**NEURODEGENERATIVE DISORDERS**

C9orf72 poly(GR) aggregation induces TDP-43 proteinopathy

Casey N. Cook1,2*, Yanwei Wu1*, Hana M. Odeh3*, Tania F. Gendron1,2, Karen Jansen-West1, Giulia M. Avendano1, Monica Castanedas-Casey1, Wei Shao1, Björn Oskarsson4, Giulio S. Tomassy5, Alexander McCampl6, Frank Rigo6, Dennis W. Dickson1,2, James Shorter3†, Yong-Jie Zhang1,2†, Leonard Petrucelli1,2†

TAR DNA-binding protein 43 (TDP-43) inclusions are a pathological hallmark of frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), including cases caused by G4C2 repeat expansions in the C9orf72 gene (c9FTD/ALS). Providing mechanistic insight into the link between C9orf72 mutations and TDP-43 pathology, we demonstrated that a glycine-arginine repeat protein [poly(GR)] translated from expanded G4C2 repeats was sufficient to promote aggregation of endogenous TDP-43. In particular, toxic poly(GR) proteins mediated sequestration of full-length TDP-43 in an RNA-independent manner to induce cytoplasmic TDP-43 inclusion formation. Moreover, in GFP-(G4C2)80 mice, poly(GR) caused the mislocalization of nucleocytoplasmic transport factors and nuclear pore complex proteins. These mislocalization events resulted in the aberrant accumulation of endogenous TDP-43 in the cytoplasm where it co-aggregated with poly(GR). Last, we demonstrated that treating G4C2 repeat–expressing mice with repeat-targeting antisense oligonucleotides lowered poly(GR) burden, which was accompanied by reduced TDP-43 pathology and neurodegeneration, including lowering of plasma neurofilament light (NFL) concentration. These results contribute to clarification of the mechanism by which poly(GR) drives TDP-43 proteinopathy, confirm that G4C2-targeted therapeutics reduce TDP-43 pathology in vivo, and demonstrate that alterations in plasma NFL provide insight into the therapeutic efficacy of disease-modifying treatments.

INTRODUCTION

Frontotemporal dementia (FTD) is characterized clinically by changes in personality, behavior, and/or language, whereas amyotrophic lateral sclerosis (ALS) is characterized by motor neuron signs. Nevertheless, these two fatal neurodegenerative diseases share pathologic features and genetic causes, and frequently co-occur within the same individuals. FTD and ALS exist on the same disease spectrum and are caused by common pathogenic mechanism(s). A hallmark feature of the vast majority of ALS cases and about half of FTD cases is TAR DNA-binding protein 43 (TDP-43) pathology. TDP-43, a predominantly nuclear DNA/RNA-binding protein with a prion-like domain, plays many roles in RNA metabolism. In patients with FTD/ALS-TDP-43, however, TDP-43 forms cytoplasmic aggregates, which result in its nuclear depletion (1, 2). The underlying factors that initiate these events are unknown, but the discovery of a G4C2 repeat expansion in chromosome 9 open reading frame 72 (C9orf72) as the most common genetic cause of FTD and ALS (c9FTD/ALS) (3, 4) has provided important insight into potential mechanisms that drive TDP-43 pathology.

In addition to the aggregation of TDP-43, pathologic features uniquely derived from the C9orf72 repeat expansion are observed in c9FTD/ALS. These include C9orf72 haploinsufficiency, as well as the accumulation of nuclear RNA foci composed of sense G4C2 and antisense G3C3 repeat transcripts (3, 5, 6) and dipeptide repeat (DPR) proteins [poly(GA), poly(GP), poly(GR), poly(PR), or poly(PA)] translated from these transcripts through unconventional means (7–11). Several lines of evidence suggest that TDP-43 pathology in c9FTD/ALS results from repeat expansion products rather than loss of C9orf72. For instance, knockout of C9orf72 in mice does not lead to TDP-43 pathology, neurodegeneration, or motor deficits (12). In contrast, mice that express G4C2 repeat expansions not only develop RNA foci and express DPR proteins but also exhibit phosphorylated and aggregated endogenous TDP-43 (13–16), indicating that the TDP-43 pathology is likely triggered by repeat RNA and/or DPR proteins produced from the expanded G4C2 repeats. Repeat expansion products, especially arginine-rich poly(GR) and poly(PR) proteins, have been implicated in causing nucleocytoplasmic transport defects (17–22) and impairing stress granule (SG) dynamics (13, 23–26), two pathomechanisms linked to TDP-43 mislocalization and aggregation, in c9FTD/ALS. For instance, poly(GR) immunoprecipitates with TDP-43, nucleocytoplasmic transport factors, and SG-associated proteins (26, 27). These findings, along with evidence that essential protein components and key regulators of SG assembly, such as ataxin-2 and T’ cell intracellular antigen-1 (TIA-1), colocalize with poly(GR) and phosphorylated TDP-43 (pTDP-43) in (G4C2)149 mice (13), suggest that poly(GR) plays a key role in inducing TDP-43 pathology through the disruption of nucleocytoplasmic transport and SG biology. That poly(GR) colocalizes with TDP-43 inclusions in the motor cortex of patients with c9ALS supports this notion (28). Nonetheless, the mechanistic links between poly(GR) and TDP-43 remain unclear. Here, we used pure protein biochemistry, cell and animal models, as well as human postmortem
brain tissue to investigate the relationship between poly(GR) and TDP-43 proteinopathy.

**RESULTS**

**Poly(GR) accelerates and enhances TDP-43 aggregation**

To investigate the relationship between poly(GR) and TDP-43, we first assessed their behavior at the pure protein level. Thus, we incubated purified recombinant maltose-binding protein (MBP)–tagged TDP-43 with or without (GR)_{20} (a protein with 20 GR repeats) or (GA)_{20} (a protein with 20 GA repeats). When tagged with MBP, wild-type TDP-43 (TDP-43WT) remained soluble in the absence or presence of (GR)_{20} or (GA)_{20} (Fig. 1A). However, after selective removal of the MBP tag by tobacco etch virus (TEV) protease, TDP-43WT formed aggregates in a time-dependent manner (Fig. 1A). Sedimentation analysis at the end point of the turbidity measurements confirmed that the increased absorbance in our aggregation assay is due to the formation of TDP-43WT aggregates (fig. S1A). Upon near completion of MBP cleavage, TDP-43WT in the absence or presence of (GR)_{20} or (GA)_{20} fully sedimented into the pellet fraction, whereas free MBP remained soluble (fig. S1A). Of note, co-incubation with (GR)_{20}, but not (GA)_{20}, significantly accelerated and enhanced TDP-43WT aggregation ($P < 0.0001$), although (GR)_{20} and (GA)_{20} alone did not aggregate (Fig. 1, A and B). More specifically, (GR)_{20} elicited the rapid formation of visible, solid-like TDP-43 aggregates at times where no TDP-43 aggregates formed in the presence of buffer or (GA)_{20} (Fig. 1, A and C). Transmission electron microscopy (TEM) revealed that whereas uncleaved TDP-43WT showed no aggregate formation in the absence or presence of (GR)_{20} or (GA)_{20} (fig. S1B), (GR)_{20} induced the formation of very dense TDP-43WT aggregates upon MBP cleavage (Fig. 1D). In contrast, (GA)_{20} had no obvious effect on TDP-43, which displayed a typical granulo-filamentous morphology (Fig. 1D). Quantitative analysis revealed that whereas uncleaved TDP-43WT showed no aggregate formation in the absence or presence of (GR)_{20} or (GA)_{20} (fig. S1B), (GR)_{20} induced the formation of very dense TDP-43WT aggregates upon MBP cleavage (Fig. 1D).

