORDS

Amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), C9orf72, ETF1, nonsense mediated decay (NMD), UPF1, nucleocytoplasmic proteome, iPSCs, motor neurons.

INTRODUCTION

A hexanucleotide repeat expansion (HRE) in the first intron of *C9orf72* is the most common genetic cause of ALS and FTD[1, 2]. It accounts for as much as 25% and 40% of familial ALS and FTD respectively, as well as 5-7% of all sporadic ALS cases. Healthy individuals typically carry less than 30 GGGGCC (G_4C_2) repeats, while ALS and FTD patients carry hundreds or thousands of repeats[3, 4]. The *C9orf72* HRE is proposed to drive neurodegeneration through three non-exclusive mechanisms[5-8]. Several reports have shown that *C9orf72* expression is decreased in patients suggesting that haploinsufficiency contributes to pathogenesis[9, 10]. Gain-of-function neurotoxic effects occur through the production of aberrant C9-HRE RNA molecules[11], as well as through dipeptide repeat proteins (DPRs) translated by a non-ATG dependent mechanism from the HRE[12, 13]. The repeat is bi-directionally transcribed and corresponding RNA sense and antisense molecules are detected by fluorescence in situ hybridization (FISH) as predominantly nuclear RNA foci in patient cells and tissue[11, 13, 14]. Moreover, five C9-HRE DPRs (GA, GP, GR, PR, PA) have been detected with antigen-specific antibodies in patient samples where they accumulate in cytoplasmic and nuclear aggregates in the frontal and motor cortex as well as the spinal cord[15, 16].

While the relative contribution of loss-of-function effects, as well as the dominant source of toxic gain-of-function effects (RNA or DPRs) is still an open-ended question, substantial and converging evidence suggests that the C9-HRE disrupts the balance of proteins that are transported between the nucleus and the cytoplasm. The long C9-RNA that is transcribed from the repeat expansion binds and sequesters RNA-binding proteins in the nucleus[11]. The arginine-rich GR and PR C9-DPRs bind and sequester a number of proteins with low complexity domains that assemble into membrane-less organelles through liquid-liquid phase separation[17, 18]. These include RNA granules, nucleoli and various components of the nuclear pore complex (NPC). Consequently, a series of groundbreaking genetic studies in C9-HRE *Drosophila* and yeast model systems identified nucleo-cytoplasmic (N/C) trafficking as a critical mediator of C9-HRE toxicity[19-21]. They specifically identified NPC proteins, as well as proteins that are involved in both nuclear import and export as significant modifiers of C9-HRE-associated toxicity. Importantly, the nature of the proteins that become redistributed or mislocalized as a result of these mechanisms remains unknown.

Abnormal N/C localization of proteins could also have broader relevance for sporadic disease, as cytoplasmic accumulation of TDP-43, a protein typically found in the nucleus, is a neuropathological hallmark in most ALS patients[22]. Recent studies suggest that cytoplasmic aggregation of TDP-43 interferes with N/C transport, and TDP-43 protein that accumulates in the cytoplasm is more toxic than TDP-43 that accumulates in the nucleus[23, 24]. At the same time, NPCs have been shown to exhibit an age-dependent deterioration in postmitotic neurons, highlighting a potential interaction between mutations, such as the C9-HRE, and age-dependent neurodegenerative mechanisms, such as NPC dysfunction[25-27].

In this study we sought to identify the proteins that are altered in their N/C distribution in C9-HRE carriers cells. We hypothesized that these alterations would highlight disease-relevant mechanisms. We reasoned that proteins can become translocated as a result of pathogenic sequestration within C9-RNA foci and C9-DPR aggregates, or as a result of protective protein redistribution to counteract the cellular defects driven by the mutation. To test this hypothesis, we developed an unbiased approach to interrogate proteome-wide N/C protein distribution, and determined how it is affected by the C9-HRE. We identified a number of proteins that become bidirectionally redistributed in C9-HRE cells, which are primarily implicated in RNA metabolism and protein synthesis. Amongst the proteins that become more nuclear was ETF1, which regulates the balance between protein translation termination and nonsense mediated decay

(NMD)-based mRNA degradation. In response to the C9 mutation, ETF1 triggers a protective shift from protein translation to UPF1-dependent NMD, and targets the expanded C9-HRE transcript for degradation. Our work strengthens the hypothesis that a major component of mutant *C9orf72* toxicity is related to mRNA processing defects, provides a novel link between the C9-HRE and the NMD pathway, and highlights ETF1 and UPF1 as novel therapeutic targets in *C9orf72*-associated ALS/FTD.

RESULTS

The *C9orf72* Repeat Expansion Leads to Proteome Wide Nucleo-Cytoplasmic Redistribution

In order to determine whether the HRE in *C9orf72* affects the N/C distribution of proteins, we developed a cellular model of C9-HRE toxicity. We used a series of expression constructs[19] that include either a control or disease-associated number of G_4C_2 repeat sequences (8 or 58 respectively) fused to GFP (Fig. 1a). The expression of (G_4C_2)x58 in mammalian cells has been shown to produce both RNA foci and DPRs[19]. We validated the production of the GP DPR in transfected HEK-293 cells by ELISA (Supplementary Fig. 1a) and that the size of the G_4C_2-GFP fusion protein corresponded to the respective number of repeats by Western blot (WB) analysis (Supplementary Fig. 1b). We introduced the (G_4C_2)x8, (G_4C_2)x58, or a GFP control vector in HEK-293 cells and used fluorescent activated cell sorting (FACS) to purify transfected cells (Fig. 1a and Supplementary Fig. 1c-e). We then performed biochemical subcellular fractionation to isolate nuclear (N) and cytosolic (C) protein extracts, which we analyzed by semi-quantitative mass spectrometry (MS)-based proteomics to evaluate the N/C ratio of proteins (Fig. 1a). We corroborated the precision of the subcellular fractionation in two ways: a) we performed WB analysis and confirmed that the nuclear marker TAFII and the cytoplasmic marker VINCULIN were each confined to the correct fraction (Fig. 1b); and b) we selected well-characterized nuclear (n=39) and cytosolic (n=68) resident proteins and analyzed their localization in our MS dataset (Supplementary Table 1). Cytosolic protein markers were enriched in the cytosolic fraction and exhibited a cytosolic ratio that was close to 1 (calculated as the ratio of cytosolic abundance over total abundance, i.e. cytosolic ratio = C / (C + N); see methods section), while nuclear markers were enriched in the nuclear fraction and exhibited a low cytosolic ratio (Fig. 1c and Supplementary Fig. 1f). Importantly, we found that at 48 hours after transfection the expression of (G_4C_2)x58 was not toxic, and thus selected this time point for all MS analysis (Supplementary Fig. 1g-h).

We identified approximately 10,000 redundant proteins in each one of the three samples (GFP, (G_4C_2)x8, (G_4C_2)x58), with no statistically significant differences in the average protein number (n=4 independent biological replicates, m=8 of N and C fractions per sample, $p = 0.116$) (Fig. 1d). To assess subcellular protein distribution, we calculated the average N/C ratio for all the proteins identified across experiments in the three samples. We then asked whether the expression of the C9-HRE affected the proteome-wide N/C distribution at a global level. We found that the average cytosolic ratio of the proteome shifted to a higher level in the presence of 58 repeats, significantly more so than in the presence of 8 repeats or control GFP expression (GFP *vs*. (G_4C_2)x58: $p < 0.0001$ and (G_4C_2)x8 *vs*. (G_4C_2)x58: $p < 0.0001$) (Fig. 1e). To investigate the effects of the C9-HRE on protein distribution by an alternative method we used a tdTomato reporter construct containing both a nuclear localization signal (NLS), and a nuclear export signal (NES) (Fig. 1f, top right). We first validated that we could effectively measure changes in the N/C ratio of the tdTomato signal in transfected HEK-293 cells. Under control conditions the signal was predominantly nuclear, and became even more nuclear (2-fold higher) after treatment with the nuclear export inhibitor KPT-330 (Supplementary Fig. 1i). We next asked how the C9-HRE

would affect the localization of the reporter and found that in the presence of 58 G_4C_2 repeats the tdTomato signal became significantly more cytosolic relative to both controls (n=3 independent biological replicates; GFP *vs*. (G_4C_2)x58: $p = 0.01$, 23% increase ± 8; and (G_4C_2)x8 *vs*. (G_4C_2)x58: $p = 0.03$, 36% increase ± 7) (Fig. 1f, left and bottom right). These results suggest that (G_4C_2)x58-expressing cells undergo global protein redistribution, with more proteins accumulating in the cytoplasm.

Repeat Expansion Mutations Lead to Proteome Wide Nucleo-Cytoplasmic Redistribution of Distinct Proteins

The shift in N/C ratio of the proteome that we detected in C9-HRE expressing cells prompted us to determine whether it represented a specific effect of this particular repeat expansion, or a nonspecific phenomenon that also occurs in response to other neurodegeneration-associated repeat expansion mutations. To address this question, we used constructs expressing a polyglutamine (CAG, polyQ) expansion in the Ataxin 1 gene (*ATX1*-84Q), or in the first exon of the Huntingtin gene (*HTT*-72Q) that cause spinocerebellar ataxia type 1 (SCA1) and Huntington's disease (HD), respectively. As controls, we used corresponding constructs with polyQ expansions below the threshold of disease (32Q and 25Q, respectively), as a well as a GFP control vector (Supplementary Fig. 1j, top). Following FACS purification, subcellular fractionation and MS analysis (Supplementary Fig. 1j, bottom, and k), we found that expression of polyQ repeats also caused a significant shift towards a more cytosolic proteome (n=3 independent biological replicates, m=6 of N and C fractions per sample, $p < 0.0001$) relative to controls (Fig. 1g). As in the case of the C9-HRE, there were no significant differences in the total number of proteins detected between each one of the five samples ($p = 0.979$) (Supplementary Fig. 1l).

We next examined the nature of the proteins that become significantly redistributed and found that there was very little overlap amongst all three repeat expansions, where only 2.3% and 2.2% of the proteins that shifted in mutant C9-(G_4C_2)x58-expressing cells were also respectively affected by polyQ-ATX1 or HTT (Fig. 1h and Supplementary Table 1). Moreover, although both the *ATX1* and *HTT* repeats represent polyQ sequences, the degree of overlap between redistributed proteins was similarly small (3.4%), possibly due to their distinct subcellular localization pattern (Fig. 1h, Supplementary Fig. 1j, bottom and Supplementary Table 1). While the nature for this differential effect is not clear, it could be associated with the biophysical or functional properties of each repeat expansion, its subcellular localization, or even the size of the repeat itself. Nevertheless, these results suggest that impaired N/C trafficking could be broadly implicated in repeat expansion-related toxicity, corroborating recent studies for HD and AD models[24, 28-30].

The *C9orf72* Repeat Expansion Causes a Redistribution of Proteins Involved in RNA Metabolism and Protein Translation

We next performed comprehensive statistical analysis and determined that only 126 unique proteins were significantly and consistently affected in their N/C ratio in cells expressing 58 G_4C_2 repeats (Kruskal–Wallis rank test; (G_4C_2)x58 *vs*. GFP and (G_4C_2)x8: $p < 0.05$) (Fig. 2a and Supplementary Table 1). These 126 proteins represented 1.2% of the total number of proteins that we identified (Fig. 2b, left). As expected, all 126 proteins showed a bidirectional change, with 56% relocating from the nucleus to the cytosol, and 44% shifting in the opposite direction (Fig. 2b, right). Approximately 40% of the proteins that relocated to the cytosol normally reside in the nucleus, whereas 30% of the proteins that relocated to the nucleus, normally reside in the cytosol[31], suggesting that these changes might have critical functional ramifications (Supplementary Fig. 2a). Analysis of proteome-wide N/C distribution in all three samples revealed that approximately 32% of all proteins were highly cytosolic (cytosolic ratio > 0.9), 34%

were highly nuclear (cytosolic ratio < 0.1), and the remaining 34% were found in both compartments (cytosolic ratio = 0.1 - 0.9) (Supplementary Fig. 2b). The proteins that were redistributed in cells expressing (G_4C_2)x58 repeats were not significantly enriched within any one specific cytosolic ratio bin. Importantly, there was also no clear correlation between protein abundance and subcellular redistribution upon the expression of (G_4C_2)x58 (Supplementary Fig. 2c).

We next asked whether the proteins affected by the C9-HRE were biased for ones that engage in active nuclear pore trafficking. Nucleocytoplasmic exchange of proteins can occur by active transport or passive diffusion through the NPC based on protein mass (Fig. 2c, left)[32]. Interrogating the 126 proteins for size revealed that 64% were larger than 40 kDa, and hence were more likely actively transported through the nuclear pore (Fig. 2c, right). Along the same lines, the majority (79%) contained either a putative NLS or NES or both (Supplementary Fig. 2d). To elucidate whether redistributed proteins comprised cargoes associated with specific karyopherins, we compared our dataset with previously published proteomic datasets describing transporter cargoes[33-35]. We specifically examined cargoes associated with CRM1, which is the major nuclear export receptor, as well as importin-α and transportin, which are two key nuclear import receptors. We found very little overlap between the proteins identified in our N/C analysis and cargoes associated with these carriers, probably on account of differences in methodology (Supplementary Fig. 2e and Supplementary Table 2). To determine whether our protein targets were translocated as a result of sequestration, we compared them with proteins known to interact with the C9-HRE-DPRs and C9-HRE-RNA[11, 17, 18, 36-40]. We found that 31% (43 proteins) overlapped with proteins previously identified as interacting with repeat expansion RNA and DPRs (Supplementary Fig. 2f and Supplementary Table 2), suggesting that sequestration by toxic C9-HRE products contributes in part, to the protein redistribution events we identified.

In order to gain further insight into the nature of the proteomic shift we asked whether there was a unifying functional component within the 126 proteins that are significantly redistributed. Gene ontology (GO) analysis revealed a striking enrichment for proteins involved in RNA metabolism, protein synthesis/translation and cell cycle dynamics (Fig. 2d-e and Supplementary Table 2). Collectively our findings indicate that a significant and highly specific redistribution of proteins takes place in cells carrying the C9-HRE that comprise major cellular functions including RNA and protein metabolism.

Redistributed Proteins are Genetic Modifiers of *C9orf72* Repeat Expansion Toxicity in *Drosophila* Models of Disease

Alterations in the N/C localization of proteins in the context of the C9-HRE could reflect aberrant protein translocation due to sequestration or impaired trafficking, but they could also represent a defensive cellular response. To assess the functional significance of some of the redistributed proteins in the context of an intact nervous system, we utilized established *Drosophila* models of *C9orf72* toxicity[37, 41]. We specifically used transgenic flies expressing 30 or 36 G_4C_2 repeats under the eye-specific GMR-GAL4 driver that results in a prominent degeneration of the eye. We rationally selected six proteins that exhibited a significant magnitude of change in their localization (>1.2 fold shift in (G_4C_2)x58-expressing cells relative to both controls), and were representative of the enriched cellular pathways identified in the GO analysis. These included: the mRNA transport protein ENY2, the arginine methyltransferase PRMT1, the RNA-binding protein SRSF1, which was recently linked to *C9orf72* mRNA trafficking[42], the nuclear import receptor TNPO3, the protein folding chaperone CCT8, and the translation termination factor ETF1 (Fig. 3a). We obtained RNAi lines for the fly homologues of these proteins and performed individual crosses into the G_4C_2 x 30 or x 36 fly (Fig. 3b). As controls we monitored the magnitude of target gene expression knockdown in the eye by RT-

qPCR (Supplementary Fig. 3a), as well as whether reducing these proteins alone, led to an eye degeneration phenotype (Supplementary Fig. 3b, top). We found that all six selected proteins were strong genetic modifiers of *C9orf72* toxicity *in vivo,* some acting as enhancers and others as suppressors of toxicity (n=12-18 flies/genotype, $p < 0.01$) (Fig. 3a-c and Supplementary Fig. 3b). While this analysis was limited to only six proteins, there was no clear indication that the direction of protein redistribution (i.e. more nuclear or more cytosolic) correlated with suppression or enhancement of toxicity in the fly models (Fig. 3a-c). Collectively, these experiments suggest that the proteome-wide analysis of protein localization in the heterologous C9-HRE expression model highlighted proteins that act as genetic interactors of *C9orf72* toxicity *in vivo.*

The eukaryotic translation termination factor 1 (ETF1), which plays an essential role in RNA metabolism and protein translation[43-46], was the strongest enhancer of *C9orf72* toxicity in the G_4C_2 x 30 fly (n=18 flies, $p < 0.0001$) (Fig. 3c). Given the enrichment of both of these pathways within the proteins affected by the C9-HRE in the N/C proteomic analysis (Fig. 2d-e), we next focused on the role of ETF1 in *C9orf72* toxicity. We first tested whether ETF1 acted as an enhancer of toxicity in the more potent G_4C_2 x 36 fly model. Reduction in *ETF1* expression significantly exacerbated eye degeneration (n=18 flies, $p < 0.0001$) (Fig. 3d-e), and also led to substantial motor dysfunction as compared to controls, which is indicative of neurodegeneration (n=43-83 flies, $p < 0.01$) (Supplementary Fig. 3c). An enhancer of toxicity in this assay likely mediates a protective effect in disease pathophysiology, as when its expression is ablated the toxicity related to the repeat expansion becomes more pronounced. In line with this hypothesis, when we conversely overexpressed ETF1 protein into the G_4C_2 x 36 fly (Supplementary Fig. 3d) we found that it significantly repressed eye degeneration (n=18-20 flies, $p < 0.0001$) (Fig. 3d-e). Intriguingly, these effects on eye degeneration were associated with a concurrent relative increase (after ETF1 reduction), or decrease (after ETF1 overexpression), in the abundance of the C9-GR dipeptide produced in the fly eye, as measured in a dot blot assay (Fig. 3f). These results suggest that ETF1 modulates C9-HRE toxicity *in vivo* by affecting either directly or indirectly the abundance of C9-DPRs.

ETF1 is Redistributed in *C9orf72* iPSC-Derived Motor Neurons and Postmortem *C9orf72* ALS Tissue

In order to validate the shift in ETF1 protein localization by an alternative approach, we next used subcellular fractionation coupled to WB analysis. We found that HEK-293 cells expressing the (G_4C_2)x58, indeed exhibited a significant shift in the N/C ratio of ETF1 protein (Supplementary Fig. 4a). To determine whether our findings on ETF1 redistribution hold significance in a more physiological disease model system, we used established protocols to differentiate human spinal motor neurons (MNs)[47] from iPSCs derived from three individual ALS patients with a *C9orf72* repeat expansion and from three healthy controls (Supplementary Fig. 4b and Supplementary Table 3). Subcellular N/C fractionation and WB analysis from differentiated day 50 MNs revealed a stronger nuclear enrichment of ETF1 protein in C9 cases than controls (Supplementary Fig. 4c). We then used immunocytochemistry (ICC) coupled to confocal microscopy to confirm the ETF1 redistribution at single cell resolution. We found that ETF1 exhibited a significantly higher N/C ratio in C9-ALS MNs relative to controls (day 50, n=3 patients per group, n=3 independent biological replicates; controls *vs* C9: $p<0.003$, 114% increase $\pm$ 34) (Fig. 4a-b). Consistent with the role of ETF1 in mRNA surveillance and translation, the majority of ETF1 in control MNs was localized in the cytosol and in close proximity to the nuclear membrane, while in mutant C9 MNs a proportion of ETF1 appeared nuclear (Fig. 4a-b and Supplementary Fig. 4c). Using 3D reconstruction of confocal images of neuronal nuclei we specifically found that the nuclear ETF1 protein amassed within discrete puncta (Fig. 4a). We

therefore tested the hypothesis that the nuclear accumulation of ETF1 was the result of aberrant sequestration by the nuclear C9-HRE products (RNA and/or DPRs). Using FISH coupled to ICC, we found that only 18% of nuclear C9-HRE RNA foci colocalized with ETF1 puncta and less than 1% of nuclear ETF1 puncta colocalized with C9-HRE RNA foci, suggesting that protein sequestration in RNA foci was at best a minor contributing factor to the accumulation of ETF1 in neuronal nuclei (Fig. 4c). At the same time, there was a strongly negative correlation between the nuclear ETF1 signal and GR or PR staining, also excluding the possibility that the C9-DPRs sequestered ETF1 protein in the nucleus (Fig. 4d).

Given that neither C9-DPRs nor C9-RNA foci co-localized with the nuclear ETF1 puncta, we examined whether ETF1 was closely associated with the nuclear membrane, which would result in it being fractionated within the nuclear extract in our MS analysis. Intriguingly, we found that the majority of ETF1 puncta (>70%) strongly colocalized with highly complex LMNB1+ nuclear envelope structures (n=3 patients, m > 15 neurons per patient) (Fig. 4e). Superresolution imaging by structured illumination microscopy (SIM) (Fig. 4f) and serial confocal imaging (Fig. 4g) revealed a stunning architecture of the neuronal nuclear envelope with LMNB1+ protrusions within the nucleus and ETF1 puncta within these structures. Importantly, all iPSC-derived MNs exhibited some level of nuclear LMNB1+ invaginations irrespective of their genotype.

In order to further substantiate these findings in the context of ALS patient neurons *in vivo*, we performed immunohistochemistry (IHC) in postmortem cortical and spinal cord tissue from three *C9orf72* ALS patients and three age-matched controls (Supplementary Table 3). We found a significantly higher N/C ratio of ETF1 in C9-ALS neurons (n=3 patients per group, m=28-68 neurons per patient; controls *vs* C9: $p < 0.0001$) (Fig. 5a,b and Supplementary Fig. 5). Intriguingly, we also found a significantly higher abundance of complex LMNB1+ nuclear envelope invaginations in C9 cases relative to controls (n=3 patients per group, m=31-76 neurons per patient; controls *vs* C9: $p < 0.0001$) (Fig. 5c,d). Importantly, the patient tissue reflected examples of neurons that exhibited a dramatic accumulation of ETF1 within their nuclei (Fig. 5a, bottom panel), as well as neurons that were reminiscent of patient-iPSC derived MNs, with focal ETF1 staining within LMNB1+ invaginations (Fig. 5c). Collectively these experiments show that a proportion of ETF1 is redistributed from the cytosol within nuclear envelope protrusions in mutant *C9orf72* patient neurons *in vitro* and *in vivo*.

ETF1 Mediates a Shift From Protein Translation to mRNA Degradation Through Nonsense Mediated Decay in C9-HRE Expressing Cells

ETF1 plays an essential role in two closely related cellular mechanisms, termination of protein translation, and nonsense mediated decay (NMD) of mRNA molecules[43-46]. During regular termination of protein synthesis, ETF1 recognizes stop codons, binds to the translating ribosome, and cleaves the nascent peptide[48-50]. In cases where ETF1 detects a premature stop codon that is upstream of an exon junction complex or an irregularly spliced mRNA, it triggers its degradation by recruiting UPF1, which becomes phosphorylated (pUPF1) and is considered the master regulator of NMD[44, 51-53].

