HHS Public Access

Author manuscript

*Neuron*. Author manuscript; available in PMC 2021 July 22.

Published in final edited form as:

*Neuron*. 2020 July 22; 107(2): 292–305.e6. doi:10.1016/j.neuron.2020.04.011.

# Chimeric peptide species contribute to divergent dipeptide repeat pathology in c9ALS/FTD and SCA36

**Zachary T. McEachin**[1,2,3,*], **Tania F. Gendron**[4,5], **Nisha Raj**[1,2], **María García-Murias**[6,7], **Anwesha Banerjee**[1], **Ryan Purcell**[1], **Patricia J. Ward**[1], **Tiffany W. Todd**[4], **Megan E. Merritt-Garza**[2], **Karen Jansen-West**[4], **Chadwick M. Hales**[8], **Tania García-Sobrino**[9], **Beatriz Quintáns**[6,7], **Christopher J. Holler**[10], **Georgia Taylor**[10], **Beatriz San Millán**[11,12], **Susana Teijeira**[11,12], **Toru Yamashita**[13], **Ryuichi Ohkubo**[14], **Nicholas M. Boulis**[15], **Chongchong Xu**[16], **Zhexing Wen**[1,2,8,16], **Nathalie Streichenberger**[17,18,19], **Neuro-CEB Neuropathology Network**, **Brent L. Fogel**[20], **Thomas Kukar**[10], **Koji Abe**[13], **Dennis W. Dickson**[4], **Manuel Arias**[6,7], **Jonathan D. Glass**[8], **Jie Jiang**[1], **Malú G. Tansey**[21,22,23], **María-Jesús Sobrido**[6,7], **Leonard Petrucelli**[4,5], **Wilfried Rossoll**[4,5], **Gary J. Bassell**[1,2,3,*,†]

[1]Department of Cell Biology, Emory University, Atlanta, GA 30322, USA

[2]Laboratory for Translational Cell Biology, Emory University, Atlanta, GA 30322, USA

[3]Wallace H. Coulter Graduate Program in Biomedical Engineering, Georgia Institute of Technology & Emory University, Atlanta, GA 30332, USA

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

[5]Mayo Clinic Graduate School of Biomedical Sciences, Mayo Clinic, Jacksonville, FL 32224, USA

[6]Centro de Investigación Biomédica en red de Enfermedades Raras (CIBERER), Santiago de Compostela, Spain

[7]Neurogenetics Research Group, Instituto de Investigación Sanitaria (IDIS), Hospital Clínico Universitario, SERGAS, Santiago de Compostela, Spain

[8]Department of Neurology, Emory University, Atlanta, GA 30322, USA

[9]Department of Neurology, Hospital Clínico Universitario, SERGAS, Santiago de Compostela, Spain

[10]Department of Pharmacology, Emory University, Atlanta, GA 30322, USA

[*]**Correspondence:** zmceach@emory.edu, gbassel@emory.edu. [†]**Lead Contact:** gbassel@emory.edu.

Author contributions:

Z.T.M. conceived and coordinated project, Z.T.M., T.F.G., L.P., W.R., G.J.B. designed experiments. Z.T.M., T.F.G., J.J., W.R. and G.J.B. analyzed and interpreted data. Z.T.M. generated iPS cells, performed iPS differentiations, immunohistochemistry in patient tissue, ASO experiments, prepared protein lysates, western blot, qPCR, immunofluorescence staining, luciferase assays, RP-PCR and sequencing. T.F.G. performed MSD immunoassays. T.F.G., D.D. performed poly(PR) immunohistochemistry. Z.T.M., K.J-W. made constructs. Z.T.M., N.R., M.E.M-G., M.G-M. performed cell culture. C.X., Z.W. performed organoid differentiations. A.B. performed ICV injections. Z.T.M., A.B., P.J.W. processed mice tissue and performed immunostaining. Z.T.M. and J.J. performed poly(GP)/poly(GA) co-localization experiment. Z.T.M., J.J., and M.G.T. analyzed DSP data. R.P., N.B., C.H., G.T., T.T., T.K., M.G.T., J.D.G. provided technical support. C.M.H., B.S.M., T.G-S., B.Q., S.T., T.Y., R.O., B.L.F., N.S., K.A., M.A., J.D.G., M-J.S. provided patient samples and associated genetic and clinical data. Z.T.M. wrote manuscript. Z.T.M., T.F.G., N.R., J.J., M.G.T., M.A., J.D.G., M-J.S., W.R., G.J.B. revised and edited manuscript.

Declaration of Interests:

The authors declare no competing interests.

[11]Rare Diseases and Pediatric Medicine Research Group, Galicia Sur Health Research Institute (IIS Galicia Sur). SERGAS-UVIGO, Vigo, Spain

[12]Pathology Department, Complexo Hospitalario Universitario de Vigo (CHUVI), SERGAS. Vigo, Spain

[13]Department of Neurology, Okayama University, Okayama, Japan

[14]Department of Neurology, Fujimoto General Hospital, Miyazaki, Japan

[15]Department of Neurosurgery, Emory University, Atlanta, GA 30322, USA

[16]Department of Psychiatry & Behavioral Sciences, Emory University, Atlanta, GA 30322, USA

[17]Hospices Civils de Lyon, Lyon, France

[18]Université Claude Bernard Lyon, Lyon, France

[19]Institut NeuroMyogène CNRS UMR 5310

[20]Department of Neurology & Human Genetics, University of California Los Angeles, Los Angeles, CA, 90095, USA

[21]Department of Neuroscience, University of Florida, Gainesville, FL 32607, USA

[22]Center for Translational Research in Neurodegenerative Disease, University of Florida, Gainesville, FL 32607, USA

[23]Norman Fixel Institute for Neurological Diseases, University of Florida, Gainesville, FL 32607, USA

## Summary:

GGGGCC hexanucleotide repeat expansions (HRE) in *C9orf72* cause amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), and lead to the production of aggregating dipeptide repeat proteins (DPRs) via repeat associated non-AUG (RAN) translation. Here, we show the similar intronic GGCCTG HRE that causes spinocerebellar ataxia type 36 (SCA36) is also translated into DPRs, including poly(GP) and poly(PR). We demonstrate that poly(GP) is more abundant in SCA36 compared to c9ALS/FTD patient tissue due to canonical AUG-mediated translation from intron-retained GGCCTG repeat RNAs. However, the frequency of the antisense RAN translation product poly(PR) is comparable between c9ALS/FTD and SCA36 patient samples. Interestingly, in SCA36 patient tissue, poly(GP) exists as a soluble species and no TDP-43 pathology is present. We show that aggregate-prone chimeric DPR (cDPR) species underlie the divergent DPR pathology between c9ALS/FTD and SCA36. These findings reveal key differences in translation, solubility and protein aggregation of DPRs between c9ALS/FTD and SCA36.

### Keywords

SCA36; *C9orf72* Expansion; ALS; FTD; Dipeptide repeats; RAN Translation; Chimeric DPRs; Antisense oligonucleotide therapy; Neurodegeneration

## Introduction:

Spinocerebellar ataxia type 36 (SCA36) is a late-onset, progressive neurodegenerative disease characterized chiefly by truncal ataxia (Arias et al., 2017). First described clinically by two independent groups, SCA36 was initially referred to as *Costa da Morte* ataxia in the Galician region of Spain (Arias et al., 2008), and as *Asidan* in the Chugoku region of Japan (Ohta et al., 2007). Subsequently, it was discovered that both of these ataxic syndromes result from a massively expanded GGCCTG hexanucleotide repeat expansion (HRE), alternatively denoted as TGGGCC ($TG_3C_2$), in the first intron of the nucleolar protein 56 (*NOP56*) gene (García-Murias et al., 2012; Kobayashi et al., 2011). SCA36 has since been reported in individuals from Italy, France, China, Taiwan, Poland, and the U.S.A. (Lee et al., 2016b; Obayashi et al., 2015; Sarto et al., 2013; Sulek et al., 2013; Valera et al., 2017; Zeng et al., 2016). SCA36 patients present with other neurological symptoms to varying degrees, including appendicular ataxia, sensorineural hearing loss, and tongue fasciculations and atrophy (Ikeda et al., 2012; Ikeda et al., 2013).

Although the underlying pathomechanisms by which the $TG_3C_2$ repeat expansion causes SCA36 remain understudied, proposed disease mechanisms include haploinsufficiency of the *NOP56* gene and RNA-mediated toxicity resulting from the accumulation of repeat RNA (Liu et al., 2014; Matsuzono et al., 2017). These disease mechanisms are also implicated in other microsatellite repeat expansion disorders. A highly related GC-rich HRE, GGGGCC ($G_4C_2$), in the first intron of the gene *C9orf72* is the most common cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (DeJesus-Hernandez et al., 2011; Renton et al., 2011), which are commonly referred to as c9ALS/FTD. In addition to RNA-mediated toxicity, an atypical form of translation – repeat-associated non-AUG (RAN) translation – is proposed to play a key role in c9ALS/FTD pathogenesis (Freibaum and Taylor, 2017). RAN translation occurs in a number of repeat expansion diseases (Banez-Coronel et al., 2015; Mori et al., 2013b; Todd et al., 2013; Zu et al., 2017; Zu et al., 2011), and results in the production of potentially toxic proteins (Cleary and Ranum, 2017). In c9ALS/FTD, expanded $G_4C_2$ repeats are RAN translated in all six reading frames from sense and antisense transcripts to produce five different dipeptide repeat proteins (DPRs; polyGA, -GP, -GR, -PR, -PA) (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013a; Mori et al., 2013b). *In vitro* (Kwon et al., 2014; Lin et al., 2016) and *in vivo* (Mizielinska et al., 2014; Zhang et al., 2018; Zhang et al., 2016; Zhang et al., 2019) studies in models of c9ALS/FTD indicate that some DPRs are toxic, in particular poly(PR) and poly(GR), in c9ALS/FTD. Given the strikingly similar repeat expansions causative of SCA36 and c9ALS/FTD, we sought to determine whether $TG_3C_2$ HREs, like $G_4C_2$ HREs, are also translated into DPRs and to identify key pathological differences between these two clinically distinct diseases.

## Results:

### The SCA36-associated $TG_3C_2$ expansion is translated into dipeptide repeat proteins

Translation across expanded $TG_3C_2$ repeats results in six putative DPR proteins: poly(GP), poly(WA), and poly(GL) from sense transcripts, and poly(GP), poly(PR), and poly(AQ) from antisense transcripts (Figure 1A). Two of these putative DPRs, poly(GP) and poly(PR),

are produced via RAN translation in c9ALS/FTD (Gendron et al., 2013; Mori et al., 2013b). To determine whether the $TG_3C_2$ repeat could undergo RAN translation *in vitro*, we generated *NOP56* minigenes containing various numbers of $TG_3C_2$ repeats (Figures S1A and S1B). Using a quantitative poly(GP) immunoassay (Gendron et al., 2017; Gendron et al., 2015), we found that translation of $TG_3C_2$ repeats is length-dependent, and requires a certain repeat-length threshold before it can occur (Figure S1C). In addition to poly(GP), we could also detect poly(WA) and poly(GL) generated from sense $TG_3C_2$ transcripts by immunoblotting or immunocytochemistry using antibodies to frame-specific C-terminal tags (Figures S1D and S1E). These findings demonstrate that the intronic $TG_3C_2$ repeat can undergo RAN translation to produce DPRs in SCA36. To test this hypothesis in a disease-relevant context, we developed multiple patient-derived cellular models of SCA36. Quantitative immunoassays revealed that poly(GP) is detectable in SCA36 patient cells. Strikingly, compared to c9ALS/FTD patient fibroblasts, poly(GP) expression was significantly higher in SCA36 patient fibroblasts (Figure 1B). Poly(GP) abundance was similarly higher in SCA36 lymphoblastoid cell lines (LCLs), induced pluripotent stem cells (iPSC), and 3D forebrain organoids (day 105) in comparison to c9ALS/FTD cellular models (Figures 1C–1E). However, poly(GP) abundance in iPSC-derived motor neurons (iPSC-MNs) was comparable between c9ALS/FTD and SCA36 samples (Figure 1F).

## DPR proteins common to c9ALS/FTD and SCA36 exhibit distinct localization and solubility properties resulting in divergent DPR pathology

Given that poly(GP) DPRs are expressed in cellular models of SCA36, we assessed whether poly(GP) is present in SCA36 post-mortem tissue. Poly(GP) was detected in all evaluated regions of the central nervous system (CNS) by immunoassay, and its abundance was largely comparable across the majority of neuroanatomical regions tested (Figure 2A). However, poly(GP) expression was markedly higher in the dentate nucleus of the cerebellum, one of the primarily affected regions in SCA36 (Arias et al., 2017; Liu et al., 2014). We next compared poly(GP) burden in disease-relevant CNS regions from c9ALS/FTD and SCA36 cases (Figures 2B–2E). Although poly(GP) abundance in the cerebellar cortex of SCA36 cases was comparable to that of c9ALS/FTD cases, poly(GP) was higher in the frontal cortex of SCA36 patients, consistent with our observations in iPSC-derived 3D forebrain organoid models (Figure 1E). Additionally, poly(GP) abundance in the spinal cord and the cerebellar dentate nucleus of SCA36 patients was greater than in c9ALS/FTD.