**Poly(GR) recruits cytoplasmic TDP-43 and SG-associated proteins to inclusions**

To investigate the effects of poly(GR) aggregates on TDP-43, we first examined the morphology of poly(GR) aggregates using...
immuno-electron microscopy (IEM). We observed gold labeling of poly(GR) antibody on both granular and filamentous materials within cytoplasmic aggregates (Fig. 2A). Following confirmation that poly(GR) forms filaments and granular deposits in transfected cells, human embryonic kidney (HEK) 293T cells expressing green fluorescent protein (GFP) or GFP-tagged (GR)100 were cotransfected with various Myc-tagged TDP-43 constructs. Consistent with our previous findings (29), we observed that cytoplasmic GFP-(GR)100 inclusions were immunopositive for the SG markers TIA-1, eukaryotic translation initiation factor 3 (eIF3), and ataxin-2 (Fig. 2B and fig. S2A). TDP-43WT, however, was absent from poly(GR)/TIA-1-positive inclusions, instead remaining in the nucleus (Fig. 2, B and C). In contrast, TDP-43 with a mutated nuclear localization signal (TDP-43NLSm) was confined to the cytosol where it co-aggregated with poly(GR) and TIA-1 (Fig. 2, B and C). To confirm that co-aggregation was not merely caused by TDP-43NLSm overexpression, we examined the distribution and expression of TDP-43NLSm in the absence of poly(GR) (fig. S2B). TDP-43NLSm remained diffusely distributed throughout the cytosol in GFP-transfected cells (Fig. 2C and fig. S2B), despite the fact that TDP-43NLSm expression was higher in GFP–expressing cells than in GFP-(GR)100–expressing cells (fig. S2, C and D) (likely because poly(GR) inhibits translation (26, 29)).

Unlike poly(GR), poly(GA) inclusions in GFP-(GA)100–expressing cells failed to sequester cytoplasmic TDP-43NLSm (fig. S2E). Together, these data indicate that both the cytoplasmic localization of TDP-43 and its specific recruitment to poly(GR) inclusions promote TDP-43 aggregation.

Given that TDP-43 is an RNA-binding protein, we examined whether cytoplasmic TDP-43 recruitment to GFP-(GR)100 inclusions was RNA dependent by mutating five phenylalanine residues within the RNA-binding motifs of TDP-43NLSm to leucine residues (TDP-43NLSm-5FL). Although these mutations disrupt the RNA-binding ability of TDP-43 (30), TDP-43NLSm-5FL nonetheless colocalized to poly(GR) aggregates (Fig. 2, B and C). Next, because the C-terminal prion-like domain of TDP-43 is prone to aggregate (31, 32), we co-expressed TDP-43 C-terminal fragments (CTFs) or TDP-43NLSm lacking the C-terminal domain (TDP-431-273-NLSm) with GFP-(GR)100. Whereas TDP-431-273-NLSm maintained the ability to be recruited to poly(GR) inclusions (Fig. 2, B and C), TDP-43 CTF formed cytoplasmic aggregates devoid of poly(GR) (fig. S2F). To study the poly(GR)–TDP-43 interaction in more detail, we used a proximity ligation assay (PLA) in which a fluorescent PLA signal is indicative of GFP-(GR)100 binding to a given Myc-tagged TDP-43 species. Consistent with a lack of colocalization, no PLA signal was detected in cells expressing GFP-(GR)100 and TDP-43NLSm (fig. S2E).

To confirm that poly(GR) aggregation, in the absence of other G4 C2-associated pathologies, is sufficient to promote the aggregation of endogenous, wild-type TDP-43 in vivo.

Poly(GR) aggregates sequester nucleocytoplasmic transport factors and nuclear pore complex proteins in vivo

Because the cytoplasmic localization of TDP-43 is required for its sequestration to poly(GR) inclusions, we next determined how endogenous TDP-43 becomes mislocalized in GFP-(GR)200 mice. Given that poly(GR) disrupts nucleocytoplasmic transport in yeast, flies, and cultured cells (17, 18, 20), which could disrupt TDP-43 transport to and from the nucleus, we evaluated whether poly(GR) influences nucleocytoplasmic transport factors. We first examined importin α5
and karyopherin α2 (KPNA2), two nuclear import factors that mediate TDP-43 nuclear import (33). Co-immunofluorescence staining demonstrated that the colocalization of poly(GR) and importin α5 or KPNA2 was observed in both diffuse and aggregated poly(GR)-positive cells (Fig. 4A). However, compared to non-transduced cells, aggregated but not diffuse poly(GR) affected the distribution pattern of importin α5 and KPNA2 by recruiting these proteins in poly(GR) inclusions (Fig. 4, A and B). We next assessed various components of the nuclear pore complex (NPC), which consists of about 30 nucleoporins. Using an NPC antibody that detects four nucleoporins (NUP62, NUP153, NUP214, and NUP358), we observed that NPCs were homogeneously distributed throughout the nuclear envelope in nontransduced and diffuse poly(GR)-positive cells (Fig. 4C). In contrast, in poly(GR) inclusion–bearing cells, NPCs were irregularly distributed around the nuclear envelope and partially colocalized with poly(GR) aggregates, and this was similarly observed for the NPC components NUP98 and pore membrane protein of 121 kDa (POM121) (Fig. 4C and fig. S4). Quantitative analysis revealed that more than 90% of poly(GR) inclusion–bearing cells contained abnormal KPNA2 and NUP98 staining (Fig. 4, B and D). Last, we evaluated whether TDP-43 or eIF3η distribution was altered in poly(GR)-positive cells with NPC abnormalities. As anticipated, aggregated poly(GR) was associated with aberrant NPC distribution, loss of nuclear TDP-43, and recruitment of cytoplasmic TDP-43 and eIF3η to poly(GR) inclusions.
These results provide in vivo evidence that poly(GR) expression alone is sufficient to disturb nucleocytoplasmic transport factors and NPC proteins, contributing to the mislocalization of endogenous, wild-type TDP-43 and ultimately driving TDP-43 proteinopathy.