Given the importance of ETF1 in adjusting the balance of these two cellular pathways, we next asked whether the redistribution of ETF1 protein affected the state of protein translation and NMD in mutant C9-expressing cells. We first used ICC to measure protein translation by analyzing the amount of the amino acid analog homopropargylglycine (HPG) incorporated into newly translated proteins at single-cell resolution (Supplementary Fig. 6a). We found a significant reduction in the relative amount of *de novo* protein synthesis in heterologous C9-(G_4C_2)x58 expressing cells relative to controls (n=2 independent biological replicates; GFP *vs*. (G_4C_2)x58: $p < 0.001$, 42% decrease $\pm$ 3; and (G_4C_2)x8 *vs*. (G_4C_2)x58: $p = 0.004$, 41% decrease $\pm$ 4) (Supplementary Fig. 6b-c). We next measured the basal rate of protein synthesis in C9

patient and control iPSC-derived MNs and at the same time asked whether ETF1 could in fact regulate this process in human neurons. ALS-C9 patient MNs exhibited significantly reduced levels of *de novo* protein synthesis relative to control MNs in accordance with previous reports[17, 54-56] (day 50, n=2 patients *vs* 1 control, n=3 independent biological replicates; siSCR control *vs* siSCR C9: $p < 0.0001$, 44% decrease ± 4) (Supplementary Fig. 6f). Critically, siRNA-mediated knockdown of ETF1 led to a significant reduction but not complete ablation of translation only in control MNs (siSCR *vs* siETF1 in controls $p < 0.029$) (Supplementary Fig. 6d-f).

In contrast to reduced protein synthesis, C9-ALS patient-derived MNs exhibited a pronounced induction of NMD as measured by the levels of pUPF1 normalized to total UPF1 protein, and evaluated by ICC at single cell resolution (day 50, n=3 patients per group, n=4 independent biological replicates; controls *vs* C9: $p < 0.007$, 38% increase ± 13) (Fig. 6a-b). Critically, C9-ALS cultures showed a striking increase in the proportion of MNs with high levels of pUPF1 relative to controls, indicating actuated NMD (from 25% in controls to 41% in patients, n > 217 MNs per individual) (Fig. 6c). We also examined the protein levels of SMG1, which is the canonical kinase responsible for phosphorylating UPF1, and observed a significant increase in C9-ALS MN cultures relative to controls (n=6 independent biological replicates; controls *vs* C9: $p < 0.05$, 31% increase ± 11) (Fig. 6d). Critically, the activation of NMD was ETF1-dependent as knockdown of ETF1 in either HEK-293 cells (Supplementary Fig. 6g-h) or patient-derived MNs (Fig. 6e) significantly diminished the levels of pUPF1 ($p < 0.0001$). These results indicate that C9 patient neurons exhibit significant changes in the two primary functional pathways that the ETF1 protein coregulates. They specifically display a reduction in protein translation and a concomitant ETF1-dependent increase in NMD of mRNA transcripts.

In order to directly assess whether NMD was indeed functional in C9-HRE expressing cells, we used a reporter system[57] that can effectively monitor NMD efficiency based on the ratio of a transcript with or without a premature termination codon (PTC) by WB analysis (Fig. 6f). We found that the expression of C9-(G_4C_2)x58 repeats hindered the efficiency of NMD as the ratio of PTC to control transcript was significantly increased relative to cells expressing 8 repeats or GFP alone (n=4 independent biological replicates; GFP *vs*. (G_4C_2)x58: $p < 0.0001$; and (G_4C_2)x8 *vs*. (G_4C_2)x58: $p = 0.004$) (Fig. 6g and Supplementary Fig. 6i). Critically, the degradation of the transcript with the premature stop codon was ETF1-dependent as this ratio significantly increased after effectively knocking down ETF1 (n=4 independent biological replicates; (G_4C_2)x58: siSCR *vs*. siETF1 $p = 0.003$) (Fig. 6h). Thus, while NMD is activated in an ETF1-dependent manner in C9-HRE cells, the repeat expansion hinders its efficiency.

***C9orf72* mRNA is a Targeted for Degradation by Nonsense Mediated Decay**

Collectively, our data suggest that in response to the expression of the C9-HRE, ETF1 mediates a shift from protein translation to NMD-dependent mRNA degradation (Figs. 5 and 6). This shift likely confers a protective effect as ETF1 reduction enhances C9 toxicity, while ETF1 overexpression suppresses C9 toxicity and the production of GR DPR *in vivo* (Fig. 3). It is also accompanied by the accumulation of ETF1 protein in elaborate nuclear envelope invaginations (Fig. 4). Importantly, we found evidence for protein synthesis based on the detection of puromycin, as well as a strong colocalization of ETF1 and pUPF1 within these LMNB1+ nuclear invaginations (Supplementary Fig. 7a-b), suggesting that a proportion of mRNA surveillance takes place within these structures.

Previous reports have shown that the C9-HRE increases the retention of the first intron of *C9orf72*[58]. Thus it is likely that this intron retention, in combination with the high GC content of the HRE, earmarks the expanded G_4C_2 C9 transcript as a target for NMD-dependent degradation[59, 60]. To test this hypothesis, we employed two approaches. We first established that we could effectively measure the ratio of the mutant to wild type C9 transcript based on intron

retention by semi-quantitative RT-PCR (Fig. 7a-b and Supplementary Fig. 7c-d). *C9orf72* patient-derived MNs exhibited a significantly higher intron retention ratio than control MNs (day 50, n=3 patients per group, n=1-3 independent biological replicates; controls *vs* C9: $p = 0.0046$) (Fig. 7a and Supplementary Fig. 7c-d). This ratio was further enhanced when *ETF1* expression was reduced (n=3 patients, n=6 independent biological replicates; siSCR *vs*. siETF1: $p = 0.014$) (Fig. 7b and Supplementary Fig. 7c-e). Given that NMD degrades mRNAs as they exit the nuclear pore complex[52], we next asked whether the absence of ETF1 would affect the number or localization of C9-G_4C_2 RNA foci, a prototypical pathological marker of *C9orf72*-associated ALS. Indeed, we found that knocking down *ETF1* expression in C9-MNs led to a dramatic reduction of nuclear C9 RNA foci and a concomitant increase in the number of cytosolic foci (day 50, n=3 patients, n=3 independent biological replicates; siSCR *vs*. siETF1 nuclear foci: $p < 0.0001$; cytosolic foci: $p < 0.0001$) (Fig. 7b-c). These findings suggest that in the absence of ETF1-dependent degradation, the expanded *C9orf72* transcript exits the nucleus more efficiently and accumulates in the cytoplasm. They are also in accordance with the effects of ETF1 on the production of the C9-GR DPR *in vivo* as the increased availability of cytosolic expanded *C9orf72* would allow for more potent DPR translation.

Lastly, to directly assess the role and significance of NMD and specifically UPF1 in *C9orf72* toxicity *in vivo,* we crossed *UPF1* overexpressing and RNAi lines with the G_4C_2 x 30 and x 36 fly models. RNAi mediated knockdown or overexpression of UPF1 had minimal effects on eye degeneration in control flies. However, we found that UPF1 reduction significantly enhanced eye degeneration (n=2-18 flies/genotype, $p < 0.01$), while UPF1 overexpression significantly ameliorated (n=18 flies/genotype, $p < 0.0001$) eye degeneration in both C9-HRE fly models (Fig. 7d-e and Supplementary Fig. 7f). In accordance with these effects, and a putative role for UPF1 in targeting the expanded *C9orf72* RNA, we found that a reduction in UPF1 expression led to an increased abundance of the C9-GR dipeptide produced in the fly eye, similar to what we found after ETF1 knockdown (Fig. 7f). In summary, these results indicate that ETF1 mediates an UPF1-dependent degradation of the expanded *C9orf72* transcript.

DISCUSSION

The transport of mRNA and proteins between the nucleus and the cytoplasm has emerged as a critical cellular pathway in neurodegenerative diseases[24, 28, 30, 61, 62]. Here, we developed an approach to assess how the repeat expansion in *C9orf72* impacts the subcellular N/C distribution of proteins. We found a small but highly specific redistribution of proteins in C9-HRE expressing cells, with more proteins accumulating in the cytosol. The majority of proteins affected were ontologically associated with RNA metabolism and translation. To address the functional relevance of these alterations, we focused on the nuclear accumulation of ETF1, a protein that regulates two interlinked mechanisms: the termination of translation of physiological transcripts, and the degradation of aberrant transcripts by the NMD pathway[43-46]. We found that ETF1 becomes enriched within an elaborate network of nuclear envelope invaginations in *C9orf72* patient neurons *in vitro* and *in vivo*. This redistribution coincided with a reduction in protein translation and a converse ETF1-dependent increase in NMD. While the full spectrum of mRNAs degraded by NMD remains unclear, we found that the expanded *C9orf72* allele itself is likely an NMD substrate. Overexpression of ETF1 suppressed toxicity, while reduction of ETF1 expression enhanced toxicity and DPR production *in vivo* and led to a dramatic increase of cytosolic RNA foci *in vitro*. Lastly UPF1, which is a master regulator of NMD, modified C9-induced toxicity *in vivo,* suppressing or enhancing eye degeneration in *Drosophila* when overexpressed or when its levels were reduced, respectively.

Our work highlights ETF1 and UPF1 as strong modulators of *C9orf72*-associated pathophysiology and provides a functional link between the C9-HRE and the NMD pathway. Overexpression of UPF1 has previously been shown to have beneficial effects in other ALS model systems, including mutant *TDP-43* and *FUS* rodent neurons *in vitro,* where UPF1 was protective in a helicase- and NMD-dependent manner[63-65]. NMD is a highly conserved mRNA surveillance pathway that controls the quality of newly synthesized transcripts, and is widely accepted as having a broad role in mRNA homeostasis[53, 67]. In fact, it is estimated that NMD might degrade as many as 10% of all physiological transcripts and thus be used by cells not only to prevent the translation of aberrant mRNAs, but also as an effective way to modulate gene expression[52]. We found that C9-ALS patient-derived MNs activate the NMD pathway but with compromised efficiency. This is in agreement with a recent report suggesting that *FUS* mutations also compromise NMD[68].

These findings underscore several critical points related to *C9orf72* toxicity. The activation of the NMD pathway could be a result of broad mRNA processing defects, and/or of the expression of the expanded *C9orf72* RNA, which we show is targeted for NMD-dependent degradation. The genetics of ALS clearly highlight RNA metabolism as one of the principle pathophysiological disease mechanisms with causal mutations identified in a number of RNA-binding proteins such as TDP-43, FUS, MATR3 and HNRNPA1[1]. Moreover, a number of studies have extensively characterized RNA metabolism defects in ALS[69]. These include RNA instability[70], defective RNA splicing in mutant TDP-43 and FUS models[71-73], as well as within postmortem patient brain samples from *C9orf72* and sporadic ALS cases[74].

Moreover, our findings further strengthen the hypothesis that a major component of mutant *C9orf72* toxicity is related to mRNA processing defects. The sequestration of mRNA processing proteins within C9-HRE foci and C9-HRE DPR aggregates might result in broad splicing errors that are known to trigger NMD[59]. The fact that the efficiency of NMD is perturbed in C9-HRE expressing cells, likely further exacerbates the defects in mRNA metabolism and compromises neuronal function as NMD has been shown to enable compartmentilized gene expression[75], and to play a role in synaptic activity and axon guidance in mammalian neurons[76]. It is noteworthy that both GR and PR have been shown to directly interact with UPF1[17], suggesting that UPF1 sequestration by the C9 DPRs could also contribute to the impaired efficiency of NMD that we observed.

The canonical site of NMD has been a subject of debate[52]. While some studies have indicated that NMD occurs in close association with the nucleus[77-80], most models indicate that NMD-dependent mRNA degradation takes place in the cytoplasm within the vicinity of the nuclear membrane and soon after mRNA exits the nuclear pore[81, 82]. Our findings provide novel insight to this debate and suggest that, at least in human neurons, NMD can occur within LMNB1+ nuclear envelope invaginations as we detected the concurrent presence of ETF1 and pUPF1, as well as evidence of translation within these structures.

Tubular invaginations of the nuclear envelope, such as those we describe here, have been observed before in a wide range of cell types and are particularly prominent in dividing cells and cancer tissue[83]. While the mechanistic relevance of these protrusions remains unclear, they have been associated with DNA repair[84] and mRNA transport[85]. These invaginations are not an exclusive feature of patient *C9orf72* neurons, nor do they represent an *in vitro* artifact. We propose that they serve to extend cytosolic function related to RNA metabolism and translation in closer proximity to the nucleus, perhaps to increase the efficiency of these critical cellular processes by shortening the distance between the nucleus and the cytosol, or by increasing the N/C interaction area. This functional role would be particularly relevant to postmitotic neurons, which might heavily rely on the ability to rapidly and dynamically regulate

RNA metabolism and compartmentalized protein translation. Future investigations will need to address what these invaginations contain, their functional relevance and whether they truly represent a beneficial neuronal response or a pathological consequence of aging and disease.

The accumulation of ETF1 and pUPF1 within of LMNB1 nuclear envelope invaginations suggest a model where NMD safeguards neurons from toxic mRNA molecules, such as the expanded *C9orf72,* by preventing transcript exit from the nucleus and ensuing cytosolic translation into toxic C9-DPRs (Supplementary Fig. 8). Although the repeat expansion in *C9orf72* is by far the largest genetic cause of ALS and FTD[3, 4], the extent to which C9 pathophysiological mechanisms are implicated in other familial and sporadic ALS cases remains unclear. We discovered that NMD is protective in C9 model systems. At the same time, the cytosolic accumulation of TDP-43 and thus erroneous mRNA metabolism is a unifying pathological feature amongst ALS patients[1, 70], suggesting that the NMD pathway and its regulatory elements could represent a therapeutic approach that could be broadly relevant for ALS and FTD.

AUTHOR CONTRIBUTIONS

JAO, ELD and EK conceived the study. JAO and ELD designed and performed experiments, analyzed and interpreted all data presented, assembled figures and wrote the manuscript. SK and UBP performed and analyzed all *Drosophila* experiments. MS, LT and YT performed imaging and data analysis. EAH and JNS performed MS proteomics experiments. TYE and TS provided postmortem patient tissue and assisted with IHC. TFG performed DPR ELISA analysis. CJD provided reagents and assisted with FISH experiments. EK designed experiments, interpreted the data, provided guidance and supervision and wrote the manuscript. All authors helped to edit the manuscript.

ACKNOWLEDGEMENTS

We are grateful to the following funding sources: US National Institutes of Health (NIH) / National Institute on Neurological Disorders and Stroke (NINDS) and National Institute on Aging (NIA) R01NS104219 (EK), NIH/NINDS grant R21NS107761 (EK/JNS), Les Turner ALS Foundation (EK), Muscular Dystrophy Association (EK and UBP), ALS Association (UBP), Robert Packard Center for ALS at Johns Hopkins (UBP), NIH R01NS081303 (UBP), R21 NS098379, NS101661, NS100055 (UBP) and NIH/NINDS R01AG061787 (JNS). We also thank the: Center for Advanced Microscopy and Nikon Imaging Center, Flow Cytometry–Cancer Center, NUSeq Core Facility and Biostatistics Collaboration Center. We thank Alfred W Rademaker (Professor in the Division of Biostatistics at Northwestern University) for advice on statistical analysis. The $(G_4C_2)x$ plasmids, the NLS-tdTomato-NES plasmid and the pBS-TCR(PTC)-TCR(WT) NMD reporter were generously provided by Paul Taylor, Jeffrey Rothstein and Zhongsheng You respectively. The HTT and ATXN1 constructs were provided by Dimitri Krainc and Puneet Opal respectively. EK is a member of the Simpson Querrey Institute for BioNanotechnology and a Les Turner ALS Research Center Investigator.

DECLARATION OF INTERESTS

The authors have no conflict to report.

FIGURES AND LEGENDS

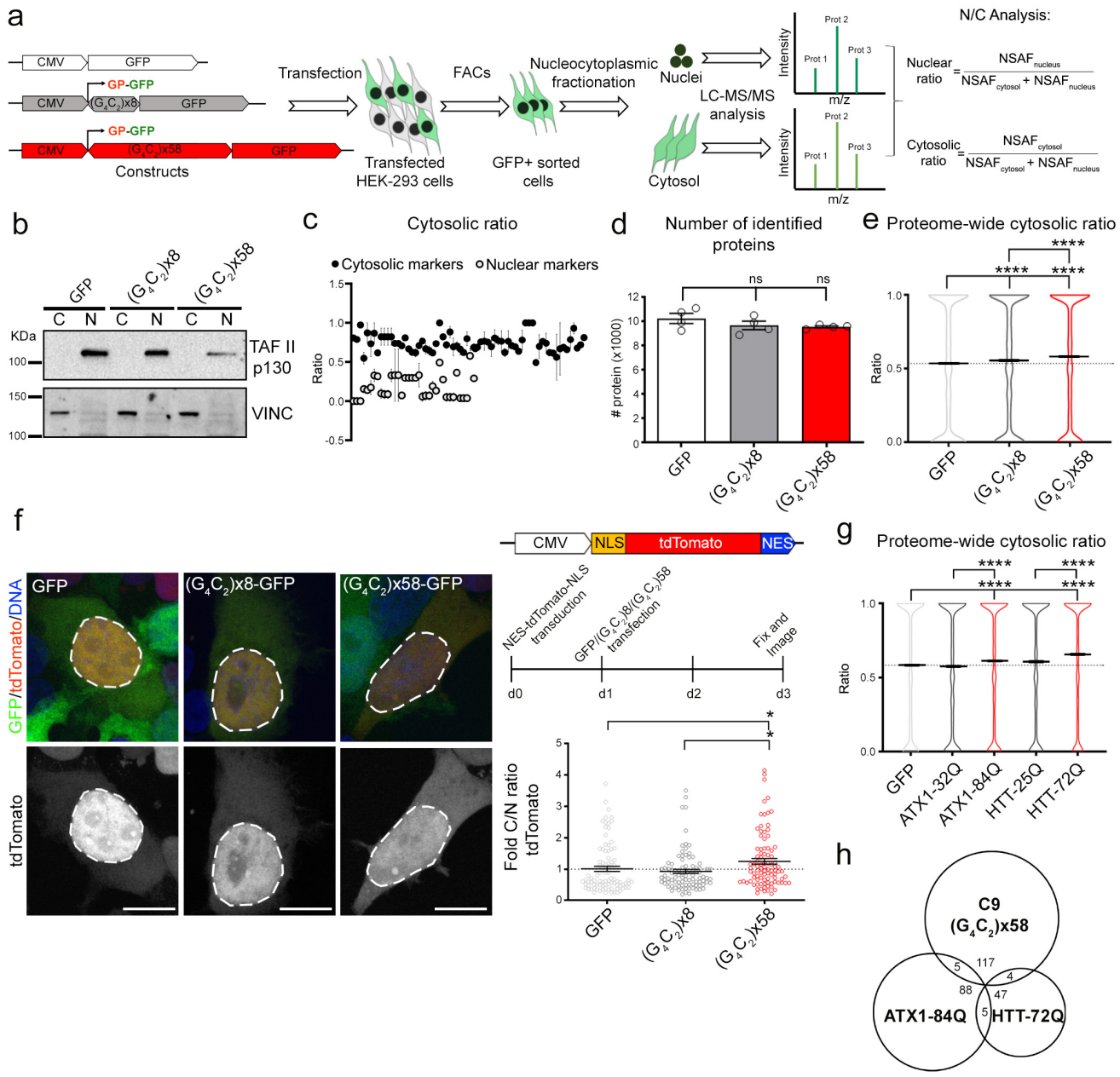


Fig. 1 | The *C9orf72* Repeat Expansion Leads to Proteome-Wide Nucleo-Cytoplasmic Redistribution. a, Experimental schematic used to identify the effects of the G_4C_2 repeat expansion on the subcellular proteome. **b**, Western blot analysis for a nuclear (TAFII p130) and a cytosolic (VINCULIN) marker in the cytosolic (C) and nuclear (N) fractions of GFP, (G_4C_2)x8 and (G_4C_2)x58 transfected cells. **c**, Dot plots representing the cytosolic ratios obtained from MS analysis in GFP-transfected cells of *bona fide* cytosolic (black) and nuclear (gray) markers (n=3 independent biological replicates; data are represented as mean ± SD). **d**, Bar plots representing the average number of redundant proteins identified in cells transfected with GFP, (G_4C_2)x8 and (G_4C_2)x58 by LC-MS/MS (n=4 independent biological replicates; data are represented as mean ± SEM; ANOVA, p = ns, not significant). **e**, Violin plots showing the proteome-wide cytosolic ratio of the ~10,000 proteins identified in cells transfected with GFP (0.528±0.003; mean ± SEM), (G_4C_2)x8 (0.548±0.003), and (G_4C_2)x58 (0.575±0.003) (n=4 independent biological replicates; bars represent mean ± SEM of proteome-wide cytosolic ratio; Kruskal-Wallis, p ****<0.0001). **f**, Left: Cells expressing GFP, (G_4C_2)x8 and (G_4C_2)x58 were infected with a N/C shuttling tdTomato reporter shown in the top right. Dashed lines mark the nuclear circumference. Scale bar: 15 μm. Right-bottom: Dot plot graph showing the fold change in the cytosolic ratio of the tdTomato signal in (G_4C_2)x58 transfected cells compared to controls. Individual cells are displayed as dots and dashed line marks the mean level in control samples (n=3 independent biological replicates, data are represented as mean ± SEM; two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, p *<0.05, **<0.01, p values reflect data pooled from independent biological replicates after testing negative for differential variance by the Brown-Forsythe test). **g**, Violin plots showing the proteome-wide cytosolic ratio of the ~8,000 proteins identified in cells transfected with GFP (0.583±0.004), ATX1-32Q (0.574±0.004), ATX1-84Q (0.612±0.004), HTT-25Q (0.607±0.004) and HTT-72Q (0.656±0.004) (n=3 independent biological replicates; bars represent mean ± SEM of proteome-wide cytosolic ratio; Kruskal-Wallis, p ****<0.0001). **h**, Venn diagram showing the level of overlap of redistributed proteins in cells transfected with (G_4C_2)x58, ATX1-84Q and HTT-72Q.

