Translation of dipeptide repeat proteins in \textit{C9orf72} ALS/FTD through unique and redundant AUG initiation codons

Yoshifumi Sonobe $^{1,2,3}$, Soojin Lee $^{4-5}$, Gopinath Krishnan $^{4-5}$, Yuanzheng Gu $^6$, Deborah Y. Kwon $^6$
Fen-Biao Gao $^{4-5}$, Raymond P. Roos $^{1,2,3,*}$, Paschalis Kratsios $^{3,7,*}$

* These authors contributed equally.

Correspondence:
R.P.R (roos@neurology.bsd.uchicago.edu), P.K (pkratsios@uchicago.edu)

Affiliations:
1 University of Chicago Medical Center, Chicago, IL 60637, USA
2 Department of Neurology, University of Chicago Medical Center, Chicago, IL 60637, USA
3 Neuroscience Institute, University of Chicago, Chicago, IL 60637, USA
4 RNA Therapeutics Institute, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA
5 Department of Neurology, University of Massachusetts Chan Medical School, Worcester, MA 01605, USA
6 Neuromuscular & Movement Disorders, Biogen, Cambridge, MA 02142, USA
7 Department of Neurobiology, University of Chicago, Chicago, IL, USA
ABSTRACT

A hexanucleotide repeat expansion in C9ORF72 is the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). A hallmark of ALS/FTD pathology is the presence of dipeptide repeat (DPR) proteins, produced from both sense GGGGCC (poly-GA, poly-GP, poly-GR) and antisense CCCCCG (poly-PR, poly-PG, poly-PA) transcripts. Translation of sense DPRs, such as poly-GA and poly-GR, depends on non-canonical (non-AUG) initiation codons. Here, we provide evidence for canonical AUG-dependent translation of two antisense DPRs, poly-PR and poly-PG. A single AUG is required for synthesis of poly-PR, one of the most toxic DPRs. Unexpectedly, we found redundancy between three AUG codons necessary for poly-PG translation. Further, the eukaryotic translation initiation factor 2D (EIF2D), which was previously implicated in sense DPR synthesis, is not required for AUG-dependent poly-PR or poly-PG translation, suggesting that distinct translation initiation factors control DPR synthesis from sense and antisense transcripts. Our findings on DPR synthesis from the C9ORF72 locus may be broadly applicable to many other nucleotide-repeat expansion disorders.
INTRODUCTION

The hexanucleotide GGGGCC repeat expansion in the first intron of C9ORF72 is the most common monogenic cause of inherited amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) 1, 2. This mutation is predicted to cause ALS/FTD via three non-mutually exclusive mechanisms: (1) a loss-of-function mechanism due to reduced C9ORF72 protein expression 3, 4, 5, 6, (2) a gain-of-function mechanism due to toxicity from repeat-containing sense (GGGGCC) and antisense (CCCGGG) RNA 7, 8, and (3) toxicity from dipeptide repeat (DPR) proteins produced from these transcripts 9, 10, 11, 12. However, loss of C9ORF72 protein by itself does not cause neurodegeneration 13. On the other hand, DPRs produced from both sense (poly-GA, poly-GP, poly-GR) and antisense (poly-PR, poly-PG, poly-PA) transcripts are present in the central nervous system of ALS/FTD patients 14, 15. Strong evidence from experimental model systems suggests DPRs are toxic 16, underscoring the importance of uncovering the molecular mechanisms responsible for DPR synthesis.

To design therapies that reduce DPR levels, it is valuable to identify initiation codons used in DPR translation. To date, the synthesis of sense DPRs has been a major focus in the ALS/FTD field, resulting in the identification of translation initiation codons for poly-GA and poly-GR 17, 18, 19, 20. As previously shown, non-canonical codons (CUG for poly-GA, AGG for poly-GR) initiate DPR synthesis from the sense strand 17, 18, 19, 20, 21. Interestingly, studies in Drosophila and cultured cells showed that the presence of an expanded GGGGCC repeat alone, without flanking sequences, can result in DPR production, suggesting an unconventional form of translation 14. However, deletion analysis of cis-regulatory elements upstream of the GGGGCC repeats and ribosome profiling revealed that translation initiation in the poly-GA and poly-GR frames does depend on flanking intronic sequences surrounding the repeats 21, 22, 23. Moreover, a recent study proposed that a canonical AUG initiation codon is used for poly-PG synthesis from the antisense CCCGGG transcript 19, suggesting conventional translation is involved in the synthesis of at least one DPR. However, the initiation
codons for other DPRs (e.g., poly-PR, poly-PA) from the antisense transcript remain unknown. Hence, it is unclear which mode of translation is utilized for DPR synthesis from the antisense transcript.

Although both sense and antisense transcripts produce GP-containing dipeptides (sense: poly-GP, anti-sense: poly-PG), the antisense transcript seems to be the primary source of poly-PG/poly-GP inclusions in the brain of C9ORF72 ALS/FTD patients. Further, two recent ALS clinical trials that specifically targeted the production of DPRs from the sense transcript failed. Therefore, studying the mechanisms responsible for DPR synthesis from the antisense transcript is important, and this is the focus of the present study.

An additional challenge in ALS/FTD is the identification of regulatory factors necessary for DPR synthesis. Research efforts have uncovered a number of proteins that act at different steps of DPR synthesis: RNA helicases (eIF4A, DHX36, and DDX3X), proteins of the eIF4F complex (eIF4A, eIF4B, eIF4E, eIF4H), small ribosomal protein subunit 25 (RPS25), ribosome quality control protein ZNF598, and eukaryotic translation initiation factors (DAP5, eIF2A, eIF3F, eIF2D, and eIF2D co-factors DENR and MCTS-1). Except for RPS25, all remaining factors have only been assessed for their effects on DPRs produced from the sense GGGGCC transcript. Hence, it remains unknown whether any of these factors is used for DPR synthesis from the antisense transcript. Furthermore, the role of these factors on DPR synthesis in induced pluripotent stem cell (iPSC)-derived neurons from C9ORF72 ALS/FTD patients remains largely untested.

Here, we employ cell-based models of C9ORF72 ALS/FTD to identify translation initiation codons for DPRs produced from the antisense transcript. Transfection into cultured cells of constructs carrying 35 CCCCGG repeats (preceded by 1,000 bp of human intronic C9ORF72 sequence) leads to DPR production (poly-PR, poly-PG) and reduced cell survival. We find that a canonical AUG initiation codon located 273 base pairs (-273 bp) upstream of the CCCCGG repeats is necessary for poly-PR synthesis. Further, we provide evidence for redundancy in usage of canonical initiation codons for poly-PG synthesis. Although an AUG at -194 bp is the main start codon for poly-PG, two
other AUG codons (at -212 bp and at -113 bp) can also function as alternative translation initiation sites. These findings suggest that DPR synthesis from the antisense transcript occurs via AUG-dependent translation, contrasting with the mode of DPR synthesis from the sense transcript, which depends on near-cognate start codons (CUG for poly-GA, AGG for poly-GR). Finally, we show that the translation initiation factor eIF2D, which is necessary for CUG-dependent poly-GA synthesis from the sense transcript 35, is not involved in AUG-dependent antisense DPR (poly-PG, poly-PR) synthesis. Hence, distinct translation initiation sites and factors are employed for DPR synthesis from sense GGGGCC and antisense CCCCGG transcripts.
RESULTS

Transfection of constructs carrying 35 CCCCCG repeats leads to antisense DPR synthesis and reduced cell survival

To study DPR synthesis from the antisense transcript, we engineered three constructs with 35 CCCCCG repeats preceded by 1,000bp-long intronic sequence from human C9ORF72 (Figure 1A) and then followed by nanoluciferase (nLuc) in frame of poly-PR, poly-PG, or poly-PA (see Materials and Methods). Forty eight hours (48h) after transfection of poly-PR::nLuc or poly-PG::nLuc into HEK293 and NSC34 cells, robust expression of poly-PR and poly-PG was detected both in luciferase assays (Figure 1B-C) and Western blotting for poly-PR, poly-PG, and nLuc (Figure 1D-E, Figure 1-figure supplement 1, Figure 1 - source data 1), suggesting the luciferase signal is an accurate readout for DPR production. Protein isolation of soluble and insoluble fractions showed that both DPRs (poly-PG and poly-PR) are predominantly detected in the soluble fraction under these experimental conditions (Figure 1 - figure supplement 2). Further, production of poly-PR and poly-PG in transfected NSC34 cells was confirmed with immunofluorescence staining (Figure 1F-G). Finally, transfection of either poly-PR::nLuc or poly-PG::nLuc into NSC34 cells led to reduced cell survival (Figure 1H-I).