A neuropathological hallmark of c9ALS/FTD is the aggregation of both sense and antisense DPRs in patient post-mortem tissue (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013b; Zu et al., 2013). Since poly(GP) was abundant in lysates from post-mortem tissues from SCA36 patients, we investigated whether poly(GP) aggregation pathology was also present. Surprisingly, immunohistochemical analysis of poly(GP) in disease-relevant regions of SCA36 did not reveal the characteristic poly(GP) aggregates seen in c9ALS/FTD (Figure 2F; Figure S2). The lack of poly(GP) aggregation in SCA36 was supported by the absence of pathology for ubiquitin-binding protein p62/SQSTM1, a common component of protein aggregates (Bjorkoy et al., 2006) (Figure S2). Although detectable by immunoassay, poly(GP) did not form insoluble inclusions in other CNS regions including the ventral horn of the spinal cord, hypoglossal nucleus of the medulla, inferior olive of the medulla,

temporal cortex, occipital cortex, or parietal cortex (Figure S3). To further investigate the lack of poly(GP) aggregation in SCA36, we sequentially fractionated cerebellar and frontal cortex tissue samples using reagents of increasing solubilizing power. Dotblot analysis for poly(GP) in the urea soluble fraction further supports the presence of insoluble poly(GP) in c9ALS/FTD but not SCA36 (Figure 2G). Quantitative analysis of poly(GP) in each fraction revealed that a much greater percentage of poly(GP) was insoluble in c9ALS/FTD than in SCA36 (Figure 2H; Figure S4). Highly multiplexed, spatially resolved protein analysis (Digital Spatial Profiling; DSP) of the cerebellar granule, Purkinje and molecular layers also revealed divergent pathologies between c9ALS/FTD and SCA36 (Figures 2I and 2J; Figures S5A–S5C), such as the presence of poly(GA) and p62 pathology only in c9ALS/FTD cases. Although poly(GP) is detectable by immunoassay in SCA36, poly(GP) was not detected with DSP, which is consistent with the lack of apparent poly(GP) inclusions by immunohistochemistry. Notably, in various cerebellar regions in SCA36, there is a decrease in the abundance of structural and synaptic proteins, e.g., synaptophysin (SYP) in the Purkinje layer, supporting degeneration of the cerebellum in SCA36 compared to control and c9ALS/FTD patients. Moreover, in SCA36, the Purkinje and molecular layers show a reduction in isocitrate dehydrogenase (IDH1) protein abundance (Figure S5A), a protein known to be important for neuroprotection against oxidative stress (Yang et al., 2017).

The chemical properties and lack of immunogenicity of the other putative sense DPRs, poly(GL) and poly(WA), precluded us from generating antibodies targeting them. However, using previously established polyclonal antibodies to poly(PR), an antisense DPR in c9ALS/FTD, we determined that poly(PR) formed aggregates in the granule cell layer of the cerebellar cortex in SCA36 (Figure 3A) and, similar to c9ALS/FTD, poly(PR) aggregates were rare. In contrast to c9ALS/FTD, immunohistochemistry and immunoblots for phosphorylated TDP-43 in frontal cortex sections or ures fractions, respectively, demonstrated that SCA36 patients do not exhibit TDP-43 pathology characteristic of c9ALS/FTD (Figures 3B and 3C).

### *NOP56* intron 1 is retained in a repeat-dependent manner in SCA36

Given the intronic nature of the SCA36 repeat expansion, we hypothesized that the expanded $TG_3C_2$ repeat impairs splicing of *NOP56* and results in retention of the repeat-containing intron 1. To determine whether expanded $TG_3C_2$ repeats could impair splicing, we adapted a splicing luciferase reporter (“Luc-In”) that was previously developed to screen small-molecule modulators of splicing (Younis et al., 2010). The full intron 1 sequence of *NOP56* was modified to contain 2, 37, or 82 $TG_3C_2$ repeats, and was inserted into the firefly luciferase ORF at an optimized splice site such that defective splicing of the splicing luciferase reporter results in reduced luciferase activity (Figure S6A). Relative to a *NOP56* intron harboring two $TG_3C_2$ repeats, we observed a ~14% and ~25% reduction in luciferase activity in *NOP56* introns harboring 37 and 82 $TG_3C_2$ repeats, respectively (Figure S6B). This repeat length-dependent reduction in luciferase activity suggests that expanded $TG_3C_2$ repeats in the context of the *NOP56* intron 1 can impair splicing. To validate whether intron 1 of *NOP56* is retained in mRNA from SCA36 patient samples, we designed custom Taqman qPCR probes spanning the exon 1–exon 2 (ex1–2) splice junction (Figure S6C). We reasoned that retention of intron 1 would prevent hybridization of the probe spanning the

junction, thus resulting in a reduced qPCR signal. We found a significant reduction in ex1–2 *NOP56* transcript levels (~40%) in the cerebellum of SCA36 patients (Figure 4A). This reduction was also observed in patient-derived cellular models of SCA36 including fibroblast, iPSCs, and iPSC-motor neurons (Figures S6D–S6F). When we assessed *NOP56* mRNA levels using probes spanning the downstream ex3–4 and ex6–7 splice junctions, we found a marked increase in *NOP56* transcripts in SCA36 compared to controls (Figure 4A; Figures S6G and S6H). Similar to total *NOP56* mRNA expression, we see an increase in NOP56 protein as detected by western blot (Figure 4B). Although unexpected, this finding is consistent with a recent study that demonstrated that the NOP56 protein is autoregulated by modulating the available pool of protein-coding *NOP56* mRNA (Lykke-Andersen et al., 2018). Next, we assessed whether a retained intron could be directly detected. Using a custom Taqman probe spanning the intron 1-exon 2 junction, we found a significant increase in *NOP56* intron 1-exon 2 expression in SCA36 compared to controls (Figure 4C). Additionally, using custom Nanostring probes that target RNA directly, we detected robust expression of a *NOP56* intron 1 retained RNA species in SCA36 cerebellar cortex samples compared to control; in contrast, retention of intron 4 or intron 10 is not observed in SCA36 samples (Figure 4D). Furthermore, we observed a higher correlation of poly(GP) abundance with intron 1 retained *NOP56* transcripts than total *NOP56* transcripts (Figures S6I–S6L). These findings provide evidence that expanded $TG_3C_2$ repeats impair splicing and result in retention of the repeat-containing intron 1 of *NOP56* in SCA36 patients.

## AUG-mediated translation preferentially drives production of poly(GP) DPRs in SCA36

Retention of *NOP56* intron 1 in SCA36 results in a new mRNA variant. Unlike *C9orf72* intron 1 that is flanked by untranslated exons, *NOP56* intron 1 is flanked by coding exons. Intron 1 of *NOP56* is a phase 0 intron positioned immediately after the ATG start codon in exon 1. This type of exon-intron structure is referred to as a “start-codon intron” and is overrepresented in genes encoding highly conserved proteins such as the ribosomal proteins (Nielsen and Wernersson, 2006; Yoshihama et al., 2002). Given the ATG start site in the first exon, we hypothesized that retention of intron 1 would allow for AUG-mediated translation of the expanded $TG_3C_2$ repeats in SCA36. Interestingly, analysis of the *NOP56* Refseq gene sequence (NM_006392) reveals that AUG-mediated translation of a retained intron transcript variant would produce a polypeptide containing repeating Gly-Pro units (i.e., poly(GP); Figure 4E). Additionally, the retained intron transcript variant maintains a Kozak consensus sequence (GCCATG|g, where | denotes the exon-intron boundary). Therefore, we sought to determine whether the ORF of a retained intron transcript variant specifically on the allele harboring expanded $TG_3C_2$ repeats (“expanded allele”) was also in frame with poly(GP) in SCA36 patients. We performed repeat-primed PCR (RP-PCR) on gDNA isolated from SCA36 patient cerebellum. RP-PCR amplicons were separated via gel electrophoresis and high-molecular weight amplicons were gel extracted, TOPO-cloned, and Sanger sequenced (Figure S7A). Amplicons were considered to be from the expanded allele if the amplicon: 1) contained a greater number of $TG_3C_2$ repeat units than the corresponding WT allele, as determined by sequencing of the WT specific allele; and 2) contained the rs28970277 and rs6138678 alleles in the intron region and 5’ UTR region, respectively, specific to the expanded allele haplotype (Figures S7B–S7E). Similar to the *NOP56* Refseq gene sequence, we found AUG translation of the expanded allele would produce a poly(GP)-

containing peptide (Figure 4E). In light of this finding, it stands to reason that canonical AUG translation upstream of the retained intron could drive production of the poly(GP) DPR in SCA36. Therefore, we next investigated how canonical AUG-mediated translation in the GP frame (+0 frame) would affect translation of the other putative SCA36 specific sense DPRs, poly(GL) and poly(WA). To this end, we generated a number of translation reporter vectors that contained 23 bp of *NOP56* exon 1 and intron 1 containing 82 $TG_3C_2$ repeats followed by NanoLuc luciferase in either the +0, +1, or +2 frame (Figure 4F). As expected, translation in the +0 (GP) frame was most efficient and resulted in robust luciferase signal driven by the canonical AUG; however, signal in the +1 and +2 frames was significantly lower (Figure 4G). Given that expanded $TG_3C_2$ repeats can undergo RAN translation *in vitro* (Figure S1), we compared the translational efficiency of the +1 and +2 frame-associated DPRs in both a canonical AUG-mediated and noncanonical RAN translation-mediated context. The translational reporter constructs were modified by replacing the 23 bp exon 1 of *NOP56* with a 23 bp sequence containing six stop codons (two stop codons per frame). Notably, translation of the +1 and +2 frame is significantly reduced in a canonical AUG-mediated context compared to the noncanonical RAN translation context (Figure 4H). Collectively, these data suggest that the retention of the repeat-containing intron 1 of *NOP56* in SCA36 results in a novel mRNA variant that promotes canonical AUG driven translation of poly(GP) DPRs in SCA36. Moreover, canonical translation preferentially produces poly(GP) relative to the putative poly(GL) and poly(WA) DPRs.

## Aggregation-prone poly(GA) contributes to the divergent poly(GP) pathology in c9ALS/FTD and SCA36

Next, we sought to determine the underlying mechanism that drives divergent poly(GP) pathology in c9ALS/FTD and SCA36. Transient expression of $GP_{80x}$-GFP in HEK293T cells suggests that poly(GP) is soluble (Figures S8A and S8B). To assess whether poly(GP) remains soluble after long-term expression in disease-relevant cell types both *in vitro* and *in vivo*, we generated AAV vectors to express GFP alone or $GP_{80x}$-GFP. Transduction of iPSC-MNs with AAV-$GP_{80x}$-GFP for 21 days revealed that poly(GP) remained diffuse throughout the cytoplasm, neurites, and nucleus of cells (Figure S8C). Since poly(GP) forms perinuclear inclusions in c9ALS/FTD post-mortem tissue, it is reasonable to posit that expression and processing of poly(GP) *in vivo* is required for its aggregation. To address this, we expressed GFP or $GP_{80x}$-GFP in the mouse CNS via intracerebroventricular delivery of AAV vectors to post-natal day 0 mice. Immunohistochemical analysis at 2 months of age revealed diffuse poly(GP) staining throughout the CNS including the cell bodies and dense dendritic neurites of Purkinje cells in the cerebellum as well as neurons of the frontal cortex (Figure 5A; Figure S8D). Notably, poly(GP) immunoreactivity was robust throughout the neuropil, which is consistent with poly(GP) being diffuse in neurites of iPSC-MNs. The c9ALS/FTD-specific poly(GA) DPR is aggregation-prone and, unlike poly(GP), it forms perinuclear inclusions when transiently expressed in HEK293T cells or using AAV in the mouse brain. Moreover, poly(GA) inclusions are highly abundant in c9ALS/FTD post-mortem tissues (Mackenzie et al., 2015). Thus, we hypothesized that poly(GA) may promote the aggregation of poly(GP) in c9ALS/FTD and underlie divergent pathology with SCA36. Indeed, poly(GA) and poly(GP) colocalized in human c9ALS/FTD post-mortem tissue and c9BAC mice expressing an expanded $G_4C_2$ repeat (Figure 5B). We found that >90% of