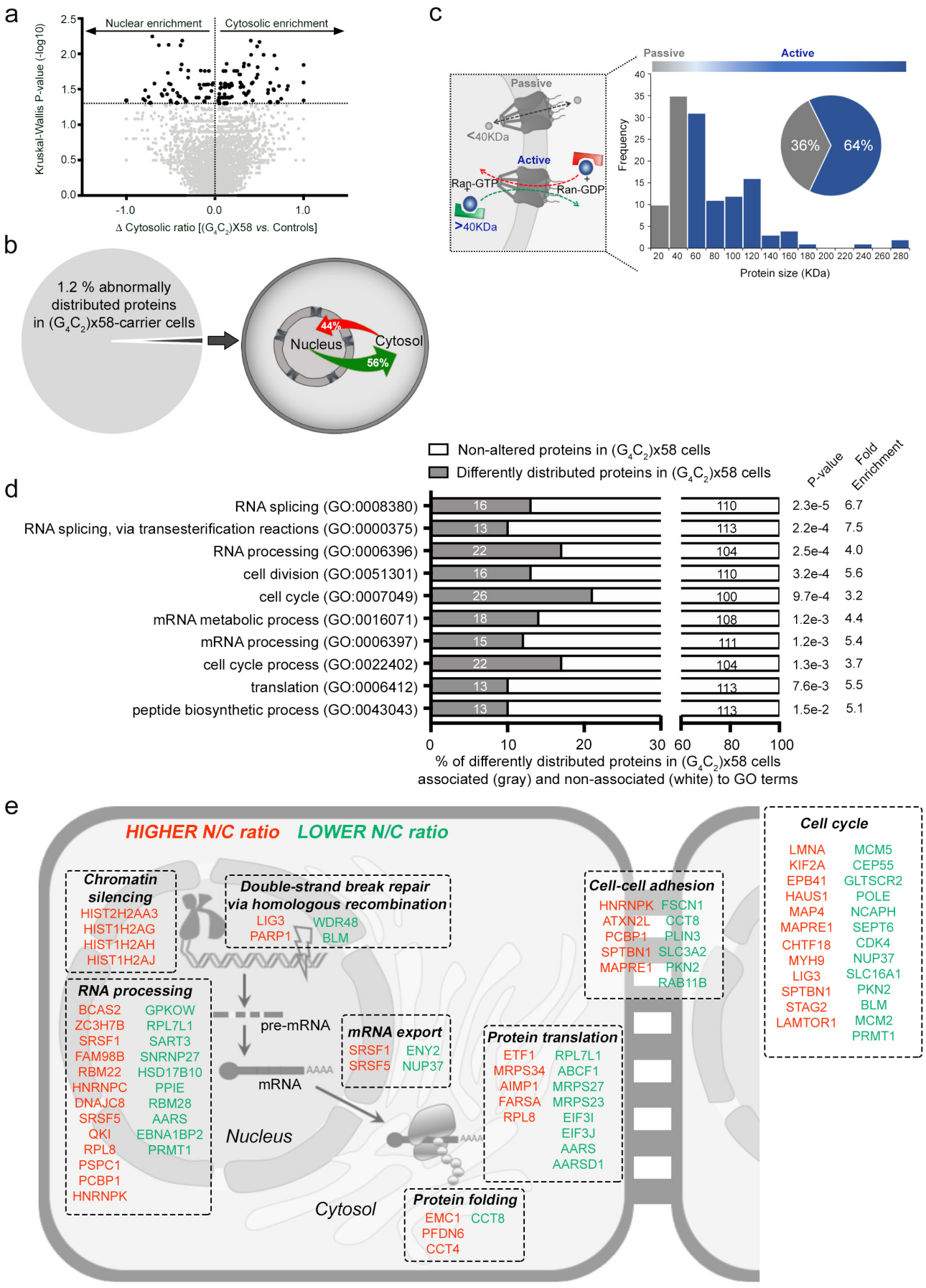

a
Nuclear enrichment
Cytosolic enrichment
Kruskal-Wallis P-value (-log10)
Δ Cytosolic ratio [(G4C2)X58 vs. Controls]
b
1.2 % abnormally distributed proteins in (G4C2)x58-carrier cells
Nucleus
Cytosol
44%
56%
c
Passive
Active
<40KDa
>40KDa
Rap-GTP
Ran-GDP
Frequency
Protein size (KDa)
36%
64%
d
Non-altered proteins in (G4C2)x58 cells
Differently distributed proteins in (G4C2)x58 cells
P-value
Fold Enrichment
RNA splicing (GO:0008380) 16 110 2.3e-5 6.7
RNA splicing, via transesterification reactions (GO:0000375) 13 113 2.2e-4 7.5
RNA processing (GO:0006396) 22 104 2.5e-4 4.0
cell division (GO:0051301) 16 110 3.2e-4 5.6
cell cycle (GO:0007049) 26 100 9.7e-4 3.2
mRNA metabolic process (GO:0016071) 18 108 1.2e-3 4.4
mRNA processing (GO:0006397) 15 111 1.2e-3 5.4
cell cycle process (GO:0022402) 22 104 1.3e-3 3.7
translation (GO:0006412) 13 113 7.6e-3 5.5
peptide biosynthetic process (GO:0043043) 13 113 1.5e-2 5.1
% of differently distributed proteins in (G4C2)x58 cells associated (gray) and non-associated (white) to GO terms
e
HIGHER N/C ratio
LOWER N/C ratio
Chromatin silencing
HIST2H2AA3
HIST1H2AG
HIST1H2AH
HIST1H2AJ
Double-strand break repair via homologous recombination
LIG3
PARP1
WDR48
BLM
RNA processing
BCAS2
ZC3H7B
SRSF1
FAM98B
RBM22
HNRNPC
DNAJC8
SRSF5
QKI
RPL8
PSPC1
PCBP1
HNRNPK
GPKOW
RPL7L1
SART3
SNRNP27
HSD17B10
PPIE
RBM28
AARS
EBNA1BP2
PRMT1
pre-mRNA
mRNA
AAAA
mRNA export
SRSF1
SRSF5
ENY2
NUP37
Nucleus
Cytosol
Protein translation
ETF1
MRPS34
AIMP1
FARSA
RPL8
RPL7L1
ABCF1
MRPS27
MRPS23
EIF3I
EIF3J
AARS
AARSD1
Protein folding
EMC1
PFDN6
CCT4
CCT8
Cell-cell adhesion
HNRNPK
ATXN2L
PCBP1
SPTBN1
MAPRE1
FSCN1
CCT8
PLIN3
SLC3A2
PKN2
RAB11B
Cell cycle
LMNA
KIF2A
EPB41
HAUS1
MAP4
MAPRE1
CHTF18
MYH9
LIG3
SPTBN1
STAG2
LAMTOR1
MCM5
CEP55
GLTSCR2
POLE
NCAPH
SEPT6
CDK4
NUP37
SLC16A1
PKN2
BLM
MCM2
PRMT1

Fig. 2 | The C9orf72 Repeat Expansion Causes a Redistribution of Proteins Involved in RNA Metabolism and Protein Translation. a, Volcano plot representing the proteins that exhibited significant (Kruskal-Wallis, $p < 0.05$, represented by -log10 values) nuclear or cytosolic differences in (G_4C_2)x58 expressing cells compared to both controls (GFP and (G_4C_2)x8 expressing cells). **b**, Left: pie chart displaying the percentage of redundant proteins identified by LC-MS/MS that demonstrated statistically significant subcellular redistribution in (G_4C_2)x58-expressing cells compared to controls (GFP and (G_4C_2)x8-expressing cells). Right: Schematic showing the percentage of redistributed proteins that become more nuclear (red) or more cytosolic (green) in (G_4C_2)x58-expressing cells. **c**, Left: graphical representation of size-dependent nucleocytoplasmic transport of proteins by passive diffusion or active transport. Right: histogram showing the size distribution of proteins significantly redistributed in (G_4C_2)x58-expressing cells. Proteins likely to cross the nuclear envelope by passive or active transport are highlighted in gray and blue respectively in the histogram and in the integrated pie chart. **d**, Gene ontology (GO) analysis of redistributed proteins in (G_4C_2)x58-expressing cells relative to controls. 126 unique proteins were assessed for significant enrichment in biological processes. Gray bars represent the number of proteins enriched for each pathway relative to the total number of redistributed proteins. P-values and fold enrichment values for each cellular process are displayed on the right. **e**, Graphical representation of the most significant (p *<0.05) GO terms associated with redistributed proteins in (G_4C_2)x58-expressing cells. Proteins with higher (red) or lower (green) N/C ratios are depicted with their corresponding GO term and subcellular localization.

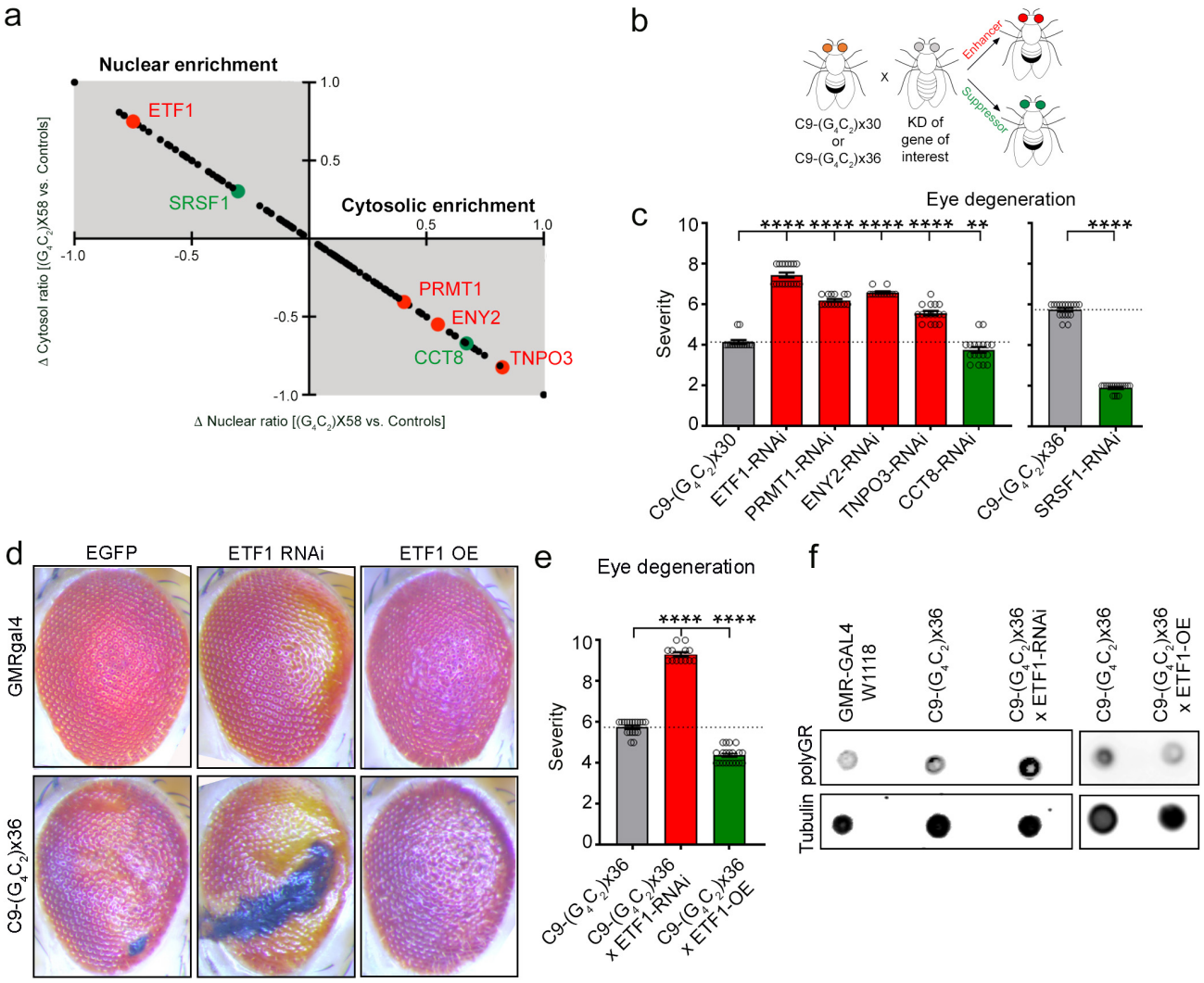
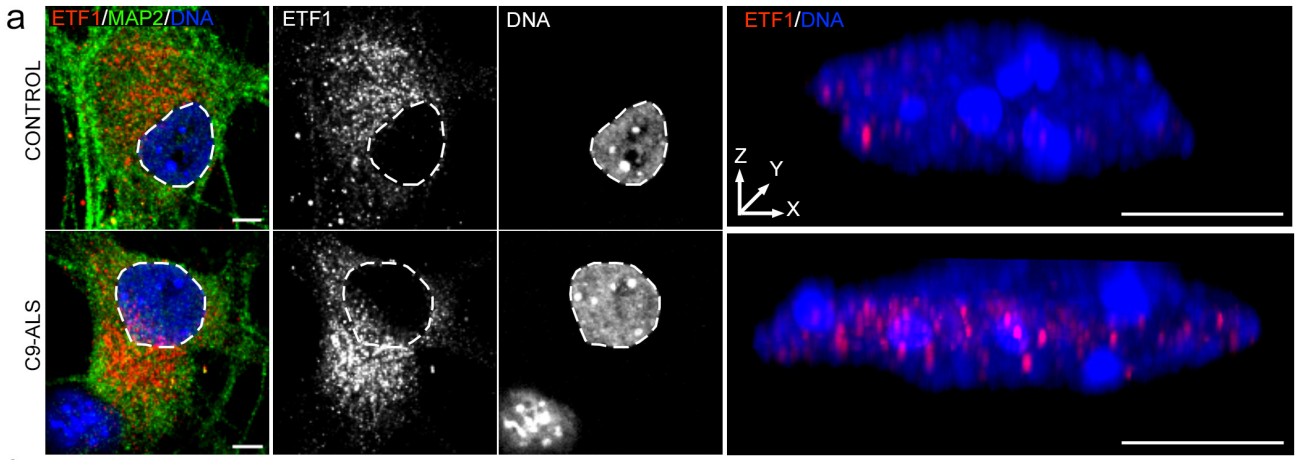
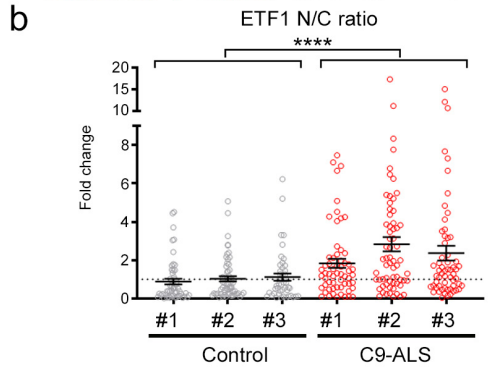
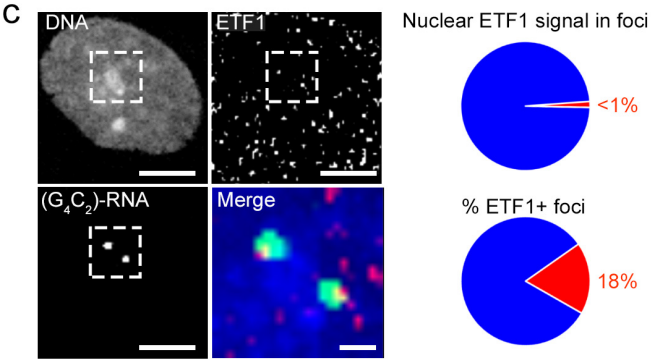
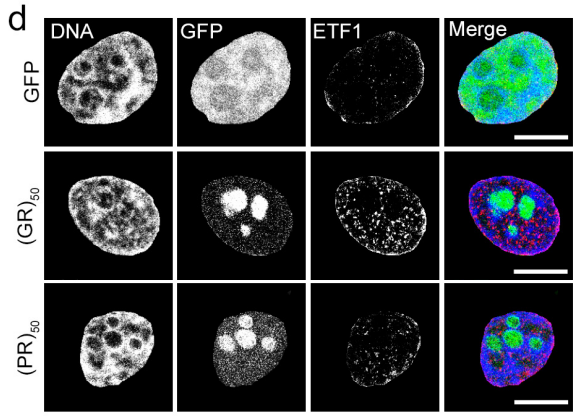

a
Nuclear enrichment
Cytosolic enrichment
ETF1
SRSF1
PRMT1
ENY2
CCT8
TNPO3
Δ Cytosol ratio [(G4C2)X58 vs. Controls]
Δ Nuclear ratio [(G4C2)X58 vs. Controls]
b
C9-(G4C2)x30 or C9-(G4C2)x36
KD of gene of interest
Enhancer
Suppressor
c
Eye degeneration
Severity
C9-(G4C2)x30
ETF1-RNAi
PRMT1-RNAi
ENY2-RNAi
TNPO3-RNAi
CCT8-RNAi
C9-(G4C2)x36
SRSF1-RNAi
d
EGFP
ETF1 RNAi
ETF1 OE
GMRgal4
C9-(G4C2)x36
e
Eye degeneration
Severity
C9-(G4C2)x36
C9-(G4C2)x36 x ETF1-RNAi
C9-(G4C2)x36 x ETF1-OE
f
GMR-GAL4 W1118
C9-(G4C2)x36
C9-(G4C2)x36 x ETF1-RNAi
C9-(G4C2)x36
C9-(G4C2)x36 x ETF1-OE
polyGR
Tubulin

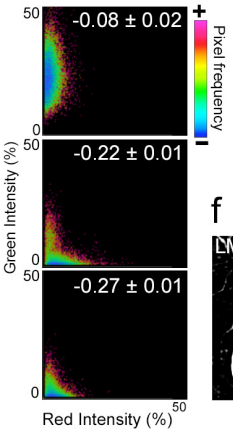
Fig. 3 | Representative Redistributed Proteins are Genetic Modifiers of *C9orf72* Repeat Expansion Toxicity in *Drosophila* Disease Models. a, Graph displaying the significantly redistributed proteins, as well as the direction and magnitude of the change observed in (G_4C_2)x58-expressing cells. The y-axis represents the change in cytosolic ratio and the x-axis represents the change in nuclear ratio of proteins identified in (G_4C_2)x58-expressing cells compared to controls. The upper left gray quadrant includes proteins with higher N/C ratios, while the lower right quadrant proteins with lower N/C ratios. Target proteins selected for genetic interaction studies in *Drosophila* are highlighted in green (suppressors of toxicity) and red (enhancers of toxicity) fonts. **b**, Schematic illustrating the approach for assessing genetic interaction in *Drosophila* models of *C9orf72* toxicity. Flies expressing RNAi for selected proteins were crossed with C9-HRE-expressing flies and assessed for the level of suppression (green eye) or enhancement (red eye) of eye degeneration. **c**, Bar plots displaying the level of eye degeneration in flies carrying C9-HRE (30 or 36 G_4C_2 repeats, gray bars) with RNAi for ETF1, PRMT1, ENY2, TNPO3, CCT8, SRSF1. Enhancers and suppressors of toxicity are represented by red and green bars/symbols, respectively. Individual flies are represented as dots. The dashed line marks the mean level of eye degeneration in RNAi control 30 or 36 G_4C_2 repeats flies (n ≥12 flies; bars represent mean ± SEM; Mann-Whitney U test, p **<0.01, ****<0.0001). Of note, knockdown of target proteins alone did not lead to significant eye degeneration (see Supplementary Fig. 3b). **d**, Representative images of fly eyes of wild-type (GMRgal4 x EGFP) or C9-HRE (30 or 36 G_4C_2 repeats) mutant flies, with normal (EGFP), reduced (ETF1 RNAi) or increased (ETF1 OE) levels of ETF1. **e**, Bar plots displaying the level of eye degeneration in flies carrying 30 or 36 G_4C_2 repeats with basal (gray), reduced (red), or increased (green) levels of ETF1. Individual flies are represented as dots and the dashed line marks the mean level of eye degeneration in RNAi control 36 G_4C_2 repeats flies (n ≥13 flies; bars represent mean ± SEM; Mann-Whitney U test, p ****<0.0001). **f,** Dot blot of GR DPR in control (GMR-GAL4) *Drosophila* eyes, as well as in a (G_4C_2)x36 carrier, and (G_4C_2)x36 carriers with decreased (ETF1 RNAi) or increase (ETF1 OE) levels of ETF1. Tubulin was used as a loading control.

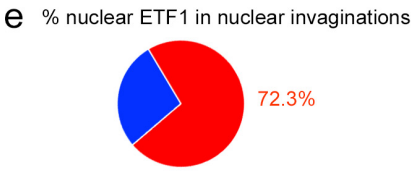

a
ETF1/MAP2/DNA
ETF1
DNA
ETF1/DNA
CONTROL
C9-ALS
Z
Y
X

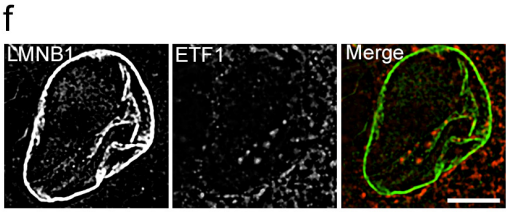

b
ETF1 N/C ratio
****
Fold change
20
15
10
8
6
4
2
0
#1
#2
#3
#1
#2
#3
Control
C9-ALS


c
DNA
ETF1
(G4C2)-RNA
Merge
Nuclear ETF1 signal in foci
<1%
% ETF1+ foci
18%


d
DNA
GFP
ETF1
Merge
GFP
(GR)50
(PR)50


-0.08 ± 0.02
-0.22 ± 0.01
-0.27 ± 0.01
Pixel frequency
Green Intensity (%)
Red Intensity (%)
50
0


e
% nuclear ETF1 in nuclear invaginations
72.3%


f
LMNB1
ETF1
Merge

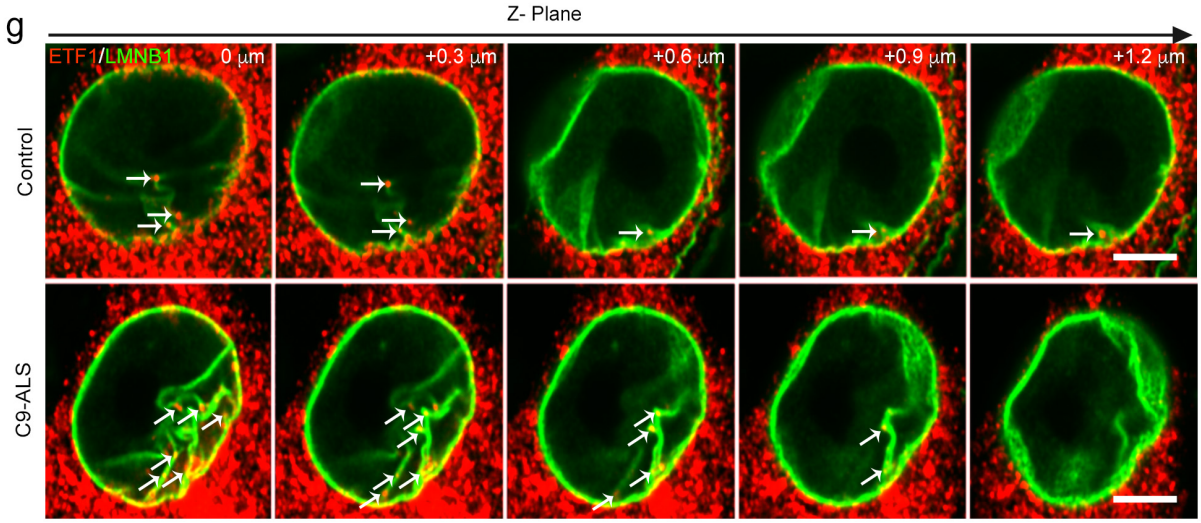

g
Z- Plane
ETF1/LMNB1
0 µm
+0.3 µm
+0.6 µm
+0.9 µm
+1.2 µm
Control
C9-ALS

Fig. 4 | ETF1 is Redistributed in *C9orf72* iPSC-Derived Motor Neurons. a, Left: Representative ICC confocal images of control and C9-ALS patient iPSC-derived MNs immunolabeled with ETF1 (red), MAP2 (green) and DNA (blue). Dashed line marks the nuclear circumference, scale bar: 10 μm. Right: 3D reconstruction of the nucleus of a control (top) and a C9-ALS (bottom) iPSC-derived MN, scale bar: 5 μm. **b**, Dot plot graph displaying the fold change in the N/C ratio of ETF1 signal quantified in MNs derived from three controls and three C9-ALS patient iPSC lines. Individual cells are represented as dots, and the dashed line marks the mean N/C ratio in control samples (n=3 independent differentiations; horizontal bars represent mean ± SEM; Mann-Whitney U test, p ****<0.0001, p value reflects data pooled from independent biological replicates after testing negative for differential variance by the Brown-Forsythe test). **c**, Left: Representative FISH/ICC confocal images for G_4C_2 RNA foci (green), ETF1 (red) and DNA (blue). Scale bar: 5 μm. Dashed insets mark the region magnified in the right bottom merged image. Scale bar: 1 μm. Right: pie chart showing the percentage of nuclear ETF1 signal colocalizing with C9-RNA foci (top) and the percentage of C9-RNA foci containing ETF1 (bottom). **d**, Left: Representative ICC confocal images of HEK-293 cells nuclei expressing GFP, $(GR)_{50}$-GFP and $(PR)_{50}$-GFP (green), and immunolabeled for ETF1 (red) and DNA (blue). Right: Representative plots wherein each point reflects the signal intensity in the green channel (DPR-GFP) on the y-axis, and the signal intensity in the red channel (ETF1) on the x-axis. The mean ± SEM of the Pearson's correlation coefficient for the entire dataset is displayed within the corresponding plot. **e**, Pie chart displaying the percentage of nuclear ETF1 signal that colocalizes with LMNB1+ nuclear invaginations in C9-ALS MNs. **f**, Structured-illumination microscopy (SIM) showing colocalization of nuclear ETF1 and LMNB1+ nuclear envelope invaginations. Scale bar: 20 μm. **g**, Representative ICC serial confocal imaging showing colocalization of ETF1 puncta in the nucleus (red) and LMNB1+ nuclear envelope invaginations (green) in control and C9-ALS MNs. Arrows indicate ETF1+ puncta located in LMNB1+ nuclear invaginations. Scale bar: 5 μm.