Consistent with a previous study, we did not detect poly-PA with luciferase assays (Figure 1B-C) and Western blotting (Figure 1-figure supplement 3) upon poly-PA::nLuc transfection. We surmise that the initiation codon for poly-PA may lie outside the 1,000 bp intronic sequence used in our construct, or that the specific regulatory machinery needed for poly-PA synthesis is lacking in the cellular context examined here (HEK293 and NSC34 cells). Altogether, our cell-based model of C9ORF72 (construct with 35 CCCCCG repeats and 1,000 bp of human intron) produces two antisense DPRs (poly-PR, poly-PG) and displays reduced cell survival.
A canonical AUG initiation codon located 273 bp upstream of CCCCCCCG repeats is required for poly-PR synthesis

The poly-PR::nLuc and poly-PG::nLuc constructs offer an opportunity to identify the initiation codons for poly-PR and poly-PG synthesis. We initially focused on poly-PR, one of the most toxic DPRs based on in vitro \(^{11, 37, 38}\) and in vivo studies in worms\(^ {39}\), flies \(^ {12, 37, 40}\) and mice \(^ {40, 41, 42}\). Using our recently developed machine-learning algorithm for initiation codon prediction \(^ {43}\), we identified a CUG at -366bp (Kozak sequence: guaCUGa) and an AUG at -273bp (Kozak sequence: cggAUGc) as putative initiation codons for poly-PR (Figure 2A). We then mutated these codons either to CCC or the termination codon UAG (Figure 2A). Western blotting and luciferase assays showed that mutation of the CUG at -366bp to CCC or UAG did not affect poly-PR expression (Figure 2B-G, Figure 2-source data 1). However, mutation of the AUG at -273bp to CCC or UAG completely abolished poly-PR expression both in HEK293 and NSC34 cells, as shown by Western blotting (Figure 2B-E), luciferase assays (Figure 2F-G), and immunofluorescence staining against poly-PR (Figure 2H). Importantly, the reduced survival of NSC34 cells upon poly-PR::nLuc transfection was partially rescued when the -273bp AUG codon was mutated into the UAG termination codon, suggesting poly-PR production is toxic under these experimental conditions (Figure 2I). These results strongly suggest that the AUG at -273 bp is the start codon for translation of poly-PR, one of the most toxic DPRs in C9ORF72 ALS/FTD. This AUG is predicted to be included in the endogenous antisense CCCCCCCG transcript based on 5’ Rapid Amplification of cDNA Ends (RACE) analysis on brain samples of C9ORF72 ALS/FTD patients\(^ {14}\).

Evidence for redundancy of AUG initiation codon usage in poly-PG translation

We next investigated poly-PG, which is less toxic than poly-PR \(^ {12, 37, 44, 45}\), and has been proposed as a biomarker for C9ORF72-ALS/FTD \(^ {46, 47}\). Using the same machine-learning algorithm \(^ {43}\), we identified four putative initiation codons (AUG at -212bp, AUG at -194bp, CUG at -182bp, AUG at -113bp)
(Figure 3A), all with relatively good Kozak sequences (gaaAUGa at -212bp, aaaAUGc at -194bp, gctCUGa at -182bp, aggAUGc at -113bp). Of note, a prior publication previously identified the AUG at -194bp as an initiation codon \(^{19}\). Simultaneous mutation of all four of these codons to CCC completely blocked poly-PG expression (Figure 3B-D, Figure 3 - source data 1), suggesting one or more of these codons is required. Next, we simultaneously mutated three codons to CCC, but left intact the AUG at -212bp. We refer to this construct as “-212 AUG”. Upon transfection of -212 AUG, we observed poly-PG expression, suggesting poly-PG translation can start at the AUG at -212bp. Intriguingly, when we followed a similar approach to mutate 3 codons to CCC but leave intact the AUG at -194bp or at -113bp, we also observed poly-PG production, but this time at an expected lower molecular weight (Figure 3B-D, Figure 3 - source data 1). Of note, when we mutated to CCC all three AUG codons (-212bp, -194bp, -113bp) but left intact the CUG at -182bp, we observed no poly-PG expression (Figure 3B-D, Figure 3 - source data 1). These results suggest that any of these three AUGs, but not the CUG at -182bp, can function as a start codon for poly-PG, indicating redundancy in the translation initiation codon for poly-PG.

We observed a strong (higher molecular weight) band and a fainter (lower molecular weight) band for poly-PG when the intact version of the poly-PG::NanoLuc plasmid was translated (Figure 3B, Figure 3 – figure supplement 1, Figure 3 - source data 1). The strong band is likely to result from translation initiation at the AUG at -194bp, whereas the faint band is likely initiated at the AUG at -113bp (Figure 3B). Hence, the AUG at -194bp appears to be the main initiation codon for poly-PG synthesis from the antisense transcript of 35 C4G2 repeats (Figure 3B), which is consistent with mass-spectrometry results from a previous report \(^{19}\).

Interestingly, selective mutation of the AUG at -194 to CCC did not abolish poly-PG expression (Figure 4A-D, Figure 4 – figure supplement 1). Instead, it led to the production of two poly-PG products: a high molecular weight product (strong band) resulting from use of the AUG at -212bp as well as a lower molecular weight product (faint band) resulting from AUG at -113bp (Figure 4A-D, Figure 4 – figure supplement 1).
4B, Figure 4 - source data 1). Altogether, these results suggest that the AUG at -194bp is mainly used for poly-PG expression from antisense C4G2 repeats. However, when this AUG is mutated, two other AUG codons (at -212bp and -113bp) can also function as translation initiation sites, again revealing redundancy in the start codon usage for poly-PG synthesis.

Mutation of the -113 AUG abolishes poly-PG production

We further corroborated redundant initiation for poly-PG translation by separately mutating each of the AUG codons to a termination UAG codon (Figure 5A-D, Figure 5 – figure supplement 1, Figure 5 - source data 1). Mutation of the AUG at -212bp to UAG (construct name: -212 UAG) did not affect poly-PG expression, most likely because the AUG at -194bp became the start codon as shown by Western blotting (Figure 5B-D, Figure 5 - source data 1). Similarly, mutation of the AUG at -194bp to UAG (construct name: -194 UAG) did not affect poly-PG expression because the AUG at -113bp became the start codon (Figure 5B-D). However, mutation of AUG at -113bp to UAG (construct name: -113 UAG) completely blocked poly-PG expression, as shown by Western blotting (Figure 5B, Figure 5 – figure supplement 1), luciferase assays (Figure 5C-D) and immunofluorescence staining (Figure 5E). Finally, the reduced survival of NSC34 cells was not rescued upon transfection of the -113 UAG construct, suggesting poly-PG production is not toxic under these experimental conditions (Figure 5F).

Altogether, these findings strongly suggest that the AUG at -194bp is primarily used for poly-GP translation, but the other two AUG codons at -212bp and -113bp can also function as translation initiation sites under certain experimental conditions.

EIF2D does not control poly-PR and poly-PG synthesis from the antisense transcript

Following the identification of AUG codons for translation initiation of poly-PR and poly-PG, we next sought to identify translation initiation factors necessary for synthesis of these antisense DPRs. We
focused on EIF2D because we previously found it to be necessary for poly-GA synthesis from the sense transcript in *C. elegans* and cell-based models (HEK293 and NSC34 cell lines)\(^{35}\). To this end, we generated an *EIF2D* knockout HEK293 line using CRISPR/Cas9 gene editing (see Materials and Methods) (**Figure 6A-C, Figure 6 - source data 1**). Next, we transfected the poly-PR::nLuc reporter construct into control and *EIF2D* knockout HEK293 cells. We found that knockout of *EIF2D* did not affect the expression levels of the poly-PR::nLuc reporter (**Figure 6E**). We obtained similar results upon knockdown of *EIF2D* with a short hairpin RNA (shRNA) (**Figure 6H**), again suggesting that eIF2D is not required for poly-PR synthesis from antisense CCCGG transcripts. Lastly, knock-out or knock-down (shRNA) of *EIF2D* in HEK293 cells transfected with poly-PG::nLuc did not decrease poly-PG expression based on a luciferase assay (**Figure 6D, G**). Hence, knock-out or knock-down of *EIF2D* does not affect the production of two antisense DPR (poly-PR, poly-PG). On the other hand, knock-down of *EIF2D* did reduce the levels of poly-GA (**Figure 6I**), a DPR generated from sense RNA. The poly-GA reduction is consistent with our previous observations in a *C. elegans* model of *C9ORF72* ALS/FTD\(^{35}\), albeit more modest - likely due to a technical reason (see legend of **Figure 6I**).

**Knockdown of EIF2D in human iPSC-derived motor neurons**

We next tested whether EIF2D is required for DPR synthesis in a cellular context that maintains the endogenous human *C9ORF72* gene locus. We initially used one published iPSC line from a *C9ORF72* carrier (line 26#6), as well as an isogenic control line (26Z90) which had CRISPR/Cas9-mediated deletion of expanded GGGGCC repeats\(^{48}\). The iPSC lines were differentiated into motor neurons as previously described\(^{49}\). Repeated transfection of a small interfering RNA (siRNA) against *EIF2D* (*EIF2D*-siRNA-1), but not of a control scrambled siRNA, resulted in robust downregulation of *EIF2D* mRNA as assessed by RT-PCR (**Figure 7A**) and eIF2D protein analysis (**Figure 7 – figure supplement 1**). The mRNA levels of eIF2A, a related initiation factor, remained unaltered, suggesting specificity in the siRNA effect. Despite this knockdown, an immunoassay (conducted in a blinded
manner) failed to show any differences in the steady-state levels of soluble poly-PG (Figure 7B), suggesting eIF2D is not necessary for poly-PG translation from the antisense transcript. We caution though that our immunoassay does not distinguish between poly-PG produced from the antisense transcript and poly-GP from the sense transcript (Figure 7B). Hence, a mild effect upon EIF2D knockdown on poly-PG (from antisense transcript) can potentially be masked by poly-GP (from sense transcript). Of note, PG/GP inclusions in brain tissue of C9ORF72 ALS/FTD patients contain ~80% of poly-PG from the antisense transcript and ~20% of poly-GP from the sense transcript\(^\text{14}\). However, other studies indicate that the exact contribution of sense poly-GP and antisense poly-PG C9ORF72 ALS/FTD has not been resolved\(^\text{25, 26, 46}\). Hence, our data hint that eIF2D may not affect poly-PG synthesis from the antisense CCCCGG transcript.