poly(GP) inclusions co-localized with poly(GA) but only ~10% of poly(GA) aggregates co-localized with poly(GP) (Figures S9A and S9B). Colocalization of poly(GP) inclusions with poly(GA) supports the hypothesis that poly(GA) mediates aggregation of poly(GP) in c9ALS/FTD. This may result from at least two mechanisms that are not mutually exclusive: 1) poly(GA) DPRs cross-seed the aggregation of poly(GP); and 2) aberrations in the transcription or translation of the c9ALS/FTD-associated $G_4C_2$ repeat result in the production of a chimeric DPR ("cDPRs") species consisting of GA and GP DPRs (Figure 5C). In support of the latter, a recent study demonstrated ribosomal frameshifting during translation of $G_4C_2$ repeats *in vitro* (Tabet et al., 2018), and long-read sequencing across expanded $G_4C_2$ repeats uncovered that genetic interruptions may occur in c9ALS/FTD patients (Ebbert et al., 2018). To explore whether chimeric DPRs containing GA and GP DPRs could explain poly(GA)-mediated poly(GP) aggregation, we generated a number of tagged constructs expressing DPRs alone [$GA_{80x}$-GFP ("GA"), $GA_{80x}$-HA ("GA-HA"), $GP_{80x}$-GFP ("GP")] or a chimeric DPR (cDPR) construct expressing $GA_{50x}GP_{50x}$-GFP ("GA:GP") (Figure S9C). Immunofluorescence analysis for GFP revealed that chimeric GA:GP peptides form perinuclear inclusions in HEK293T cells similar to inclusions in cells expressing only GA, and contrasting with the diffuse localization of poly(GP) in cells expressing only GP (Figure 5D). Contrary to cells expressing chimeric GA:GP, co-transfection of GA and GP in HEK293T cells does not promote robust formation of poly(GP) aggregates (Figure S9D). Additionally, immunoblots for GFP revealed that chimeric GA:GP peptides but not co-transfected GA and GP peptides exist in the triton-insoluble, SDS-soluble fraction (Figure 5E). To further assess the relationship between poly(GA) and poly(GP), we developed an immunoassay that uses an antibody to capture poly(GA)-containing DPRs and a detection antibody towards poly(GP) (Figure 5F). Using this assay, we observed complexes of poly(GA) and poly(GP) in lysates from cells expressing chimeric GA:GP proteins but not from cells co-expressing GA and GP (Figure S9E–S9G). These results are consistent with our findings that chimeric GA:GP proteins aggregate in HEK293T cells but that GA does not cross-seed GP aggregation in co-transfected cells. Of note, when we employed this same immunoassay on lysates from detergent-insoluble, urea soluble fractions of control, c9ALS/FTD, and SCA36 frontal cortex, we observed that pulling-down poly(GA) in c9ALS/FTD lysates also resulted in the co-immunoprecipitation of poly(GP) (Figure 5G; Figures S9I and S9J). We further extended these findings by performing our chimeric immunoassay on lysates from detergent-insoluble, urea soluble fractions of control and c9ALS/FTD cerebellar cortex. Again, we observe a robust GA:GP chimeric DPR signal in c9ALS/FTD cases (Fig. 5H). Taken together, these findings demonstrate that chimeric GA:GP peptides are formed in c9ALS/FTD and thus underlie the divergent DPR pathology between c9ALS/FTD and SCA36.

### Antisense oligonucleotides reduce poly(GP) abundance in SCA36 cellular models

Finally, we evaluated the therapeutic potential of antisense oligonucleotides (ASOs) to mitigate DPR burden in SCA36. RNase H-dependent ASOs were designed to target various regions of the *NOP56* locus (Figure S10A). A dose-response experiment was performed in SCA36 fibroblasts (n=3 independent patient cell lines) transfected with increasing doses (0.1nM to 1000nM; 10-fold serial dilution) of ASO-#1 or ASO-#2, targeting the intron or

expanded repeats, respectively, for 48 hours. We observed a dose-dependent reduction in poly(GP) abundance with increasing ASO concentrations ($EC_{50}$=9.1nM and 4.5nM for ASO-#1 and ASO-#2, respectively) (Figure 6A). NOP56 is an essential cellular protein and knockdown of NOP56 results in increased apoptosis (Lykke-Andersen et al., 2018); therefore, we aimed to find an effective ASO concentration that minimizes loss of NOP56 protein. 10nM ASO-#2 treatment significantly reduced intron 1 retained *NOP56* transcripts (Figure S10B), but not total *NOP56* mRNA transcripts (Figure 6B). Whereas 100nm and 1000nM ASO-#2 notably reduced both *NOP56* mRNA expression and NOP56 protein abundance (Figures 6B and 6C; Figure S10C). Since delivery of ASO-#2 at 10nM resulted in reduction of poly(GP) without compromising total NOP56, we next validated the efficacy of both ASO-#1 and ASO-#2 at a 10nM concentration to reduce poly(GP) abundance relative to a scrambled control ASO and observed a ~69% and ~80% reduction, respectively, in poly(GP) abundance (Figure 6D; Figure S10D). Given the limitations of lipid-based delivery methods as a viable *in vivo* therapeutic option, we tested gymnotic delivery of ASOs targeting the expanded repeat and found that they could robustly reduce poly(GP) abundance in patient fibroblasts and LCLs relative to a scrambled control (Figures 6E and 6F). Lastly, we demonstrate that poly(GP) is detectable in the cerebrospinal fluid (CSF) of SCA36 patients as observed in c9ALS/FTD patients (Gendron et al., 2017) (Figure 6G). Together, our data provide proof-of-principle that ASOs targeting the expanded $TG_3C_2$ repeat can be a viable therapeutic strategy for reducing repeat-RNA and DPRs in SCA36 and that detection of poly(GP) in the CSF may prove to be a useful pharmacodynamic marker for ASO target engagement.

## Discussion

Here we show that the intronic SCA36-associated $TG_3C_2$ repeat expansion is translated in both the sense and antisense direction to produce DPR proteins. We found that poly(GP) levels were higher in SCA36 than in c9ALS/FTD, but that inclusions of poly(PR), reported to be highly toxic in several model systems (Kwon et al., 2014; Mizielinska et al., 2014), were rare in both SCA36 and c9ALS/FTD. Mechanistically, we demonstrate that expanded $TG_3C_2$ repeats in SCA36 impair splicing of *NOP56* and result in retention of the repeat-containing intron 1. Previous studies provided evidence that the c9ALS/FTD-associated $G_4C_2$ repeat expansion similarly promotes retention of *C9orf72* intron-1 (Niblock et al., 2016; Sznajder et al., 2018). However, unlike the *C9orf72* intron 1, the *NOP56* intron 1 is located between two coding exons and thus produces a novel coding transcript variant. Indeed, we demonstrate that the increased abundance of poly(GP) in SCA36 compared to c9ALS/FTD is likely due to canonical AUG-mediated translation across the $TG_3C_2$ repeat in SCA36. While canonical translation preferentially drives production of poly(GP) in the sense direction, the presence of the antisense DPR, poly(PR), suggests that RAN translation also occurs in SCA36.

Despite poly(GP) abundance being higher in SCA36 than in c9ALS/FTD, poly(GP) is predominately soluble in SCA36 but present in both soluble and insoluble forms in c9ALS/FTD. Both *in vitro* and *in vivo* expression of poly(GP) confirms the inherent solubility of poly(GP) as seen in SCA36 patient samples. These data suggest that secondary mechanisms initiate poly(GP) aggregation in c9ALS/FTD. Specific to c9ALS/FTD is the production of

the inclusion-forming amyloidogenic poly(GA) DPR. Neuropathologically, poly(GA) is arguably the most abundant DPR in c9ALS/FTD post-mortem brain tissue (Mackenzie et al., 2015). Furthermore, it has been shown that a near-cognate CUG start codon upstream of the $G_4C_2$ repeat is in the poly(GA) frame and drives production of poly(GA) DPRs in c9ALS/FTD (Cheng et al., 2018; Green et al., 2017; Sonobe et al., 2018; Tabet et al., 2018). A recent study demonstrated that expanded $G_4C_2$ repeat can promote ribosomal frameshifting during translation *in vitro* (Tabet et al., 2018). In addition, long-read sequencing across the $G_4C_2$ repeat suggests that genetic interruptions may occur within the $G_4C_2$ repeats in c9ALS/FTD patients (Ebbert et al., 2018). Together, these studies imply that more complex DPR species could be produced in c9ALS/FTD. In contrast, a recent study, using an engineered c9ALS/FTD iPS cell line, has demonstrated that while the deletion of a large piece of intronic sequence containing the near-cognate CUG start codon ablates poly(GA) production, poly(GP) and poly(GR) are not affected (Almeida et al., 2019). While this study provides further evidence that the non-canonical CUG in intron 1 of C9orf72 drives translation of poly(GA) in c9ALS/FTD, it is not possible to make additional conclusions on how loss of translation in the GA frame may directly affect the GP frame. It remains to be determined what are other consequences of a large deletion upstream of the $G_4C_2$ repeats on translation initiation, for example, introduction of an alternative near cognate start codon. It is further unclear whether these results are specific to the one individual cell line used in this study or whether these findings remain in additional c9ALS/FTD cellular models. Indeed, our data support the hypothesis that poly(GA) DPRs exist as a chimeric species with poly(GP) in c9ALS/FTD patient samples. Interestingly, we show that the ratio of GA to GP DPRs determined the solubility of cDPR species *in vitro*; GA-rich DPRs promote aggregation of chimeric DPRs whereas GP-rich chimeric species remain diffuse (Figures 11A and 11B). Thus, in c9ALS/FTD, initiation at the near-cognate CUG start codon followed by frameshifting would lead to the production of poly(GA)-rich cDPRs that produce inclusions composed of different DPR species. While it remains to be determined if cDPRs are produced in SCA36 patients, AUG-mediated translation in the poly(GP) frame would produce GP-rich and thus soluble cDPR species (Figures S11C and 11D). Our findings demonstrate that aggregation-prone, trans-frame, chimeric poly(GA:GP) DPRs exist in c9ALS/FTD and provide a compelling and parsimonious explanation for the divergent poly(GP) pathology seen between c9ALS/FTD and SCA36 (Figure 7).

Putative chimeric DPRs could have several functional consequences. As shown, GA-containing chimeras alter the localization and solubility of poly(GP) species. It is believed that poly(GP) is non-toxic; however, studies to date have been limited to studying the toxicity of soluble poly(GP), which does not recapitulate the neuropathology of c9ALS/FTD. Moreover, it is plausible that other cDPR species such as poly(GA:GR) or more complex cDPR species containing two or more contiguous stretches of DPRs exist. Notably, poly(GR) localizes to the nucleolus or is soluble in the cytoplasm in overexpression model systems (Bennion Callister et al., 2016; Yang et al., 2015); this is in contrast to the characteristic insoluble poly(GR) inclusions in c9ALS/FTD post-mortem tissue (Sakae et al., 2018). Additionally, it has been shown that poly(GR) inclusions co-localize with poly(GA) inclusions in c9ALS/FTD (Mori et al., 2013b). Poly(GR) is toxic *in vitro* (Kwon et al., 2014; Lopez-Gonzalez et al., 2016; Yang et al., 2015) and *in vivo* (Lopez-Gonzalez et

al., 2016; Zhang et al., 2018) and thus it has been implicated in the pathogenesis of c9ALS/FTD. The formation of chimeric GR species could alter its localization, solubility, and interactome. Functionally, chimeric GR species could modify the propensity of poly(GR) to undergo liquid-liquid phase separation, to perturb stress granule formation, or to impair protein translation – proposed mechanisms by which poly(GR) induces toxicity (Boeynaems et al., 2017; Kanekura et al., 2016; Lee et al., 2016a; Zhang et al., 2018). Lastly, chimeric DPRs would have important implications for therapies targeting DPRs. It is reasonable to posit that therapies targeting the dominant DPR species in a chimera, e.g., poly(GA) in c9ALS/FTD, would simultaneously modulate poly(GP) and possibly poly(GR).

A hallmark of c9ALS/FTD as well as sporadic and other genetic forms of ALS and FTD is the nuclear clearing of TDP-43 and the accumulation of phosphorylated and cleaved TDP-43 into cytoplasmic aggregates. Previous studies demonstrated that TDP-43 pathology correlates with neurodegeneration and clinical phenotypes (Davidson et al., 2016). Although it was demonstrated that c9ALS/FTD-associated DPRs can cause TDP-43 mislocalization in *Drosophila* (Solomon et al., 2018), work in mice found that overexpressing poly(GA) (Zhang et al., 2016), poly(GR) (Zhang et al., 2018), and poly(PR) (Zhang et al., 2019) alone cannot promote the formation of TDP-43 inclusions. However, it remains to be determined whether chimeric DPRs are able to induce TDP-43 pathology *in vitro* or *in vivo* and thus similar approaches used for modeling homogenous DPRs can be employed for chimeric DPR species. Nevertheless, expression of the c9ALS/FTD-associated expanded $G_4C_2$ repeat in mice does cause TDP-43 pathology (Chew et al., 2015), suggesting that TDP-43 dysfunction in c9ALS/FTD results from the presence of multiple DPRs (be they homogenous and/or chimeric species), from RNA-mediated toxicity or from some combination of these mechanisms. Consistent with these data, we show that the presence of soluble poly(GP) and poly(PR) are insufficient to cause TDP-43 pathology in SCA36.

Overall, this study reveals key similarities and differences in the translation, solubility, and aggregation of DPRs between SCA36 and c9ALS/FTD and implicates alternate mechanisms for how highly similar repeat expansions can cause markedly distinct forms of neurodegeneration. Our approach to compare these related yet different diseases provides us with a unique opportunity to address key questions about HRE-dependent neurodegeneration, and to expand our understanding of DPR properties and their role in neurological disorders.