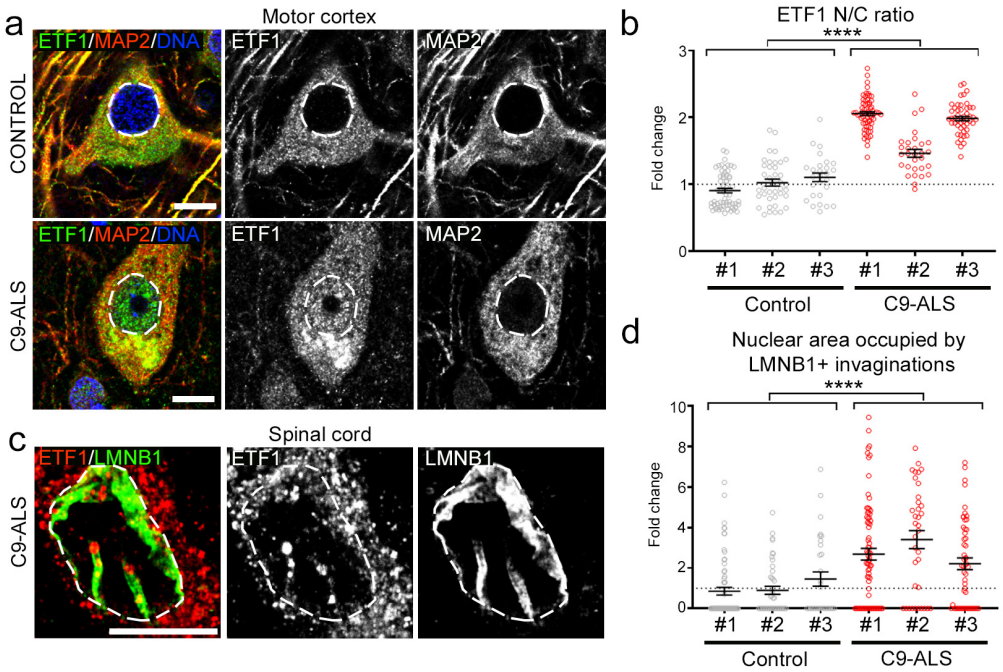


Fig. 5 | ETF1 is Redistributed in Postmortem *C9orf72* ALS Tissue. a, Representative IHC confocal images of layer V neurons immunolabeled for ETF1 (green), MAP2 (red) and DNA (blue) in motor cortex tissue from non-neurological controls and C9-ALS patients. Scale bar: 10 μm. **b**, Dot plot graph displaying the fold change in the N/C ratio of ETF1 signal observed in cortical neurons of three non-neurological age-matched controls and three C9-ALS patients. Individual cells are represented as dots and dashed line marks the mean N/C ratio in control samples (n=3 biological replicates; horizontal bars represent the mean ± SEM; Mann-Whitney U test, p ****<0.0001). **c**, Representative IHC confocal image of a C9-ALS spinal motor neuron located in the ventral horn showing pronounced LMNB1+ invaginations. Scale bar: 5 μm. **d**, Dot plot graph displaying the fold change in the nuclear area occupied by LMNB1+ invaginations in cortical neurons from three controls and three C9-ALS patients. Individual cells are represented as dots and dashed line marks the mean LMNB1+ nuclear area in control samples (n=3 biological replicates; horizontal bars represent the mean ± SEM; Mann-Whitney U test, p ****<0.0001).

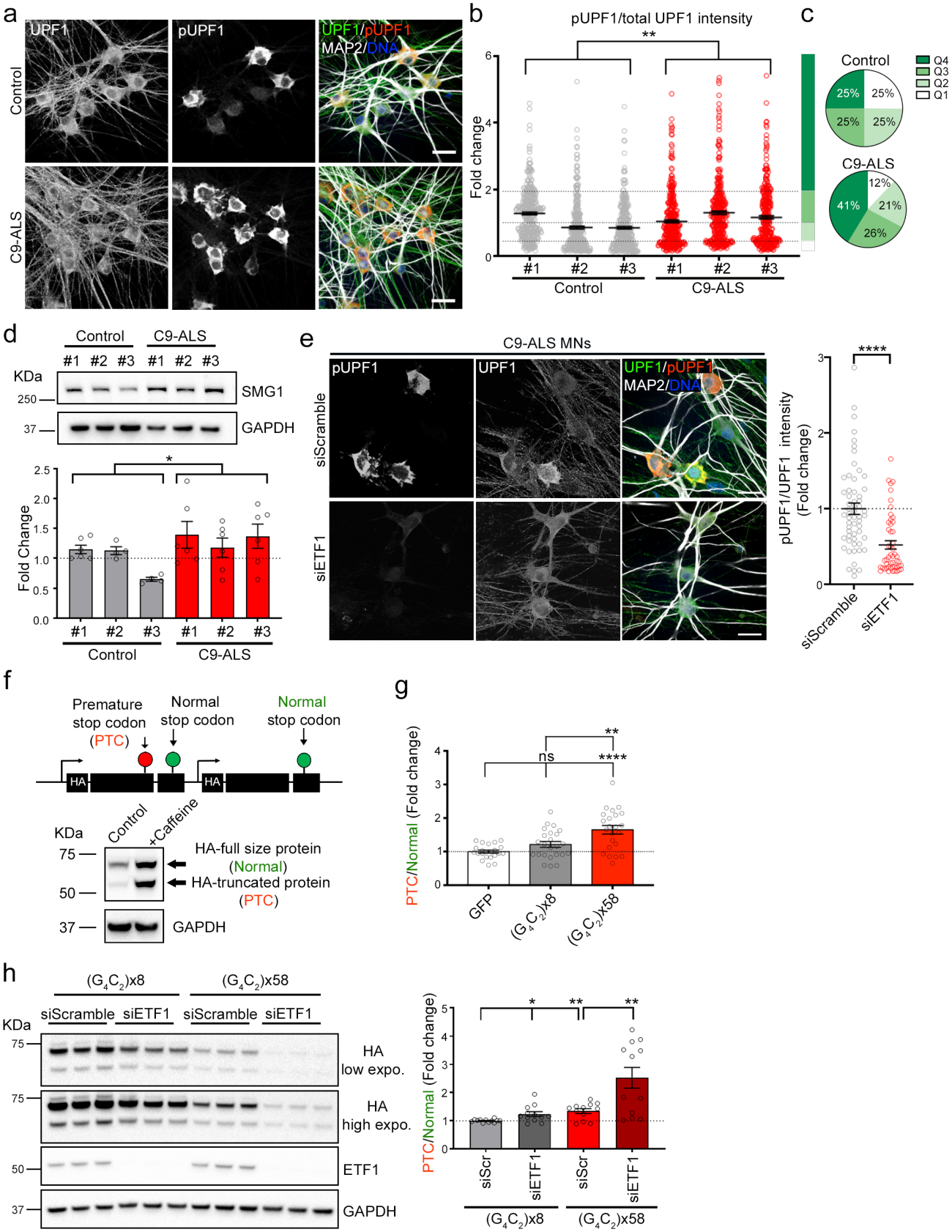

a
UPF1
pUPF1
UPF1/pUPF1
MAP2/DNA
Control
C9-ALS
b
pUPF1/total UPF1 intensity
**
Fold change
#1
#2
#3
Control
C9-ALS
c
Control
Q4
Q3
Q2
Q1
25%
25%
25%
25%
C9-ALS
12%
41%
21%
26%
d
Control
C9-ALS
KDa
250
37
SMG1
GAPDH
*
Fold Change
e
C9-ALS MNs
siScramble
siETF1
****
pUPF1/UPF1 intensity (Fold change)
f
Premature stop codon (PTC)
Normal stop codon
Normal stop codon
HA
Control
+Caffeine
75
50
HA-full size protein (Normal)
HA-truncated protein (PTC)
g
ns
**
****
PTC/Normal (Fold change)
GFP
(G4C2)x8
(G4C2)x58
h
(G4C2)x8
(G4C2)x58
siScramble
siETF1
HA low expo.
HA high expo.
ETF1
GAPDH
*
**
**
siScr
siETF1

Fig. 6 | Protein Translation and mRNA Degradation Through Nonsense mediated decay in C9orf72 repeat expansion carrier cells. a, Confocal images of control and C9-ALS MNs immunolabeled for total UPF1 (green), pUPF1 (red) and DNA (blue). **b**, Dot plot graph displaying the fold change in pUPF1/total UPF1 signal in control and C9-ALS MNs. Individual cells are represented as dots and dashed lines mark the limits of data quartiles (Q1-4) related to pUPF1/total UPF1 ratio in control samples (n=3 independent differentiations; horizontal bars represent the mean ± SEM; p values were calculated for each independent biological replicate by Mann-Whitney U test and all replicates displayed p **<0.01). **c**, Top: Pie chart displaying NMD activation levels defined by pUPF1/UPF1 quartiles observed in control MNs, and (bottom) the corresponding shift in the proportion of C9 MNs with NMD activation. **d**, Top: western blot of SMG1 in three control and three C9-ALS iPSC-derived MN cultures. GAPDH was used as a loading control. Bottom: bar plot graph showing the fold change in the SMG1/GAPDH ratio in C9-ALS respect to control MN samples. Individual samples are displayed as dots and dashed line marks the mean ratio value in control samples (n=6 independent differentiations; bars represent the mean ± SEM; Mann-Whitney U test, p ****<0.0001). **e**, Left: ICC confocal image analysis of NMD activation assessed by pUPF1/total UPF1 intensity in C9-ALS iPSC-derived MNs transfected with scramble (SCR) and ETF1 siRNAs. Right: dot plots displaying pUPF1/UPF1 ratios in C9-ALS MNs transfected with scrambled or ETF1 siRNA. Individual cells are represented as dots and dashed line marks the mean pUPF1/total UPF1 ratio in siSCR transfected MNs (n=3 independent differentiations; horizontal bars represent the mean ± SEM; Mann-Whitney U test, p ****<0.0001). **f**, Top: Schematic of the reporter construct utilized in this study to assess NMD activity *in vitro*. The construct encodes two TRB (T cell receptor-β) minigenes sharing >99% homology and fused to an HA tag; one of them contains a premature stop codon (PTC) that produces a truncated protein. Bottom: Western blot analysis of HEK-293 cells treated with caffeine, an established NMD inhibitor, validating the efficacy of the reporter construct. Full-length and truncated protein products were detected by an anti-HA antibody. GAPDH was used as a loading control. **g**, Bar plot graph displaying the quantification of the average NMD activity, calculated by the expression levels of the truncated product (PTC) relative to the full-length product (Normal) of the TRB minigenes in cells expressing GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP. Independent replicates are displayed as dots, and dashed line marks the mean PTC/normal ratio in control samples (n=4 biological replicates; horizontal bars represent the mean ± SEM; ANOVA test, p **<0.01, ****<0.0001, p value reflects data pooled from independent biological replicates after testing negative for differential variance by the Brown-Forsythe test). **h**, Left: Western blot analysis displaying the level of expression of both TRB minigenes in cells expressing GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP and treated with scrambled or ETF1 siRNA. ETF1 blot demonstrates the efficiency of the ETF1-KD. GAPDH was used as a loading control. Right: bar plot displaying the quantification of NMD activity in in cells expressing (G_4C_2)x8-GFP and (G_4C_2)x58-GFP with normal or reduced levels of ETF1. Individual samples are displayed as dots, and dashed line marks the mean PTC/normal ratio in siSCR-(G_4C_2)x8-transfected cells (n=4 biological replicates; horizontal bars represent the mean ± SEM; ANOVA test, p *<0.05, **<0.01)

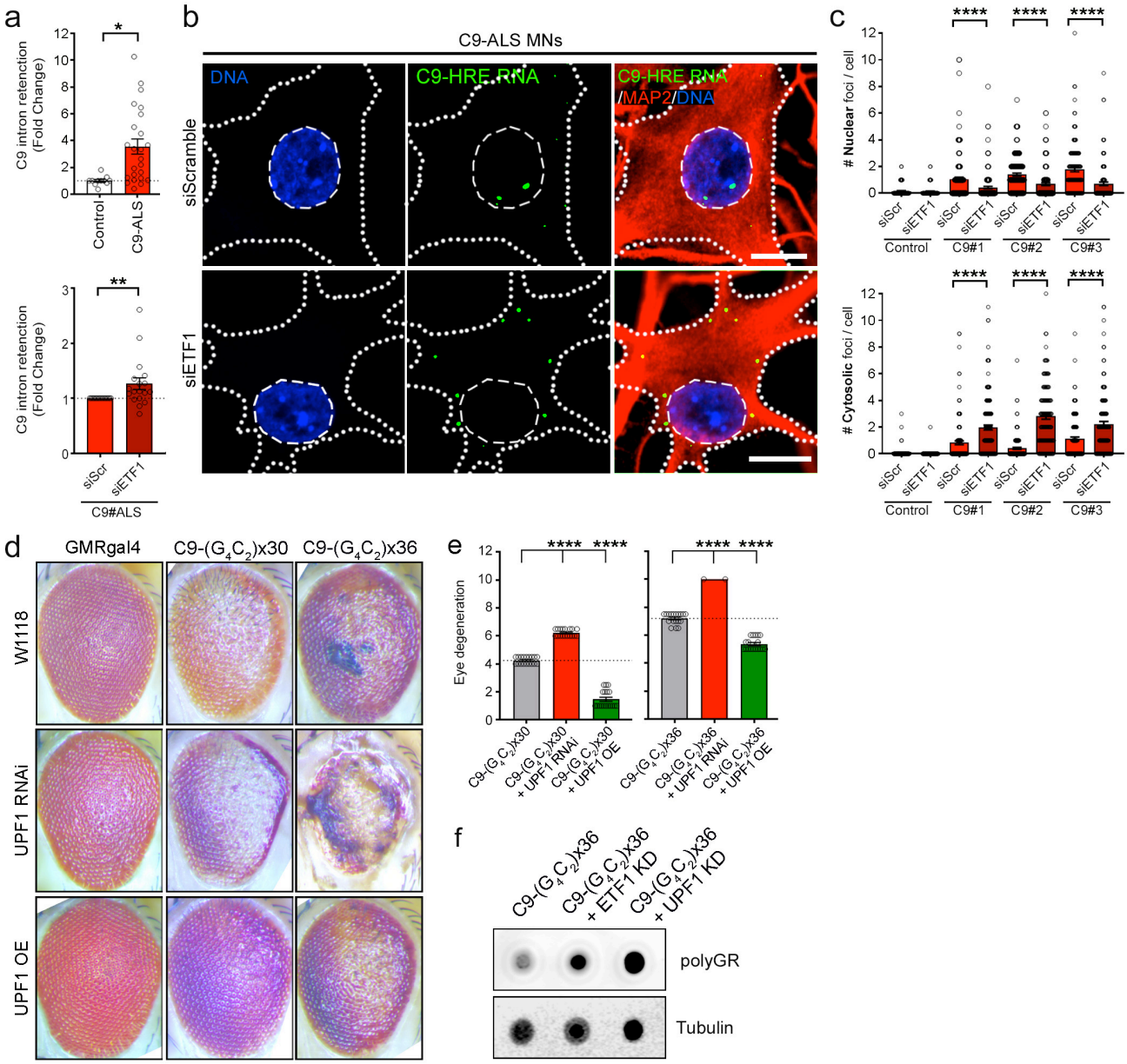
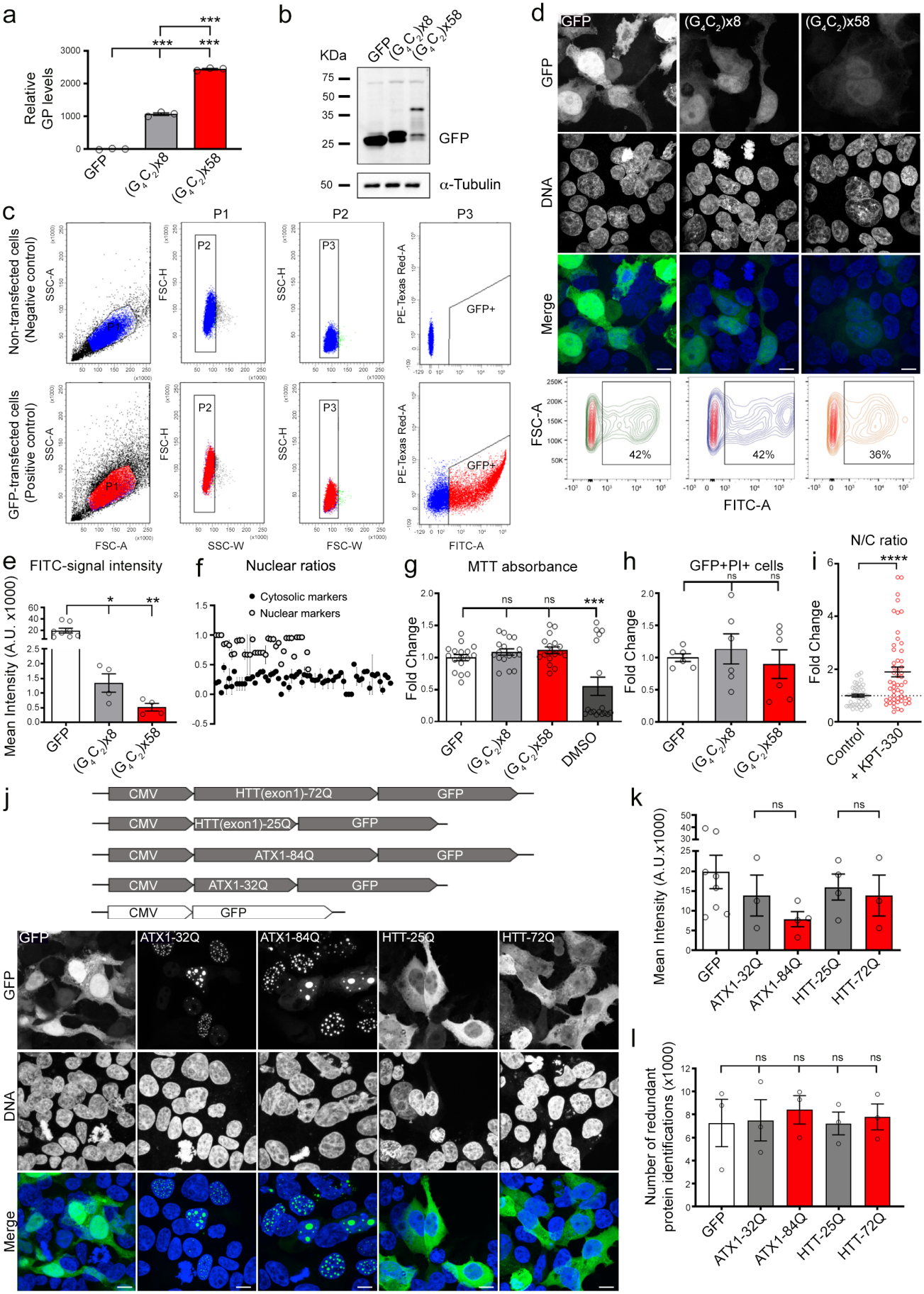
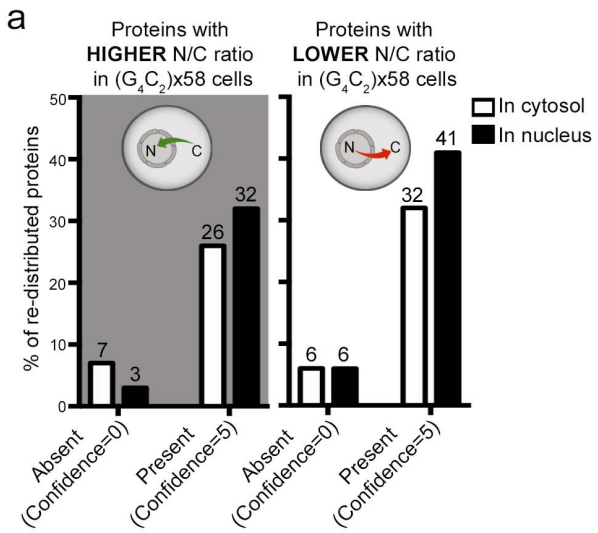
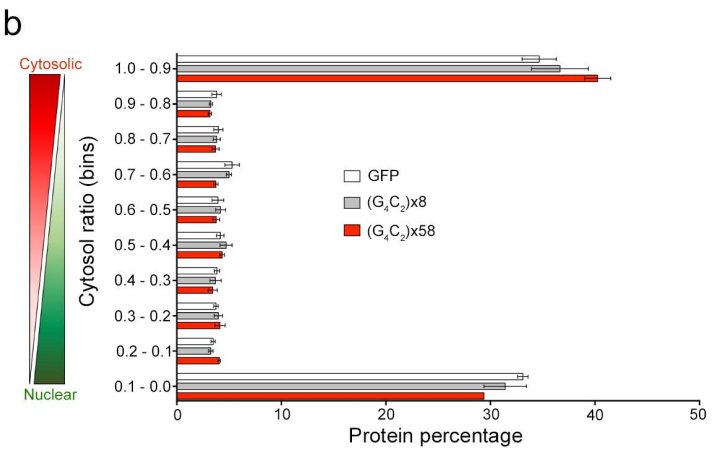

a
C9 intron retention (Fold Change)
Control
C9-ALS
siScr
siETF1
C9#ALS
b
C9-ALS MNs
DNA
C9-HRE RNA
C9-HRE RNA /MAP2/DNA
siScramble
siETF1
c
Nuclear foci / cell
Cytosolic foci / cell
Control
C9#1
C9#2
C9#3
d
GMRgal4
C9-(G4C2)x30
C9-(G4C2)x36
W1118
UPF1 RNAi
UPF1 OE
e
Eye degeneration
C9-(G4C2)x30
C9-(G4C2)x30 + UPF1 RNAi
C9-(G4C2)x30 + UPF1 OE
C9-(G4C2)x36
C9-(G4C2)x36 + UPF1 RNAi
C9-(G4C2)x36 + UPF1 OE
f
C9-(G4C2)x36
C9-(G4C2)x36 + ETF1 KD
C9-(G4C2)x36 + UPF1 KD
polyGR
Tubulin

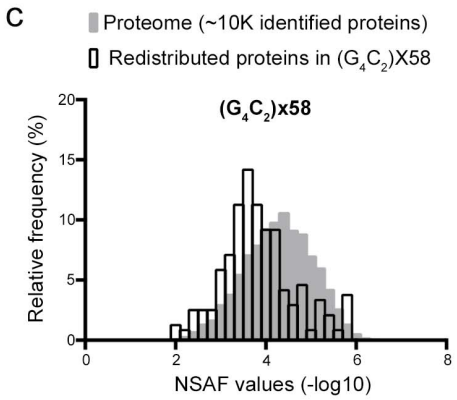
Fig. 7 | C9orf72 mRNA is Targeted for Degradation by Nonsense Mediated Decay. a, Top: Bar plot showing fold change levels of C9-intron retention obtained from MN cultures derived from three control and three C9-ALS iPSC lines as measured by RT-PCR. Individual samples are displayed as dots and dashed line marks the mean C9-intron retention in control MNs (n=8 independent differentiations; horizontal bar represent the mean ± SEM; Mann-Whitney, p *<0.05)**.** Bottom: Bar plot showing fold change levels of C9-intron retention in C9-ALS iPSC-derived MNs transfected with scramble (SCR) or ETF1 siRNA. Individual samples are displayed as dots and dashed line marks the mean C9-intron retention in siSCR-transfected C9-ALS MNs (n=6 independent differentiations; horizontal bars represent the mean ± SEM; Mann-Whitney, p **<0.01)**. b**, Representative FISH/ICC confocal images of iPSC-derived MNs transfected with scramble or ETF1 siRNA and immunolabelled with G_4C_2 RNA foci (green), MAP2 (red) and DNA (blue). Dashed line marks the nuclear circumference and dotted lines the neuronal soma. Scale bar: 10 μm. **c**, Bar plot showing the number of nuclear (top) and cytosolic (bottom) C9-HRE RNA foci per cell in MNs derived from one control and three C9-ALS iPSC lines, transfected with scramble or ETF1 siRNAs. Individual cells are represented as dots (n=3 independent differentiations; horizontal bars represent the mean ± SEM; Wilcoxon signed rank, p ****<0.0001)**. d**, Representative brightfield images of eye phenotype of WT and G_4C_2 (30 or 36 repeats) mutant flies, with normal (W1118), increased (UPF1 OE) or reduced (UPF1 RNAi) levels of UPF1 expression. **e**, Bar plot showing the degree of external eye degeneration in flies that carry G_4C_2 (30 or 36 G_4C_2 repeats) after knockdown or overexpression of UPF1, respectively. Individual flies are displayed as dots and a dashed line marking the mean level in RNAi control 30 or 36 G_4C_2 repeats flies (bars represent mean ± SEM eye degeneration severity; Mann-Whitney U test, p ****<0.0001). **f**, Dot blot showing levels of polyGR DPR in C9-(G_4C_2)36, C9-(G_4C_2)36 x ETF1-RNAi and C9-(G_4C_2)36 x UPF1-RNAi flies. Tubulin was used as a loading control.