A G_{4C2}-targeted therapeutic mitigates poly(GR) and TDP-43 aggregation and rescues neurodegeneration in vivo

We previously reported that expressing expanded G_{4C2} repeats in the brain of mice leads to the accumulation of repeat-containing transcripts, DPR proteins, and pTDP-43 pathology (13, 14). Data from the present study strongly suggest that poly(GR) is directly responsible for inducing the TDP-43 pathology detected in G_{4C2}\_{66} and G_{4C2}\_{149} mice, and this role of poly(GR) is further supported by our previous observation that poly(GR) aggregates in G_{4C2}\_{66} mice are immunopositive for pTDP-43 and TIA-1 (13). Although we have shown that treating G_{4C2}\_{66} mice with antisense oligonucleotides targeting G_{4C2} repeats (c9ASO) decreases G_{4C2} repeat RNA and DPR protein expression (34), it remains unclear whether c9ASOs, which are currently being tested in clinical trials, will also alleviate pTDP-43 pathology. This issue is critical as it is likely imperative to combat TDP-43 proteinopathy and restore functional TDP-43 to the nucleus to rescue degenerating neurons. To investigate...
this issue, c9ASOs were injected into the brain of 3-month-old \( (G_4C_2)_{149} \) mice, which develop robust RNA foci and DPR protein pathology by this age (13). As anticipated, \( G_4C_2 \) repeat–containing RNA expression, assessed by both fluorescence in situ hybridization (FISH) and quantitative reverse transcription polymerase chain reaction (qRT-PCR), was reduced 3 months after treatment (Fig. 5A and fig. S5, A and B). To assess DPR burden, we used both immunohistochemistry of fixed brain tissue and immunoassays of cortical brain lysates, which confirmed that poly(GR) was reduced by c9ASO treatment, along with the other sense DPR proteins poly(GA) and poly(GP) (Fig. 5B and fig. S5C). Moreover, the number of inclusions immunopositive for pTDP-43 or ataxin-2 was also significantly reduced (Fig. 5, C to E; \( P < 0.0001 \) and \( P = 0.0407 \) for
To assess the neuroprotective effect of c9ASO-mediated alleviation of the abovementioned pathologies, we examined both neuronal number and amount of plasma neurofilament light (NFL), a biomarker of neuronal injury (35–37). c9ASO treatment prevented the reduction in NeuN-positive cortical neurons otherwise observed in (G4C2)149 mice (Fig. 5F). In addition, we observed increased plasma NFL concentrations in (G4C2)149 mice compared to control (G4C2)2 mice, which were attenuated in c9ASO-treated (G4C2)149 mice (Fig. 5G).

To examine the pathological meaning of alterations in plasma NFL in the (G4C2)149 model, we evaluated the relationship between plasma NFL and neuronal counts,
as well as TDP-43 pathology. Of note, plasma NFL negatively correlated with the number of NeUN-positive neurons (Fig. 5H) and positively correlated with pTDP-43 pathology (Fig. 5I) in (G4C2)149 mice, suggesting that plasma NFL concentrations are reflective of neuronal viability and TDP-43 burden in the brain. Together, these findings demonstrate that c9ASO treatment mitigates TDP-43 proteinopathy, which is associated with a reduction in neuronal loss and stabilization of plasma NFL concentrations in vivo.

**DISCUSSION**

In this study, we uncovered a key role for G4C2 repeat–derived poly(GR) in promoting TDP-43 proteinopathy in vitro and in vivo. Specifically, we established that poly(GR) directly accelerates and enhances TDP-43 aggregation. Moreover, poly(GR) sequesters cytoplasmic full-length TDP-43 through an RNA-independent mechanism. In doing so, poly(GR) induces the formation of dense TDP-43 aggregates in vitro. The formation of poly(GR) inclusions immunopositive for TDP-43 and SG-resident proteins in GFP-(GR)200 mice demonstrates that poly(GR) aggregation, in the absence of other G4C2-associated pathologies, is sufficient to induce endogenous TDP-43 aggregation. Combined with corroborating findings of inclusions containing both poly(GR) and TDP-43 in patients with c9FTD/ALS, these results provide in vivo validation in both mouse and human postmortem brain tissues that poly(GR) accumulation directly induces TDP-43 proteinopathy.

A key feature of TDP-43 pathology in humans is that TDP-43 inclusions are accompanied by the loss of nuclear TDP-43. Thus, our observation that a subset of neurons with cytoplasmic poly(GR)/TDP-43–positive inclusions exhibited reduced nuclear TDP-43 in GFP-(GR)200 mice increases pathological meaning. Given that importins and NPC proteins interact with poly(GR) (20, 26, 27), we asked whether the aberrant cytoplasmic distribution of endogenous TDP-43 in poly(GR)-positive cells might be caused by poly(GR)-induced nucleocytoplasmic transport defects. In support of this idea, we present evidence that several importins and nucleoporins are abnormally distributed in poly(GR)-positive cells and partially colocalize with poly(GR) inclusions in GFP-(GR)200 mice. These features were accompanied by both the loss of nuclear TDP-43 and its recruitment to cytoplasmic poly(GR) inclusions. These results support a temporal mechanism in which poly(GR) impairs nucleocytoplasmic transport by sequestering key nucleocytoplasmic transport factors and nucleoporins, which subsequently drives cytoplasmic accumulation and ultimately co-aggregation of TDP-43 with poly(GR) through protein-protein interactions.

Given our observations that poly(GR) alone is sufficient to induce endogenous TDP-43 mislocalization and aggregation, it is noteworthy that both TDP-43 and poly(GR) burden correlate with neurodegeneration in patients (28, 38, 39). Consistent with this notion, we observed an age-dependent loss of poly(GR)-positive cells in GFP-(GR)200 mice, suggesting that expression of poly(GR) is toxic to neurons. Moreover, the more marked reduction in aggregated poly(GR) compared to diffuse poly(GR) is likely due to the combination of poly(GR) and TDP-43 abnormalities in cells with poly(GR) inclusions. Therefore, therapeutics targeting poly(GR), TDP-43, or both may be needed to modulate neurotoxicity in c9FTD/ALS. Although the current study demonstrates that c9ASOs reduce TDP-43 pathology in an AAV model of c9FTD/ALS, future studies will be needed to explore the benefit of therapeutic strategies specifically targeting poly(GR) rather than all G4C2 repeat–associated pathologies. Although our results indicate that poly(GA) alone does not cause TDP-43 aggregation at either the pure protein level, in cultured cells, or in the mouse brain, it remains possible that additional products generated from the C9orf72 repeat expansion (RNA foci and other DPR proteins), either on their own or in combination, contribute to disease pathogenesis. TDP-43 has been shown to colocalize with poly(GA) and poly(GP) inclusions in patients with c9ALS, albeit less frequently than with poly(GR) (28). These findings are consistent with our observations in patients with c9FTD/ALS, suggesting that poly(GR) plays a more direct role in driving TDP-43 inclusion formation. In contrast to poly(GR), poly(GA)-induced TDP-43 pathology may require the presence of other C9orf72-associated pathology. It also bears mentioning that because poly(PR) shares similar biophysical properties and functions as poly(GR) (26), it too may also cause TDP-43 aggregation. However, poly(PR) inclusions are rare in patients with c9FTD/ALS (10, 40) and thus unlikely to be a major driver of TDP-43 pathology.