## STAR Methods:

### Resource Availability

**Lead Contact—**Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gary Bassell (gbassel@emory.edu).

**Materials Availability—**All unique reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

**Data and Code Availability:** All data are available from the corresponding author upon request.

## Experimental Model and Subject Details

**Animals—**The animal protocol was approved by the Institutional Animal Care and Use Committees of Emory University and complied with the Guide for the Care and Use of Laboratory Animals. Wildtype pregnant mice (E13) were maintained on a C57BL/6 background. C57BL/6 wild type (WT) mice were obtained from The Jackson Laboratory. Since no behavioral assessments were conducted, the influence of sex was not considered for neuropathology studies. All mice were housed in plastic boxes in a climate-controlled facility and were provided food and water *ad libitum*.

**Human tissues—**Post-mortem brain tissues from c9ALS/FTD patients and controls were obtained from the Emory Neuropathology Core and Brainbank Neuro-CEB. SCA36 patient tissues were obtained from Brainbank Neuro-CEB, Tissu-Tumorotheque Est, and Brain Bank of Biobank Galicia Sur Health Research Institute. Patient information is provided in Table S1. Biological samples were obtained with Institutional Review Board approval.

**Induced pluripotent stem cells (iPS) generation and maintenance—**iPSCs were generated as previously described (Holler et al., 2016) using the Cytotune 2.0 Kit (Life Technologies) per the manufacturer's protocol. In brief, early passage fibroblast (<P10) were grown in fibroblast medium consisting of 10% ES-qualified FBS (Life Technologies), 0.1 mM NEAA, 55μM β-mercaptoethanol, high glucose DMEM with Glutamax (Life Technologies). On Day 0, fibroblasts were transduced with Sendai virus encoding KOS, hc-Myc, hKlf4, each at a MOI of 5. Cells were fed with fibroblast medium every other day for one week. On day 7, cells were passaged onto vitronectin (Life Technologies) coated dishes at a density of 250,000 to 500,000 cells/well. Beginning on day 8, cells were fed every day in Essential 8 medium (Life Technologies). iPS colonies were manually picked and transferred to a dish coated with either vitronectin or Matrigel (BD). iPSCs were maintained on Matrigel coated dishes and mTesR1 medium (Stem Cell Technologies). iPSCs were passaged every 5–7 days using ReLeSR (Figure S12).

**Generation of 3D forebrain organoid—**iPSCs colonies were detached from Matrigel coated plates with collagenase (1mg/ml; Invitrogen) treatment for 1hr and suspended in EB medium, comprising of FGF-2-free hESC medium supplemented with Dorsomorphin (2μM; Tocris) and A-83 (2μM; Tocris), in non-treated polystyrene plates for 4 days with a daily medium change. On days 5–6, half of the medium was replaced with induction medium consisting of DMEM:F12, N2 Supplement (1X; Invitrogen), 10μg/ml Heparin (Sigma), NEAA (1X), Glutamax (1X), WNT-3A (4 ng/ml; R&D Systems), CHIR99021 (1μM; Cellagentech), and SB-431542 (1μM; Cellagentech). On day 7, organoids were embedded in Matrigel (Corning) and continued to grow in induction medium for 6 more days. On day 14, embedded organoids were mechanically dissociated from Matrigel by pipetting up and down onto the plate with a 5mL pipette tip. Typically, 10 – 20 organoids were transferred to each well of a 12-well spinning bioreactor (SpinΩ) (Qian et al., 2016) containing differentiation medium, consisting of DMEM:F12, N2 (1X), B27 (1X; Invitrogen), 2-Mercaptoenthanol (100μM), NEAA (1X), Insulin (2.5μg/ml; Sigma). At day 71, differentiation medium was exchanged with maturation medium, consisting of Neurobasal (Invitrogen), B27 (1X), 2-Mercaptoenthanol (100 μM), 0.2mM Ascorbic Acid (Sigma), 20 ng/ml BDNF (Peprotech),

20ng/ml GDNF (Peprotech), 1 ng/ml TFGβ (Peprotech), and 0.5 mM cAMP (Sigma). Media was changed every other day until day 105 when organoids were collected for protein lysates.

**iPS-motor neuron differentiation—**Prior to differentiation, iPS colonies were treated with 10μM ROCK inhibitor, Y-27632 (Stem Cell Technologies), for ~1 hour. iPSCs were treated with Accutase (Stem Cell Technologies) for ~8 min to obtain a single cell suspension. Cells were spun out of Accutase and resuspended in N2B27 differentiation medium (1:1 Advanced DMEM-F12/Neurobasal, 1x N2, 1x B27, 0.2% Penstrep, 1x Glutamax, 110μM β-mercaptoethanol) and seeded in 10cm Ultra-Low Attachment dishes (Corning) in order to form embryoid bodies. Cells were maintained as embryoid bodies throughout the differentiation and were fed every 2 days. For the first 2 days, the differentiation medium contained 3μM CHIR99021 (Stem Cell Technologies), 10μM SB431542 (Stemgent), 5μM DMH1 (Tocris), and 10μM Y-27632 (Tocris). Starting on day 2, 1μM Retinoic Acid (Sigma) and 500nM Smoothened Agonist (Millipore) were added to the medium and Y-27632 was removed. On Day 8, CHIR99021 was removed from the medium. On Day 8, SB and DMH1 were removed from the medium. On day 15, 10μM DAPT (Tocris) was added. 10ng/mL BDNF (Peprotech) and 10ng/mL GDNF (Peprotech) were added to the media on day 16. On day 20, embryoid bodies were disassociated to single cells using papain/DNase (Worhtington Bio) and plated on polyornithine/laminin (3.3 μg/mL) coated glass coverslips or cell culture plates. Disassociated neurons were maintained in N2B27 maintainence medium (Neurobasal, 1x N2, 1x B27 plus, 0.2% Penstrep, 1x Glutamax, 110μM β-mercaptoethanol) supplemented with 10ng/mL BDNF and 10ng/mL GDNF. Media was replenished every 2–3 days (Figure S13).

## Method Details

**Lysate preparation—**Adherent cells (HEK293T, fibroblast, iPS, iPS-motor neurons) were directly lysed in ice-cold co-IP buffer (50mM Tris–HCl, pH 7.4, 300mM NaCl, 1% Triton X-100, 5mM EDTA, 2% SDS) with 1x HALT protease and phosphatase inhibitors (Pierce). Lysates were sonicated at 25% amplitude for 3 cycles of 5 seconds on/5 seconds off. Cells in suspension (lymphoblastoid cell lines, 3D forebrain organoids) were first pelleted by centrifugation for 5 minutes at 300xg and subsequently lysed as was done for adherent cells.

**Immunoblotting—**For western blots, protein lysates were prepared in 4x loading buffer, heat-denatured at 70°C for 15 minutes. Samples were normalized to equal amounts and resolved on 10% Bis-Tris gels (Thermo Fisher), transferred to 0.44 μm PVDF membranes, and blocked in Licor Odyssey blocking buffer (Licor) for 1 hour. The membrane was incubated overnight at 4°C with primary antibodies diluted in blocking buffer. Secondary antibodies were diluted in blocking buffer and applied to membrane for 1 hour at R.T. Primary antibodies used in immunoblotting experiments included anti-pTDP43 (1:3000, Cosmo Bio), anti-TDP (1:1000, Protein Tech), anti-FLAG (1:1000, Sigma), anti-HA (1:1000, Cell Signaling), and anti-cMyc (1:1000, Roche) For dot blots, 1.5 μL of urea soluble protein fractions were applied to a pre-wetted PVDF membrane and allowed to air dry for ~1.5 hours at R.T. The membrane was blocked in 5% non-fat dry milk for 1 hour and

subsequently probed with purified anti-poly(GP) (Rb9258) overnight at 4°C. Raw immunoblot images provided in Figure S14.

**Sequential biochemical fractionation—**Tissue lysates were fractionated according to previous published protocols (Neumann et al., 2006). In brief, ~200mg tissue was homogenized in low salt buffer (10mM Tris, 5mM EDTA, 1mM DTT, 10% sucrose, 1x HALT protease/phosphatase inhibitor). Lysates were pelleted at 25,000g for 30 min. Supernatants were collected as the “low salt” fraction. The resulting pellet was solubilized in Triton-X buffer (low salt buffer + 1% triton X-100 + 0.5M NaCl). Lysates were subsequently pelleted at 180,000g for 30 min. Supernatants were collected as the “Triton-X” fraction. The resulting pellet was solubilized in Triton-X buffer + 30% sucrose. Lysates were subsequently pelleted at 180,000g for 30 min. Supernatants from this fraction were not used for analysis. The resulting pellet was solubilized in Sarkosyl buffer (low salt buffer + 1% sarkosyl + 0.5M NaCl). Lysates were subsequently pelleted at 180,000g for 30 min. Supernatants were collected as the “Sark” fraction. The resulting insoluble pellet was resolubilized in 8M urea.

**Immunofluorescence—**Cells were fixed with 4% paraformaldehyde for 10 minutes at R.T. Cells were washed 3x with 1x PBS and permeabilized for 10 min in 1x PBST (0.1% Triton-X). Cells were then blocked in 5% normal donkey serum (Jackson Labs) for 1 hour at R.T. Cells were incubated with primary antibodies (Table S3) in 5% normal donkey serum overnight at 4°C. The next day cells were washed 3x in PBS and incubated with secondary antibodies (Alexa Fluor, 1:1500) for 1 hour at R.T.

**Meso-Scale Discovery (MSD) immunoassay—**The expression of poly(GP) in brain, organoid or cell lysates, as well as CSF, was measured blinded using a previously described sandwich immunoassay that utilizes MSD electrochemiluminescence detection technology, and an affinity purified rabbit polyclonal poly(GP) antibody (Rb9259) for both the capture and the detection antibody (Gendron et al., 2017; Gendron et al., 2015). For detection of chimeric DPR proteins, the capture antibody was a monoclonal poly(GA) antibody (Millipore MABN889) and the detection antibody was a monoclonal poly(GP) antibody (TALS828.179). In brief, lysates were diluted in Tris-buffered saline (TBS) such that all samples of a given type were made up to the same concentration and an equal amount of protein for samples was tested in duplicate wells. CSF samples were tested in duplicate wells using 90 µl per well. Response values corresponding to the intensity of emitted light upon electrochemical stimulation of the assay plate using the MSD QUICKPLEX SQ120 were acquired. These response values were background corrected by subtracting the average response values from corresponding negative controls (e.g., lysates from tissues or cells lacking a repeat expansion).

**Intracerebroventricular injection (ICV) of AAV in newborn pups:** Neonates at postnatal day 0 after birth are collected from the cage and cryoanesthesized at 0°C for ca.5 min until movement stops before injection. A solution of recombinant AAV virus ($6\times10^{13}$ vg/mL) diluted in sterile saline containing the opaque tracer Evans blue (0.1%; to visualize the borders of the lateral ventricle after injection) is injected stereotactically into the ventricles using a 10 µL Nanofil syringe with a 33 gauge needle. The injection site is located

1 mm lateral to the sagittal suture, halfway between lambda and bregma. The needle was held perpendicular to the skull surface during insertion into the lateral ventricle. Once the needle was in place, 1–2 μL of viral solution were injected slowly (33 nl/sec) with a MicroSyringe Pump Controller into the lateral ventricle. One minute after the injection is complete, the needle was withdrawn slowly, and the injection procedure was repeated for the other ventricle. Post injection the pups were placed on another polystyrene weight boat with cage bedding on a warming pad (Deltaphase Isothermal Pad) until they regained normal color and resumed movement. Post recovery, the whole litter was returned to the mother together.

**Tissue Processing—**Mice were transcardially perfused with ice cold PBS followed by cold 4% Paraformaldehyde (Sigma). The brains were post fixed overnight in PFA followed by cryoprotection in 30% sucrose for ~ 48 hours.

**Immunohistochemistry—**Paraffin embedded sections (5μm or 8μm) were deparaffinized in Histo-clear (National Diagnostics) and rehydrated in ethanol. Antigen retrieval was performed in dd$H_2O$ by steam heat for 30 minutes. Endogenous peroxidase activity was quenched using hydrogen peroxide and washed 3x in TBS-T. Tissue sections were blocked using serum-free protein block (Dako) for 1 hour. Primary antibodies were applied for 45 minutes at R.T. and washed 3x in TBS-T. HRP-conjugated secondaries (Dako) were applied for 30 minutes at R.T. Peroxidase labeling was visualized with 3,3'-diaminobenzidine (DAB). Sections were subsequently counterstained with Gill's hematoxylin and blued in Scott's tap water substitute.