Despite the lack of an effect on poly-PG/GP, we found that EIF2D knockdown reduced poly-GA synthesis from the sense GGGGCC transcript in neurons derived from iPSC line 26#6 (Figure 7B), critically extending previous observations made in C. elegans and cell-based models \(^\text{35}\). Consistent with the latter study, EIF2D knockdown had no effect on poly-GR synthesis from the sense transcript based on an immunoassay that measures soluble poly-GR (Figure 7B). Altogether, these findings from one patient line (26#6) suggest that eIF2D is required for CUG start codon-dependent poly-GA synthesis from the sense transcript in human iPSC-derived neurons, but is dispensable for poly-GA (from sense transcript) and poly-PG synthesis, albeit our immunoassay cannot distinguish between poly-PG and poly-GP. However, when we repeated this experiment with two additional iPSC lines (27#11 and 40#3) from C9ORF72 carriers with two siRNAs (EIF2D-siRNA-1 and 2), we did not achieve robust EIF2D knockdown (Figure 7C-D). We note that the same siRNA (EIF2D- siRNA-1) led to robust EIF2D knockdown in the first patient line (26#6) (compare Figure 7A with 7C-D). Hence, the issue of variable siRNA knockdown efficiency prevents us from drawing any general conclusions on the role of EIF2D in DPR synthesis in the context of motor neurons derived from different iPSC lines of C9ORF72 carriers (Figure 7B, E).
**DISCUSSION**

Here, we show that canonical AUG codons on the antisense CCCCCG transcript serve as translation initiation codons for two DPRs - poly-PR and poly-PG. This finding may inform the design of future therapies for ALS/FTD, especially since poly-PR is a highly toxic DPR and poly-PG is thought to be primarily translated from the antisense transcript. Our finding of canonical AUG codons serving as translation initiation codons for antisense DPRs (poly-PR, poly-PG) differs from the proposed mode of translation of sense DPRs (e.g., poly-GA, poly-GR). In the latter case, it is thought that repeat-associated non-AUG (RAN) translation of poly-GA and poly-GR occurs via non-canonical CUG and AGG initiation codons, respectively, located in the intronic sequence upstream of the GGGGCC repeats. Interestingly, studies in *Drosophila* and cultured cells showed that the presence of an expanded GGGGCC repeat alone, without flanking sequences, can result in DPR production. Hence, our findings together with these previous studies suggest that DPR synthesis may involve at least three different modes of translation: (a) near-cognate start codon (e.g., CUG, AGG) dependent-translation for poly-GA and poly-GR from sense GGGGCC transcripts, (b) canonical AUG-dependent translation for poly-PR and poly-PG synthesis from antisense CCCCCG transcripts, and (c) DPR synthesis may also occur through RAN translation mechanisms that solely utilize the repeat. It is conceivable that all three modes of translation may occur simultaneously in disease, and that the use of non-canonical and canonical initiation codons may be the primary contributors of DPR production.

A notable finding is the presence of redundancy in start codon usage for poly-PG synthesis. Our data suggest that the AUG at -194bp is primarily used for poly-GP translation from antisense CCCCCG transcripts, consistent with a previous investigation. However, when this AUG is mutated, two other canonical AUG codons (at -212bp and -113bp) can also function as translation initiation sites under the experimental conditions described herein. Although it is unclear whether such redundancy in DPR translation initiation occurs in the central nervous system of *C9ORF72* ALS/FTD patients, these findings nevertheless suggest that targeting only one translation initiation site may be
insufficient to prevent poly-PG synthesis. Redundancy in start codon usage may also apply to other DPRs, such as poly-PR synthesis from the antisense transcript. Although we identified an AUG at -273 bp as necessary for poly-PR synthesis, a previous study detected poly-PR when only 100bp downstream of the GGGGCC repeats were included in an adeno-associated viral (AAV) vector\textsuperscript{50}. It is important to note that this intronic 100bp-long sequence was placed next to a 589 bp regulatory element of the woodchuck hepatitis virus (WPRE), which contains several putative start codons. The AUG initiation codons we identified as necessary for either poly-PR or poly-PG synthesis are predicted to be included in the endogenous antisense CCCCGG transcript based on 5’ Rapid Amplification of cDNA Ends (RACE) analysis on brain samples of C9ORF72 ALS/FTD patients\textsuperscript{14}. Nevertheless, endogenous mutagenesis of these codons - in the native genomic context of the C9ORF72 locus – is needed in the future to further test the validity of our findings.

Emerging evidence suggests distinct proteins affect translation initiation of DPRs from sense and antisense transcripts in C9ORF72 ALS/FTD. For example, the RNA helicase DDX3X directly binds to sense (GGGGCC), but not antisense (CCCCCGG) transcripts, thereby selectively repressing the production of sense DPRs (poly-GA, poly-GP, poly-GR)\textsuperscript{28}. Here, we provide evidence that the translation initiation factor EIF2D is not involved in DPR (viz., poly-PG, poly-PR) synthesis from antisense (CCCCCGG) transcripts. In a previous study\textsuperscript{35}, we showed in C. elegans and in vitro cellular systems (HEK293 and NSC34 cells) that EIF2D is required for poly-GA production from sense (GGGGCC) transcripts. These findings are important because they indicate that not only distinct translation initiation codons, but also different regulatory proteins are involved in DPR synthesis from sense and antisense transcripts, perhaps a reflection that different modes of DPR translation (e.g., RAN translation, AUG-dependent translation) occur simultaneously in C9ORF72 ASL/FTD. Consistent with this idea, translation initiation is the most heavily regulated step in protein synthesis because it is the rate-limiting step\textsuperscript{51}. Hence, we favor a model where distinct regulatory factors are necessary for translation initiation of different DPRs. In striking contrast, the transcriptional control of sense and
antisense transcripts appears coordinated. For example, a single protein – the transcription elongation
factor Spt4 – controls production of both sense and antisense transcripts\textsuperscript{52}.

In addition to \textit{C9ORF72} ALS/FTD, nucleotide repeat expansions are present in various genes, causing more than 30 neurological diseases\textsuperscript{53, 54}. In many of these, products translated from the expanded repeat sequences have been detected in the nervous system of affected individuals. Hence, our findings may also apply to this large group of genetic disorders in the following ways. First, translation of peptides from the same nucleotide repeat expansion may require different modes of translation (RAN- and AUG-dependent translation), as previously proposed\textsuperscript{55}. Second, the surprising redundancy in canonical AUG initiation codon usage for DPR (poly-PG) synthesis may also apply to proteins translated from nucleotide repeat expansions in other genes. Lastly, our results support the idea that distinct translation initiation factors are involved in the synthesis of individual DPRs produced from the same nucleotide repeat expansion. Future studies focused on transcriptional and translational mechanisms of expanded nucleotide repeats may critically contribute to the design of therapies for these diseases.
Acknowledgements

This work was supported by a grant from the Lohengrin Foundation (R.P.R, P.K), a basic science pilot grant from the Association for Frontotemporal Degeneration (AFTD) (R.P.R, P.K), and two NIH grants (R37NS057553 and R01NS101986) to F.B.G. The antibodies used to measure GA levels were discovered by Neurimmune AG (Zurich, Switzerland).

Competing Interests

The authors declare no competing or financial interests.

Data availability

All data generated or analyzed during this study are included in the manuscript and supporting files.
Materials and Methods

Generation of the plasmid constructs

All oligonucleotides were obtained from Integrated DNA Technologies. Oligonucleotide I-F/R (Supplementary file 1) contains part of a HindIII site followed by 113 nucleotides that are normally upstream of the G4C2 repeats and then by three G₄C₂ repeats. Oligonucleotide II-F/R contains 10 G₄C₂ repeats followed by part of a NotI site. These two oligonucleotides were phosphorylated, annealed, and then ligated into restriction sites of HindIII and NotI of a pAG plasmid. The plasmid was then digested with HindIII and BamHI. The HindIII-BamHI fragment was digested with BanII, and the resultant HindIII-BanII fragment was then ligated with oligonucleotide II-F/R into the pAG plasmid. This approach was repeated three times with similar digestions and ligations of oligonucleotide II. Finally, the HindIII-BanII fragment was ligated with oligonucleotide III-F/R (which contains 2 G₄C₂ repeats followed by a 99 bp flanking sequence and then followed by part of the NotI site) into the pAG plasmid (referred to as 113bp-35RG4C2-99bp plasmid). To delete stop codons after the C₄G₂ repeats, the plasmid was treated with BfaI and NotI, and the digested fragment was ligated with oligonucleotide IV-F/R. To add sequence upstream from the C₄G₂ repeats, a 543 bp portion (408-950 of NCBI reference sequence, NC_000009.12) of the C₉ORF72 gene from HEK293 genomic DNA was amplified by PCR using the primer shown in Supplementary file 1. The amplified construct was then ligated with the BtgI/NotI-digested fragment of the 113bp-35RG4C2-99bp plasmid into XbaI and NotI sites of pcDNA6/V5-His A plasmid (referred to as 609bp-35RC4G2 plasmid). To further increase the length of sequence upstream from C₄G₂ repeats, a 392 bp portion (951-1342 of NCBI reference sequence, NC_000009.12) of C₉ORF72 gene from HEK293 genomic DNA was amplified by PCR using the primer shown in Supplementary file 1. The amplified construct was then ligated with the XbaI/NotI fragment of 609bp-35RC4G2 plasmid into HindIII and NotI sites of the pAG plasmid (referred to as AS-C9 plasmid). The ΔC9 plasmid was generated as previously described.
To mutate sequences, a 560bp portion upstream from the repeats in the AS-C9 plasmid was amplified by PCR using a primer shown in Supplementary file 1. The amplified portion was then ligated into the HindIII and NotI sites of pcDNA6/V5-His A plasmid. Mutations were made with Q5® Site-Directed Mutagenesis Kit (New England Biolabs) using primer sets (Supplementary file 1). The Stul/BtgI portion of the resultant mutants was then cloned back into the StuI and NotI sites of AS-C9 plasmid with BtgI/NotI portion of AS-C9 plasmid using the primer sets in Supplementary file 1.