Although our findings reveal a new mechanism to explain TDP-43 pathology in poly(GR)-positive cells, it is important to note that there are other mechanisms underpinning TDP-43 proteinopathy. For example, our results indicate that the aggregation of pathological TDP-43 CTFs is independent of poly(GR) in cultured cells. Moreover, despite the abundance of TDP-43 pathology in the spinal cord of patients with c9ALS, poly(GR) aggregates in this region are rare (41). In addition, two recent studies provide insight into the underlying causes of TDP-43 pathology in sporadic ALS and FTD and/or poly(GR)-negative cells by demonstrating that cytoplasmic TDP-43 aggregation is mediated by aberrant phase transitions and occurs independently of SGs (42, 43). These data suggest that the observed recruitment of SG proteins to poly(GR) aggregates in GFP-(GR)200 mice occurs through their direct interaction with poly(GR) rather than by binding TDP-43. This mechanism is also consistent with the identification of the SG proteins RasGAP SH3 domain binding protein 1 (G3BP1) and ataxin-2 as poly(GR)-interacting proteins (26, 27), and with our finding that endogenous TIA-1 is recruited to poly(GR) inclusions in cultured cells expressing nuclear TDP-43. Confined to the nucleus, TDP-43WT remains absent from these poly(GR) aggregates. Last, the fact that blocking the RNA-binding ability of TDP-43 disrupts its interaction with ataxin-2 (44) and its localization to SGs (43), but does not prevent its sequestration to poly(GR)-induced aggregates immunopositive for SG proteins, provides compelling evidence that the recruitment of TDP-43 to poly(GR) inclusions is mediated by a direct interaction, rather than its co-aggregation, with cytoplasmic SGs.

Considering that c9ASOs targeting G4C2 repeat transcripts are currently being tested in clinical trials, our data showing that c9ASO treatment in (G4C2)149 mice mitigates poly(GR) burden, TDP-43 pathology, the aberrant SG response, and neurodegeneration are especially relevant. Moreover, we demonstrate that plasma NFL, which is abnormally elevated in (G4C2)149 mice and reduced with c9ASO treatment, also correlates with neuronal loss and TDP-43 pathology in (G4C2)149 mice. These results indicate that alterations in plasma NFL concentrations are both reflective of the severity of endogenous TDP-43 pathology in the (G4C2)149 model and responsive to treatment.

Along with the insights provided by our findings, it is important to acknowledge the limitations of the current study. Although we and others have observed substantial colocalization between TDP-43

**8 of 11**
and poly(GR) in patients with c9FTD/ALS using TDP-43 antibodies targeting the N-terminal region and phosphorylated serine-409/410 sites (28), additional studies using an array of antibodies targeting different epitopes of the TDP-43 protein in multiple brain regions from a large cohort of c9FTD/ALS patients are warranted. Next, we demonstrated the promising effects of c9ASO treatment in (G₄C₂)₁₄₉ mice. However, as all G₄C₂ repeat–associated pathologies are decreased by c9ASO treatment, additional studies are needed to determine whether the protective effect is mediated by reductions in poly(GR), poly(GA), or G₄C₂ repeat–containing RNA transcripts. In addition, considering that c9ASOs were administered relatively early in the disease course, before substantial pTDP-43 deposition in the (G₄C₂)₁₄₉ mouse model (13), it is unclear whether similar treatment efficacy would be observed if c9ASOs were delivered at a later time point.

In conclusion, we established that poly(GR) potently promotes TDP-43 aggregation through protein-protein interactions (independent of RNA binding and its C-terminal region), sequesters SG-resident proteins and nucleocytoplasmic transport factors, and thus drives TDP-43 proteinopathy in vivo. These pathological features, as well as neuronal loss, can be alleviated by c9ASO treatment. Together, the results of this study establish poly(GR) as a mechanistic link between the C9orf72 repeat expansion and TDP-43 proteinopathy, and show that decreasing poly(GR) with G₄C₂-targeted therapeutics may also be used to assess efficacy of disease-modifying therapeutics in preclinical models to ultimately guide and facilitate the successful translation of future therapies to humans.

**SUPPLEMENTARY MATERIALS**

**Table S1.** cDNA sequence of the GFP-(GR)₂₀₀ and mCherry-(GR)₁₀₀ plasmids.

**Table S2.** Characteristics of patients with c9FTD/ALS.

**Table S3.** Analysis of plasma NFL concentrations with disease severity and brain atrophy.

**Table S4.** Primers for qPCR.

**Figure S1.** Polyclonal antibody for mouse TDP-43.

**Figure S2.** Polyclonal antibody for human TDP-43.

**Figure S3.** Representative examples of in vivo experiments to assess the ability of poly(GR) to drive accumulation of endogenous mouse TDP-43.

**Figure S4.** Western blot analysis of endogenous mouse TDP-43.

**Figure S5.** G4C2 repeat would affect TDP-43 pathology and neurodegeneration in vivo.

**Figure S6.** G4C2 repeat–associated pathologies, as well as nucleocytoplasmic transport factors and NPC protein abnormalities were examined by biochemical, electron microscopy, immunochemistry, immuno-fluorescence, FISH, PLA, and immunoassays. Mouse and human samples were randomly selected for analyses. Image analysis was performed in a blinded or unblinded fashion, as detailed in individual experiments. Quantification was performed in a blinded manner. Data from cell culture studies are based on three independent experiments. Group sizes for in vivo studies vary, but are noted in the figure legend. Sample sizes were adequately powered to observe the effects based on previous reports (13, 14, 45–47).

**Statistics**

All statistical analyses were performed in GraphPad Prism. Data are presented as mean ± SEM. Data were analyzed by one-way or two-way analysis of variance (ANOVA) followed by Tukey’s post hoc analysis, except analyses with two groups were analyzed with two-tailed unpaired t test (Figs. 3B and 5E and fig. S3F), and association analyses were performed by determining the Pearson’s correlation coefficient (Fig. 5, H and I). Specific tests are also noted in each figure legend. P < 0.05 is considered statistically significant.

**REFERENCES AND NOTES**


Acknowledgments: We are grateful to all patients who agreed to donate postmortem tissue. We thank W.-L. Lin for IEM study. We thank S. Boevoena for providing the pGR12 and pGR10 peptides. Funding: This work was supported by the NIH (R35NS097273 to LP; P01NS084974 to D.W.D., T.F.G., Y.-J.Z., and LP; P01NS099114 to T.F.G. and LP; R01NS086869 to LP; and R21AG065854 and RO1GM099836 to J.S.). Mayo Clinic Foundation (to LP), Amoyotic Lateral Sclerosis Association (to T.F.G., L.P., J.S., and Y.-J.Z.), Robert Packard Center for ALS Research at Johns Hopkins (to LP and J.S.), Target ALS Foundation (to T.F.G., L.P., J.S., and Y.-J.Z.), Biogen Idec (to LP), and AstraZeneca postdoctoral fellowship (to H.M.O.).