**Digital Spatial Profiling—**An automated setup capable of imaging and sample collection was developed by modifying a standard microscope. For protein detection, a multiplexed cocktail of 90 primary antibodies (Table S4), each with a unique, UV photocleavable indexing oligo, or 1fluorescent markers (antibodies and/or DNA dyes) was applied to a slide-mounted FFPE tissue section. The tissue slide was placed on the stage of an inverted microscope. A custom gasket was then clamped onto the slide, allowing the tissue to be submerged in 1.5 mL of buffer solution. The microcapillary tip is connected to a syringe pump primed with buffer solution, allowing for accurate aspiration of small volumes (<2.5 μL). Under the microscope, wide field fluorescence imaging was performed with epi-illumination from visible LED light engine. The tissue area of interest was then located using fluorescence imaging. 20x image corresponds to 650μm x 650μm of tissue area with a CMOS camera. The 20x images were stitched together to yield a high-resolution image of the tissue area of interest. The regions of interest (ROIs) were then selected based on the fluorescence information and sequentially processed by the microscope automation. The steps performed for each ROI by the microscope automation were as follows: First, the microcapillary tip was washed by dispensing clean buffer out the capillary and into a wash station. Next, the tissue slide was bulk washed by exchanging the buffer solution on the slide via the inlet and outlet wash ports on the gasket clamp. The microcapillary tip was then moved into position directly over the ROI with a distance of 50 μm from the tissue. The local region of tissue around the ROI was washed by dispensing 100 μL of buffer solution from the microcapillary. Then, the area of tissue within the ROI was selectively illuminated

with UV light to release the indexing oligos. UV LED light was collimated to be reflected from the DMD surface into the microscope objective and focused at the sample tissue. Each micro mirror unit in the DMD corresponds to ~1μm$^2$ area of sample and reflects the UV light in controlled pattern based on the ROI selection in the image. Following each UV illumination cycle, the eluent was collected from the local region via microcapillary aspiration and transferred to an individual well of a microtiter plate or strip tubes. Once all ROIs were processed, indexing oligos were hybridized to NanoString optical barcodes for *ex situ* digital counting and subsequently analyzed with an nCounter® Analysis System (Nanostring Technologies).

**Genomic DNA (gDNA) extraction and repeat-primed PCR (RP-PCR)—**Genomic DNA was extracted from patient cerebellum using the Puragene Kit (Qiagen) as per the manufacturer's protocol. RP-PCR was performed as previously described (DeJesus-Hernandez et al., 2011) with slight modifications. PCR reactions were performed in 50uL total reaction volumes containing 100ng sample gDNA, 1x SuperFi Buffer (Invitrogen), 0.2mM dNTPs (Invitrogen), 0.5μM primer mix, 1x SuperFi GC enhancer, and 0.02 U/μL Platinum SuperFi DNA polymerase (Invitrogen) PCR cycling conditions were as follows: 95°C for 2 minutes, 34 cycles of 98°C for 10 seconds, 64.7°C for 10 seconds, 72°C for 40 seconds, followed by a final elongation step at 72°C for 5 minutes.

**RNA extraction, quantitative reverse transcription with PCR (qRT-PCR), and Nanostring Gene Expression—**RNA was extracted using the Quick RNA kit (Zymo Research) with a combined on-column DNase I digestion step. Adherent cells were directly lysed in the culture dish and RNA extraction proceeded per the manufacture's protocol. For RNA extraction from post-mortem tissue, ~30mg of tissue was homogenized in lysis buffer using a bullet blender tissue homogenizer (Next Advance). RNA lysates were cleared by spinning samples at 10,000g for 1 minute. Cleared lysates were used for RNA extraction as per the manufacture's protocol. cDNA was obtained via RT-PCR using the High Capacity cDNA Reverse Transcription Kit (Thermo). To quantify relative mRNA expression of *NOP56*, qPCR was performed for each sample using Taqman gene expression assays (Thermo) on a Quantstudio 6 Flex system (Applied Biosystems). For *NOP56* qPCR analysis of tissue samples, the following FAM labelled Taqman assays were used Hs0197340 (*NOP56* Ex3–4), Hs00975373 (*NOP56* Ex1–2), AI89MBZ (custom *NOP56* Ex1–2), APXGVEM (custom *NOP56* Ex1–2), or AID1V09 (custom *NOP56* Intron1-Ex2). Relative tissue RNA expression was normalized to the *RPLP0*. For qPCR analysis of ASO treated fibroblast a custom FAM labelled Taqman assay spanning the exon 1-exon 2 splice junction of *NOP56* was used (APXGVEM) or the intron 1 – exon 2 junction (AID1V09). Relative RNA expression was normalized to *RPLP0*. For Nanostring gene expression assays, 75–100ng total RNA was mixed with reporter and capture probesets as per the manufacturers protocol and quantified on the nCounter® Analysis System (Nanostring Technologies). The nCounter Advanced Analysis software was used for normalization of RNA counts.

## Quantification and Statistical Analysis

All data are presented as mean ± standard error of the mean (SEM). The statistical test used and value of n samples for each experiment are indicated in the respective figure legends. In

general, data were analyzed using unpaired, two-tailed, Student's *t*-test, Kruskal-Wallis analysis of variance, one-way ANOVA followed by Dunn's *post hoc* or Dunnett's *post hoc* analysis, or two-way ANOVA followed by Sidak's *post hoc* analysis (GraphPad Prism 7). Data obtained from poly(GP) immunoassays were considered to be non-normally distributed. Data from qPCR were assumed to be normally distributed. Statistical analysis of qPCR data was performed on delta Ct values.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments:

We are extremely grateful to the patients and their families for participating in this study. We would also like to acknowledge the technical expertise provided by Marla Gearing, Deborah Cooper, and Kaylor Kelly of the Emory Neuropathology Core. We would like to acknowledge Alexandra Dürr of Pitié-Salpêtrière Hospital - Sorbonne Université for her assistance in obtaining patient samples. Additionally, we would like to thank Dr. Dieter Edbauer for kindly providing the monoclonal clone 5F2 poly(GA) antibody (provided to J.J.). We would like to acknowledge Yan Liang, Jingjing Gong, and Liang Zhang for their assistance with DSP data collection and analysis. Human biological samples and associated data were obtained from the Emory Neuropathology Core (P30 NS055077), Brain Bank of Biobank Galicia Sur Health Research Institute (PT17/0015/0034), Brainbank Neuro-CEB, Hôpital de la Pitié-Salpêtrière, and Tissu-Tumorotheque Est, Hospices Civils de Lyon Biobank (CRB-HCL, BB-0033-00046). This work was supported by an Emory University Research Committee grant (G.J.B.), Emory University School of Medicine Laboratory for Translational Cell Biology (G.J.B), National Institute for Neurological Disorders and Stroke R21NS114908 (J.J. and G.J.B), the Center for Neurodysfunction and Inflammation at Emory (M.G.T. and G.J.B.), the National Ataxia Foundation (W.R.), National Institute for Neurological Disorders and Stroke R01NS082094 (B.L.F.), the National Institutes of Health/National Institute of Neurological Disorders and Stroke [R35NS097273 (L.P.); P01NS084974 (L.P., D.D.); P01NS099114 (T.F.G., L.P.)], and the ISCIII grants PI12/01013 and PI17/01789 (M-J.S.), cofounded by FEDER. Z.T.M. was supported by a Training and Research in Translational Neurology Training Grant (2T32 NS007480-15).

## References:

Almeida S, Krishnan G, Rushe M, Gu Y, Kankel MW, and Gao F-B (2019). Production of poly(GA) in C9ORF72 patient motor neurons derived from induced pluripotent stem cells. Acta Neuropathologica 138, 1099–1101. [PubMed: 31624870]

Arias M, Arias-Rivas S, Blanco-Arias P, Dapena D, Vazquez F, and Rossi M (2008). SCA from the Costa da Morte: A new SCA. Description of the phenotype. Neurologia 23, 628–629.

Arias M, Garcia-Murias M, and Sobrido MJ (2017). Spinocerebellar ataxia 36 (SCA36): <<Costa da Morte ataxia>>. Neurologia 32, 386–393. [PubMed: 25593102]

Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, Dejesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, et al. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron 77, 639–646. [PubMed: 23415312]

Banez-Coronel M, Ayhan F, Tarabochia AD, Zu T, Perez BA, Tusi SK, Pletnikova O, Borchelt DR, Ross CA, Margolis RL, et al. (2015). RAN Translation in Huntington Disease. Neuron 88, 667–677. [PubMed: 26590344]

Bennion Callister J, Ryan S, Sim J, Rollinson S, and Pickering-Brown SM (2016). Modelling C9orf72 dipeptide repeat proteins of a physiologically relevant size. Hum Mol Genet 25, 5069–5082. [PubMed: 27798094]

Bjorkoy G, Lamark T, and Johansen T (2006). p62/SQSTM1: a missing link between protein aggregates and the autophagy machinery. Autophagy 2, 138–139. [PubMed: 16874037]

Boeynaems S, Bogaert E, Kovacs D, Konijnenberg A, Timmerman E, Volkov A, Guharoy M, De Decker M, Jaspers T, Ryan VH, et al. (2017). Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. Mol Cell 65, 1044–1055 e1045. [PubMed: 28306503]

Cheng W, Wang S, Mestre AA, Fu C, Makarem A, Xian F, Hayes LR, Lopez-Gonzalez R, Drenner K, Jiang J, et al. (2018). C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2α phosphorylation. Nature Communications 9, 51.

Chew J, Gendron TF, Prudencio M, Sasaguri H, Zhang YJ, Castanedes-Casey M, Lee CW, Jansen-West K, Kurti A, Murray ME, et al. (2015). Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. Science 348, 1151–1154. [PubMed: 25977373]

Cleary JD, and Ranum LP (2017). New developments in RAN translation: insights from multiple diseases. Curr Opin Genet Dev 44, 125–134. [PubMed: 28365506]

Davidson Y, Robinson AC, Liu X, Wu D, Troakes C, Rollinson S, Masuda-Suzukake M, Suzuki G, Nonaka T, Shi J, et al. (2016). Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9orf72 is linked to TDP-43 pathology and not associated with aggregated forms of dipeptide repeat proteins. Neuropathol Appl Neurobiol 42, 242–254. [PubMed: 26538301]

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72, 245–256. [PubMed: 21944778]

Ebbert MTW, Farrugia SL, Sens JP, Jansen-West K, Gendron TF, Prudencio M, McLaughlin IJ, Bowman B, Seetin M, DeJesus-Hernandez M, et al. (2018). Long-read sequencing across the C9orf72 'GGGGCC' repeat expansion: implications for clinical use and genetic discovery efforts in human disease. Molecular Neurodegeneration 13, 46. [PubMed: 30126445]

Freibaum BD, and Taylor JP (2017). The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. Front Mol Neurosci 10, 35. [PubMed: 28243191]

García-Murias M, Quintáns B, Arias M, Seixas AI, Cacheiro P, Tarrío R, Pardo J, Millán MJ, Arias-Rivas S, Blanco-Arias P, et al. (2012). 'Costa da Morte' ataxia is spinocerebellar ataxia 36: clinical and genetic characterization. Brain 135, 1423–1435. [PubMed: 22492559]

Gendron TF, Bieniek KF, Zhang YJ, Jansen-West K, Ash PE, Caulfield T, Daughrity L, Dunmore JH, Castanedes-Casey M, Chew J, et al. (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. Acta Neuropathol 126, 829–844. [PubMed: 24129584]

Gendron TF, Chew J, Stankowski JN, Hayes LR, Zhang YJ, Prudencio M, Carlomagno Y, Daughrity LM, Jansen-West K, Perkerson EA, et al. (2017). Poly(GP) proteins are a useful pharmacodynamic marker for C9ORF72-associated amyotrophic lateral sclerosis. Sci Transl Med 9.