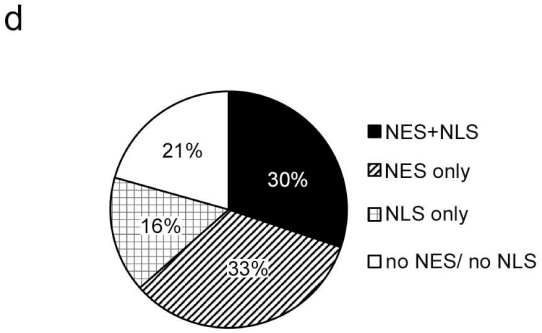

a
Relative GP levels
GFP
(G4C2)x8
(G4C2)x58
b
KDa
GFP
α-Tubulin
c
Non-transfected cells (Negative control)
GFP-transfected cells (Positive control)
P1
P2
P3
SSC-A
FSC-H
SSC-H
PE-Texas Red-A
GFP+
FSC-A
SSC-W
FSC-W
FITC-A
d
GFP
DNA
Merge
42%
42%
36%
e
FITC-signal intensity
Mean Intensity (A.U. x1000)
f
Nuclear ratios
Cytosolic markers
Nuclear markers
g
MTT absorbance
Fold Change
DMSO
h
GFP+PI+ cells
i
N/C ratio
Control
+ KPT-330
j
CMV
HTT(exon1)-72Q
HTT(exon1)-25Q
ATX1-84Q
ATX1-32Q
k
Mean Intensity (A.U.x1000)
l
Number of redundant protein identifications (x1000)

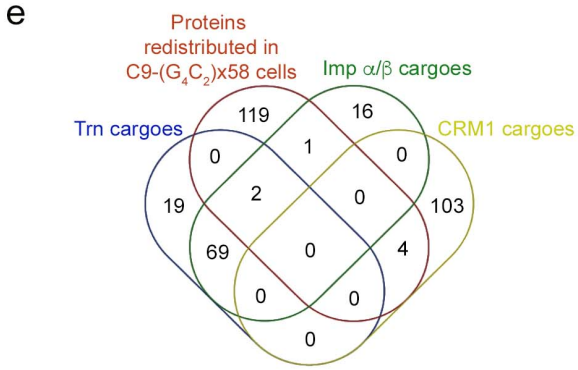
Supplementary Figure 1, related to Figure 1: Unbiased Proteomic Analysis of Repeat Expansion Mutations Associated With Neurodegeneration. a, Bar plots representing the average levels of GP DPR detected by ELISA in cells transfected with GFP, (G_4C_2)x8-GFP, and (G_4C_2)x58-GFP. Every dot represents an independent sample (n=3); data are represented as mean ± SEM; ANOVA, p ***<0.001. **b**, Western blot analysis of cells transfected with GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP and probed with an anti-GFP antibody. Tubulin was used as a loading control. **c**, Example of the gating strategy utilized to FACS-purify GFP+ cells. **d**, Top: representative images of HEK-293 cells transfected with GFP, (G_4C_2)x8-GFP, and (G_4C_2)x58-GFP. Scale bar: 20 μm. Bottom: representative FACS plots displaying effective sorting for GFP+ cells. **e**, Bar plots displaying the mean ± SEM intensity levels of GFP fluorescence acquired by FACS analysis in GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP-transfected cells. Individual samples are displayed as dots (n=4-7 biological replicates); ANOVA, p *<0.05, **<0.01. **f**, Dot plot showing the mean± SD of the nuclear ratio obtained from LC-MS/MS analysis in GFP-transfected cells of bona fide cytosolic (black) and nuclear (gray) markers (n=3 independent experiments). **g**, Bar plots displaying MTT viability assay 48hrs after transfection of GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP in HEK293 cells. Bars represent the mean ± SEM, and each dot represents an independent sample (n=3 biological replicates; ANOVA, p =ns, not significant; ***<0.001). **h**, Bar plots displaying FACS-based quantification of double GFP/Propidium Iodide+ cells from HEK-293 cell cultures transfected with GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP. Individual samples are displayed as dots (n=2 biological replicates; ANOVA, p =ns). **i**, Dot plot graphs displaying the fold change in the N/C ratio of the tdTomato reporter in cells treated with the nuclear export inhibitor KPT-330. Individual cells are represented as dots, and the dashed line marks the mean N/C ratio in control samples (n=3 independent differentiations; horizontal bars represent mean ± SEM; Mann-Whitney U test, p ****<0.0001). **j**, Top: Schematic of the polyglutamine (CAG, polyQ) expansion in *ATX1* or *HTT* (exon 1) used in a complementary subcellular screening. Bottom: Representative images of HEK-293 cells transfected with GFP, ATX1-32Q, ATX1-84Q, HTT-25Q, or HTT-72Q. Scale bar: 20 μm. **k**, Bar plots displaying the mean intensity levels of GFP acquired by FACS analysis in GFP, ATX1-32Q, ATX1-84Q, HTT-25Q, or HTT-72Q transfected cells. Individual experimental samples are displayed as dots (n=3-7 biological replicates; ANOVA, p =ns). **l**, Bar plots representing the average number of proteins detected in cells transfected with GFP, ATX1-32Q, ATX1-84Q, HTT-25Q, or HTT-72Q. Independent biological replicates represented as dots (n=3 independent biological replicates; data are represented as mean ± SEM; ANOVA, p = ns, not significant).

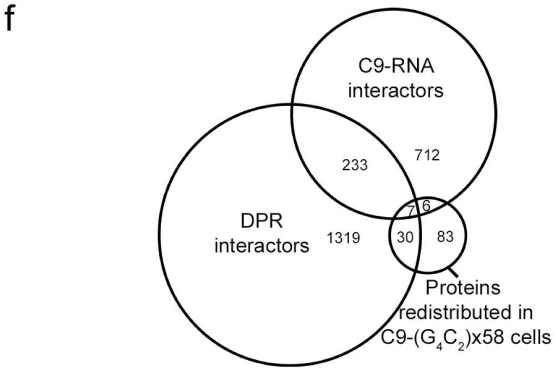

a
Proteins with HIGHER N/C ratio in (G4C2)x58 cells
Proteins with LOWER N/C ratio in (G4C2)x58 cells
In cytosol
In nucleus
% of re-distributed proteins
N
C
7
3
26
32
6
6
32
41
Absent (Confidence=0)
Present (Confidence=5)
Absent (Confidence=0)
Present (Confidence=5)


b
Cytosolic
Nuclear
Cytosol ratio (bins)
1.0 - 0.9
0.9 - 0.8
0.8 - 0.7
0.7 - 0.6
0.6 - 0.5
0.5 - 0.4
0.4 - 0.3
0.3 - 0.2
0.2 - 0.1
0.1 - 0.0
GFP
(G4C2)x8
(G4C2)x58
Protein percentage

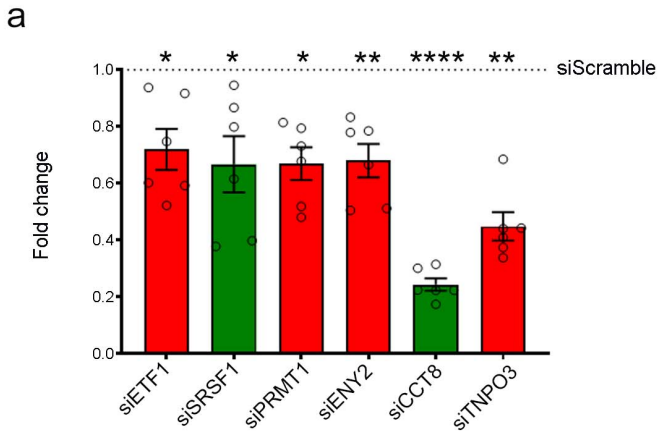

c
Proteome (~10K identified proteins)
Redistributed proteins in (G4C2)X58
(G4C2)x58
Relative frequency (%)
NSAF values (-log10)

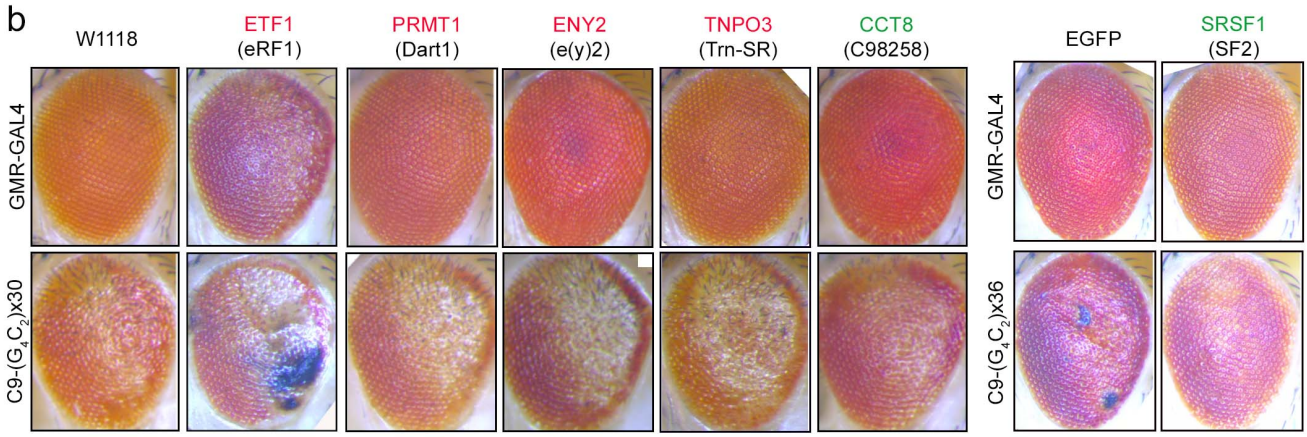

d
21%
30%
16%
33%
NES+NLS
NES only
NLS only
no NES/ no NLS

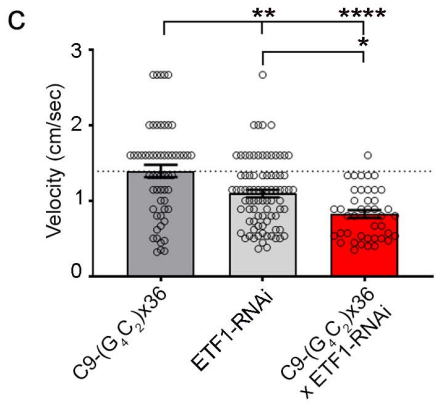

e
Proteins redistributed in C9-(G4C2)x58 cells
Imp α/β cargoes
Trn cargoes
CRM1 cargoes
119
16
0
1
0
19
2
0
103
69
0
4
0
0
0

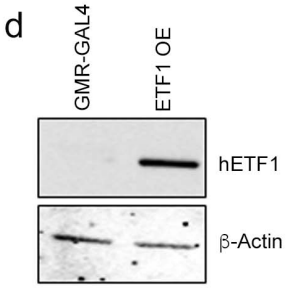

f
C9-RNA interactors
DPR interactors
233
712
76
1319
30
83
Proteins redistributed in C9-(G4C2)x58 cells

Supplementary Figure 2, related to Figure 2: Analysis of Subcellular Proteomic Changes Identified in Cells Carrying the C9-HRE. a, Subcellular protein localization analysis of the proteins found redistributed in cells expressing (G_4C_2)x58-GFP. The presence of these proteins in the nucleus or the cytosol is ranked based on confidence values established by the https://compartments.jensenlab.org database. Bars represent the percentage of proteins that are absent (Confidence=0) or present (Confidence=1 to 5) in the cytosol (white) or nucleus (black). The number of proteins of each group is shown above the bars. Left/gray shows the analysis of proteins with higher N/C ratio, and right/white, shows the analysis of proteins with lower N/C ratio in (G_4C_2)x58-GFP expressing cells compared to controls. **b**, Analysis of subcellular proteome-wide distribution showing the percent of proteins across cytosolic ratio bins in cells transfected with GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP. **c**, Frequency distribution of proteins identified by LC-MS/MS based on their respective NSAF values in cells expressing (G_4C_2)x58-GFP. Grey bars represent all proteins identified and clear bars represent all proteins that become redistributed in (G_4C_2)x58-GFP expressing cells. **d**, Pie chart representing the percentage of redistributed proteins in (G_4C_2)x58-GFP cells, that contain both an NLS and NES (black), an NLS (checks), an NES (stripes), or neither (white). **e**, Venn diagram displaying the overlap between redistributed proteins in C9-(G_4C_2)x58-GFP cells and published karyopherin cargoes. **f,** Venn diagram displaying the overlap between redistributed proteins in C9-(G_4C_2)x58-GFP cells and previously reported C9-HRE-RNA and C9-HRE-DPR protein interactors.

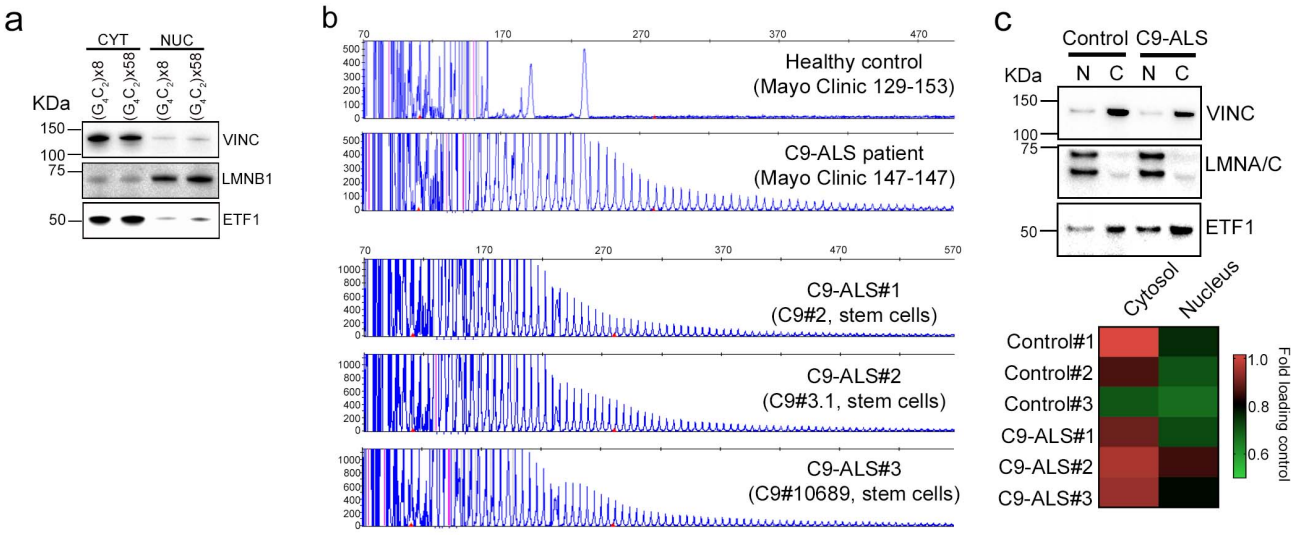

a
siScramble
Fold change
siETF1
siSRSF1
siPRMT1
siENY2
siCCT8
siTNPO3


b
W1118
ETF1 (eRF1)
PRMT1 (Dart1)
ENY2 (e(y)2)
TNPO3 (Trn-SR)
CCT8 (C98258)
EGFP
SRSF1 (SF2)
GMR-GAL4
C9-(G4C2)x30
C9-(G4C2)x36