Author contributions: L.P., Y.-J.Z., J.S., C.N.C., Y.W., and H.M.O. contributed to the conception and design. C.N.C., Y.W., G.d.R., M.Y., J.T., N.M.A., M.C.-C., W.S., G.S.T., A.M., and Y.-J.Z. contributed to mouse and cell studies. P.J. contributed to IEM study. H.M.O. and E.G. conducted the in vitro prion-like domains. M.O. and D.W.D. contributed to the collection of human samples. We thank W.-L. Lin for IEM study. We thank S. Boevoena for providing the pGR12 and pGR10 peptides. Funding: This work was supported by the NIH (R35NS097273 to LP; P01NS084974 to D.W.D., T.F.G., Y.-J.Z., and LP; P01NS099114 to T.F.G. and LP; R01NS086869 to LP; and R21AG065854 and RO1GM099836 to J.S.). Mayo Clinic Foundation (to LP), Amoyotic Lateral Sclerosis Association (to T.F.G., L.P., J.S., and Y.-J.Z.), Robert Packard Center for ALS Research at Johns Hopkins (to LP and J.S.), Target ALS Foundation (to T.F.G., L.P., J.S., and Y.-J.Z.), Biogen Idec (to LP), and AstraZeneca postdoctoral fellowship (to H.M.O.).

Competing interests: B.O. has consulted for Biogen, Medicinova, Mitsubishi, Amylyx, and Tsumura. F.R. provided c9ASO. B.O. and D.W.D. contributed to the collection of human samples. L.P., Y.-J.Z., C.N.C., Y.W., G.d.R., M.O., H.M.O., and J.S. analyzed data and wrote the manuscript.

**Supplementary Materials for**

*C9orf72* poly(GR) aggregation induces TDP-43 proteinopathy

Casey N. Cook, Yanwei Wu, Hana M. Odeh, Tania F. Gendron, Karen Jansen-West, Giulia del Rosso, Mei Yue, Peizhou Jiang, Edward Gomes, Jimei Tong, Lillian M. Daughrity, Nicole M. Avendano, Monica Castanedes-Casey, Wei Shao, Björn Oskarsson, Giulio S. Tomassy, Alexander McCampbell, Frank Rigo, Dennis W. Dickson, James Shorter*, Yong-Jie Zhang*, Leonard Petrucelli*

*Corresponding author. Email: petrucelli.leonard@mayo.edu (L.P.); zhang.yongjie@mayo.edu (Y.-J.Z.); jshorter@penne medicine.upenn.edu (J.S.)

DOI: 10.1126/scitranslmed.abb3774

The PDF file includes:

Materials and Methods
Fig. S1. Poly(GR) accelerates and enhances TDP-43 aggregation.
Fig. S2. Poly(GR) mediates sequestration of cytosolic full-length TDP-43 into the inclusions.
Fig. S3. TDP-43 and DPR burden in mouse and human.
Fig. S4. Poly(GR) aggregates sequester nuclear pore POM121 protein in vivo.
Fig. S5. c9ASO reduces G4C2 repeat–containing RNA and sense DPR protein burden.
Table S1. cDNA sequence of the GFP-(GR)200 and mCherry-(GR)100 plasmids.
Table S2. Primary antibodies for Western blot, immunohistochemistry, and immunofluorescence staining.
Table S3. Characteristics of patients with c9FTD/ALS.
Table S4. Primers for qPCR.
Legend for Data file S1
References (48, 49)

Other Supplementary Material for this manuscript includes the following:

(available at stm.scientemag.org/cgi/content/full/12/559/eabb3774/DC1)

Data file S1 (Microsoft Excel format). Raw data for all the quantitative figures where *n* < 20.
Materials and Methods

Reagents


TDP-43-MBP protein purification

Plasmid encoding human TDP-43 with a C-terminal MBP tag (TDP-43-TEV-MBP-6×His) was purchased from Addgene (Plasmid # 104480). Recombinant TDP-43 was purified as described (46). Briefly, protein was expressed in E. coli BL21-CodonPlus (DE3)-RIL cells (Agilent). Cell cultures were grown to an OD

In vitro aggregation assay

To measure aggregation kinetics, purified TDP-43 was first thawed and buffer exchanged into 20 mM HEPES-NaOH (pH 7.4), 150 mM NaCl and 1 mM DTT using a Micro Bio-Spin P-6 Gel Column (Bio-Rad). Protein concentration was measured by nanodrop, and TDP-43 was then diluted to a final concentration of 5 µM [in 20 mM HEPES (pH 7.0), 150 mM NaCl, 1 mM DTT], with addition of 2 µM poly(GR), poly(GA) or equivalent volume of 1× PBS as a control. At time = 0 minutes, aggregation was initiated by cleavage of the MBP tag using 1 µg/ml TEV protease, and monitored via turbidity measurements at an absorbance of 395 nm using a TECAN M1000 plate reader. Values were normalized to TDP-43 + TEV protease alone to determine the relative extent of aggregation. The area under each turbidity curve was calculated using GraphPad Prism. To visualize aggregate formation, aggregation was initiated by adding 10
µg/ml TEV protease to samples. After 30 minutes, 10 µl of sample was mounted onto a glass slide and imaged by differential interference contrast (DIC) microscopy.

**In vitro sedimentation assay**

At the end point of the turbidity assay described above (t = 16 hours), samples were sedimented by centrifugation for 10 minutes at 21,130 x g. Pellet and supernatant were immediately separated and equal volumes of each fraction were analyzed by SDS-PAGE. Proteins were visualized by Coomassie stain.

**Transmission electron microscopy**

Transmission electron microscopy (TEM) was performed as previously described (45). Briefly, at the start of the aggregation assay described above (t = 0), and at the end point of the turbidity assay (t = 16 hours), 10 µl of each sample was adsorbed onto a 300-mesh Formvar/carbon-coated copper grids (Electron Microscopy Sciences) and stained with 2% (w/v) uranyl acetate. Excess uranyl acetate was washed with water and the grids were allowed to air dry. Samples were viewed and imaged using JEOL 1010 transmission electron microscope. Electron micrographs were quantified using ImageJ. Images were inverted to have a black background. Scale was set to 110 pixels/µm, based on the scale bar of the micrographs. Threshold was used to determine the region of interest (ROI), which was considered as the total area occupied by TDP-43 aggregates. Total area (µm²), mean gray value (a.u.), and integrated density were measured, limited to thresholded area. The integrated density, which is the product of the area of the ROI of TDP-43 aggregates and mean gray value of pixel intensity, was reported. For each condition, 9 representative micrographs were quantified, collected from 3 independent experiments.