To generate the vector to induce expression of poly-PA, the fragment AUG-PA-F/R (Supplementary file 1) were phosphorylated, annealed, and then ligated into restriction sites of HindIII and BtgI of the AS-C9 plasmid.

**Cell culture**

HEK293 and NSC34 cells were cultured in DMEM supplemented with 10% FBS, 2 mM L-Glutamine, 100 U/ml Penicillin and 100 μg/ml Streptomycin. The cell lines were checked for mycoplasma contamination by DAPI staining but were not authenticated.

**Luciferase Assay**

The cells were plated in 24-well plates at $5 \times 10^4$ per well and then cotransfected using Lipofectamine LTX (Thermo Fisher Scientific) with 100 ng of the plasmid along with 100 ng fLuc plasmid as a transfection control. After 48h, the cells were lysed with $1\times$ passive lysis buffer (Promega). Levels of nLuc and fLuc were assessed with the Nano-Glo Dual-Luciferase Reporter assay system (Promega) and a Wallac 1420 VICTOR 3V luminometer (Perkin Elmer) according to the manufacturer’s protocol.

**Western blotting**
The cells were plated in 6-well plates at $2 \times 10^5$ per well and then cotransfected with 2.5 μg of plasmids using Lipofectamine LTX (ThermoFisher Scientific). After 48h, cell lysates were prepared using RIPA buffer (50 mM Tris-HCl, pH 7.5; 150 mM NaCl; 0.1% SDS; 0.5% sodium deoxycholate; 5 mM EDTA containing 1 × HaltTM Protease inhibitor Cocktail). The RIPA-insoluble pellet was lysed in 8M urea and used as the RIPA-insoluble fraction. H3K4me2 was used as marker for RIPA-insoluble fraction, as previously described. Lysates were subjected to electrophoresis on Mini-PROTEAN TGX Gels (BIO-RAD), and then transferred to Amersham Hybond P 0.45 μm PVDF membranes (GE Healthcare). The membrane was blocked with 5% non-fat skim milk in Tris-buffered saline containing 0.05% Tween-20 for 1 h at room temperature, and then incubated overnight at 4 °C with primary antibodies against poly-PR (1:1000, ABN1354, EMD Millipore), poly-GP (1:1000, TALS 828.179, Target ALS), eIF2D (1:1000, 12840-1-AP, Proteintech), poly-PA (1:1000, ABN1356, EMD Millipore), nLuc (1:500, N700A, Promega), α-tubulin (1:5000, YL1/2, Abcam), and dimethyl-Histone H3 (H3K4me2 ) (1:2000, 07-030, EMD Millipore). Following washing, the membrane was incubated for 1 h at room temperature with anti-mouse (1:5000, GE Healthcare), anti-rabbit (1:5000, GE Healthcare), or anti-rat horseradish peroxidase–conjugated secondary antibodies (1:1000, Cell Signaling Technology). The signal was detected using SuperSignal West Dura Extended Duration Substrate (ThermoFisher Scientific) and analyzed using ChemiDoc MP Imaging System and Image Lab software (version 6.0.1, Bio-Rad).

**Cell viability assay**

Cell viability assay was performed using Cell counting kit-8 (Dojindo) according to the manufacturer’s protocol. In brief, NSC34 cells were plated in 96-well plates at $2.5 \times 10^3$ per well and then transfected using Lipofectamine LTX with 100 ng of the indicated plasmid. After 48h, 10 μl of the CCK-8
solution was added to the well and incubated for 2h in a CO₂ incubator. The reaction was stopped by adding 0.1M HCl and the absorbance at 450 nm was measured.

Immunocytochemistry

The cells were plated in 4 well Lab-Tek™ II Chamber Slide (Nunc) coated with 50 μg/ml Poly-D-Lysine (Sigma) at 5 x 10⁴ per well and transfected using Lipofectamine LTX with 500 ng of the indicated plasmid. After 48h, the cells were fixed with 4% paraformaldehyde for 15 min at room temperature. Then, the cells were permeabilized with phosphate buffered saline (PBS) with 0.2% Tween-20 for 20 min at room temperature. The samples were incubated with blocking buffer (2% BSA in PBS) for 1h at room temperature and then incubated overnight at 4 °C with antibodies against poly-PR (1:250, ABN1354, EMD Millipore) or poly-GP (1:100, TALS 828.179, Target ALS). After rinsing with PBS, cells were incubated with Alexa 488-conjugated chicken anti-mouse IgG (1:2000, Thermo Fisher Scientific) or Alexa 488-conjugated goat anti-rabbit IgG (1:2000, Thermo Fisher Scientific) for 1h at room temperature, and then counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Images were captured using a confocal laser microscope system (Leica TCS SP5, Leica Microsystems) and processed using ImageJ2 software (version 2.9.0/1.53t).

Generation of EIF2D knockout cells by CRISPR/Cas9 gene editing

A single guide RNA (sgRNA) (GCAGTGACTGTGTACGTGAG) that targets exon 2 of eIF2D was cloned into lentiCRISPR v2 plasmid (Addgene). HEK293 cells were plated into 6-well plates at 4 x 10⁵ cells per well, and then transfected using Lipofectamine LTX with 2.5 μg lentiCRISPR v2 plasmids containing the sgRNA sequence. Transfected cells were selected using 3 μg/ml puromycin for 3 days. EIF2D knockout cell clones were obtained by limited dilution. The resulting EIF2D knockout cells carry allele-specific mutations, as follows. Compared to the WT
GGATGCAGTGACTGTGTACGTGAGTGGTGG sequence, one allele
GGATGCAGTGACTGTGTACGTTAGTGGTGAGTGGTGG has a single nucleotide insertion shown bolded
while the other allele contains a two-nucleotide deletion GGATGCAGTGACTGTGTA—
TGAGTGGTGG. Both alleles lead to a premature stop codon, likely resulting in two different
truncated eIF2D proteins with the following respective sequence:

MFAKAFRVKSKNTAIGSDRRKLRADVTTAFPTLGTQVSELVPGKEELNIVKLYAHKGDAVT
VYEW and MFAKAFRVKSKNTAIGSDRRKLRADVTTAFPTLGTQVSELVPGKEELNIVKLY
AHKGDAVT
VYEWW.

Knockdown of eIF2D in HEK293 cells
shRNA plasmids against human eIF2D were prepared using previously published methods. In brief,
oligonucleotides with an siRNA sequence were cloned into the BamHI and HindIII sites of pSilencer
2.1-U6 neo Vector (ThermoFisher Scientific) according to the manufacturer’s protocol. The latter kit
also contained a control shRNA vector. For luciferase assays (shown above), the cells were plated in
24-well plates at $5 \times 10^4$ per well and cotransfected with 50 ng of the AS-C9 plasmids and 50 ng of the
fLuc plasmids along with 500 ng of either control shRNA or anti-eIF2D shRNA using Lipofectamine
LTX (ThermoFisher Scientific).

Motor Neuron Differentiation from human iPSC lines
Human motor neurons were differentiated as previously described from a published iPSC line obtained
from a $C9ORF72$ carrier (FTD26-6), as well as an isogenic control line that had a CRISPR/Cas9-
mediated deletion of expanded GGGGCC repeats. Briefly, iPSCs were plated and expanded in
mTSER1 medium (Stem Cell Technologies) in Matrigel-coated wells. Twenty-four hours after plating,
the culture medium was replaced every other day with neuroepithelial progenitor (NEP) medium,
DMEM/F12 (Gibco), neurobasal medium (Gibco) at 1:1, 0.5X N2 (Gibco), 0.5X B27 (Gibco), 0.1 mM
ascorbic acid (Sigma), 1X Glutamax (Invitrogen), 3 μM CHIR99021 (Tocris Bioscience), 2 μM DMH1 (Tocris Bioscience), and 2 μM SB431542 (Stemgent) for 6 days. NEPs were dissociated with accutase, split 1:6 into Matrigel-coated wells, and then cultured for 6 days in motor neuron progenitor induction medium (NEP with 0.1 μM retinoic acid and 0.5 μM purmorphamine, both from Stemgent). Motor neuron progenitors were dissociated with accutase to generate suspension cultures, and the cells were cultured in motor neuron differentiation medium (NEP with 0.5 μM retinoic acid and 0.1 μM purmorphamine). After 6 days, the cultures were dissociated into single cells, and seeded on Matrigel-coated plates in motor neuron medium, 0.5X B27 supplement, 0.1 mM ascorbic acid, 1X Glutamax, 0.1 μM Compound E (Calbiochem), 0.26 μg/ml cAMP, 1 μg/ml Laminin (Sigma), 10 ng/ml GDNF (R&D Systems), and 10 ng/ml GDNF (R&D Systems), and 10 ng/ml BDNF. Motor neurons were cultured for 5 weeks.