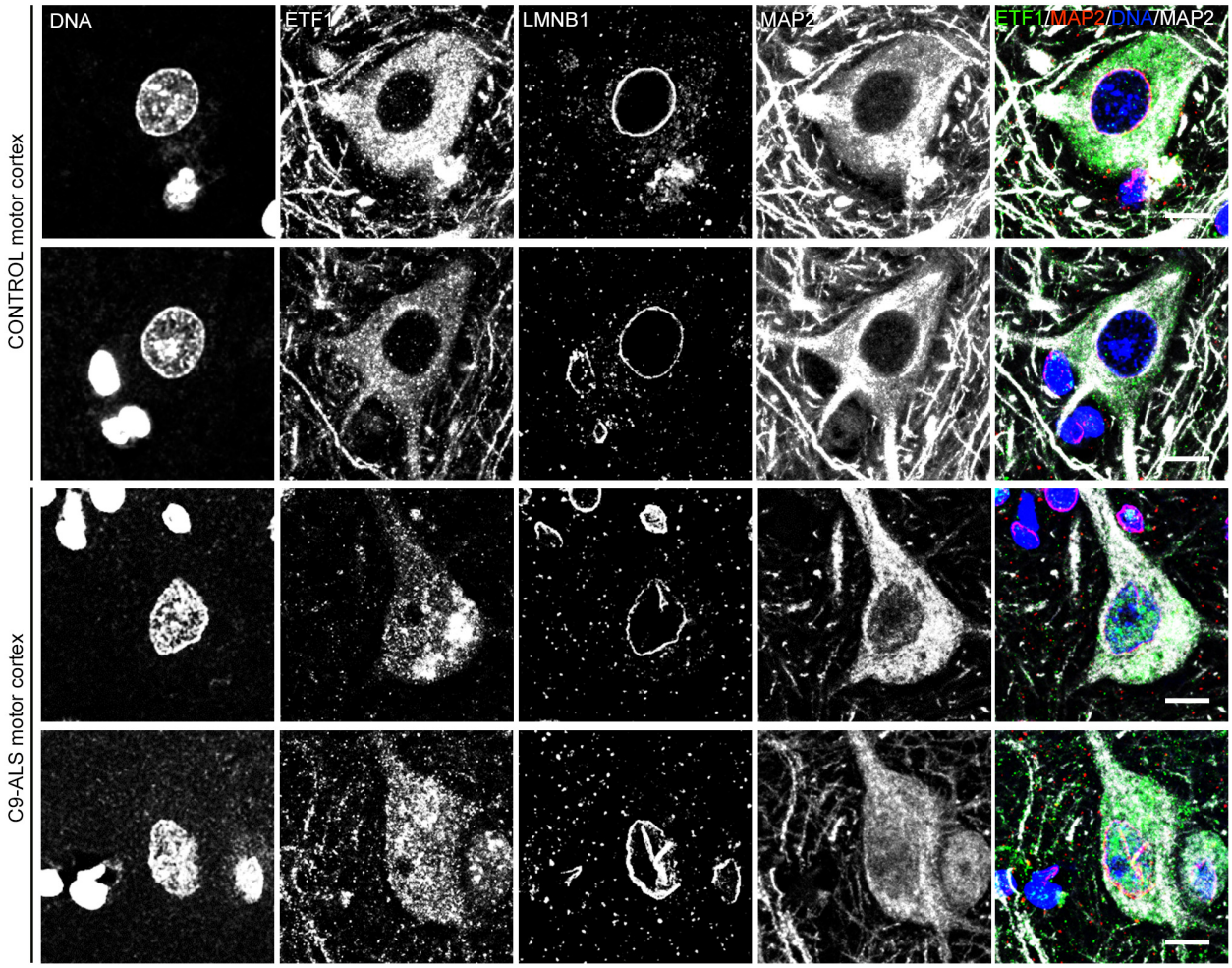

c
Velocity (cm/sec)
C9-(G4C2)x36
ETF1-RNAi
C9-(G4C2)x36 x ETF1-RNAi


d
GMR-GAL4
ETF1 OE
hETF1
β-Actin

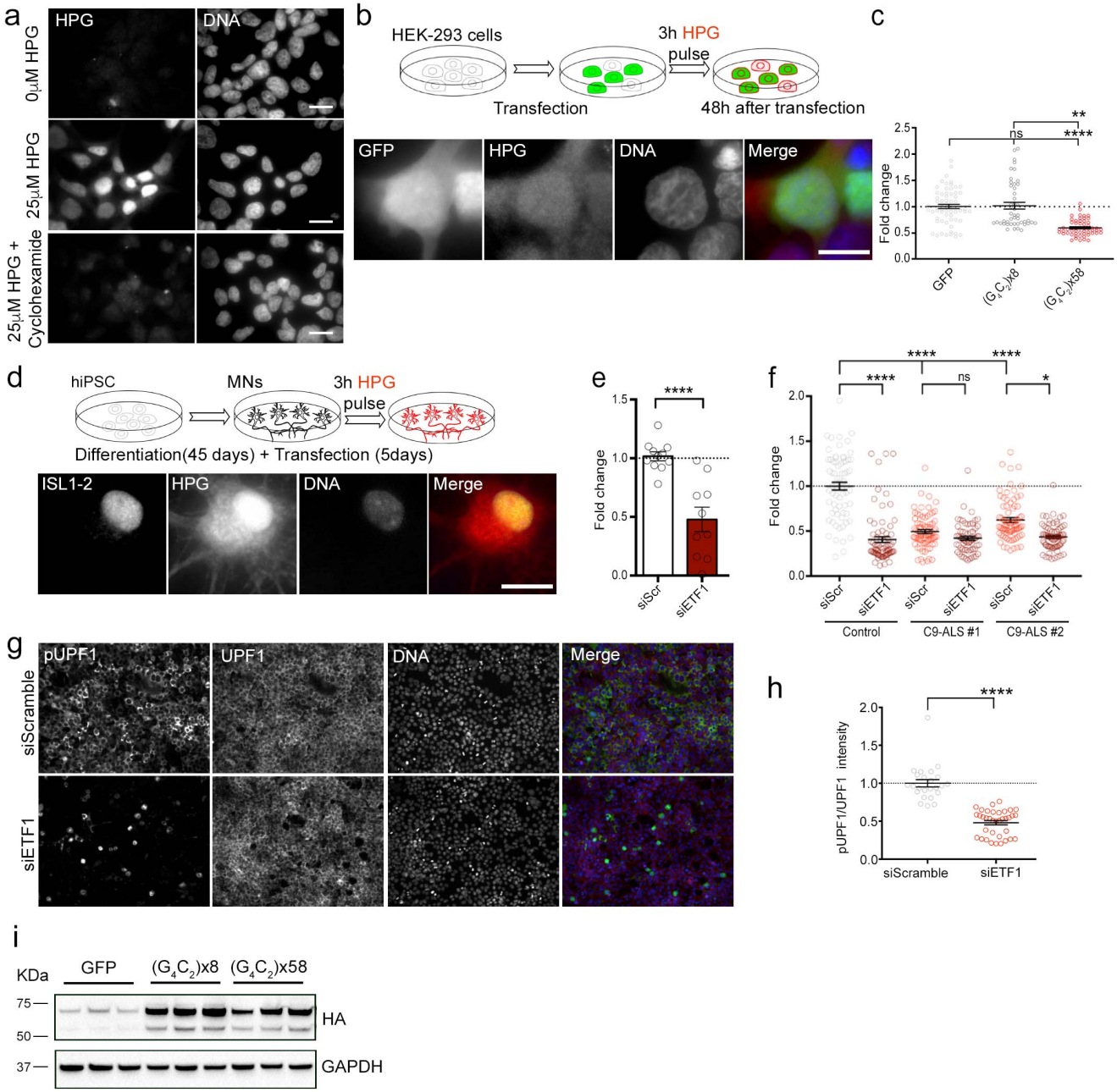
Supplementary Figure 3, related to Figure 3: Representative Redistributed Proteins are Genetic Modifiers of *C9orf72* Repeat Expansion Toxicity in *Drosophila* Disease Models. **a**, Bar plots displaying RT-qPCR results of target genes in RNAi *Drosophila* models. Individual flies are displayed as dots, and dashed line marks the respective mean levels in control flies (bars represent mean ± SEM of the respective gene expression fold change; ANOVA, p *<0.05, **<0.01, ****<0.0001). **b**, Expression of GMR-GAL4 driven C9-(G_4C_2)x30 or C9-(G_4C_2)x36 in the *Drosophila* eye led to a neurodegenerative eye phenotype. Top: *Drosophila* homologue genes of redistributed proteins in (G_4C_2)x58-expressing cells, were knock-down in GMR-GAL4 control flies without exerting any obvious eye degeneration. Bottom: However, reduction of *Drosophila* homologue genes encoding the redistributed proteins C9-HRE carrier cells, significantly modified the eye phenotype in C9-(G_4C_2)x30 and/or C9-(G_4C_2)x36 flies. ETF1, PRMT1, ENY2 and TNPO3 acted as enhancers, while CCT8 and SRSF1 acted as suppressors of toxicity. **c**, Bar plots displaying motor function alterations in C9-(G_4C_2)36, ETF1-RNAi and C9-(G_4C_2)36 x ETF1-RNAi flies. Individual flies are displayed as dots, bars represent mean ± SEM of the climbing velocity of each fly group, and dashed line represent the mean velocity in control flies; ANOVA, p *<0.05, **<0.01, ****<0.0001. **d**, Western blot showing the expression of human ETF1 protein in the ETF1-OE fly. β-Actin was used as a loading control.



Supplementary Figure 4, related to Figure 4: ETF1 is Redistributed in *C9orf72* iPSC-Derived Motor Neurons. a, Western blot analysis showing the protein levels of ETF1 in the nuclear and cytosolic fractions of cells expressing (G_4C_2)x8-GFP and (G_4C_2)x58-GFP. Cells were FACS-purified based on GFP expression. VINCULIN and LMNB1 were used as cytosolic and nuclear markers respectively. **b**, Repeat prime PCR analysis in C9-ALS patient iPSC lines utilized in this study. **c**, Top: Western blot analysis showed the ETF1 protein expression in nuclear and cytosolic fractions of iPSC-derived MNs derived from healthy controls and C9-ALS patients. Bottom: Heat map showing the average expression levels of ETF1 corrected by the respective subcellular fraction-loading controls; VINCULIN for cytosolic, and LMNA/C for nuclear fractions (n= 5-7 independent differentiations).



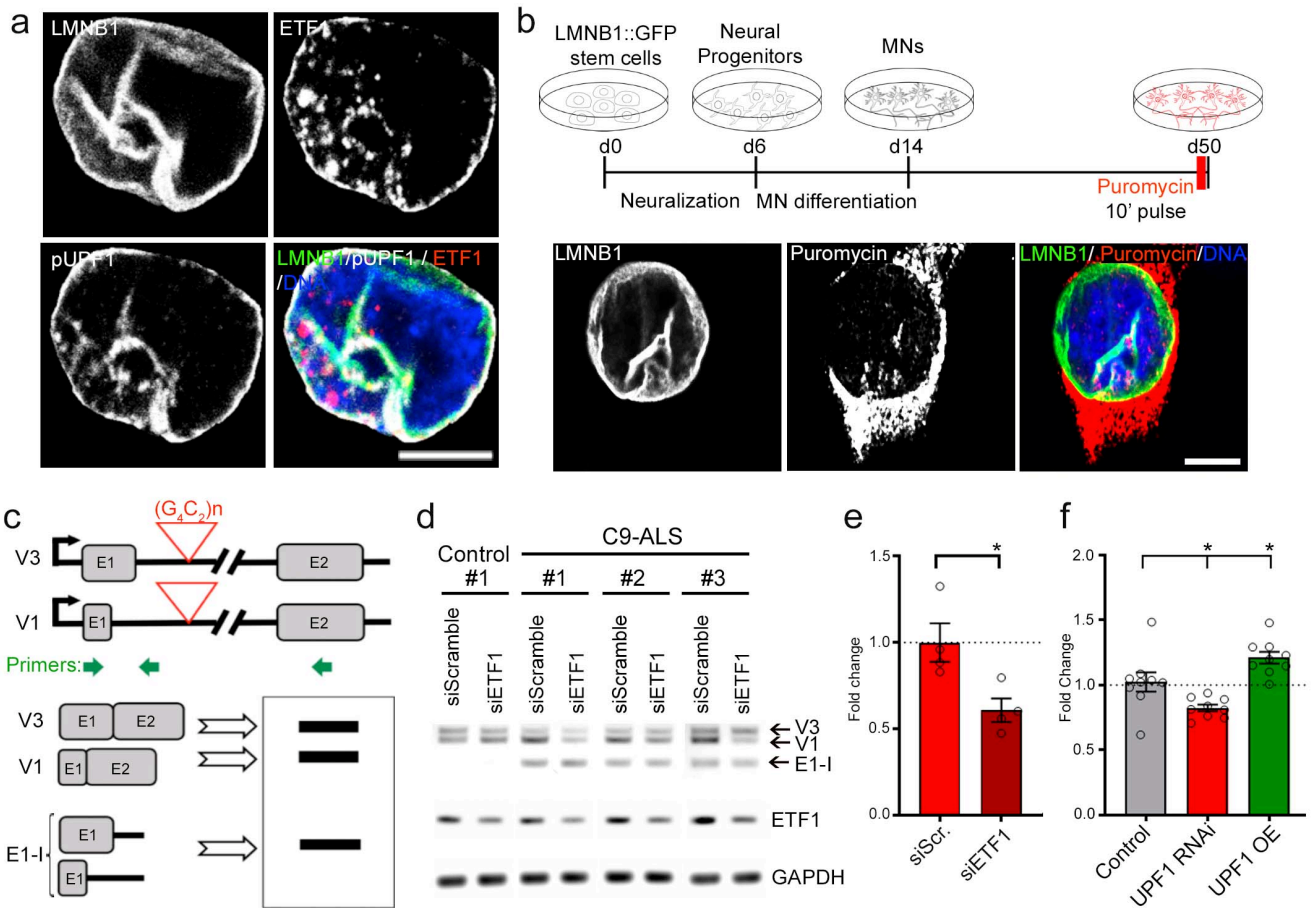
Supplementary Figure 5, related to Figure 5: ETF1 Accumulates in Neuronal Nuclei in Postmortem Tissue of C9orf72 ALS Patients. Representative IHC confocal images of postmortem cortical tissue from C9-ALS patients exhibiting significant changes in N/C ETF1 signal, and in nuclear area occupied by LMNB1+ nuclear invaginations compared to healthy controls. Scale bar: 10μm.


a
HPG
DNA
0μM HPG
25μM HPG
25μM HPG + Cyclohexamide
b
HEK-293 cells
3h HPG pulse
Transfection
48h after transfection
GFP
HPG
DNA
Merge
c
Fold change
ns
**
****
GFP
(G4C2)x8
(G4C2)x58
d
hiPSC
MNs
3h HPG pulse
Differentiation(45 days) + Transfection (5days)
ISL1-2
HPG
DNA
Merge
e
Fold change
****
siScr
siETF1
f
Fold change
****
ns
*
siScr
siETF1
Control
C9-ALS #1
C9-ALS #2
g
pUPF1
UPF1
DNA
Merge
siScramble
siETF1
h
pUPF1/UPF1 intensity
****
siScramble
siETF1

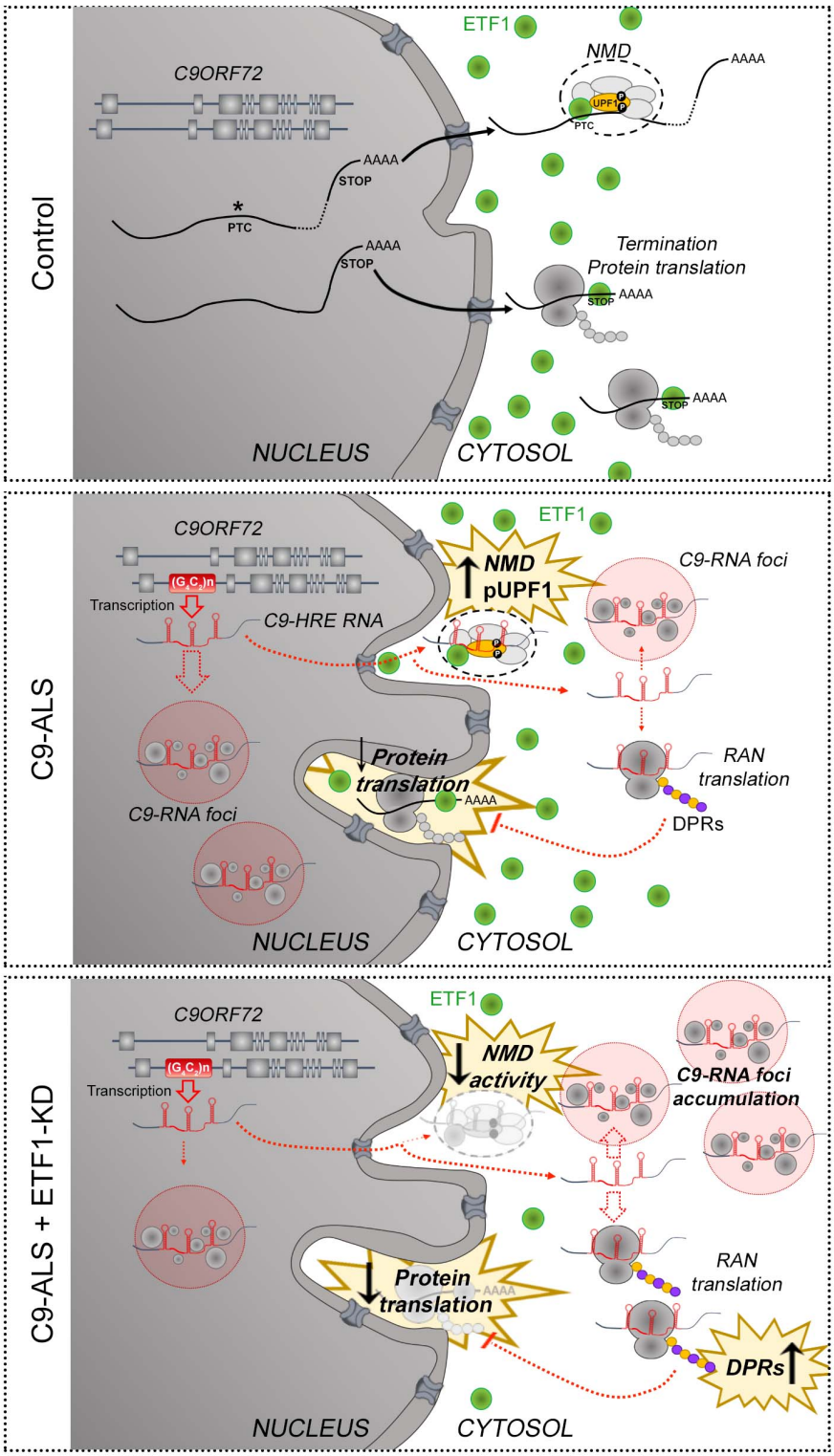
i

KDa
GFP
(G4C2)x8
(G4C2)x58
75
50
37
HA
GAPDH

Supplementary Figure 6, related to Figure 6: ETF1 is a Key Regulator of Protein Translation and NMD. a, Quality control analysis of Click-iT HPG method utilized in this study to measure *de novo* protein translation by fluorescence intensity. Top: Cells treated with 0μM HPG demonstrated low HPG signal. Middle: Cells treated with 25 μM HPG demonstrated significant HPG signal. Bottom: Cells treated with the translation inhibitor cyclohexamide demonstrated expected reduction in HPG signal. **b**, Top: schematic of *de novo* protein translation assay performed in transfected HEK-293 cells. Bottom: representative image of Click-iT HPG fluorescent labeling of *de novo* protein synthesis in GFP-expressing cells. Scale bar: 20 μm. **c**, Dot plot graph displaying the levels of *de novo* protein translation in GFP, (G_4C_2)x8- and (G_4C_2)x58-expressing cells with individual cells displayed as dots and dashed line indicating the mean level in the GFP control condition (n=2 independent biological replicates, data are represented as mean ± SEM; Kruskal-Wallis, p-values were calculated for each independent biological replicate by two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli; p =ns, not significant, **<0.01, ****<0.0001). **d**, Top: schematic of *de novo* protein translation assay performed in iPSC-derived MNs. Bottom: representative image of Click-iT HPG fluorescent labeling of *de novo* protein synthesis in an ISL1-2+ MN. Scale bar: 20 μm. **e**, RT-qPCR of ETF1 mRNA in control and C9-ALS iPSC-derived MNs upon treatment with scrambled or ETF1 siRNA. Each dot represents an independent sample (n=3 independent differentiations; data are represented mean ± SEM, and dashed line indicates ETF1 levels in siSCR-transfected MNs; T-test, p ****<0.0001). **f**, Dot plot graph displaying the levels of *de novo* protein translation in control and C9-ALS MNs transfected with scramble and ETF1 siRNA. Individual cells are represented as dots and dashed line marks the mean level in siSCR-transfected control MNs (n=3 independent differentiations; horizontal bars represent mean ± SEM; Kruskal-Wallis, p-values were calculated for each independent biological replicate by two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli, p=ns, not significant; *<0.05; ****<0.0001). **g**, ICC analysis of NMD activation, assessed by the pUPF1/UPF1 ratio upon ETF1 knockdown in HEK-293 cells. Scale bar: 50 μm. **h**, Quantification of pUPF1/UPF1 levels visualized in (g), with each dot representing a field of view, and dashed line marks the ratio mean in the siSCR control condition (n=2-3 biological replicates; horizontal bars represent mean ± SEM; Mann-Whitney U test, p ****<0.0001). **i**, Western blot analysis of HA-tagged NMD reporter products in GFP, (G_4C_2)x8- and (G_4C_2)x58-transfected cells. GAPDH was used as a loading control.



Supplementary Figure 7, related to Figure 7: NMD Targets the Expanded *C9orf72* Transcript. a, Representative ICC confocal image (nucleus only) of ETF1 (red), pUPF1 (white), LMNB1 (green), and DNA (blue). A proportion of ETF1 and pUPF1 colocalize within LMNB1+ nuclear invaginations. Scale bar: 5 μm. **b**, Top: schematic of translation analysis performed in iPSC-derived MNs. Bottom, representative SIM image of puromycin (red), LMNB1 (green) and DNA (blue). Bottom: Representative ICC confocal image of puromycin (red), LMBN1 (green) and DNA (blue). Puromycin+ puncta labels newly synthesized proteins found within LMNB1+ nuclear invaginations. Scale bar: 5 μm. **c,** Schematic of the *C9orf72* intron retention assay. **d**, Semi-quantitative RT-PCR analysis of the ratio of the mutant relative to wild type transcript in one control and three C9-ALS patient-derived MNs treated with scrambled or ETF1 siRNA. RT-PCR was also performed for ETF1 to validate its knock-down when transfecting MNs with siRNAs, and for GAPDH as an internal control. **e**, Bar plots showing the quantification of ETF1 levels visualized in (**d**). Individual samples are displayed as dots, and dashed line marks the mean ETF1 levels in scramble siRNA-transfected MNs. Bars represented the mean ± SEM; T- test, p *<0.05. **f,** qPCR of UPF1 levels in the Drosophila models of UPF1-KD (RNAi) and UPF1-OE. Every dot represent individual flies, and bars represent mean ± SEM of UPF1 mRNA levels. Dashed line marks the mean level in control flies. T-test; p *<0.05.



Supplementary Figure 8: Proposed Model for the Role of ETF1, NMD and LMNB1+ Nuclear Envelope Invaginations in *C9orf72*-Patient Motor Neurons.

Supplementary Table 1, related to Figure 1. LC-MS/MS data acquired in HEK-293 cells transfected with plasmids expressing GFP, (G_4C_2)x8, (G_4C_2)x58, ATX1-32Q, ATX1-84Q, HTT-25Q, or HTT-72Q. The table includes six sheets with information related to figure 1 panels:
Fig. 1c: Cytosolic and nuclear ratios of *bona fide* cytosolic and nuclear markers.
Fig. 1d,e: LC-MS/MS raw data obtained from GFP, (G_4C_2)x8-, and (G_4C_2)x58-transfected cells.
Fig. 1g: LC-MS/MS raw data obtained from GFP, ATX1-32Q- , ATX1-84Q- , HTT-25Q- , or HTT-72Q-transfected cells.
Fig. 1h: Kruskal-Wallis significance rank test for all the redundant proteins identified in cells transfected with plasmids expressing GFP, (G_4C_2)x8, (G_4C_2)x58, ATX1-32Q, ATX1-84Q, HTT-25Q, or HTT-72Q. List of all proteins that were significantly altered in each one of the three repeat expansion mutations (C9-ATX-HTT comparison).

Supplementary Table 2, related to Figure 2. Comprehensive analysis of the subcellular proteomic differences detected in (G_4C_2)x58-carrier cells. The table includes seven sheets with information related to figure 2 panels:
Fig. 2a: Mean ± SD N/C ratio changes of proteins that are redistributed proteins in cells expressing (G_4C_2)x58.
Fig. 2c: Molecular weights of redistributed proteins in cells expressing (G_4C_2)x58.
Fig. 2d: Gene Ontology analysis of redistributed proteins in cells expressing (G_4C_2)x58.
Supplementary Fig. 2a: Comparative subcellular localization analysis of proteins redistributed in cells expressing (G_4C_2)x58.
Supplementary Fig. 2d: NES and NLS sequences in proteins redistributed in cells expressing (G_4C_2)x58.
Supplementary Fig. 2e: List of proteins previously identified as karyopherin cargoes.
Supplementary Fig. 2f: List of proteins previously identified as C9-RNA and C9-DPR interactors.

Supplementary Table 3, Information related to iPSC lines utilized in the present study.

METHODS

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture conditions

HEK-293FT cells were grown in DMEM (Corning) supplemented with Glutamax (Gibco) and 10 % fetal bovine serum (FBS, VWR). HEK-293FT cells were dissociated by incubating for 5 min with Trypsin-EDTA (Gibco) at 37°C. Stem cells were maintained on Matrigel (BD Biosciences) with mTeSR1 media (Stem Cell Technologies) and passaged on a weekly basis using 1mM EDTA or Accutase (Sigma). All cell cultures were maintained at 37°C, 5% CO_2 without antibiotics and tested on a monthly basis for mycoplasma. All cell lines utilized in the study were mycoplasma-free.

Stem cell cultures and motor neuron differentiation

iPSCs were dissociated with Accutase and plated with mTESR1 and 10μM ROCK inhibitor (Y-27632, DNSK International) at a density of 100,000 cells/cm^2. Next day, media was replaced with N2B27 medium (50% DMEM:F12, 50% Neurobasal, supplemented with NEAA, Glutamax, N2 and B27; Gibco) supplemented with a cocktail of small molecules that induces the generation of spinal neural progenitors: 10μM SB431542 (DNSK International), 100nM LDN-193189 (DNSK International), 1 μM Retinoic Acid (RA, Sigma) and 1μM of Smoothened-Agonist (SAG, DNSK International). Culture medium was changed daily for 6 days and then was switched to N2B27 medium supplemented with 1μM RA, 1μM SAG, 5μM DAPT (DNSK International) and 4μM SU5402 (DNSK International) to generate postmitotic spinal MNs. Cells were fed daily with this media for seven days and then dissociated using TrypLE Express (Gibco) supplemented with DNase I (Worthington). Cells were plated onto pre-coated Matrigel-coated surfaces (BD Biosciences) and cultured in Neurobasal medium supplemented with NEAA, Glutamax, N2, B27, Ascorbic acid (0.2 μg/ml) and BDNF, CNTF and GDNF (10 ng/mL, R&D systems). For imaging analysis, MNs were seeded on pre-plated mouse glia cells.

METHOD DETAILS

Overexpression and knockdown transfections

For overexpression experiments, 40% confluent HEK-293 cell cultures were transfected with HilyMax transfection reagent (Dojindo Molecular Technologies) according to manufacturer guidelines. Briefly, exogenous DNA was mixed with HilyMax (1ug DNA:3uL HilyMax ratio) in Opti-MEM medium (Gibco) and incubated for 15 min at room temperature (RT) before being added to cells. Cells were incubated with transfection mixture for 4 hours at 37°C and then media was replaced. Analyses made on transfected cells were performed 48 to 72 hours after transfection. For transient knock down experiments, pre-designed siRNAs (Ambion Silencer Select) were transfected using Lipofectamine RNAiMAX (Invitrogen) according to manufacturer guidelines. Briefly, for 2 wells of a 24-well plate with 0.5-2 x 10^5 cells/well, 10 pmol of siRNA were mixed with 3 μL of Lipofectamine RNAiMAX reagent in 100 μL Opti-MEM medium and incubated for 15 min at RT. 400 μL of cell medium were then added and 250 μL distributed per well. Analyses of siRNA transfection experiments were performed 3-4 days after transfection.