**Generation of plasmids**

To generate GFP-(GR)200 and mCherry-(GR)100 plasmids, a gene fragment containing 50 repeats of the dipeptide GR was synthesized by GeneArt and used as a template for PCR to generate fragments containing one or two Type IIS restriction enzyme sites. These fragments were ligated together sequentially to generate 100 or 200 GR repeats. The coding sequence for EGFP was cloned into the modified AAV packaging vector [pAM/CBA-pl-WPRE-BGH (“pAAV”)] containing the CMV-enhanced chicken β-actin promoter, and (GR)200 was ligated in frame downstream of EGFP to generate pAAV EGFP-(GR)200. The mCherry coding sequence was amplified using PCR and cloned into the AgeI and HindIII sites of pEGFP-C1, creating the vector mCherry-C1. A (GR)100 fragment was ligated in frame downstream of the mCherry to generate mCherry-(GR)100. The sequence of GFP-(GR)200 and mCherry-(GR)100 are listed in Table S1, respectively. To generate TDP-43-Myc plasmids, complementary DNA (cDNA) from constructs for wild-type TDP-43 or for TDP-43 with mutations in the nuclear localization signal (NLS) and/or the RNA recognition motifs (RRMs) (47, 48) was used as the PCR template to generate TDP-43 fragments. These fragments were cloned into the HindIII and XhoI sites of
pAG3-Myc vector. The sequences of all plasmids were verified by sequence analysis.

Cell culture and treatments

HEK293T cells were grown in Opti-Mem plus 10% FBS and 1% penicillin–streptomycin. Cells grown in 6-well plates or on glass coverslips in 24-well plates were transfected with the indicated plasmids using Lipofectamine 2000 (Thermo Fisher Scientific). Cells were harvested or fixed for Western blot and immunofluorescence staining, respectively, 24 hours post-transfection.

Immunofluorescence staining and quantification in cultured cells

Fixed cells were permeabilized with 0.5% Triton X-100 for 10 minutes, blocked with 5% nonfat dry milk in PBS for 1 hour, then incubated with primary antibody (Table S2) overnight at 4°C. After washing, cells were incubated with corresponding Alexa Fluor 488-, 568- or 647-conjugated donkey anti-species antibodies (1:500 or 1:1000, Molecular Probes) for 2 hours. Hoechst 33258 (1 µg/ml, H3569, Thermo Fisher Scientific) was used to stain cellular nuclei. Images were obtained on a Zeiss LSM 880 laser scanning confocal microscope. To quantify the percentage of cells with cytoplasmic TDP-43 inclusions in GFP- or poly(GR)-positive cells, the number of TDP-43-positive cells containing diffuse or aggregated TDP-43 was counted in a blinded fashion from 3 independent experiments (~500–1000 cells were counted for each group in each independent experiment). To quantify the percentage of cells with cytoplasmic TDP-43 inclusions in poly(GA)-positive cells, the number of TDP-43-positive cells containing diffuse or aggregated TDP-43 was counted in a blinded fashion from 3 independent experiments (~110–190 poly(GA) inclusions were counted for each group in each independent experiment).

Preparation of cell lysates

Cell pellets were lysed in co-immunoprecipitation (co-IP) buffer (50 mM Tris–HCl, pH 7.4, 300 mM NaCl, 1% Triton X-100, 5 mM EDTA) plus 2% SDS, and both protease and phosphatase inhibitors, sonicated on ice, and then centrifuged at 16,000 × g for 20 minutes. Supernatants were saved as cell lysates. The protein concentration of lysates was determined by BCA assay (Thermo Fisher Scientific), and samples were then subjected to Western blot analysis.

Western blot analysis

Cell lysates were diluted with 2× SDS-loading buffer at a 1:1 ratio (v/v), and then heated at 95°C for 5 minutes. Afterwards, equal amounts of protein were loaded into 10-well 4–20% Tris-glycine gels (Novex). After transferring proteins to PVDF membranes, membranes were blocked with 5% nonfat dry milk in TBS plus 0.1% Tween 20 (TBST) for 1 hour, and then incubated with primary antibody (Table S2) overnight at 4°C. Membranes were washed in TBST and incubated with donkey anti-rabbit or anti-mouse IgG antibodies conjugated to horseradish peroxidase (1:5000; Jackson ImmunoResearch) for 1 hour. Protein expression was
visualized by enhanced chemiluminescence treatment and exposure to film. The intensity of bands was quantified by FUJI FILM MultiGauge Software, and then normalized to the corresponding controls.

**Proximity ligation assay (PLA)**

HEK293T cells grown in an 8-well chamber slide (ibidi) were transfected with the indicated plasmids using Lipofectamine 2000 (Life Technologies). Twenty-four hours after transfection, cells were fixed for the PLA study using the Duolink In Situ kit (DUO92004, DUO92002, Sigma-Aldrich) per the manufacturer’s protocol. In brief, fixed cells were permeabilized with 0.1% Triton X-100 for 10 minutes, blocked using Duolink blocking buffer for 1 hour at room temperature, and then incubated with rabbit polyclonal anti-GR (Rb7810, 1:2000) (10) and mouse monoclonal anti-Myc (MA1-980, 1:1000, Invitrogen) overnight at 4°C. After washing with PBS plus 0.05% Tween 20, cells were incubated with the PLA probes MINUS and PLUS (1:5 dilution) for 1 hour at 37°C. After washing in 1x wash buffer A, cells were incubated with the ligase (1:40 dilution) in ligation buffer for 30 minutes at 37°C. Cells were washed with 1x wash buffer A, and then inoculated with polymerase (1:80 dilution) in amplification buffer for 90 minutes at 37°C. Cells were washed in 1x wash buffer B, followed by 0.01x wash buffer B. Hoechst 33258 (1 µg/ml, H3569, Thermo Fisher Scientific) was used to stain cellular nuclei. Images were obtained on a Zeiss LSM 880 laser scanning confocal microscope. To quantify the percentage of cells with PLA signal, the number of GFP-(GR)100-positive cells containing PLA signal was counted in a blinded fashion from 14-18 images.

**Immuno- electron microscopy**

To examine the ultrastructure of poly(GR) aggregates, immuno- electron microscopy (IEM) was performed as previously described (21). Rabbit polyclonal anti-poly(GR) antibody (7810, 1:20) (10) was used as a primary antibody and goat anti-rabbit IgG conjugated with 18 nm colloidal gold particles (1:20, Jackson ImmunoResearch Laboratories) was used as the secondary antibody. Thin sections stained with uranyl acetate and lead citrate were examined with a Philips 208S electron microscope (FEI) fitted with a Gatan 831 Orius CCD camera (Gatan).