Gendron TF, van Blitterswijk M, Bieniek KF, Daughrity LM, Jiang J, Rush BK, Pedraza O, Lucas JA, Murray ME, Desaro P, et al. (2015). Cerebellar c9RAN proteins associate with clinical and neuropathological characteristics of C9ORF72 repeat expansion carriers. Acta Neuropathol 130, 559–573. [PubMed: 26350237]

Green KM, Glineburg MR, Kearse MG, Flores BN, Linsalata AE, Fedak SJ, Goldstrohm AC, Barmada SJ, and Todd PK (2017). RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. Nat Commun 8, 2005. [PubMed: 29222490]

Holler CJ, Taylor G, McEachin ZT, Deng Q, Watkins WJ, Hudson K, Easley CA, Hu WT, Hales CM, Rossoll W, et al. (2016). Trehalose upregulates progranulin expression in human and mouse models of GRN haploinsufficiency: a novel therapeutic lead to treat frontotemporal dementia. Mol Neurodegener 11, 46. [PubMed: 27341800]

Ikeda Y, Ohta Y, Kobayashi H, Okamoto M, Takamatsu K, Ota T, Manabe Y, Okamoto K, Koizumi A, and Abe K (2012). Clinical features of SCA36: a novel spinocerebellar ataxia with motor neuron involvement (Asidan). Neurology 79, 333–341. [PubMed: 22744658]

Ikeda Y, Ohta Y, Kurata T, Shiro Y, Takao Y, and Abe K (2013). Acoustic impairment is a distinguishable clinical feature of Asidan/SCA36. J Neurol Sci 324, 109–112. [PubMed: 23140984]

Kanekura K, Yagi T, Cammack AJ, Mahadevan J, Kuroda M, Harms MB, Miller TM, and Urano F (2016). Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. Hum Mol Genet 25, 1803–1813. [PubMed: 26931465]

Kobayashi H, Abe K, Matsuura T, Ikeda Y, Hitomi T, Akechi Y, Habu T, Liu W, Okuda H, and Koizumi A (2011). Expansion of intronic GGCCTG hexanucleotide repeat in NOP56 causes SCA36, a type of spinocerebellar ataxia accompanied by motor neuron involvement. Am J Hum Genet 89, 121–130. [PubMed: 21683323]

Kwon I, Xiang S, Kato M, Wu L, Theodoropoulos P, Wang T, Kim J, Yun J, Xie Y, and McKnight SL (2014). Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. Science 345, 1139–1145. [PubMed: 25081482]

Lee KH, Zhang P, Kim HJ, Mitrea DM, Sarkar M, Freibaum BD, Cika J, Coughlin M, Messing J, Molliex A, et al. (2016a). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. Cell 167, 774–788 e717. [PubMed: 27768896]

Lee YC, Tsai PC, Guo YC, Hsiao CT, Liu GT, Liao YC, and Soong BW (2016b). Spinocerebellar ataxia type 36 in the Han Chinese. Neurol Genet 2, e68. [PubMed: 27123487]

Lin Y, Mori E, Kato M, Xiang S, Wu L, Kwon I, and McKnight SL (2016). Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. Cell 167, 789–802 e712. [PubMed: 27768897]

Liu W, Ikeda Y, Hishikawa N, Yamashita T, Deguchi K, and Abe K (2014). Characteristic RNA foci of the abnormal hexanucleotide GGCCUG repeat expansion in spinocerebellar ataxia type 36 (Asidan). Eur J Neurol 21, 1377–1386. [PubMed: 24985895]

Lopez-Gonzalez R, Lu Y, Gendron TF, Karydas A, Tran H, Yang D, Petrucelli L, Miller BL, Almeida S, and Gao FB (2016). Poly(GR) in C9ORF72-Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. Neuron 92, 383–391. [PubMed: 27720481]

Lykke-Andersen S, Ardal BK, Hollensen AK, Damgaard CK, and Jensen TH (2018). Box C/D snoRNP Autoregulation by a cis-Acting snoRNA in the NOP56 Pre-mRNA. Mol Cell 72, 99–111 e115. [PubMed: 30220559]

Mackenzie IR, Frick P, Grasser FA, Gendron TF, Petrucelli L, Cashman NR, Edbauer D, Kremmer E, Prudlo J, Troost D, and Neumann M (2015). Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. Acta Neuropathol 130, 845–861. [PubMed: 26374446]

Matsuzono K, Imamura K, Murakami N, Tsukita K, Yamamoto T, Izumi Y, Kaji R, Ohta Y, Yamashita T, Abe K, and Inoue H (2017). Antisense Oligonucleotides Reduce RNA Foci in Spinocerebellar Ataxia 36 Patient iPSCs. Mol Ther Nucleic Acids 8, 211–219. [PubMed: 28918022]

Mizielinska S, Gronke S, Niccoli T, Ridler CE, Clayton EL, Devoy A, Moens T, Norona FE, Woollacott IOC, Pietrzyk J, et al. (2014). C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. Science 345, 1192–1194. [PubMed: 25103406]

Mori K, Arzberger T, Grasser FA, Gijselinck I, May S, Rentzsch K, Weng SM, Schludi MH, van der Zee J, Cruts M, et al. (2013a). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. Acta Neuropathol 126, 881–893. [PubMed: 24132570]

Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretzschmar HA, Cruts M, Van Broeckhoven C, et al. (2013b). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science 339, 1335–1338. [PubMed: 23393093]

Neumann M, Sampathu DM, Kwong LK, Truax AC, Micsenyi MC, Chou TT, Bruce J, Schuck T, Grossman M, Clark CM, et al. (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314, 130–133. [PubMed: 17023659]

Niblock M, Smith BN, Lee YB, Sardone V, Topp S, Troakes C, Al-Sarraj S, Leblond CS, Dion PA, Rouleau GA, et al. (2016). Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: implications for the pathogenesis of ALS/FTD. Acta Neuropathol Commun 4, 18. [PubMed: 26916632]

Nielsen H, and Wernersson R (2006). An overabundance of phase 0 introns immediately after the start codon in eukaryotic genes. BMC Genomics 7, 256–256. [PubMed: 17034638]

Obayashi M, Stevanin G, Synofzik M, Monin ML, Duyckaerts C, Sato N, Streichenberger N, Vighetto A, Desestret V, Tesson C, et al. (2015). Spinocerebellar ataxia type 36 exists in diverse populations

and can be caused by a short hexanucleotide GGCCTG repeat expansion. J Neurol Neurosurg Psychiatry 86, 986–995. [PubMed: 25476002]

Ohta Y, Hayashi T, Nagai M, Okamoto M, Nagotani S, Nagano I, Ohmori N, Takehisa Y, Murakami T, Shoji M, et al. (2007). Two cases of spinocerebellar ataxia accompanied by involvement of the skeletal motor neuron system and bulbar palsy. Internal medicine (Tokyo, Japan) 46, 751–755.

Qian X, Nguyen HN, Song MM, Hadiono C, Ogden SC, Hammack C, Yao B, Hamersky GR, Jacob F, Zhong C, et al. (2016). Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. Cell 165, 1238–1254. [PubMed: 27118425]

Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72, 257–268. [PubMed: 21944779]

Sakae N, Bieniek KF, Zhang Y-J, Ross K, Gendron TF, Murray ME, Rademakers R, Petrucelli L, and Dickson DW (2018). Poly-GR dipeptide repeat polymers correlate with neurodegeneration and Clinicopathological subtypes in C9ORF72-related brain disease. Acta Neuropathologica Communications 6, 63. [PubMed: 30029693]

Sarto E, Magri S, Mariotti C, Nanetti L, Gellera C, Di Bella D, and Taroni F (2013). SCA36 molecular analysis in patients with spinocerebellar ataxia. In International Conference on Spinocerebellar Degenerations (Paris, France).

Solomon DA, Stepto A, Au WH, Adachi Y, Diaper DC, Hall R, Rekhi A, Boudi A, Tziortzouda P, Lee Y-B, et al. (2018). A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin-α mediates C9orf72-related neurodegeneration. Brain 141, 2908–2924. [PubMed: 30239641]

Sonobe Y, Ghadge G, Masaki K, Sendoel A, Fuchs E, and Roos RP (2018). Translation of dipeptide repeat proteins from the C9ORF72 expanded repeat is associated with cellular stress. Neurobiology of disease 116, 155–165. [PubMed: 29792928]

Sulek A, Krysa W, Elert-Dobkowska E, Rajkiewicz M, STepniak I, Rakowicz M, Rola R, Dusza-Rowinska M, and Zaremba J (2013). Spinocerebellar ataxias (SCAs) and hereditary spastic paraplegias (HSP) - rare movement disorders prevalence in Poland. In International Conference on Spinocerebellar Degenerations (Paris, France).

Sznajder LJ, Thomas JD, Carrell EM, Reid T, McFarland KN, Cleary JD, Oliveira R, Nutter CA, Bhatt K, Sobczak K, et al. (2018). Intron retention induced by microsatellite expansions as a disease biomarker. Proc Natl Acad Sci U S A 115, 4234–4239. [PubMed: 29610297]

Tabet R, Schaeffer L, Freyermuth F, Jambeau M, Workman M, Lee C-Z, Lin C-C, Jiang J, Jansen-West K, Abou-Hamdan H, et al. (2018). CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. Nature Communications 9, 152.

Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basrur V, et al. (2013). CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. Neuron 78, 440–455. [PubMed: 23602499]

Valera JM, Diaz T, Petty LE, Quintans B, Yanez Z, Boerwinkle E, Muzny D, Akhmedov D, Berdeaux R, Sobrido MJ, et al. (2017). Prevalence of spinocerebellar ataxia 36 in a US population. Neurol Genet 3, e174. [PubMed: 28761930]

Yang D, Abdallah A, Li Z, Lu Y, Almeida S, and Gao F-B (2015). FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. Acta Neuropathologica 130, 525–535. [PubMed: 26031661]

Yang J, Kim MJ, Yoon W, Kim EY, Kim H, Lee Y, Min B, Kang KS, Son JH, Park HT, et al. (2017). Isocitrate protects DJ-1 null dopaminergic cells from oxidative stress through NADP+-dependent isocitrate dehydrogenase (IDH). PLOS Genetics 13, e1006975. [PubMed: 28827794]

Yoshihama M, Uechi T, Asakawa S, Kawasaki K, Kato S, Higa S, Maeda N, Minoshima S, Tanaka T, Shimizu N, and Kenmochi N (2002). The human ribosomal protein genes: sequencing and comparative analysis of 73 genes. Genome research 12, 379–390. [PubMed: 11875025]

Younis I, Berg M, Kaida D, Dittmar K, Wang C, and Dreyfuss G (2010). Rapid-response splicing reporter screens identify differential regulators of constitutive and alternative splicing. Mol Cell Biol 30, 1718–1728. [PubMed: 20123975]

Zeng S, Zeng J, He M, Zeng X, Zhou Y, Liu Z, Xia K, Pan Q, Jiang H, Shen L, et al. (2016). Genetic and clinical analysis of spinocerebellar ataxia type 36 in Mainland China. Clin Genet 90, 141–148. [PubMed: 26661328]

Zhang YJ, Gendron TF, Ebbert MTW, O'Raw AD, Yue M, Jansen-West K, Zhang X, Prudencio M, Chew J, Cook CN, et al. (2018). Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. Nat Med 24, 1136–1142. [PubMed: 29942091]

Zhang YJ, Gendron TF, Grima JC, Sasaguri H, Jansen-West K, Xu YF, Katzman RB, Gass J, Murray ME, Shinohara M, et al. (2016). C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. Nat Neurosci 19, 668–677. [PubMed: 26998601]

Zhang YJ, Guo L, Gonzales PK, Gendron TF, Wu Y, Jansen-West K, O'Raw AD, Pickles SR, Prudencio M, Carlomagno Y, et al. (2019). Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity. Science 363.

Zu T, Cleary JD, Liu Y, Banez-Coronel M, Bubenik JL, Ayhan F, Ashizawa T, Xia G, Clark HB, Yachnis AT, et al. (2017). RAN Translation Regulated by Muscleblind Proteins in Myotonic Dystrophy Type 2. Neuron 95, 1292–1305 e1295. [PubMed: 28910618]

Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MA, et al. (2011). Non-ATG-initiated translation directed by microsatellite expansions. Proc Natl Acad Sci U S A 108, 260–265. [PubMed: 21173221]

Zu T, Liu Y, Banez-Coronel M, Reid T, Pletnikova O, Lewis J, Miller TM, Harms MB, Falchook AE, Subramony SH, et al. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. Proc Natl Acad Sci U S A 110, E4968–4977. [PubMed: 24248382]

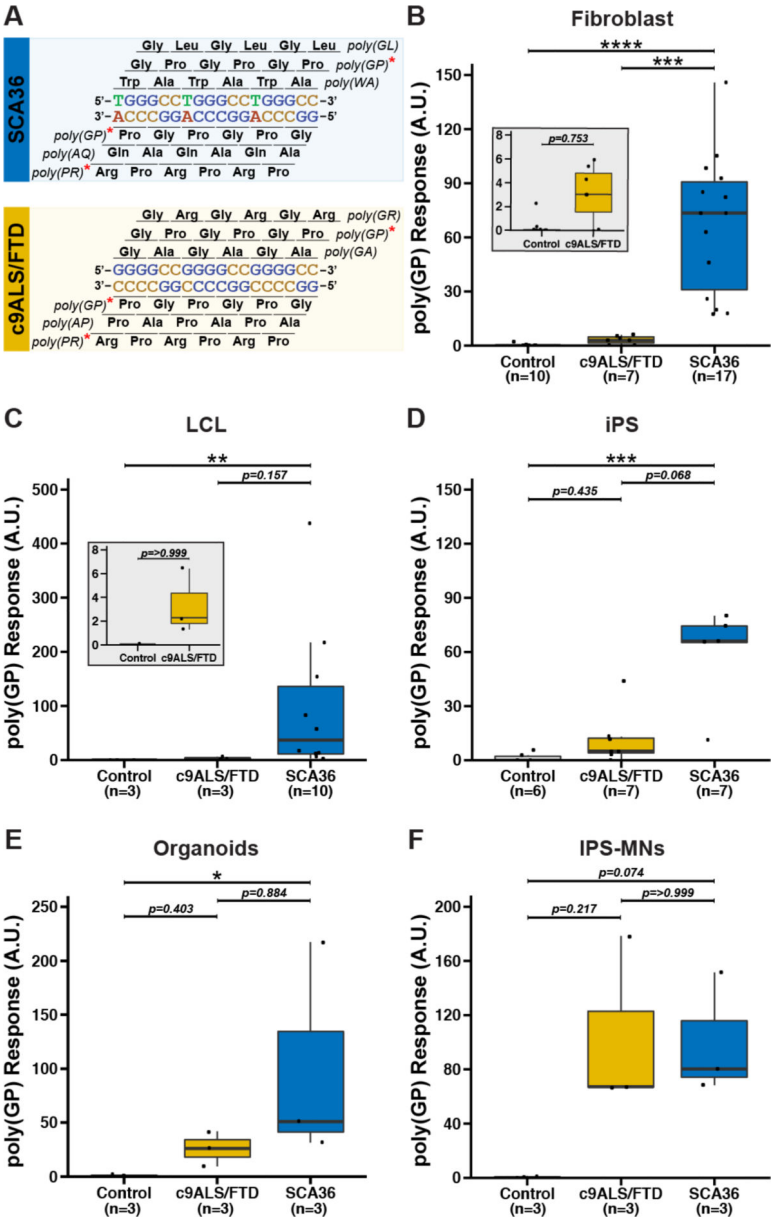


**Figure 1: Poly(GP) is more abundant in patient-derived cellular models of SCA36 compared to c9ALS/FTD.**

**(A)** Schematic of putative dipeptide repeats in SCA36 and c9ALS/FTD. Red asterisk indicates putative DPRs common to both c9ALS/FTD and SCA36. **(B-F)** Immunoassay for poly(GP) in SCA36 fibroblast, lymphoblastoid cell lines (LCL), induced pluripotent stem cell (iPS), 3D forebrain organoid (Day 105), and iPS-derived motor neurons (iPS-MNs). (B-F) Adjusted P-values displayed if not significant; *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001; Kruskal-Wallis one way analysis of variance followed by Dunn's multiple comparison test.