Plasmids

For Mass Spectrometry experiments, HEK-293 cells were transfected with the following DNA constructs: pcDNA3.1 plasmids containing GFP or (G_4C_2)x8-GFP or (G_4C_2)x58-GFP, were

provided by J. Paul Taylor's lab; pEGFP plasmids containing HTT exon1 with (CAA CAG CAG CAA CAG CAA)n repeats amounting to either 25Q or 72Q repeats, were provided by Dimitri Krainc's lab; pLEX plasmids containing Atx1-32Q or Atx1-84Q, were provided by Puneet Opal's lab. The pcDNA3.1 plasmids containing GFP, $(GA)_{50}$, $(GR)_{50}$, $(PA)_{50}$, or $(PR)_{50}$, in which alternative codons were used to generate dipeptide repeat proteins without generating the (GGGGCC)n transcript were originally made by Davide Trotti[86] and modified by Christopher Donnelly. The pLenti-c-Myc-DDK plasmid containing NLS-tdTomato-NES, was provided by Jeffrey Rothstein[21]. The pBS-TCR(PTC)-TCR(WT) NMD reporter plasmid was provided by Zhongsheng You[57].

Flow cytometry and nucleocytoplasmic fractionation

Cell cultures were dissociated with trypsin and resuspended in media containing DMEM without phenol red (Gibco) and 2% FBS before flow cytometry cell sorting on BD FacsAria SORP 6-Laser. Cells were collected in PBS supplemented with protease inhibitors (Millipore). Cell suspensions were immediately spun down and subcellular fractionation was performed using NE-PER nuclear and cytoplasmic extraction reagents (Thermo Scientific). Briefly, cells were harvested using a hypertonic detergent solution that disrupted plasma membranes, leaving nuclear membranes intact. This lysate was spun down and supernatant was collected as the cytosolic protein fraction. Nuclear proteins were isolated from the remaining pellet using high salt extraction buffer. The obtained nuclear and cytosolic fractions were used for Western blot and mass spectrometry analysis.

Liquid Chromatography Mass Spectrometry (LC-MS/MS)

Nuclear and cytosolic extracts were subjected to methanol and chloroform precipitation, the precipitated protein pellets were solubilized first in 100 μl of 8 M urea for 30 min, and next in 100 μl of 0.2% ProteaseMAX (Promega) for an additional 2 h. The protein extracts were reduced and alkylated as described by[87], followed by the addition of 300 μl of 50 mM ammonium bicarbonate, 5 μl of 1% ProteaseMAX, and 20 μg of sequence-grade trypsin (Promega). Samples were digested overnight in a 37°C thermomixer (Eppendorf). We used 3 μg of peptides in each Orbitrap Fusion MS analysis.

For Orbitrap Fusion Tribrid MS analysis, the tryptic peptides were purified with Pierce C18 spin columns and fractionated with two ACN concentrations (30% and 70%). Three micrograms of each fraction were auto-sampler loaded with a Thermo Fisher EASY nLC 1000 UPLC pump onto a vented Acclaim Pepmap 100, 75 μm x 2 cm, nanoViper trap column coupled to a nanoViper analytical column (Thermo Fisher 164570, 3 μm, 100 Å, C18, 0.075 mm, 500 mm) with stainless steel emitter tip assembled on the Nanospray Flex Ion Source with a spray voltage of 2000 V. Buffer A contained 94.785% H2O with 5% ACN and 0.125% FA, and buffer B contained 99.875% ACN with 0.125% FA. The chromatographic run was for 4 h in total with the following profile: 0–7% for 7 min, 10% for 6 min, 25% for 160 min, 33% for 40 min, 50% for 7, 95% for 5 min, and 95% again for 15 min, respectively. Additional MS parameters include: ion transfer tube temp 300°C, Easy-IC internal mass calibration, default charge state = 2 and cycle time = 3 s. Detector type set to Orbitrap, with 60 K resolution, with wide quad isolation, mass range normal, scan range = 300–1500 m/z, max injection time = 50 ms, AGC target = 200,000, microscans = 1, S-lens RF level = 60, without source fragmentation, and datatype = positive and centroid. MIPS was set as on, included charge states = 2–6 (reject unassigned). Dynamic exclusion enabled with *n* =1 for 30 and 45 s exclusion duration at 10 ppm for high and low. Precursor selection decision = most intense, top 20, isolation window = 1.6, scan range = auto normal, first mass = 110, collision energy 30%, CID, Detector type = ion trap, Orbitrap resolution

= 30K, IT scan rate = rapid, max injection time = 75 ms, AGC target = 10,000, Q = 0.25, inject ions for all available parallelizable time.

Tandem mass spectra analysis

Samples of 4 independent biological replicates, which included both nuclear and cytosolic fractions of each of the 3 transfected conditions (GFP, 8 and 58 x G_4C_2), were analyzed independently by LC-MS/MS. The spectral files from all replicates were pooled for a single database search. Spectrum raw files were extracted into MS1 and MS2 files using in-house program RawXtractor or RawConverter (http://fields.scripps.edu/downloads.php)[88] and the tandem mass spectra were searched against UniProt human protein database (downloaded on 12-19-2015; UniProt Consortium, 2015) and matched to sequences using the ProLuCID/SEQUEST algorithm (ProLuCID version 3.1;[89, 90] with 50 ppm peptide mass tolerance for precursor ions and 600 ppm for fragment ions. The search space included all fully and half-tryptic peptide candidates that fell within the mass tolerance window with no miscleavage constraint, assembled, and filtered with DTASelect2 (version 2.1.3)[91, 92] through Integrated Proteomics Pipeline IP2 version 3, Integrated Proteomics Applications (http://www.integratedproteomics.com). In order to estimate peptide probabilities and false-discovery rates (FDR) accurately, we used a target/decoy database containing the reversed sequences of all the proteins appended to the target database[93]. Each protein identified was required to have a minimum of one peptide of minimal length of six amino acid residues; however, this peptide had to be an excellent match with a FDR = 0.001 and at least one excellent peptide match. After the peptide/spectrum matches were filtered, and we estimated that the protein FDRs were = 1% for each sample analysis. Resulting protein lists include subset proteins to allow for consideration of all possible protein forms implicated by a given peptide identified from the complex IE protein mixtures. Each protein identified with the IP2 pipeline was associated with several different measures of abundance used in our analyses, including: peptide counts, spectral counts, and normalized spectral abundance factor (NSAF)[94], which takes into account protein length and number of proteins identified in the experiment. When comparing ratios and abundances of a given protein across samples, we used NSAF rank rather than abundance to minimize the effects of differences in sample sizes and stochastic differences between MS analyses.

To estimate the number of proteins detected in every experimental condition, we assessed the number of protein entries that displayed detectable NSAF values in either the cytosolic, the nuclear or in both subcellular fractions.

To calculate the cytosolic or nuclear ratio of the detected proteins in the different samples, we applied the following equation:

$$\text{Cytosolic ratio (prot}_y\text{)} = \frac{\text{NSAF cytosol}_y}{\text{NSAF cytosol}_y + \text{NSAF nucleus}_y}$$

or

$$\text{Nuclear ratio (prot}_y\text{)} = \frac{\text{NSAF nucleus}_y}{\text{NSAF cytosol}_y + \text{NSAF nucleus}_y}$$

To assess the purity of the subcellular fractionation method utilized in this study, we calculated the cytosolic and nuclear ratio values of *bona fide* nuclear and cytosolic markers in GFP control cells, that were plotted as displayed in Figure 1c and Supplementary Figure 1f. To assess proteomic changes among the multiple transfected conditions tested in this study, we calculated the average ± SD of both nuclear and cytosolic ratios of every detected protein

obtained from 3-4 independent experimental replicates (see Supplementary Table 1). Obtained results were plotted in violin plots to graphically show the subcellular proteome distribution among the different experimental conditions (Figures 1e and 1g).

Gene Ontology (GO) analysis was performed using the online tools STRING, PANTHER and DAVID. The COMPARTMENTS (compartments.jensenlab.org) tool was utilized to characterize the normal protein localization patterns of the N/C perturbed proteins and assess the physiological relevance of the observed changes. Bioinformatics & Evolutionary Genomics (ioVe://bioinformatics.psb.ugent.be/webtools/Venn/) and BioVenn, are web applications utilized for the comparison and visualization of biological lists using area-proportional Venn diagrams[95]. NLS sequences detection in the proteins of interest was performed with cNLS Mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) and NES sequences with NetNES 1.1 Server (http://www.cbs.dtu.dk/services/NetNES/).

Cell viability analysis

Two different approaches were utilized to assess cell viability in HEK-293 cells transfected with GFP, (G_4C_2)x8-GFP and (G_4C_2)x58-GFP. One, Propidium Iodide (PI) incorporation by apoptotic cells was measured in cells 48 hrs after transfection. Cells were dissociated with trypsin, resuspended in DMEM without phenol red (Gibco) + 2% FBS. PI (5µg/ml) was added to the cell suspension right before flow cytometry analysis on BD LSRFortessa. Two, the colorimetric MTT assay, wherein cells were exposed to 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT, 0.5mg/ml), which is reduced to formazan (purple) depending on cellular NADPH activity. Formazan levels were measured by a spectrophotometer at 500-600 nm.

Western blot analysis

Cells were harvested in RIPA buffer (10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1mM EDTA, 0.1 % SDS), supplemented with 1% Triton X-100 (Triton, Sigma Aldrich), 1 mM phenylmethylsulphonyl fluoride (DOT Scientific), and protease inhibitors (Millipore). Lysate was sonicated and protein extracts were separated by SDS-PAGE followed by electro-transfer to a nitrocellulose membrane (Bio-Rad). The membranes were blocked in Tris-buffered saline (TBS, 50mM Tris, 150mM NaCl, HCl to pH 7.6)) + 0.1% tween (Sigma-Aldrich) + 5% non-fat dry milk (Labscientific) and then incubated overnight at 4ºC with primary antibodies: α-tubulin (mouse, 1:5000, Santa Cruz Biotechnology), TAFII p130 (mouse, 1:1000, Santa Cruz Biotechnology), ETF1 (mouse, 1:500, Santa Cruz), GAPDH (rabbit, 1:1000, Cell Signaling), GFP (goat, 1:1000, Abcam), HA (mouse, 1:2000, BioLegend,), Puromycin (mouse, 1:25000, Millipore), LAMIN A/C (mouse, 1:10000, generously provided by Robert Goldman's lab), LMNB1 (rabbit, 1:1000, Proteintech), VINCULIN (mouse, 1:500, Sigma-Aldrich). Primary antibodies were diluted in TBS + 0.1% Tween + 5% BSA (Calbiochem). After several washes in TBS + 0.1% Tween, membranes were incubated with their corresponding secondary HRP-conjugated antibodies (1:5000, LI-COR Biotechnology). Protein signals were detected by a ChemiDoc™ XRS+ (Bio-Rad), using the SuperSignal West Pico chemilluminescent system (Thermo Fisher Scientific). Densitometry analysis for the bands was performed using Fiji software (ImageJ, NIH Image).

***Drosophila* stocks, maintenance, and phenotyping**

To identify genetic modifiers of C9-HRE toxicity, *Drosophila melanogaster* lines expressing RNAi for homologues of specific targets were obtained from the Vienna Drosophila Resource Center (Erf1, stock no # 33569, SF2 # 27775, Art1 # 110391; e(y)2 # 108212, Tnpo-SR #40991, CCT8 # 103905) or Bloomington Drosophila Stock Center (Upf1 # 64519) and crossed with a C9ORF72– UAS-(G_4C_2)x30 line, which was a kind gift from Dr Peng Jin, or a C9ORF72-UAS-(G_4C_2)x36 line (Bloomington Drosophila Stock Center). Drosophila stocks utilized in this study

were maintained on standard cornmeal medium at 25°C or 18°C in light/dark-controlled incubators. External eye degenerative phenotype was used as a screening endpoint as described previously[96, 97]. Briefly, the following parameters were scored as 0 if normal or 1 if abnormal: level of discoloration, ommatidial disorganization, absence of hairs, presence of necrotic tissue, level of development, shape, and contour. The sum of the scores defined the level of eye degeneration. All data was collected and scored in a blinded manner.

The Rapid Iterative Negative Geotaxis (RING) was performed as described in[98]. Briefly, RNAi flies with or without the C9-HRE expressing in neurons (ELAV-gene switch) were placed into polystyrene vials within the RING apparatus, wherein flies can be pulled to the bottoms of vials and climbing movement recorded. Movement was recorded using a iPhone 8 camera, climbing velocity per fly was calculated.

Drosophila dot blot, Western blot and RT-qPCR analysis

Dot blot was performed by following the protocol by[41] with some modifications. Briefly, 4 day old flies were heat-shocked for 1.5 hours at 37°C, snap-frozen and homogenised in RIPA buffer (3 µl per fly head) with complete protease inhibitors cocktail (Roche). After adding 6x Laemmli sample buffer (0.5 µl per head), the lysate was heated to 95°C for 5 minutes and 2 µl of the samples were blotted onto dry nitrocellulose membrane. The blot was left to dry for 15 minutes at room temperature and blocked for an hour in 5% BLOT-QuickBlocker reagent in TBS containing 0.05% Tween-20 (TBS-T). The membrane was probed with rabbit anti-poly-(GR) antibody (1:2000 in TBS-T), or mouse anti-tubulin antibody (1:10,000) overnight at 4 °C. The following day, the membrane was washed twice with TBS-T and incubated in secondary antibody (anti-rabbit IRDye 800 and anti-mouse IRDye 680D, LICOR, 1: 10,000) for one hour and imaged on Licor imager (Odyssey CLx).

For ETF1 Western blot, three adult fly heads were collected from each cross and crushed on dry ice. The granulated heads were incubated in the lysis buffer containing 150 mM NaCl, 0.1% SDS, 1% NP40, 50 mM NaF, 2 mM EDTA, 1% sodium deoxycholate, 1mM DTT, 0.2 mM sodium orthovandate and protease inhibitor (Roche) for 10 min. The lysate was sonicated, centrifuged at 10000xg for 10 min and boiled at 95ºC for 5 min in laemmli buffer. The proteins were separated in 4-12% NuPAGE gel (Novex/ Life technologies) and transferred to nitrocellulose membrane (Invitrogen). The blot was blocked in 2.5% milk in 0.1% TBST and probed with anti-ETF1 (Santa Cruz Biotechnology) and β-actin (Cell signaling) antibodies overnight at 4ºC. The blot was washed in 0.1% TBST and incubated with Alexa Flour-tagged secondary for 1 hour. The blot was imaged in Odyssey CLx (Li-COR Biosciences).

RNA was extracted from fly heads expressing RNAi for fly homologue of specific targets in triplicate (3 heads per replicate) using Trizol (Ambion) and 1-Bromo-3-chloropropane (SIGMA) extraction[99]. RNA samples were cleaned using DNA-free kit (Invitrogen) and quantified using NanoDrop ND-1000 spectrophotometer. cDNA was synthesized from 0.4 ug of DNase free RNA by using iScript Select cDNA Synthesis Kit (BioRad). qPCR was performed on an Applied Biosystems, 7300 Real Time PCR System by using the iQ Supermix (BioRad) and α-tubulin was used as a normalizing control. The comparative C(T) method was used for analyzing the results, as previously described[100], using GraphPad Prism 6 software for statistical analyses.

Gene (*Drosophila*)	Direction	Sequence (5' to 3')	Technique
TUBULIN	Forward	TTTGAGACCTTCAACACCCCAGCC CCTCGAAATCGTAGCTCTACAC	qPCR
	Reverse	AATGTCACGCACGATTTCCCGC ACCAAGCCTGACCAACATG	
ETF1 (eRF1)	Forward	CTCTACCTCTGCGACAACAAG	qPCR

	Reverse	AATAGCGCTCCGTTACCATC	
SRSF1 (SF2)	Forward	CGGACACTTACAAGGATGGTTC	qPCR
	Reverse	ATCGGAAGCGAGAGTCGT	
PRMT1 (Art1)	Forward	CTCGTCCAAGTTCAGCCTATG	qPCR
	Reverse	AAGACGCTTGTGGCACTT	
ENY2 (e(y)2)	Forward	GCGACCGTTCCAAGATCAA	qPCR
	Reverse	CTCTACGGTGAAGCTGTTGTT	
CCT8	Forward	CAGGTGTCAAGGTGGTTGT	qPCR
	Reverse	CGGCCATCAGACCGTATTT	
TNPO3 (Tnpo-SR)	Forward	GCACACGCACGAATCTCT	qPCR
	Reverse	ACTGAGCTGCGTCACAATC	
UPF1 (Upf1)	Forward	AACATGCCAGTTCCGGTT	qPCR
	Reverse	CTGCTGGTGCTGGTTGTA	

ELISA

Protein was isolated in the same manner as above. Polyclonal anti-GP capture and detection antibodies in conjunction with Meso Scale Discovery (MSD) electrochemiluminescence detection were used as previously described[101].

RNA preparation and RT quantitative PCR (qPCR)

RNA was isolated using RNeasy Mini Kit (Qiagen). Briefly, cells were harvested in lysis buffer and gDNA was removed by passing lysate through a gDNA eliminator column. RNA was then isolated according to manufacturer guidelines and 0.5 to 1 µg of RNA was used to generate cDNA using iSCRIPT Reverse Transcription Supermix (Bio-Rad). Real-time qPCR reactions were prepared according to manufacturer guidelines for iTaq Universal SYBR Green Supermix (Bio-Rad) on the CFX system (Bio-Rad). All qPCR assays were performed in triplicate. The averaged cycle of threshold (Ct) value of the two housekeeping genes (ACTIN/GAPDH) was subtracted from the Ct value of the gene of interest to obtain the ΔCt. Relative gene expression was determined as $^{2-}$ΔCt (ΔΔCt) and expressed relative to the control sample or the highest expressed sample in the experiment.

Gene	**Direction**	**Sequence (5' to 3')**	**Technique**
Actin	Forward	TTTGAGACCTTCAACACCCCAGCC	qPCR
	Reverse	AATGTCACGCACGATTTCCCGC	
GAPDH	Forward	AGGTCGGTGTGAACGGATTTG	qPCR
	Reverse	TGTAGACCATGTAGTTGAGGTCA	
ETF1	Forward	GGAAATCTACTCGGTACCGCA	qPCR
	Reverse	AGCTTATCGAGAGCATGCCC	
C9orf72 (intron retention)	Forward	GGTGCGTCAAACAGCGACAAGTTC (exon 1a)	RT-PCR
	Reverse	GTCCATATGTGCTGCGATCC (exon 2)	
	Reverse	GCTACAGGCTGCGGTTGTTTCC (intron 1-5'ss)	

Immunocytochemistry

Cells were fixed with 4% paraformaldehyde and blocked for 1h in Phosphate-buffered saline

(PBS) containing 10% normal donkey serum (Jackson ImmunoResearch) and 0.1% Triton. Samples were then incubated overnight at 4ºC with primary antibodies: ETF1 (mouse, 1:100, Santa Cruz Biotechnology), LAMIN-B1 (rabbit, 1:500, Proteintech), MAP2 (chicken, 1:5000, Abcam), pUPF1-Ser1127 (rabbit, 1:500, Millipore), UPF1 or RENT1 (mouse, 1:200, Santa Cruz Biotechnology), ISLET1/2 (mouse, 1:200, DHSB), Puromycin (mouse, 1:10000, Millipore). Primary antibody was removed and PBS/0.1% Triton was applied for several washes. Samples were then incubated with the appropriate secondary antibodies conjugated to Alexa488, Alexa555 or Alexa647 fluorophores (1:500 or 1:1000 Molecular Probes) for 1h at RT. Cell nuclei were labeled by DNA staining using Hoechst (Life Technologies). Immunolabeled samples were blinded upon mounting.

Immunohistochemistry

Postmortem tissue samples from three ALS patients and three non-neurological controls were obtained through the Northwestern University ALS Clinic and examined by immunohistochemistry. Immunohistochemistry were performed using previously described methods[102]. Briefly, 6µm sections were cut from formalin-fixed, paraffin-embedded spinal cord and brain regions containing motor cortex. Sections were deparaffinized and rehydrated by passing slides in serial solutions: 3× for 10 min each in xylene, 3× for 5 min each in 100% ethanol, 3× for 3 min each in 95% ethanol, once for 5 min each in 75% ethanol, 50% ethanol, deionized water, and PBS. Antigen retrieval was performed using a decloaking chamber with a 1X Antigen Decloaker solution (Biocare Medical) at 125°C for 10 min. Sections were cooled to room temperature for 30 min and rinsed with deionized water. Nonspecific background was blocked with 1% BSA in PBS for 20 min at room temperature. Slides were incubated overnight at 4ºC with primary antibodies, rinsed 3x for 5 min with PBS, and then appropriate secondary antibodies conjugated with Alexa488, Alexa555 or Alexa647 fluorophores (1:250, Molecular Probes) diluted in PBS, were applied to slides and incubated at room temperature for 45 min. Slides were rinsed 3 times for 5 min. To diminish autofluorescence, 0.3% sudan black in 70% ethanol was applied to each section for 45 sec, and rinsed in deionized water for 5 min. Slides were mounted in ProLong Diamond Antifade Mountant with DAPI (Thermo Fisher Scientific) and labels were blinded for subsequent image analysis.

***De novo* protein translation analysis**

For single-cell protein translation analysis, we utilized Click-iT HPG protein synthesis kit (Invitrogen). HPG (L-homopropargylglycine, 25µM) is a methionine analog with an alkyne moiety that is incorporated into newly synthesized proteins. Cells were pulsed for 3 hours with 25 µM HPG in Methionine free DMEM (Gibco) and incorporated HPG was detected by Click chemistry following manufacturer guidelines. In short, cells were fixed in 4% PFA for 15 min, washed twice with PBS/0.25% Triton and blocked with PBS/3% BSA for 1h at RT. Cells were incubated with primary antibodies (ISL1/2, MAP2) overnight at 4°C, then washed with PBS three times, and incubated with the appropriate secondary antibodies. Cells were washed with PBS/0.25% triton twice, and block with PBS/3% BSA for 30 min. Next, the click reaction for detection of HPG incorporation was performed, based on a reaction between an azide and alkyne, wherein the alkyne-modified protein is detected with Alexa Fluor® 594 azide. Alternatively, a puromycin-based method described by[103] was also used to label newly synthesized proteins. Cell cultures were pulsed for 10 min with Puromycin (20 µM) at 37°C. Cells were then fixed and ICC was carried out as described above utilizing anti-puromycin primary antibody (Millipore).

Fluorescent *in situ* hybridization (FISH)

FISH analysis was performed as previously described[11]. Briefly, cells were cultured on 12 mm coverslips in 24 well plates, fixed in 4% PFA, and permeabilized in 0.3% Triton for 15 min. Before hybridization, cells equilibrated in 2X SSC/50% formamide for 10 min at 60°C and probe mixture was preheated for 10 min at 95°C. For hybridization, 200 μL of hybridization buffer and 27 μL of probe mixture were added to each coverslip and cells were incubated for 1h at 60°C. Following hybridization, cells were incubated with 2X SSC/50% formamide for 20 min at 65°C, then again with fresh 2X SSC/50% formamide for 15 min at 65°C, and finally with 1X SSC/40% formamide for 10 min at 60°C. Cells then underwent a series of washes at RT: 3x quick washes with 1X SSC, 2x 5 min washes with 1X SSC, a 5 min wash with Tris-buffered saline, and a 5 min wash with Tris-glycine buffer. Cells were post-fixed in 3% PFA and then incubated with blocking buffer for 1h at RT. Cells incubated with primary antibodies diluted in immunofluorescence buffer overnight at 4ºC and secondary antibodies diluted in immunofluorescence buffer for 20 min at RT. After secondary antibody incubation, cells were sequentially washed with immunofluorescence buffer, Tris-buffered saline, Tris-glycine buffer, 2mM $MgCl_2$ in PBS, and PBS. Coverslips were mounted with ProLong Gold Antifade mounting media with DAPI (Invitrogen) and blinded.