**SiRNA Knockdown**

After 3 weeks in neuron culture media, motor neurons were transfected with a siRNA specific to *eIF2D* mRNA or a scrambled control. For the transfection, lipofectamine RNAiMAX (ThermoFisher Scientific) was first diluted in Opti-MEM medium, and then both eIF2D and scrambled control siRNAs were separately diluted in Opti-MEM medium at room temperature. Diluted siRNA and diluted lipofectamine RNAiMAX (1:1 ratio) were then mixed and incubated for 20 min. The siRNA-lipid complex solution was then brought up to the appropriate volume with MN culture medium. The culture medium in the plate was aspirated and replaced with a siRNA-lipid complex at a final concentration of 60 pmol siRNA in 1.5 ml medium per 1,000,000 cells. After 24 hours, the medium was replaced with a normal motor neuron medium. This process was repeated two more times at 26 and 31 days in culture. After 36 days in culture, we measured siRNA efficiency and levels of DPRs in harvested motor neurons.

**RNA Extraction and Quantitative Real-time PCR**
Total RNA from iPSC-derived motor neurons was extracted with the RNeasy Mini Kit (Qiagen) and then reverse transcribed to cDNA with the TaqMan Reverse Transcription Kit (Applied Biosystems). Quantitative PCR was carried out with SYBR Green Master Mix (Applied Biosystems). Using primers listed in SI Appendix, Table, Ct values for each gene were normalized to actin and GAPDH. Relative mRNA expression was calculated with the double delta Ct method.

Measurement of soluble poly-GR and poly-GP in iPSC-derived neurons

Soluble poly-GR and poly-GP levels in iPSC-derived neurons were detected using the Meso Scale Discovery (MSD) Immunoassay platform as previously reported\textsuperscript{26}. In brief, cells were lysed using Tris based lysis buffer, and lysates were adjusted to equal concentrations and loaded in duplicate wells. Background subtracted electrochemiluminescence (ECL) signals were presented as percentage. The MSD assays were performed in a blinded manner.

Soluble and insoluble fractionation for measurement of poly-GA

Motor neurons were lysed in RIPA buffer (Boston BioProducts, BP-115D) with protease and phosphatase inhibitors. The lysates were rotated for 30 min at 4 °C, followed by centrifugation at 13,500 rpm for 20 min. The supernatant was removed and used as the soluble fraction. Protein concentrations of the soluble fraction were determined by the BCA assay (Thermo Fisher Scientific, Cat # 23227). To remove carryovers, the pellets were washed with RIPA buffer, and then resuspended in the same buffer with 2% SDS followed by sonication on ice. The lysates were rotated for 30 min at 4°C, then spun at 14,800 rpm for 20 min at 4°C. The supernatant was removed and used as insoluble fraction. Protein concentrations of the insoluble fraction were determined by Pierce™ 660 nm Protein Assay (Thermo Fisher Scientific, 22660).

Measurement of poly-GA in iPSC-derived neurons
Poly-GA in soluble motor neuron lysates was measured using a Meso Scale Discovery sandwich immunoassay. A human/murine chimeric form of anti-GA antibody chGA3 was used for capture, and a human anti-GA antibody GA4 with a SULFO-tagged anti-human secondary antibody was used for detection. Poly-GA concentrations were interpolated from the standard curve using 60X-GA expressed in HEK 293 cells and presented as percentage. For background correction, values from no-repeats neuron samples were subtracted from the corresponding test samples.

Statistical analysis

Statistical analysis was performed by one-way ANOVA with Tukey’s multiple comparison test and two-way ANOVA with the Šidák multiple comparison test using GraphPad Prism version 9.3.1. A $P$-value of $<0.05$ was considered significant. The data are presented as mean $\pm$ standard error of the mean.


34. Ayhan, F. et al. SCA8 RAN polySer protein preferentially accumulates in white matter regions and is regulated by eIF3F. *EMBO J* 37, (2018).


FIGURE LEGENDS

Figure 1. Poly-PR and poly-PG are translated from antisense CCCC GG repeats.

(A) Schematic diagram of the constructs with 35 CCCC GG repeats preceded by 1000bp-long intronic sequence from human C9ORF72, and then followed by nanoluciferase (nLuc). (B-C) (B) HEK293 and (C) NSC34 cells were cotransfected with fLuc along with either ΔC9 or AS-C9 plasmids. The levels of luciferase activity were assessed by dual luciferase assays (mean ± s.e.m.). The experiments were repeated 4 times. One-way ANOVA with Tukey’s multiple comparison test was performed. (D-E) HEK293 and NSC34 cells were transfected with either ΔC9 or AS-C9 plasmids. Cell lysates were processed for Western blotting, and immunostained with antibodies to (D) poly-PR, (E) poly-PG, and α-tubulin. (F-G) NSC34 cells transfected with either ΔC9, (F) poly-PR::nLuc, or (G) poly-PG::nLuc were stained with a nuclear marker (DAPI: blue) and with antibodies against poly-PR (F: green) or poly-PG (G: green). Scale bars show 20 μm. (H-I) NSC34 cells were transfected with either ΔC9, (H) poly-PR::nLuc, or (I) poly-PG::nLuc plasmids. WST-8 assay was performed to assess the cell viability. The experiments were repeated 5 times. Unpaired t test was performed.

Figure 2. An AUG at -273bp position is the start codon for poly-PR translation.

(A) Schematic diagram showing mutants with changes in the putative start codons for poly-PR. (B-E) HEK293 (B-C) and NSC34 (D-E) cells were transfected with indicated plasmids. Cell lysates were processed for Western blotting, and immunostained with antibodies to poly-PR and α-tubulin. (B, D) Representative blots are shown. (C, E) The signal intensity of the bands were quantified (mean ± s.e.m.). The experiments were repeated 4 times. One-way ANOVA with Tukey’s multiple comparison test was performed. (F-G) (F) HEK293 and (G) NSC34 cells were cotransfected with the plasmids along with fLuc. The levels of luciferase activity were assessed by dual luciferase assays (mean ± s.e.m.). The experiments were repeated 4 times. One-way ANOVA with Tukey’s multiple comparison
test was performed. (H) NSC34 cells transfected with either ΔC9, poly-PR::nLuc, or -273 AUG -> UAG plasmids were stained against nucleus (DAPI: blue) and poly-PR (green). Scale bars show 20 μm. (I) NSC34 cells were transfected with either ΔC9, wild type (WT), or -273 AUG -> UAG plasmids. WST-8 assay was performed to assess the cell viability. The experiments were repeated 5 times. One-way ANOVA with Tukey’s multiple comparison test was performed. In ΔC9 and WT, the same datasets as Fig. 1H were used (mean ± s.e.m.). The experiments were repeated 5 times. One-way ANOVA with Tukey’s multiple comparison test was performed.

Figure 3. Mutation of AUG codons to CCC fails to suppress poly-PG translation.

(A) Schematic diagram showing mutants with changes in the putative start codons for poly-PG. (B) HEK293 and NSC34 cells were transfected with indicated plasmids. Cell lysates were processed for Western blotting, and immunostained with antibodies to poly-PG and α-tubulin. (C-D) (C) HEK293 and (D) NSC34 cells were cotransfected with fLuc plasmid along with other indicated plasmids. The level of luciferase activity was assessed by dual luciferase assay (mean ± s.e.m.). The experiments were repeated 4 times. One-way ANOVA with Tukey’s multiple comparison test was performed.

Figure 4. An AUG at -194bp position is the primary start codon for poly-PG translation.

(A) Schematic diagram of the constructs. (B) HEK293 and NSC34 cells were transfected with indicated plasmids. Cell lysates were processed for Western blotting, and immunostained with antibodies to poly-PG and α-tubulin. (C, D) (C) HEK293 and (D) NSC34 cells were cotransfected with fLuc plasmid along with indicated plasmids. The level of luciferase activity was assessed by dual luciferase assays (mean ± s.e.m.). The experiments were repeated 4 times. One-way ANOVA with Tukey’s multiple comparison test was performed.
Figure 5. Redundancy of start codon usage in poly-PG translation. (A) Schematic diagram of the constructs. (B) HEK293 and NSC34 cells were transfected with indicated plasmids. Cell lysates were processed for Western blotting, and immunostained with antibodies to poly-PG and α-tubulin. (C, D) HEK293 and (D) NSC34 cells were cotransfected with fLuc plasmid along with indicated plasmids. The level of luciferase activity was assessed by dual luciferase assays (mean ± s.e.m.). The experiments were repeated 4 times. One-way ANOVA with Tukey’s multiple comparison test was performed. (E) NSC34 cells transfected with indicated plasmids were stained against nucleus (DAPI: blue) and poly-PG (green). Scale bars show 20 μm. (F) NSC34 cells were transfected with indicated plasmids. WST-8 assay was performed to assess the cell viability (mean ± s.e.m.). The experiments were repeated 5 times. One-way ANOVA with Tukey’s multiple comparison test was performed. In ΔC9 and WT, the same datasets as Fig. 1I were used.