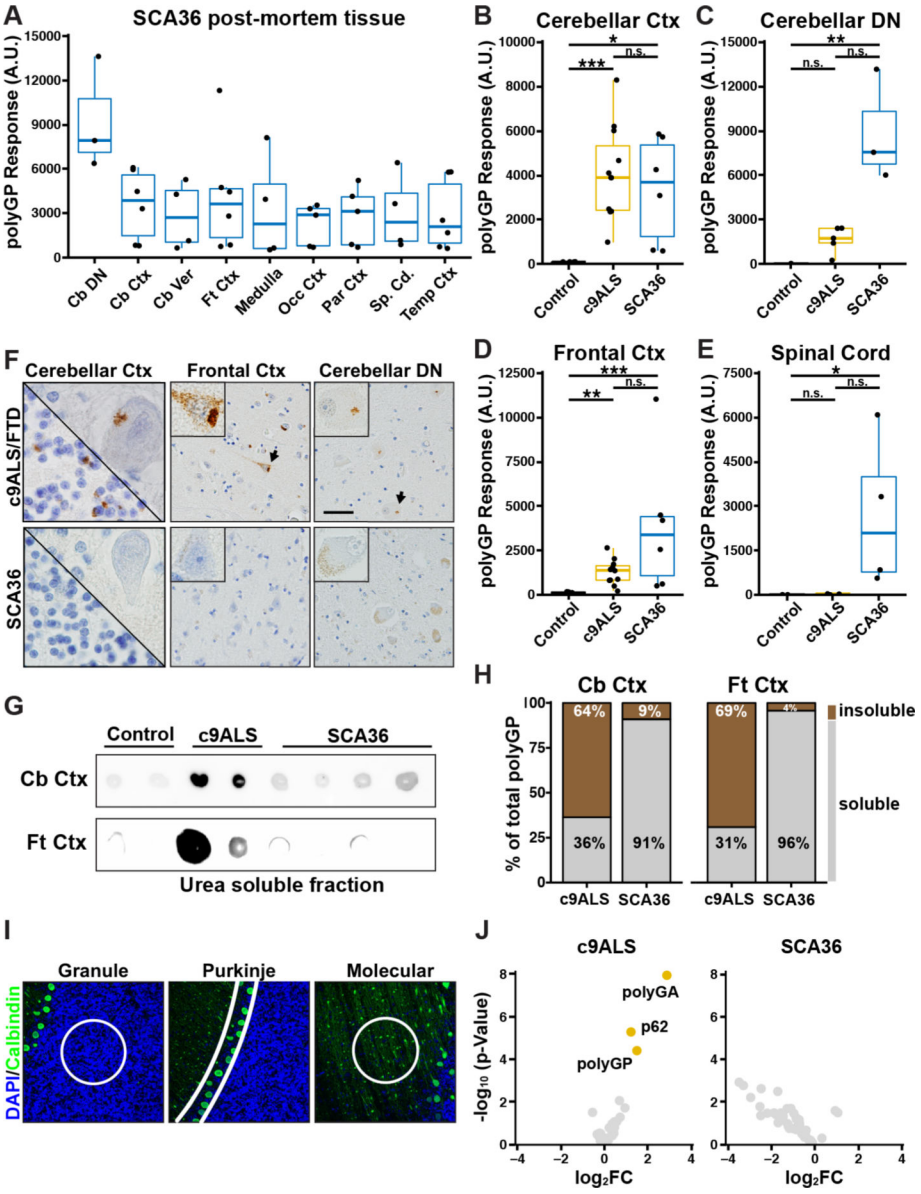


**Figure 2: Poly(GP) is abundant but does not form insoluble inclusions in SCA36 postmortem tissue.**

**(A)** poly(GP) is detectable throughout the CNS of SCA36 patients. **(B-E)** poly(GP) is expressed at similar or increased abundance in SCA36 samples compared to c9ALS/FTD samples in disease relevant CNS regions (Cb Ctx: n=7 control, n=11 c9ALS/FTD, n=6 SCA36 cases; Cb DN: n=3 control, n=5 c9ALS/FTD, n=3 SCA36 cases; Ft. Ctx: n=7 control, n=11 c9ALS/FTD, n=6 SCA36 cases; Sp Cd: n=3 control, n=3 c9ALS/FTD, n=4 SCA36 cases). **(F)** poly(GP) aggregates in c9ALS/FTD but not SCA36 patient postmortem tissue (n=5 SCA36). Scale Bar: 100 μm. **(G)** poly(GP) is detectable in the urea-soluble fraction of c9ALS/FTD but not SCA36 samples as assessed by dotblot (n=2 control, n=2 c9ALS/FTD, n=4 SCA36 cases). **(H)** Quantitative immunoassay of sequentially fractionated tissue samples reveals poly(GP) predominately exist as a soluble protein species in SCA36 whereas in c9ALS/FTD a large fraction of poly(GP) exists in the insoluble fraction (Cb Ctx:

n=4 c9ALS/FTD, n=4 SCA36 cases; Ft. Ctx: n=5 c9ALS/FTD, n=5 SCA36 cases). **(I)** Representative images of granule, Purkinje, and molecular cell layer ROI selection (white ouline) for digital spatial profiling of cerebellar samples. Calbindin and DAPI were used to delineate cerebellar architecture. **(J)** Volcano plots illustrating differentially abundant proteins as determined by digital spatial profiling in the cerebellar granule layer of c9ALS/FTD and SCA36 compared to controls (n=7 c9ALS/FTD, n=5 SCA36, n=7 control cases). Significant proteins (FDR<0.05) are colored. (B to E) *P<0.05, **P<0.005, ***P<0.0005, Kruskal-Wallis one way analysis of variance followed by Dunn's multiple comparison test.

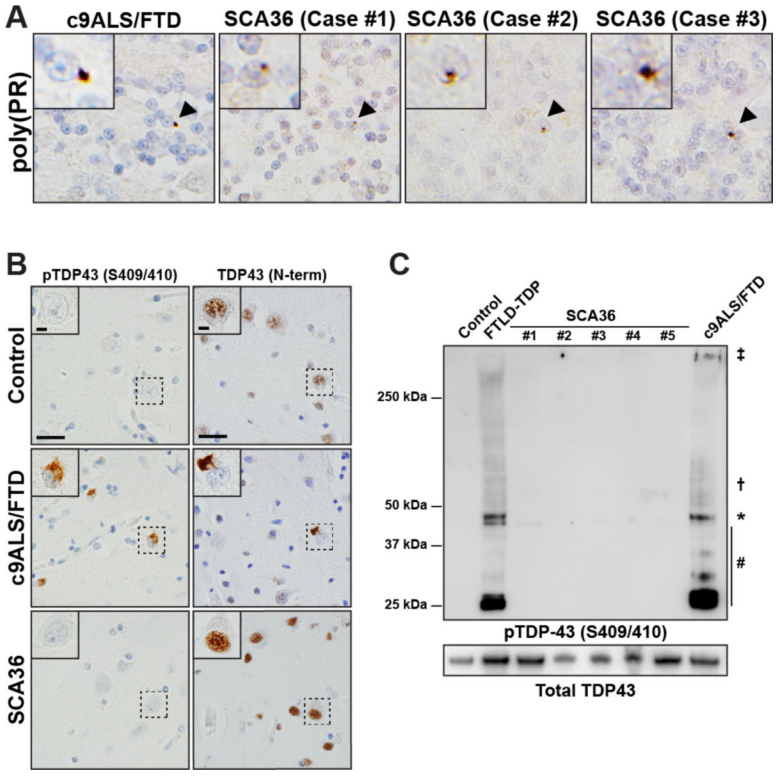


**Figure 3: SCA36 produces antisense DPRs but lacks TDP-43 pathology.**

**(A)** The antisense DPR, poly(PR) aggregates in both c9ALS/FTD and SCA36 cerebellar tissue. **(B)** SCA36 frontal cortex tissue lacks TDP43 pathology characteristic of c9ALS/FTD. Scale bars: 50μm, *inset* 10μm. **(C)** pTDP43 is not found in the detergent insoluble, urea soluble fraction of SCA36 patient frontal cortex. #cleavage products, *full-length TDP43, †oligomeric species, ‡high molecular weight species.

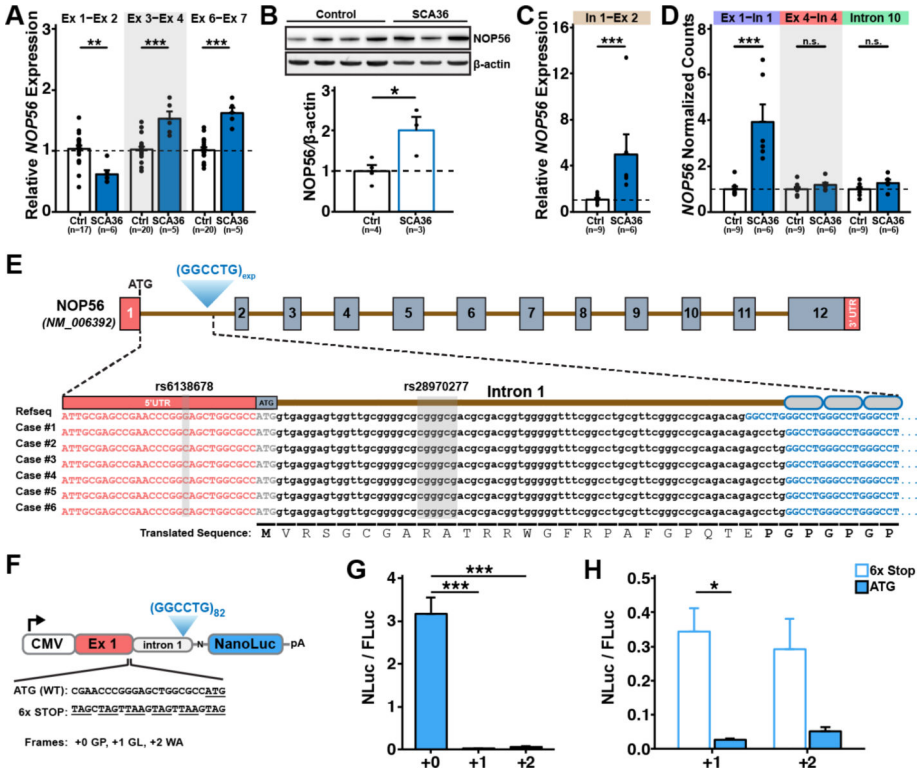


**Figure 4: Expanded $TG_3C_2$ repeats impair splicing of *NOP56* intron 1 and result in AUG-mediated translation of poly(GP) DPRs in SCA36.**

**(A)** Relative expression of *NOP56* using TaqMan probes spanning exon-exon junctions in control and SCA36 patient cerebellar cortex samples (n=20 control, n=5 SCA36). *NOP56* expression as detected by exon 3-exon 4 or exon 6-exon 7 junction probes is increased in SCA36 compared to controls. In contrast, *NOP56* expression as detected by exon 1-exon 2 junction probes is decreased in SCA36 compared to controls. **(B)** NOP56 protein is increased in SCA36 patient cerebellar cortex compared to controls (n=4 control, n=3 SCA36). *Top*: Western blot analysis of NOP56 and ß-actin. *Bottom*: Quantification of NOP56 protein in control and SCA36 samples relative to ß-actin. **(C)** Relative expression of *NOP56* using a TaqMan probe spanning the intron 1 - exon 2 junction in control and SCA36 patient cerebellar cortex samples (n=9 control, n=6 SCA36). **(D)** *NOP56* transcripts containing a retained intron 1 but not intron 4 or intron 10 is increased in SCA36 compared to control (n=9 control, n=6 SCA36) as detected by Nanostring. **(E)** Alignment of the 5' flanking sequence of intron 1 from the expanded allele of SCA36 cases. **(F)** Schematic figure of $(GGCCTG)_{82}$ NanoLuc (NLuc) reporter constructs. **(G)** Expression of ATG $(GGCCTG)_{82}$ NanoLuc reporter constructs (n=3 biological replicates). Values normalized to co-transected firefly luciferase (FLuc). **(H)** Expression of 6x Stop and ATG - $(GGCCTG)_{82}$ NanoLuc reporter constructs for the +1 (GL) and +2 (WA) frame (n=3 biological replicates). Data are presented as mean ± SEM. (A-D) *P<0.05, **P<0.005, ***P<0.0005, ****P<0.0001, Mann-Whitney test. (G) ***P<0.0005, ordinary one-way ANOVA. (H) *P<0.05, unpaired Student's t-test.