Hybridization buffer consisted of 100% formamide, BSA (20 mg/ml, Roche), 20 mM ribonucleoside vanadyl complex (Sigma-Aldrich), 100mM $NaPO_4$, and 20X SSC. Probe mixture consisted of salmon sperm (10 μg/ μL, Thermo Fisher Scientific), *E. coli* tRNA (20 μg μL, Thermo Fisher Scientific), 80% formamide, and 25 μM 5' digoxigenin (DIG)-labelled CCCCGGCCCCGGCCCC locked nucleic acid probe (Exiqon, batch 620574). Tris-buffered saline comprised a 50mM Tris and 15mM NaCl solution, pH 7.4. Tris-glycine buffer comprised a 0.75% glycine and 200mM Tris solution, pH 7.4. Blocking buffer consisted of 1% normal donkey serum (Jackson Immunoresearch) and 5% heat-shocked BSA in Tris-buffered saline. Immunofluorescence buffer consisted of 2% heat-shocked BSA in Tris-buffered saline. DIG-labelled probe was detected with a fluorescein-conjugated sheep anti-DIG antibody (1:250, Roche). All buffers were made with RNase-free water or PBS.

Quantitative image acquisition and analysis

Images used for quantitation were acquired at matched exposure times or laser settings and processed using identical settings. Quantifications were normalized within their respective experiment with n> 3 independent experiments. Image acquisition for HEK-293 experiments was performed on a Leica DMI4000B laser scanning confocal microscope (Leica, Buffalo Grove, IL). Image acquisition for MNs and post-mortem tissue experiments was performed on a Nikon W1 dual camera spinning disk confocal microscope and a Nikon A1R+ laser scanning confocal microscope (Northwestern University Center for Advanced Microscopy). Unless otherwise specified, image acquisition was performed through the z-dimension at 0.3 – 0.5 μm intervals and individual planes were projected into maximum intensity images. The nucleus was defined using Hoechst, DAPI, or LMNB1 and the cytoplasm was defined using GFP for HEK-293 experiments and MAP2 for MN experiments. The mean signal intensity of ETF1, pUPF1, and UPF1 for a region of interest was the mean pixel intensity per μm^2 as calculated by NIS-Elements Advanced Research v5 software (Nikon software, Northwestern University Center for Advanced Microscopy).

Calculation of tdTomato cytosolic ratio: Images were acquired using a 63X oil immersion objective. The mean pixel intensity per μm^2 of tdTomato for the cytoplasm and the nucleus was calculated by ImageJ (NIH) and the cytoplasmic value divided by the nuclear value was calculated.

Calculation of ETF1 N/C ratio: The N/C ratio was calculated as the mean ETF1 signal for the nucleus divided by that of the cytoplasm. Cytosolic ETF1 signal precluded the use of

projected maximum intensity images as projections caused cytosolic ETF1 signal to contaminate signal measured within the nucleus and analyses were therefore performed on single slices. For iPSC-derived MNs, images were acquired using a Plan Apo λ 100x 1.45 NA oil immersion objective and blinded computational analysis was used for this experiment. Briefly, an algorithm written in NIS-Elements identified the nuclear region, identified the cytoplasmic region, and generated values for mean pixel intensity per μm^2 of ETF1 for each region. N/C ratio was then calculated as above. 3-dimensional image reconstruction was performed using Imaris (Bitplane software, Northwestern University Center for Advanced Microscopy). For HEK-293 cells, images were acquired using a 63X oil immersion objective. For post-mortem tissue, images of layer V regions of motor cortex tissue were acquired using a Plan Apo λ 60x oil immersion objective. Corticospinal MNs were identified based on size and morphology.

ETF1-DPR colocalization analysis: The Pearson Correlation Coefficient (PCC) for ETF1 (red) and GFP, GR, or PR (green) within the nucleus of each cell was calculated by NIS-Elements software.

The % ETF1 within LMNB1 Nuclear Invaginations: Images were acquired using a Plan Apo λ 60x oil immersion objective. Blinded computational analysis was performed by using an algorithm written in NIS-Elements, which identified LMNB1 nuclear invaginations. The algorithm identified the nucleus (DAPI+) and then analyzed LMNB1+ signal that was: a) above a set threshold, and b) >0.3 μm inside the perimeter of the nucleus. Finally, the algorithm identified nuclear ETF1 as before, and generated values for the number of ETF1 puncta within invaginations and the total number of nuclear ETF1 puncta.

De novo translation and pUPF1 analyses: For *de novo* protein translation analysis, images were acquired at 100x using a Leica DMi8 microscope (Leica, Buffalo Grove, IL), photographed using a C10600-ORCA-R2 digital CCD camera (Hamamatsu Photonics, Japan) and processed with Fiji. Same procedure For single-cell pUPF1 activation analysis in MNs, images were acquired using a Plan Apo λ 60x oil immersion objective and the pUPF1/UPF1 value of each MN was calculated.

Quantification of G_4C_2 RNA foci: Images were acquired using a Plan Apo λ 60x oil immersion objective through the z dimension at 0.3 μm intervals and 3-dimensional images were analyzed by individual plane for nuclear and cytosolic foci.

Super-resolution images were acquired using the Nikon Structured Illumination Super-resolution Microscope (Northwestern University Center for Advanced Microscopy, NIH Grant 1S10OD016342-01) through the z dimension at 0.3 μm intervals to achieve ~115 nm resolution in the X-Y plane in multiple colors.

Intron retention assay

RNA was isolated and cDNA generated as previously described. PCR reaction was performed using 50ng cDNA template, 0.5μM C9ORF72-exon1a forward primer, 0.15 μM C9ORF72-intron1 reverse primer, 0.5μM C9ORF72-exon2 reverse primer, and Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific) with an annealing temperature of 55°C and 33 cycles. Products were run on 1% ethidium bromide gel.

NMD reporter assay

The pBS-TCR(PTC)-TCR(WT) NMD reporter plasmid was provided by Zhongsheng You[57]. This plasmid contains 2 minigenes of a T Cell Receptor (TCR), the first one contains a premature stop codon (PTC) and the second unit contains a wild-type (WT) TCR minigene with >99% sequence homology to the first one. The two genes are controlled by separate cytomegalovirus (CMV) promoters and each has an N-terminal hemagluttinin tag (HA). Cells were treated with 10mM caffeine (Red Dawn Energy) for 24h before protein isolation for a positive control

experiment. Western blot densitometry analysis was performed on HA bands corresponding to WT and PTC-fused TRB expressed from the NMD reporter construct. The ratio of the HA-PTC signal to the HA-WT signal was calculated to assess NMD efficiency in transfected cells.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were done with Prism 7 software (GraphPad Software) and in collaboration with Alfred Rademaker, PhD (Professor, Division of Statistics, Northwestern University). We classified an independent biological replicate as either an independent transfection experiment or an independent iPSC differentiation performed on a different day. The numbers (n) of each specific experiment are provided in the results section and the statistical test performed for each specific experiment is defined in the corresponding figure legend. For each statistical analysis, we first tested whether sample data fit the Guassian distribution using the D'Agostino-Pearson test. To compare two experimental conditions, either a student's t test (parametric) or a Mann-Whitney U test (non-parametric) was performed. To compare ≥ 3 experimental conditions, either a one-way ANOVA followed by a Bonferroni post-hoc test (parametric) or a Kruskal-Wallis rank test followed by a two-stage linear step-up procedure of Benjamini, Krieger and Yekutieli (non-parametric) was performed. For imaging experiments measuring a continuous variable at single-cell resolution, we accounted for variability among experimental replicates by normalizing all measurements to the average of the corresponding controls. We then tested all independent biological replicates for variance by the Brown-Forsythe test. In cases where biological replicates tested negative (Figures 1f and 4b), data from independent biological replicates was pooled per genotype or corresponding condition and tested for significance. In cases where biological replicates tested positive for variance (Figure 6b, and Supplementary Figures 6c, 6f), each independent biological replicate was tested for significance separately. In such cases we deemed a biological change as significant only if all independent biological replicates tested were significant and present the least significant i.e. lowest p value from all comparisons in the corresponding figure.

REFERENCES

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

2. Vucic, S., Rothstein, J.D. & Kiernan, M.C. Advances in treating amyotrophic lateral sclerosis: insights from pathophysiological studies. *Trends in neurosciences* **37**, 433-442 (2014).

3. DeJesus-Hernandez, M.*, et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245-256 (2011).

4. Renton, A.E.*, et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-268 (2011).

5. Gendron, T.F., Belzil, V.V., Zhang, Y.J. & Petrucelli, L. Mechanisms of toxicity in C9FTLD/ALS. *Acta Neuropathol* **127**, 359-376 (2014).

6. Balendra, R. & Isaacs, A.M. C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nature reviews. Neurology* **14**, 544-558 (2018).

7. Gitler, A.D. & Tsuiji, H. There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS. *Brain research* **1647**, 19-29 (2016).

8. Haeusler, A.R., Donnelly, C.J. & Rothstein, J.D. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nature reviews. Neuroscience* **17**, 383-395 (2016).

9. van Blitterswijk, M.*, et al.* Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. *Acta neuropathologica* **130**, 863-876 (2015).

10. Shi, Y.*, et al.* Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med* **24**, 313-325 (2018).

11. Donnelly, C.J.*, et al.* RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* **80**, 415-428 (2013).

12. Mori, K.*, et al.* The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science (New York, N.Y.)* (2013).

13. Zu, T.*, et al.* RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A* **110**, E4968-4977 (2013).

14. Cooper-Knock, J.*, et al.* Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta neuropathologica* **130**, 63-75 (2015).

15. Gomez-Deza, J.*, et al.* Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. *Acta neuropathologica communications* **3**, 38 (2015).

16. Mackenzie, I.R., *et al.* Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta neuropathologica* **130**, 845-861 (2015).

17. Lee, K.H., *et al.* C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* **167**, 774-788 e717 (2016).

18. Lin, Y., *et al.* Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. *Cell* **167**, 789-802.e712 (2016).

19. Freibaum, B.D., *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129-133 (2015).

20. Jovicic, A., *et al.* Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci* **18**, 1226-1229 (2015).

21. Zhang, K., *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61 (2015).

22. Neumann, M., *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (New York, N.Y.)* **314**, 130-133 (2006).

23. Barmada, S.J., *et al.* Autophagy induction enhances TDP43 turnover and survival in neuronal ALS models. *Nat Chem Biol* **10**, 677-685 (2014).

24. Woerner, A.C., *et al.* Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. *Science* **351**, 173-176 (2016).

25. D'Angelo, M.A., Raices, M., Panowski, S.H. & Hetzer, M.W. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* **136**, 284-295 (2009).

26. Savas, J.N., Toyama, B.H., Xu, T., Yates, J.R., 3rd & Hetzer, M.W. Extremely long-lived nuclear pore proteins in the rat brain. *Science (New York, N.Y.)* **335**, 942 (2012).

27. Toyama, B.H., *et al.* Identification of long-lived proteins reveals exceptional stability of essential cellular structures. *Cell* **154**, 971-982 (2013).

28. Gasset-Rosa, F., *et al.* Polyglutamine-Expanded Huntingtin Exacerbates Age-Related Disruption of Nuclear Integrity and Nucleocytoplasmic Transport. *Neuron* **94**, 48-57.e44 (2017).

29. Grima, J.C., *et al.* Mutant Huntingtin Disrupts the Nuclear Pore Complex. *Neuron* **94**, 93-107.e106 (2017).

30. Eftekharzadeh, B., *et al.* Tau Protein Disrupts Nucleocytoplasmic Transport in Alzheimer's Disease. *Neuron* **99**, 925-940.e927 (2018).

31. Binder, J.X., *et al.* COMPARTMENTS: unification and visualization of protein subcellular localization evidence. *Database : the journal of biological databases and curation* **2014**, bau012 (2014).

32. Ma, J., Goryaynov, A., Sarma, A. & Yang, W. Self-regulated viscous channel in the nuclear pore complex. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 7326-7331 (2012).

33. Kimura, M.*, et al.* Identification of cargo proteins specific for the nucleocytoplasmic transport carrier transportin by combination of an in vitro transport system and stable isotope labeling by amino acids in cell culture (SILAC)-based quantitative proteomics. *Molecular & cellular proteomics : MCP* **12**, 145-157 (2013).

34. Kimura, M., Okumura, N., Kose, S., Takao, T. & Imamoto, N. Identification of cargo proteins specific for importin-beta with importin-alpha applying a stable isotope labeling by amino acids in cell culture (SILAC)-based in vitro transport system. *The Journal of biological chemistry* **288**, 24540-24549 (2013).

35. Thakar, K., Karaca, S., Port, S.A., Urlaub, H. & Kehlenbach, R.H. Identification of CRM1-dependent Nuclear Export Cargos Using Quantitative Mass Spectrometry. *Molecular & cellular proteomics : MCP* **12**, 664-678 (2013).

36. Mori, K.*, et al.* hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta neuropathologica* **125**, 413-423 (2013).

37. Xu, Z.*, et al.* Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 7778-7783 (2013).

38. Cooper-Knock, J.*, et al.* Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain : a journal of neurology* **137**, 2040-2051 (2014).

39. Haeusler, A.R.*, et al.* C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195-200 (2014).

40. May, S.*, et al.* C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol* **128**, 485-503 (2014).

41. Mizielinska, S.*, et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science (New York, N.Y.)* **345**, 1192-1194 (2014).

42. Hautbergue, G.M.*, et al.* SRSF1-dependent nuclear export inhibition of C9ORF72 repeat transcripts prevents neurodegeneration and associated motor deficits. *Nature communications* **8**, 16063 (2017).

43. Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W. & Kulozik, A.E. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *The EMBO journal* **27**, 736-747 (2008).

44. Kashima, I.*, et al.* Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes & development* **20**, 355-367 (2006).

45. Kisselev, L., Ehrenberg, M. & Frolova, L. Termination of translation: interplay of mRNA, rRNAs and release factors? *The EMBO journal* **22**, 175-182 (2003).

46. Song, H.*, et al.* The crystal structure of human eukaryotic release factor eRF1--mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. *Cell* **100**, 311-321 (2000).

47. Ziller, M.J.*, et al.* Dissecting the Functional Consequences of De Novo DNA Methylation Dynamics in Human Motor Neuron Differentiation and Physiology. *Cell Stem Cell* **22**, 559-574 e559 (2018).

48. Lind, C., Oliveira, A. & Aqvist, J. Origin of the omnipotence of eukaryotic release factor 1. *Nature communications* **8**, 1425 (2017).

49. Brown, A., Shao, S., Murray, J., Hegde, R.S. & Ramakrishnan, V. Structural basis for stop codon recognition in eukaryotes. *Nature* **524**, 493-496 (2015).

50. Kryuchkova, P.*, et al.* Two-step model of stop codon recognition by eukaryotic release factor eRF1. *Nucleic acids research* **41**, 4573-4586 (2013).

51. Kurosaki, T.*, et al.* A post-translational regulatory switch on UPF1 controls targeted mRNA degradation. *Genes & development* **28**, 1900-1916 (2014).

52. Kurosaki, T. & Maquat, L.E. Nonsense-mediated mRNA decay in humans at a glance. *Journal of cell science* **129**, 461-467 (2016).

53. Huang, L. & Wilkinson, M.F. Regulation of nonsense-mediated mRNA decay. *Wiley interdisciplinary reviews. RNA* **3**, 807-828 (2012).

54. Hartmann, H.*, et al.* Proteomics and C9orf72 neuropathology identify ribosomes as poly-GR/PR interactors driving toxicity. *Life Sci Alliance* **1**, e201800070 (2018).

55. Kanekura, K.*, et al.* Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. *Hum Mol Genet* **25**, 1803-1813 (2016).

56. Kwon, I.*, et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345**, 1139-1145 (2014).

57. Nickless, A.*, et al.* Intracellular calcium regulates nonsense-mediated mRNA decay. *Nature medicine* **20**, 961-966 (2014).

58. Sznajder, L.J.*, et al.* Intron retention induced by microsatellite expansions as a disease biomarker. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 4234-4239 (2018).

59. Zhang, Z.*, et al.* Noisy splicing, more than expression regulation, explains why some exons are subject to nonsense-mediated mRNA decay. *BMC biology* **7**, 23 (2009).

60. Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R. & Muhlemann, O. Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. *RNA (New York, N.Y.)* **23**, 189-201 (2017).

61. Zhang, K., Grima, J.C., Rothstein, J.D. & Lloyd, T.E. Nucleocytoplasmic transport in C9orf72-mediated ALS/FTD. *Nucleus (Austin, Tex.)* **7**, 132-137 (2016).

62. Grima, J.C.*, et al.* Mutant Huntingtin Disrupts the Nuclear Pore Complex. *Neuron* **94**, 93-107.e106 (2017).

63. Barmada, S.J.*, et al.* Amelioration of toxicity in neuronal models of amyotrophic lateral sclerosis by hUPF1. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 7821-7826 (2015).

64. Ju, S.*, et al.* A yeast model of FUS/TLS-dependent cytotoxicity. *PLoS biology* **9**, e1001052 (2011).

65. Jackson, K.L.*, et al.* Preservation of forelimb function by UPF1 gene therapy in a rat model of TDP-43-induced motor paralysis. *Gene therapy* **22**, 20-28 (2015).

66. Popp, M.W. & Maquat, L.E. Organizing principles of mammalian nonsense-mediated mRNA decay. *Annual review of genetics* **47**, 139-165 (2013).

67. Schweingruber, C., Rufener, S.C., Zund, D., Yamashita, A. & Muhlemann, O. Nonsense-mediated mRNA decay - mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochimica et biophysica acta* **1829**, 612-623 (2013).

68. Kamelgarn, M.*, et al.* 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 (2018).

69. Weskamp, K. & Barmada, S.J. RNA Degradation in Neurodegenerative Disease. *Advances in neurobiology* **20**, 103-142 (2018).

70. Tank, E.M.*, et al.* Abnormal RNA stability in amyotrophic lateral sclerosis. *Nature communications* **9**, 2845 (2018).

71. Lagier-Tourenne, C.*, et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nature neuroscience* **15**, 1488-1497 (2012).

72. Reber, S.*, et al.* Minor intron splicing is regulated by FUS and affected by ALS-associated FUS mutants. *The EMBO journal* **35**, 1504-1521 (2016).

73. Polymenidou, M.*, et al.* Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nature neuroscience* **14**, 459-468 (2011).

74. Prudencio, M.*, et al.* Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat Neurosci* **18**, 1175-1182 (2015).

75. Colak, D., Ji, S.J., Porse, B.T. & Jaffrey, S.R. Regulation of axon guidance by compartmentalized nonsense-mediated mRNA decay. *Cell* **153**, 1252-1265 (2013).

76. Giorgi, C.*, et al.* The EJC factor eIF4AIII modulates synaptic strength and neuronal protein expression. *Cell* **130**, 179-191 (2007).

77. Urlaub, G., Mitchell, P.J., Ciudad, C.J. & Chasin, L.A. Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. *Molecular and cellular biology* **9**, 2868-2880 (1989).

78. Cheng, J. & Maquat, L.E. Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. *Molecular and cellular biology* **13**, 1892-1902 (1993).

79. Iborra, F.J., Jackson, D.A. & Cook, P.R. Coupled transcription and translation within nuclei of mammalian cells. *Science (New York, N.Y.)* **293**, 1139-1142 (2001).

80. Buhler, M., Wilkinson, M.F. & Muhlemann, O. Intranuclear degradation of nonsense codon-containing mRNA. *EMBO reports* **3**, 646-651 (2002).

81. Belgrader, P., Cheng, J., Zhou, X., Stephenson, L.S. & Maquat, L.E. Mammalian nonsense codons can be cis effectors of nuclear mRNA half-life. *Molecular and cellular biology* **14**, 8219-8228 (1994).

82. Trcek, T., Sato, H., Singer, R.H. & Maquat, L.E. Temporal and spatial characterization of nonsense-mediated mRNA decay. *Genes & development* **27**, 541-551 (2013).

83. Malhas, A., Goulbourne, C. & Vaux, D.J. The nucleoplasmic reticulum: form and function. *Trends in cell biology* **21**, 362-373 (2011).

84. Legartova, S.*, et al.* Nuclear structures surrounding internal lamin invaginations. *J Cell Biochem* **115**, 476-487 (2014).

85. Schoen, I., Aires, L., Ries, J. & Vogel, V. Nanoscale invaginations of the nuclear envelope: Shedding new light on wormholes with elusive function. *Nucleus* **8**, 506-514 (2017).

86. Wen, X.*, et al.* Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* **84**, 1213-1225 (2014).

87. Chen, E.I., McClatchy, D., Park, S.K. & Yates, J.R., 3rd. Comparisons of mass spectrometry compatible surfactants for global analysis of the mammalian brain proteome. *Anal Chem* **80**, 8694-8701 (2008).

88. He, L., Diedrich, J., Chu, Y.Y. & Yates, J.R., 3rd. Extracting Accurate Precursor Information for Tandem Mass Spectra by RawConverter. *Anal Chem* **87**, 11361-11367 (2015).

89. Eng, J.K., McCormack, A.L. & Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* **5**, 976-989 (1994).

90. Xu, T.*, et al. ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program* (2006).

91. Cociorva, D., D, L.T. & Yates, J.R. Validation of tandem mass spectrometry database search results using DTASelect. *Curr Protoc Bioinformatics* **Chapter 13**, Unit 13 14 (2007).

92. Tabb, D.L., McDonald, W.H. & Yates, J.R., 3rd. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. *J Proteome Res* **1**, 21-26 (2002).

93. Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J. & Gygi, S.P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* **2**, 43-50 (2003).

94. Zybailov, B.*, et al.* Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J Proteome Res* **5**, 2339-2347 (2006).

95. Hulsen, T., de Vlieg, J. & Alkema, W. BioVenn – a web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* **9**, 488 (2008).

96. Pandey, U.B.*, et al.* HDAC6 rescues neurodegeneration and provides an essential link between autophagy and the UPS. *Nature* **447**, 859-863 (2007).

97. Lanson, N.A., Jr.*, et al.* A Drosophila model of FUS-related neurodegeneration reveals genetic interaction between FUS and TDP-43. *Human molecular genetics* **20**, 2510-2523 (2011).

98. Nichols, C.D., Becnel, J. & Pandey, U.B. Methods to assay Drosophila behavior. *Journal of visualized experiments : JoVE* (2012).

99. Jensen, K.*, et al.* Purification of transcripts and metabolites from Drosophila heads. *J Vis Exp*, e50245 (2013).

100. Schmittgen, T.D. & Livak, K.J. Analyzing real-time PCR data by the comparative C(T) method. *Nature protocols* **3**, 1101-1108 (2008).

101. Su, Z.*, et al.* Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron* **84**, 239 (2014).

102. Deng, H.X.*, et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211-215 (2011).

103. Schmidt, E.K., Clavarino, G., Ceppi, M. & Pierre, P. SUnSET, a nonradioactive method to monitor protein synthesis. *Nature methods* **6**, 275-277 (2009).