Figure 6. Downregulation of EIF2D does not reduce expression levels of poly-PG and poly-PR. (A) A gRNA targeted the second exon of human EIF2D (see Materials and Methods). (B) After CRISPR/Cas9-mediated gene editing, the EIF2D knockout (EIF2DKO) HEK293 cells carried different mutations on each allele. (C) Cell lysates from WT and EIF2DKO HEK293 cells were processed for Western blotting, and immunostained with antibodies to eIF2D and α-tubulin. (D-F) WT and EIF2DKO HEK293 cells were cotransfected with fLuc plasmid along with either (D-E) AS-C9 plasmids or (F) C9 plasmids containing 75 GGGGCC repeats. The level of luciferase activity was assessed by dual luciferase assays. (G-I) WT HEK293 cells were transfected with fLuc and either (G-H) AS-C9 plasmids or (I) C9 monocistronic plasmids containing 75 GGGGCC repeats along with anti-EIF2D shRNA. The level of luciferase activity was assessed by dual luciferase assays (mean ± s.e.m.). The experiments were repeated 3 times. Unpaired t test was performed. The poly-GA reduction
upon EIF2D shRNA is consistent with our previous observations\textsuperscript{35}, albeit more modest - likely due to a technical reason (a bicistronic construct containing 75 GGGGCC repeats was used in \textsuperscript{35}).

**Figure 7. DPR levels in human iPSC-derived neurons upon eIF2D knockdown.** (A) The $\text{EIF2D}$, $\text{EIF2A}$, and $\text{actin}$ mRNA levels were assessed by real-time quantitative PCR on either isogenic control (26Z90) or $\text{C9ORF72}$ human motor neurons (patient line 26#6) upon siRNA transfection (scramble or EIF2D siRNA-1). The $\text{eIF2D}$ and $\text{eIF2A}$ mRNA levels were normalized to actin. The experiments were repeated two times. P<0.05 by one-way ANOVA with Tukey post hoc test. (B) Poly-GA, poly-GR and poly-GP levels in motor neurons differentiated independently (twice) from isogenic control and one $\text{C9ORF72}$ iPSC line. DPR levels were measured using an MSD immunoassay in a blinded manner. Data presented as mean ± S.D. P values were calculated using 2-way ANOVA with Dunnett’s multiple comparison test using Prizm (9.1) software. (C-D) The $\text{EIF2D}$ and $\text{actin}$ mRNA levels were assessed by real-time quantitative PCR on $\text{C9ORF72}$ human motor neurons (two patient lines) upon siRNAs transfection (scramble, EIF2D siRNA-1 or EIF2D siRNA-2). The $\text{eIF2D}$ mRNA levels were normalized to actin. The experiments were repeated three times. *P<0.05, ***P<0.001, ns, not significant by two-tailed unpaired $t$ tests were used for two groups and a one-way ANOVA followed by Dunnett post hoc analysis was used for more than two groups. (E) Poly-GA, poly-GR and poly-GP levels in motor neurons differentiated independently (n=3 times) from isogenic or healthy control lines and total two $\text{C9ORF72}$ patient iPSC lines (lines 27#11 and 40#3). DPR levels were measured using an MSD immunoassay in a blinded manner. For poly(GA) assay, total protein normalized poly(GA) concentrations were converted to percentage and presented as mean ± S.E. For poly(GR), poly (GP) assay, total protein normalized ECL values were converted to percentage and presented as mean ± S.E. P values were calculated using one-way ANOVA with Dunnnett’s T3 multiple comparisons test using Prizm (9.5) software.
**Figure 1 – figure supplement 1:** Nanoluciferase (nLuc) is fused to DPRs translated from anti-sense C9 plasmids containing CCCCCG repeats.

**Figure 1 – figure supplement 2:** Expression levels of poly-PR and poly-PG in the RIPA-insoluble fraction.

**Figure 1 – figure supplement 3:** Poly-PA is not detected by Western blotting upon transfection of anti-sense C9 plasmids containing CCCCCG repeats.

**Figure 3 – figure supplement 1:** Quantification of data from Figure 3B.

**Figure 4 – figure supplement 1:** Quantification of data from Figure 4B.

**Figure 5 – figure supplement 1:** Quantification of data from Figures 5B and C.

**Figure 7 – figure supplement 1:** The siRNA against eIF2D knocks down eIF2D protein levels.

**Supplementary File 1:** List of primers used for this study.
**SOURCE DATA LEGENDS**

**Figure 1 - source data 1:** Full raw unedited images of Western blots shown in Figure 1. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 1 – figure supplement 1 - source data 1:** Full raw unedited images of Western blots shown in Figure 1 – figure supplement 1. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 1 – figure supplement 2 - source data 1:** Full raw unedited images of Western blots shown in Figure 1 – figure supplement 2. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 1 – figure supplement 3 - source data 1:** Full raw unedited images of Western blots shown in Figure 1 – figure supplement 3. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 2 - source data 1:** Full raw unedited images of Western blots shown in Figure 2. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 3 - source data 1:** Full raw unedited images of Western blots shown in Figure 3. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 4 - source data 1:** Full raw unedited images of Western blots shown in Figure 4. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 5 - source data 1:** Full raw unedited images of Western blots shown in Figure 5. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 6 – source data 1:** Full raw unedited images of Western blots shown in Figure 6. Figures with the uncropped blots are clearly labelled with the relevant bands.

**Figure 7 – figure supplement 1 - source data 1:** Full raw unedited images of Western blots shown in Figure 7 – figure supplement 1. Figures with the uncropped blots are clearly labelled with the relevant bands.
<table>
<thead>
<tr>
<th>Reagent type (species) or resource</th>
<th>Designation</th>
<th>Source or reference</th>
<th>Identifiers</th>
<th>Additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>HEK293</td>
<td>ATCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell line (Mus musculus)</td>
<td>NSC34</td>
<td>Gift from Dr. Neil R. Cashman (McGill University) PMID: 1467557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>Isogenic iPS cells</td>
<td>Lopez-Gonzalez et al., 2019 PMID: 31019093</td>
<td>26z90</td>
<td>Isogenic control for patient line C926#6</td>
</tr>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>Isogenic iPS cells</td>
<td>Lopez-Gonzalez et al., 2019 PMID: 31019093</td>
<td>27m91</td>
<td>Isogenic control for patient line C927#11</td>
</tr>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>Healthy control iPS cells</td>
<td>Almeida et al., 2013 PMID: 23836290</td>
<td>Control2#20</td>
<td>Control for patient line C940#3</td>
</tr>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>C9orf72 patient iPS cells</td>
<td>Almeida et al., 2013 PMID: 23836290</td>
<td>C926#6</td>
<td>C9orf72 patient line</td>
</tr>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>C9orf72 patient iPS cells</td>
<td>Almeida et al., 2013 PMID: 23836290</td>
<td>C927#11</td>
<td>C9orf72 patient line</td>
</tr>
<tr>
<td>Cell line (Homo sapiens)</td>
<td>C9orf72 patient iPS cells</td>
<td>Freibaum et al., 2015 PMID: 26308899</td>
<td>C940#3</td>
<td>C9orf72 patient line</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-Poly-PR (Rabbit polyclonal)</td>
<td>EMD Millipore</td>
<td>ABN1354</td>
<td>WB (1:1000) IF (1:250)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Target ALS</td>
<td>Antibody target</td>
<td>Supplier</td>
<td>Catalog Number</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>----------------</td>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-Poly-PG (Mouse monoclonal)</td>
<td>EMD Millipore</td>
<td>TALS828.179</td>
<td>WB (1:1000)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-Poly-PA (Rabbit polyclonal)</td>
<td>EMD Millipore</td>
<td>ABN1356</td>
<td>WB (1:1000)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-nLuc (Mouse monoclonal)</td>
<td>Promega</td>
<td>N700A</td>
<td>WB (1:500)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-α-tubulin (Rat monoclonal)</td>
<td>Abcam</td>
<td>Ab6160</td>
<td>WB (1:5000)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-H3K4me2 (Rabbit polyclonal)</td>
<td>EMD Millipore</td>
<td>07-030</td>
<td>WB (1:2000)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-mouse horseradish peroxidase–conjugated secondary antibody (Sheep monoclonal)</td>
<td>GE Healthcare</td>
<td>NA931V</td>
<td>WB (1:5000)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-rabbit horseradish peroxidase–conjugated secondary antibody (Donkey monoclonal)</td>
<td>GE Healthcare</td>
<td>NA934V</td>
<td>WB (1:5000)</td>
</tr>
<tr>
<td>Antibody</td>
<td>Anti-rat horseradish peroxidase–conjugated secondary antibody (Goat polyclonal)</td>
<td>Cell Signaling Technology</td>
<td>7077S</td>
<td>WB (1:1000)</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------------------------------------------</td>
<td>--------------------------</td>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-ΔC9::nLuc</td>
<td>PMID: 29792928</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)-Poly-PR::nLuc (Plasmid)</td>
<td>This paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)-Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)-Poly-PA::nLuc (Plasmid)</td>
<td>This paper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) -366CUG-&gt;CCC -Poly-PR::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PR::nLuc vector with mutation of the CTG at -366bp from CCCGG repeats to CCC</td>
<td></td>
</tr>
<tr>
<td>------------------------</td>
<td>---------------------------------------------------</td>
<td>-----------</td>
<td>----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) -366CUG-&gt;UAG -Poly-PR::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PR::nLuc vector with mutation of the CTG at -366bp from CCCGG repeats to CCC</td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) -273AUG-&gt;CCC -Poly-PR::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PR::nLuc vector with mutation of the ATG at -273bp from CCCGG repeats to CCC</td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) -273AUG-&gt;UAG -Poly-PR::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PR::nLuc vector with mutation of the ATG at -273bp from CCCGG repeats to CCC</td>
<td></td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) CCC -Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -212bp, ATG at -194bp, CTG at</td>
<td></td>
</tr>
</tbody>
</table>