**Human tissues**

Post-mortem hippocampal and frontal cortical tissues from frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) patients with the C9orf72 repeat expansion were obtained from the Mayo Clinic Florida Brain Bank. Information on human patients is provided in Table S3. Written informed consent was obtained before study entry from all subjects or their legal next of kin if they were unable to give written consent, and biological samples were obtained with Mayo Clinic Institutional Review Board (IRB) approval.

**Animal studies**

All procedures using mice were performed in accordance with the National Institutes of
Health Guide for Care and Use of Experimental Animals and approved by the Mayo Clinic Institutional Animal Care and Use Committee (IACUC).

**Virus production**

rAAV9 virus was produced as previously described (13, 14, 29). Briefly, AAV vectors expressing GFP, GFP-(GR)200, (G₄C₂)₂ or (G₄C₂)₁₄₉ were co-transfected with helper plasmids in HEK293T cells using polyethylenimine (23966, Polysciences, Inc.). Cells were harvested forty-eight hours following transfection, and lysed in the presence of 0.5% sodium deoxycholate and 50 Units/mL Benzonase (Sigma-Aldrich) by freeze-thawing. The virus was isolated using a discontinuous iodixanol gradient. The genomic titer of each virus was determined by qRT-PCR, and AAV solutions were diluted in sterile phosphate-buffered saline (PBS).

**Neonatal viral injections**

Intracerebroventricular (ICV) injections of virus were performed as previously described (13, 14, 29). Briefly, 2 µl (1×10¹⁰ genomes/µl) of AAV-GFP, AAV-GFP-(GR)₂₀₀, AAV-(G₄C₂)₂ or AAV-(G₄C₂)₁₄₉ solution was manually injected into each lateral ventricle of cryoanesthetized C57BL/6J mouse pups on postnatal day 0 (P0). Pups were allowed to recover from cryoanesthesia on a heating pad, and were then returned to the home cage with the mother.

**Antisense oligonucleotide (ASO) injections and sample collection**

PBS or c9ASO was injected into the central nervous system of 3 month-old AAV-(G₄C₂)₂ or AAV-(G₄C₂)₁₄₉ mice by means of stereotactic ICV injection, as previously described (34) with some minor modifications. The c9ASO targeting the G₄C₂ repeat expansion (CGGCCCGGCCCCCGGC), developed and provided by Ionis Pharmaceuticals, was aMOE-gaper ASO with 16 nucleotides in length, wherein the central gap segment comprising eight 2’-deoxyribonucleotides (DNA) that are flanked on the 5’ and 3’ wings by four 2’-O-methoxyethyl (MOE) modified nucleotides. Internucleotide linkages are phosphorothioate interspersed with phosphodiester, and all cytosine residues are 5’-methylcytosines. Specifically, 10 µl of PBS or c9ASO solution (corresponding to 350 µg ASOs) were delivered into the right lateral ventricle using the coordinates: 0 mm anterior and 1.0 mm lateral to the right from bregma, and 1.9-2.0 mm deep as measured from the brain surface. PBS-treated AAV-(G₄C₂)₂ mice (n = 17) or AAV-(G₄C₂)₁₄₉ mice (n = 18), and c9ASO-treated AAV-(G₄C₂)₁₄₉ (n = 12) mice were compared.

**Tissue processing**

For protein, immunostaining and RNA analyses, the mice were euthanized by CO₂ or ketamine/xylazine through intraperitoneal injection. Blood samples were then collected by cardiac puncture, and mice euthanized by exsanguination followed by transcardial perfusion with saline. Then, brains were harvested and cut sagittally across the midline. The brain was rapidly removed and hemisected. Sagittal half brains were immersion fixed in in 4% paraformaldehyde, embedded in paraffin, sectioned (5 µm thick), and then mounted on glass slides for immunofluorescence or immunohistochemistry staining. The other half brains were dissected
and frozen (cortex, hippocampus, subcortex, midbrain, brainstem, and cerebellum frozen separately).

**Immunohistochemistry staining**

Sagittal half brains fixed in 4% paraformaldehyde were embedded in paraffin, sectioned at 5 μm, and mounted on positively-charged glass slides. After drying overnight, paraffin sections were deparaffinized in xylene, and rehydrated through a series of ethanol solutions, followed by washing in dH2O. Antigen retrieval was performed by steaming slides in dH2O or Tris-EDTA (DAKO), pH 9.0 for 30 minutes followed by a 5 minute incubation in Dako Peroxidase Block (S2001, DAKO) to block endogenous peroxidase activity. To detect the sense DPR proteins, sections were immunostained with primary antibody (Table S2) using the DAKO Autostainer (Universal Staining System) and the DAKO+HRP system. To detect ataxin-2 or NeuN, slides were blocked with Dako Protein Block Serum-Free (X0909, DAKO) for 1 hour, and incubated with primary antibody (Table S2) for 45 minutes. After washing, sections were incubated for 30 minutes in Dako Envision-Plus anti-rabbit (K4003, DAKO) or anti-mouse (K4001, DAKO) labeled HRP polymer, respectively. Peroxidase labeling was visualized with the Liquid DAB + Substrate Chromogen System (K3468, DAKO).

To detect pTDP-43 in AAV-(G4C2)149 mice or poly(GR) in AAV-GFP-(GR)200 mice, slides were deparaffinized and rehydrated as described above, and antigen retrieval was performed by steaming in sodium citrate buffer (10 mM sodium citrate, 0.05% Tween-20, pH 6.0) for 30 minutes. After cooling and washing with dH2O, slides were incubated with Dako Dual Endogenous Enzyme Block (DAKO), and subsequently washed in PBS. Sections were then blocked with 2% normal goat serum for 1 hour, followed by an overnight incubation with primary antibody (Table S2) at 4°C. The next day, slides were washed with PBS, incubated with biotinylated goat anti-rabbit or rabbit anti-rat secondary (1:200) for 2 hours, and again washed in PBS. Slides were then incubated with avidin-biotin complex solution for 30 minutes, washed in PBS, and reacted with 3,3′-diaminobenzidine (Acros Organics) activated with hydrogen peroxide, with the reaction stopped by rinsing slides in dH2O.

Following labeling, all sections were counterstained with hematoxylin (Thermo Fisher Scientific), dehydrated through ethanol and xylene washes, and cover-slipped with Cytoseal mounting medium (Thermo Fisher Scientific). Slides were scanned with a ScanScope AT2 (Leica Biosystems), and representative images taken with ImageScope software (v12.1; Leica Biosystems).

**Immunofluorescence staining in mouse and human brains**

Paraffin sections (5 μm) of mouse and human brain tissues were deparaffinized, rehydrated, steamed for 30 minutes in Dako antigen retrieval solution, blocked with Dako All Purpose Blocker for 1 hour, and incubated with primary antibody (Table S2). After washing, sections were incubated with corresponding Alexa Fluor 488-, 568- or 647-conjugated donkey
anti-species (1:500, Molecular Probes) for 2 hours. Hoechst 33258 (1 μg/ml, Thermo Fisher Scientific) was used to stain cellular nuclei. Images were obtained on a Zeiss LSM 880 laser scanning confocal microscope.