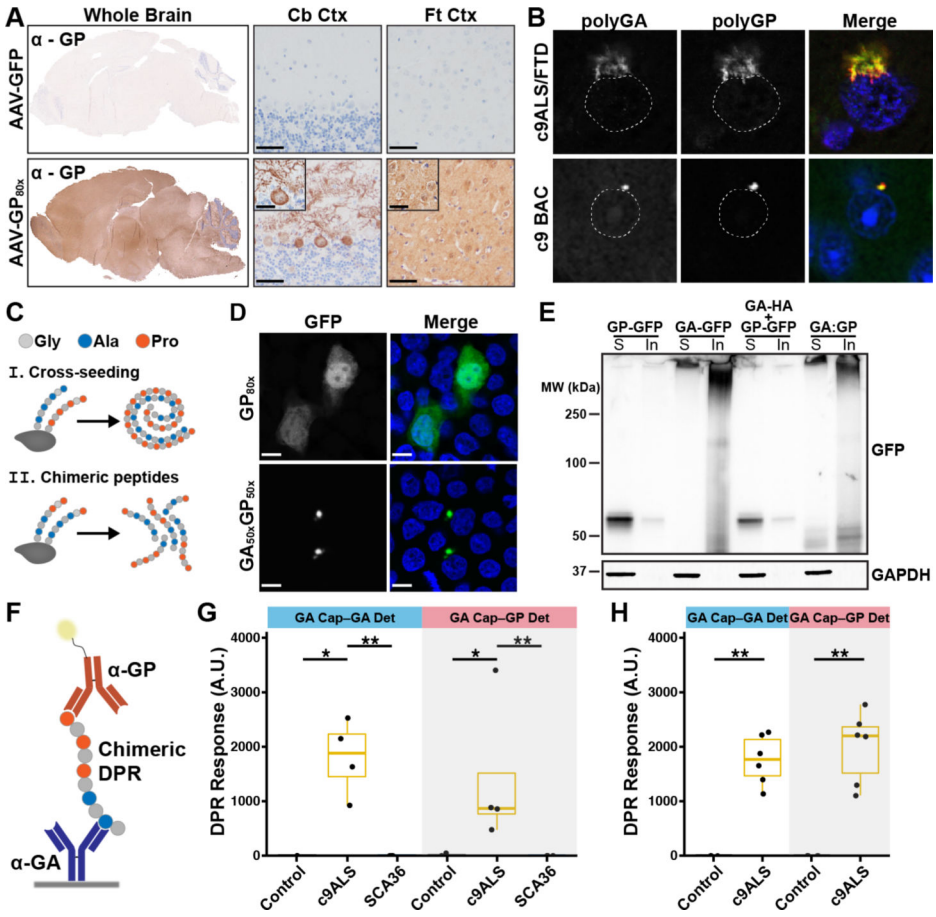


**Figure 5: Chimeric peptides underlie divergent DPR pathology in c9ALS/FTD and SCA36.**

**(A)** poly(GP) immunohistochemistry in AAV-GFP and AAV-$GP_{80x}$-GFP; poly(GP) is diffuse throughout the CNS after 2 months *in vivo*. Scale bars: 100μm, *inset* 50μm. **(B)** poly(GA) and poly(GP) colocalize in perinuclear aggregates in the frontal cortex of a c9ALS/FTD patients and c9BAC mice. **(C)** Schematic illustrating putative mechanisms of poly(GA)-mediated poly(GP) aggregation. **(D)** Immunofluorescence for GFP 48 hours after transfection in HEK293T cells expressing $GP_{80x}$-GFP or chimeric $GA_{50x}GP_{50x}$-GFP to assess the aggregation propensity of a chimeric GAGP peptide. Scale bar: 10μm. **(E)** Immunoblot for the indicated proteins to determine whether $GP_{80x}$-GFP, $GA_{80x}$-GFP, chimeric $GA_{50x}GP_{50}$-GFP, or 1:1 co-transfection of $GP_{80x}$-GFP and $GA_{80x}$-HA is present in the soluble (S) or insoluble (In) fraction. **(F)** Schematic illustrating the meso-scale discovery (MSD) immunoassay developed to detect chimeric DPRs. **(G)** A traditional GA/GA MSD assay or a chimeric peptide detecting GA/GP immunoassay was used to detect poly(GA) or chimeric poly(GA:GP) DPRs in the detergent insoluble, urea soluble fraction from control (n=3 cases), c9ALS/FTD (n=4 cases), or SCA36 (n=5 cases) patient frontal cortex and **(H)** detergent-insoluble, urea-soluble fraction from control (n=3 cases) and c9ALS/FTD (n=6 cases) patient cerebellar cortex. (G, H) *p<0.05, **p<0.005, Kruskal-Wallis one way analysis of variance followed by pairwise one-tailed Mann-Whitney Test.

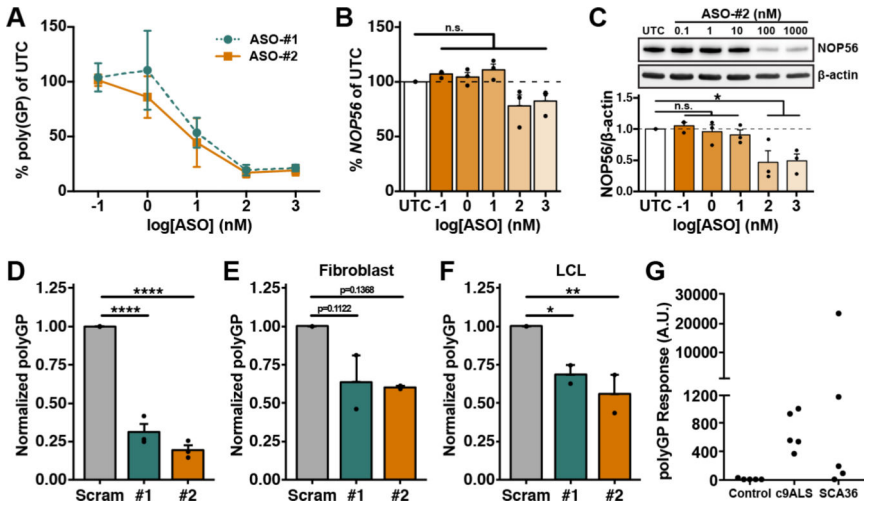


**Figure 6: Antisense oligonucleotides reduce poly(GP) in patient derived cellular models of SCA36.**

**(A-D)** Dose-response curves were generated by treating SCA36 fibroblast (n=3 patient cell lines) with described concentrations. UTC=untransfected control. **(A)** Dose-response curves show robust reduction of poly(GP) by *NOP56* intron 1 (ASO-#1) or $TG_3C_2$ (ASO-#2) targeting ASOs (see Figure S10A and Table S2). **(B)** *NOP56* mRNA levels as measured by exon 1 – exon 2 junction probes are not reduced at 1nM or 10nM ASO concentrations. **(C)** NOP56 protein is not significantly reduced at 0.1nM, 1nM, or 10nM ASO concentrations. **(D)** poly(GP) is reduced at 10nM ASO concentrations relative to a scrambled control. (A) to (D) n=3 SCA36 patient fibroblast cell lines, treated for 48 hours at respective concentrations. **(E, F)** SCA36 fibroblast (n=2 patient cell lines) and LCLs (n=2 patient cell lines) were treated with 5μM non-targeting control ASO or an ASO targeting intron 1 of *NOP56* delivered via gymnosis. Protein lysates were collected 10 days after treatment and poly(GP) levels measured. Data are presented as mean ± SEM; (B) to (E), *P<0.05, **P<0.01, ***P<0.001 one-way analysis of variance (ANOVA) followed by Dunnett's multiple comparisons test. **(G)** poly(GP) is detectable in cerebrospinal fluid (CSF) from c9ALS/FTD (n=5), and SCA36 patients (n=5) but not healthy controls (n=5).

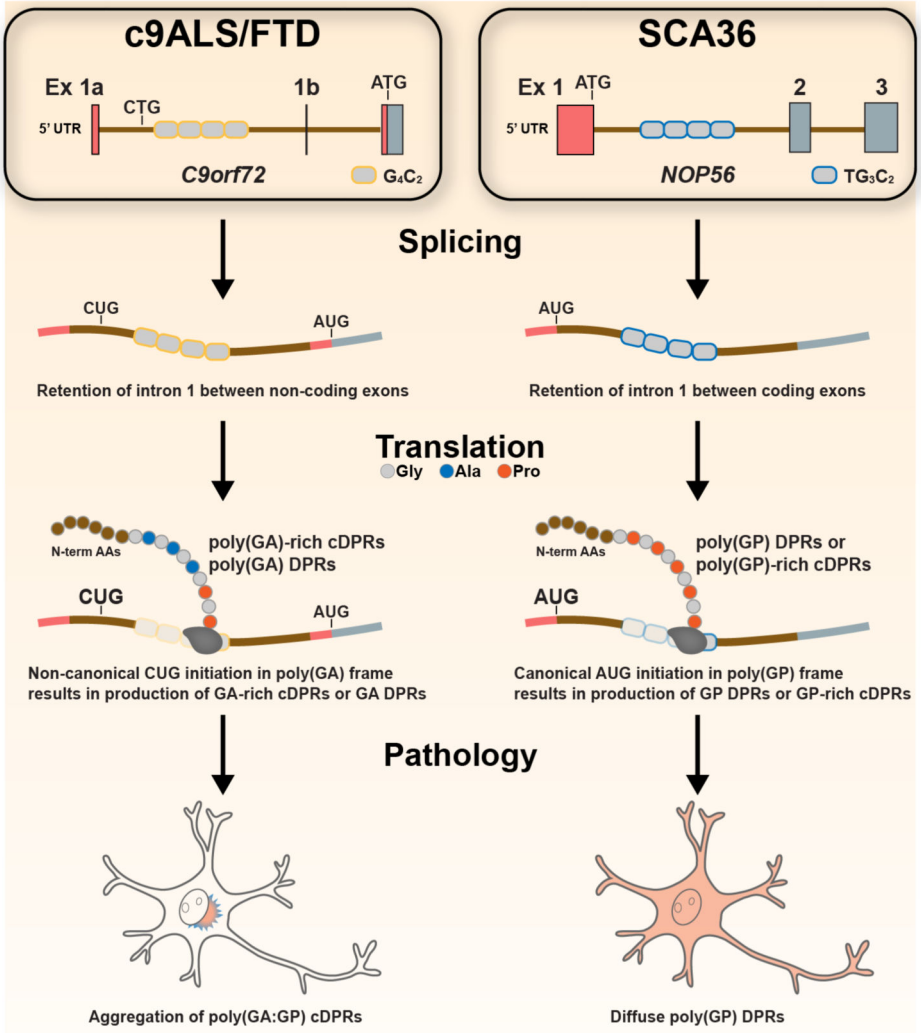


**Figure 7: Model of divergent translation and pathology of dipeptide repeats in c9ALS/FTD and SCA36.**

The c9ALS/FTD-associated $G_4C_2$ and SCA36-associated $TG_3C_2$ repeat expansions are similarly positioned in the first intron of the genes *C9orf72* and *NOP56*, respectively; however, the flanking genetic context of these repeat expansions differ. With respect to *C9orf72*, the 5' intron sequence flanking the $G_4C_2$ repeat contains a near-cognate CUG start codon. In c9ALS/FTD, translational initiation at this near-cognate start codon, either from a retained mRNA transcript variant or a spliced and exported intron 1 RNA species, would preferentially produce a poly(GA) dipeptide repeat (DPR) protein. In contrast, *NOP56* intron 1 is located between two coding exons, thus retention of intron 1 harboring a $TG_3C_2$ expansion would produce a novel *NOP56* transcript variant that would preferentially produce a poly(GP) DPR protein via canonical AUG-mediated translation. Translational frameshifting and/or genetic aberrations in the respective repeat expansions result in chimeric DPR (cDPR) species. In c9ALS/FTD, these chimeric species would results in homogenous poly(GA) DPRs and/or GA-rich poly(GA:GP) cDPRs that result in poly(GA) and poly(GP) inclusions as seen in c9ALS/FTD patient samples. In SCA36, homogenous

poly(GP) DPRs and/or GP-rich cDPRs would result in soluble poly(GP) as seen in SCA36 patient samples.