C9ORF72, and NanoLuc in frame of poly-PA
<table>
<thead>
<tr>
<th>Recombinant DNA reagent</th>
<th>pAG-AS(C9)(^{\text{-212AUG-Poly-PG::nLuc (Plasmid)}})</th>
<th>This paper</th>
<th>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -194bp, CTG at -182bp, and ATG at -113bp from CCCC GG repeats to CCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)(^{\text{-194AUG-Poly-PG::nLuc (Plasmid)}})</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -212bp, CTG at -182bp, and ATG at -113bp from CCCC GG repeats to CCC</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)(^{\text{-182CUG-Poly-PG::nLuc (Plasmid)}})</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -212bp, ATG at -194bp, and ATG at -113bp from CCCC GG repeats to CCC</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)(^{\text{-113AUG-Poly-PG::nLuc (Plasmid)}})</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -212bp, ATG at -194bp, and ATG at -113bp from CCCC GG repeats to CCC</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) $^{_{-212}}$CCC-Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -212bp from CCCGG repeats to CCC</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------</td>
<td>-------------</td>
<td>------------------------------------------------------------------</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) $^{_{-194}}$CCC-Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -194bp from CCCGG repeats to CCC</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) $^{_{-113}}$CCC-Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -113bp from CCCGG repeats to CCC</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9) $^{_{-212UAG}}$Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -212bp from CCCGG repeats to TAG</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)-194UAG-Poly-PG::nLuc</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -194bp from CCCGG repeats to TAG</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------------------------</td>
<td>------------</td>
<td>-------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pAG-AS(C9)-113UAG-Poly-PG::nLuc (Plasmid)</td>
<td>This paper</td>
<td>pAG-AS(C9)-Poly-PG::nLuc vector with mutation of ATG at -113bp from CCCGG repeats to TAG</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>lentiCRISPR v2-EIF2D (Plasmid)</td>
<td>This paper</td>
<td>Addgene (#52961) lentiCRISPR plasmid containing gRNA sequence against EIF2D</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>Sh-Control (Plasmid)</td>
<td>PMID: 34654821</td>
<td>Thermo Fisher Scientific (#AM5764) pSilencer 2.1-U6 neo plasmid containing non-specific control shRNA sequence</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>Sh-EIF2D (Plasmid)</td>
<td>PMID: 34654821</td>
<td>Thermo Fisher Scientific (#AM5764) pSilencer 2.1-U6 neo plasmid containing shRNA sequence against EIF2D</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pGL4.50 [luc2/CMV/Hygro] (Plasmid)</td>
<td>Promega</td>
<td>E131A Expression of firefly luciferase</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pNL1.1 CMV (Plasmid)</td>
<td>Promega</td>
<td>N109A Expression of NanoLuc</td>
</tr>
<tr>
<td>Classification</td>
<td>Description</td>
<td>Supplier</td>
<td>Code</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------------------------</td>
<td>---------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Recombinant DNA reagent</td>
<td>pcDNA 6/V5-His A (Plasmid)</td>
<td>Thermo Fisher Scientific</td>
<td>43-0003</td>
</tr>
<tr>
<td>Sequenced-based reagent</td>
<td>siRNA: non-targeting negative control</td>
<td>Thermo Fisher Scientific</td>
<td>4390844</td>
</tr>
<tr>
<td>Sequenced-based reagent</td>
<td>siRNA: EIF2D</td>
<td>Thermo Fisher Scientific</td>
<td>S4495</td>
</tr>
<tr>
<td>Sequenced-based reagent</td>
<td>siRNA: EIF2D</td>
<td>Thermo Fisher Scientific</td>
<td>S4496</td>
</tr>
<tr>
<td>Chemical compound, drug</td>
<td>Halt Protease Inhibitor Cocktail</td>
<td>Thermo Fisher Scientific</td>
<td>87786</td>
</tr>
<tr>
<td>Chemical compound, drug</td>
<td>SB421542</td>
<td>Stemgent</td>
<td>04-0010-10</td>
</tr>
<tr>
<td>Chemical compound, drug</td>
<td>CHIR99021</td>
<td>Stem Cell Technologies</td>
<td>72054</td>
</tr>
<tr>
<td>Chemical compound, drug</td>
<td>DMH1</td>
<td>Stem Cell Technologies</td>
<td>73634</td>
</tr>
<tr>
<td>Chemical compound, drug</td>
<td>All-Trans Retinoic Acid</td>
<td>Stem Cell Technologies</td>
<td>72262</td>
</tr>
<tr>
<td>Commercial assay or kit</td>
<td>Q5® Site-Directed Mutagenesis Kit</td>
<td>New England Biolabs</td>
<td>E0554S</td>
</tr>
<tr>
<td>Commercial assay or kit</td>
<td>Nano-Glo Dual-Luciferase Reporter assay system</td>
<td>Promega</td>
<td>N1610</td>
</tr>
<tr>
<td>Commercial assay or kit</td>
<td>Cell Counting Kit-8</td>
<td>Dojindo</td>
<td>CK-04</td>
</tr>
<tr>
<td>Type</td>
<td>Product Description</td>
<td>Supplier</td>
<td>Catalog Number</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Commercial assay or kit</td>
<td>BCA Protein Assay Kit</td>
<td>Thermo Fisher Scientific</td>
<td>23225</td>
</tr>
<tr>
<td>Commercial assay or kit</td>
<td>660 nm Protein Assay Reagent</td>
<td>Thermo Fisher Scientific</td>
<td>22660</td>
</tr>
<tr>
<td>Software, algorithm</td>
<td>Image Lab software</td>
<td>Bio-Rad</td>
<td></td>
</tr>
<tr>
<td>Software, algorithm</td>
<td>ImageJ2 software</td>
<td>PMID: 22930834</td>
<td></td>
</tr>
<tr>
<td>Software, algorithm</td>
<td>GraphPad Prism</td>
<td>Dotmatics</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5 x passive lysis buffer</td>
<td>Promega</td>
<td>E1941</td>
</tr>
<tr>
<td>Other</td>
<td>4′,6-diamidino-2-phenylindole (DAPI)</td>
<td>Thermo Fisher Scientific</td>
<td>D1306</td>
</tr>
<tr>
<td>Other</td>
<td>SuperSignal West Dura Extended Duration Substrate</td>
<td>Thermo Fisher Scientific</td>
<td>34076</td>
</tr>
<tr>
<td>Other</td>
<td>Lipofectamine LTX</td>
<td>Thermo Fisher Scientific</td>
<td>15338030</td>
</tr>
</tbody>
</table>
**Figure 1**

A. Diagram showing the expression of nLuc and BGH polyA in the Anti-sense (AS)-C9 constructs.

B. Bar graph showing relative expression (nLuc/fLuc) in HEK293 cells with 

C. Bar graph showing relative expression (nLuc/fLuc) in NSC34 cells with

D. Western blot images showing Poly-PR and Poly-PG in HEK293 and NSC34 cells with

E. Western blot images showing α-tubulin in HEK293 and NSC34 cells with

F. Immunofluorescence images showing Poly-PR and Poly-PG in HEK293 and NSC34 cells with

G. Immunofluorescence images showing DAPI and MERGE in HEK293 and NSC34 cells with

H. Bar graph showing optical density (450 nm) with

I. Bar graph showing optical density (450 nm) with
Figure 3

A

<table>
<thead>
<tr>
<th>Position</th>
<th>WT</th>
<th>CCC</th>
<th>-212 AUG</th>
<th>-194 AUG</th>
<th>-182 CUG</th>
<th>-113 AUG</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUG</td>
<td>CCC</td>
<td>AUG</td>
<td>CCC</td>
<td>CUG</td>
<td>AUG</td>
</tr>
<tr>
<td></td>
<td>AUG</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>CUG</td>
<td>CCC</td>
</tr>
<tr>
<td></td>
<td>AUG</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>CCC</td>
<td>AUG</td>
</tr>
</tbody>
</table>

poly-PG::nLuc

B

<table>
<thead>
<tr>
<th>Expression</th>
<th>HEK293 cells</th>
<th>NSC34 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative expression (nLuc/fLuc)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>50</td>
<td>100</td>
</tr>
</tbody>
</table>

C

<table>
<thead>
<tr>
<th>Expression</th>
<th>HEK293 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative expression (nLuc/fLuc)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>50</td>
</tr>
</tbody>
</table>

D

<table>
<thead>
<tr>
<th>Expression</th>
<th>NSC34 cells</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative expression (nLuc/fLuc)</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>50</td>
</tr>
</tbody>
</table>

HEK293 cells

NSC34 cells
A

ΔC9  \[\text{AG}\]  \[\text{-212 bp}\]  \[\text{-194 bp}\]  \[\text{-113 bp}\]  \[\text{poly-PG::nLuc}\]

WT  \[\text{AG}\]  \[\text{AUG}\]  \[\text{AUG}\]  \[\text{poly-PG::nLuc}\]

-212CCC  \[\text{AG}\]  \[\text{CCC}\]  \[\text{AUG}\]  \[\text{poly-PG::nLuc}\]

-194CCC  \[\text{AG}\]  \[\text{AUG}\]  \[\text{CCC}\]  \[\text{poly-PG::nLuc}\]

-113CCC  \[\text{AG}\]  \[\text{AUG}\]  \[\text{CCC}\]  \[\text{poly-PG::nLuc}\]

B

HEK293 cells

NSC34 cells

Poly-PG

α-tubulin

C

HEK293 cells

D

NSC34 cells
Figure 5

poly-PG::nLuc

WT

-212 bp

-194 bp

-113 bp

(CCCCGG)₃₅

nLuc

D

C9

-194 UAG

-113 UAG

-212 UAG

-194 UAG

HEK293 cells

NSC34 cells

Poly-PG

α-tubulin

HEK293 cells

NSC34 cells

Optical density (450 nm)

Poly-PG

DAPI

MERGE

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG

ΔC9

WT

-212 UAG

-194 UAG

-113 UAG
Figure 6

A. Schematic of human EIF2D locus with gRNA targeting site.

B. WT allele: ACTGTGTACG-TGAGTGGTGG
   EIF2D KO allele: ACTGTGTACG-TGAGTGGTGG

C. Western blots showing eIF2D and α-tubulin levels.

D. Relative expression of poly-PG::nLuc in WT and EIF2D KO alleles.

E. Relative expression of poly-PR::nLuc in WT and EIF2D KO alleles.

F. Relative expression of poly-GA::nLuc in WT and EIF2D KO alleles.

G. Relative expression of poly-PG::nLuc with control and EIF2D shRNA.

H. Relative expression of poly-PR::nLuc with control and EIF2D shRNA.

I. Relative expression of poly-GA::nLuc with control and EIF2D shRNA.

P-values: poly-PG::nLuc (WT) = 0.0451, poly-PG::nLuc (EIF2D KO) = 0.5330, poly-PR::nLuc (WT) = 0.5423, poly-PR::nLuc (EIF2D KO) = 0.2182, poly-GA::nLuc (WT) = 0.3617, poly-GA::nLuc (EIF2D KO) = 0.0022.
Figure 7

A. C9ORF72 neurons - patient line 26#6

<table>
<thead>
<tr>
<th></th>
<th>EIF2D</th>
<th>EIF2A</th>
</tr>
</thead>
<tbody>
<tr>
<td>scr-siRNA</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>EIF2D siRNA-1</td>
<td>0.75</td>
<td>0.80</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0342</td>
<td>0.4724</td>
</tr>
</tbody>
</table>

B. C9ORF72 neurons - patient line 26#6

<table>
<thead>
<tr>
<th></th>
<th>Poly-GA</th>
<th>Poly-GR</th>
<th>Poly-GP/PG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isog. Control</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>scr-siRNA</td>
<td>90</td>
<td>85</td>
<td>95</td>
</tr>
<tr>
<td>EIF2D siRNA-1</td>
<td>80</td>
<td>75</td>
<td>85</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0155</td>
<td>0.2062</td>
<td>0.7780</td>
</tr>
</tbody>
</table>

C. C9ORF72 neurons - patient line 27#11

<table>
<thead>
<tr>
<th></th>
<th>Relative eIF2D mRNA levels (eIF2D/Actin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>scr-siRNA</td>
<td>1.00</td>
</tr>
<tr>
<td>EIF2D siRNA-1</td>
<td>0.80</td>
</tr>
<tr>
<td>p-value</td>
<td>0.4811</td>
</tr>
<tr>
<td>EIF2D siRNA-2</td>
<td>0.75</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

D. C9ORF72 neurons - patient line 40#3

<table>
<thead>
<tr>
<th></th>
<th>Relative eIF2D mRNA levels (eIF2D/Actin)</th>
</tr>
</thead>
<tbody>
<tr>
<td>scr-siRNA</td>
<td>1.00</td>
</tr>
<tr>
<td>EIF2D siRNA-1</td>
<td>0.80</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0110</td>
</tr>
<tr>
<td>EIF2D siRNA-2</td>
<td>0.75</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

E. C9ORF72 neurons - average data from two C9ORF72 patient lines (27#11 and 40#3)

<table>
<thead>
<tr>
<th></th>
<th>Poly-GA</th>
<th>Poly-GR</th>
<th>Poly-GP/PG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>scr-siRNA</td>
<td>90</td>
<td>85</td>
<td>95</td>
</tr>
<tr>
<td>EIF2D siRNA-1</td>
<td>80</td>
<td>75</td>
<td>85</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0785</td>
<td>0.8898</td>
<td>0.0537</td>
</tr>
<tr>
<td>EIF2D siRNA-2</td>
<td>70</td>
<td>65</td>
<td>75</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0290</td>
<td>0.7307</td>
<td>0.0079</td>
</tr>
</tbody>
</table>

Poly-GP/PG
Figure 1 - figure supplement 1. Nanoluciferase (nLuc) is fused to DPRs translated from anti-sense C9 plasmids containing CCCCCG repeats.

(A) Schematic diagram of the construct. (B-C) (B) HEK293 and (C) NSC34 cells were transfected with either ΔC9, PG::nLuc, or positive control (Pos Ctrl) plasmids. The nanoluciferase (nLuc) expression plasmid pNL1.1. [Nluc/CMV] from Promega was used as positive control. Cell lysates were processed for Western blotting, and immunostained with antibodies to poly-PG, nLuc, and α-tubulin.
Figure 1 - figure supplement 2. Expression levels of poly-PR and poly-PG in the RIPA-insoluble fraction.

(A) Schematic diagram of the constructs. (B, D, F, H) HEK293 and (C, E, G, I) NSC34 cells were transfected with either ΔC9 (blue) or AS-C9 (red) plasmids. Cell lysates were fractionated into RIPA-soluble (S) and -insoluble (I) fraction. Western blotting was performed and immunostained with antibodies to poly-PR (B-C), poly-PG (F-G), α-tubulin, and H3K4me2. H3K4me2 is a reliable marker for the RIPA-insoluble fraction (see Materials and Methods). (D-E, H-I) The signal intensity of the bands was quantified. The expression levels of poly-PR (D-E) or poly-PG (H-I) in AS-C9 of RIPA-soluble fraction were set to 1.0. Experiments were repeated 4 times. Two-way ANOVA with Tukey's multiple comparison test was performed.
Figure 1 - figure supplement 3. Poly-PA is not detected by Western blotting upon transfection of anti-sense C9 plasmids containing CCCCCGG repeats.

(A) Schematic diagram of the constructs. (B-C) (B) HEK293 and (C) NSC34 cells were transfected with either ΔC9, AS-C9, or positive control plasmids. For positive control, we generated a plasmid that has an AUG start codon located at -67bp upstream from CCCCCGG repeats. This AUG is in poly-PA frame. Cell lysates were processed for Western blotting, and immunostained with antibodies to poly-PA and α-tubulin.
Figure 3 – figure supplement 1. Quantification of data from Figure 3B.
A) Schematic diagram of the constructs. (B-C) The signal intensity of total poly-PG and poly-PG translated from each of the AUGs (-212, -194, or -113) in (B) HEK293 and (C) NSC34 was quantified. One-way ANOVA with Tukey’s multiple comparison test was performed. The experiments were repeated 4 times. Data are presented as mean ± s.e.m.
Figure 4 – figure supplement 1. Quantification of data from Figure 4B.
(A) Schematic diagram of the constructs. (B-C) The signal intensity of poly-PG and poly-PG translated from each of the AUGs (-212, -194, or -113) in (B) HEK293 and (C) NSC34 cells was quantified. One-way ANOVA with Tukey’s multiple comparison test was performed. The experiments were repeated 4 times. Data are presented as mean ± s.e.m.
Figure 5 – figure supplement 1. Quantification of data from Figures 5B and C.
(A) Schematic diagram of the constructs. (B-C) The signal intensity of total poly-PG and poly-PG translated from each of the AUGs (-212, -194, or -113) in (B) HEK293 and (C) NSC34 cells was quantified. One-way ANOVA with Tukey’s multiple comparison test was performed. The experiments were repeated 4 times. Data are presented as mean ± s.e.m.
Figure 7 – figure supplement 1. siRNA (s4496) against eIF2D knocks down eIF2D protein levels.

3x10^5 HEK293 cells were plated into 6 well cell culture plate. 60 pmol siRNAs were transfected into the cells using 6.25 µl Lipofectamine RNAiMAX. 72h later, protein was extracted using RIPA buffer and Western blotting was performed.