Quantification of neuropathology

To quantify diffuse and aggregated poly(GR), high resolution digitized images of immunostained slides were obtained by using a ScanScope AT2 (Leica Biosystems). The cortex was annotated on mid-sagittal serial sections. The number of diffuse and aggregated poly(GR) was quantified manually in a blinded fashion. To quantify pTDP-43 inclusions, high resolution digitized images of immunostained slides were obtained by using a ScanScope AT2 (Leica Biosystems). The cortex was annotated on mid-sagittal serial sections. The number of pTDP-43 inclusions was quantified manually in a blinded fashion. To quantify TDP-43, KPNA2 and NUP98 pathology, non-transduced (NT), diffuse poly(GR) or aggregated poly(GR) cells exhibiting TDP-43, KPNA2 and NUP98 pathology were counted in a blinded fashion in the cortex of 2-week-old mice expressing GFP-(GR)200 (TDP-43: ~180–250 cells were counted per mouse; KPNA2: ~150–240 cells were counted per mouse; NUP98: ~140–170 cells were counted per mouse). To quantify abnormal co-localization of NPC with either TDP-43 or eIF3η, non-transduced (NT), diffuse poly(GR) or aggregated poly(GR) cells exhibiting abnormal co-localization of NPC with either TDP-43 or eIF3η were counted in a blinded fashion in the cortex of 2-week-old mice expressing GFP-(GR)200 (NPC and eIF3η: ~170–220 cells were counted per mouse; NPC and TDP-43: ~180–220 cells were counted per mouse). To quantify co-localization of poly(GA) with TDP-43 and eIF3η in the cortex of 3-month-old mice expressing GFP-(GA)50, ~90–120 poly(GA) inclusions were counted per mouse. To quantify co-localization of TDP-43 with poly(GA) or poly(GR), the total number of poly(GA) or poly(GR) inclusions, as well as the number of inclusions that were also positive for TDP-43 were counted in a blinded fashion in the hippocampus of c9FTD/ALS patients [poly(GA): ~90–470 inclusions were counted per patient; poly(GR): ~35–180 inclusions were counted per patient].

RNA Fluorescence In Situ Hybridization (FISH)

The RNA FISH protocol was performed as previously described (13). Briefly, tissue sections were deparaffinized and rehydrated through a series of xylene and ethanol solutions, permeabilized with ice cold 2% acetone/1× DEPC-PBS, washed with DEPC-H2O, and then dehydrated in ethanol. To detect sense RNA foci, sections were incubated with pre-hybridization buffer [50% formamide (Midsci), 10% dextran sulfate (Millipore), 2× saline-sodium citrate buffer (SSC), 50 mM sodium phosphate buffer pH 7.0] for 20–30 minutes at 66°C, and then hybridized for 24 hours at 66°C in a dark, humidified chamber with a fluorescently-labeled locked nucleic acid (LNA) probe (5) [TYE563-(CCCCGGCCCCGGGGCCCC); Exiqon product number 500150, design id: 283117] diluted to a final concentration of 40 nM. Next, sections were washed with 2× SSC/0.1% Tween-20 at room temperature for 5 minutes, and then washed
twice with pre-warmed 0.2× SSC at 60°C for 10 minutes in the dark. Following these washes, slides were coverslipped using Vectashield mounting media with DAPI (Vector Laboratories). Representative images of sense RNA foci in the cortex were taken with an AxioImager Z1 fluorescent microscope (Carl Zeiss MicroImaging). RNA foci burden was quantified in a blinded fashion by calculating the percentage of cells in the motor cortex (from 300–400 cells total) containing sense RNA foci.

**RNA extraction, reverse transcription and qPCR**

For RNA extraction, frozen hippocampi were homogenized in Trizol LS (250 μL), and total RNA extracted using the Direct-zol RNA MiniPrep kit (Zymo Research) according to manufacturer’s instructions. cDNA was then obtained following reverse transcription of 250 ng of the extracted RNA with random primers and the High Capacity cDNA Transcription Kit (Applied Biosciences). qRT-PCR was performed in triplicate for all samples using the SYBR green assay (Life Technologies) on an ABI Prism 7900HT Fast Real-Time PCR System (Applied Biosystems). As the AAV-(G₄C₂)₁₄₉ vector contains 119 base pairs of human flanking sequence 5’ of the G₄C₂ repeat, primers were designed targeting this region to assess mRNA expression of exogenous, AAV-derived G₄C₂-containing transcripts. Endogenous Gapdh and Rplp0 mRNA were also quantified. The sequences of the primers used for this study are listed in Table S4. Relative mRNA expression of the G₄C₂ 5’ flanking sequence was normalized to the geometric mean of the endogenous transcript controls, Gapdh and Rplp0.

**Immunoassay analysis of poly(GR), poly(GA), and poly(GP)**

Frozen cortex samples were homogenized in co-immunoprecipitation buffer (50 mM Tris–HCl, pH 7.4, 300 mM NaCl, 1% Triton X-100, 5 mM EDTA) containing protease and phosphatase inhibitors. After sonication, the lysates were centrifuged at 16,000 × g for 20 minutes, and BCA assay performed on the supernatant to determine protein concentration. Previously characterized Meso Scale Discovery (MSD) sandwich immunoassays were then used to detected poly (GR) (29), or poly(GP) and poly(GA) (49). In brief, lysates were diluted in Tris-buffered saline (TBS) and an equal amount of protein for all samples was tested in duplicate wells. Response values corresponding to the intensity of emitted light upon electrochemical stimulation of the assay plate using the MSD QUICKPLEX SQ120 were acquired and background corrected using the average response from lysates obtained from control (G₄C₂)₂ mice.

**Detection of neurofilament light (NfL) in plasma**

Blood samples from mice were collected in EDTA tubes, and centrifuged to obtain plasma. Plasma NfL concentrations were then determined using a Simoa NF-Light Advantage Kit (102258) run on the automated HD-1 Analyzer (Quanterix) per the manufacturer’s protocol. Briefly, plasma samples were diluted 1:4 at the bench, and subsequently transferred to 96-well plates along with calibrators, two quality control samples, and five interassay controls with a
range of known NfL concentrations. NfL concentrations were then interpolated from the standard curve using a 4 parameter logistic curve fit (1/y^2 weighted).
**Fig. S1. Poly(GR) accelerates and enhances TDP-43 aggregation.** (A) Sedimentation of TDP-43 (5 µM) in the absence or presence of 2 µM poly(GR) or poly(GA). Samples were sedimented at the end-point of the aggregation assay (t = 16 hours). (B) Representative electron micrographs of TDP-43 (5 µM) in the absence or presence of 2 µM poly(GR) or poly(GA), without TEV protease. Scale bar, 2 µm.
Fig. S2. Poly(GR) mediates sequestration of cytosolic full-length TDP-43 into the inclusions.
(A) Triple-immunofluorescence staining for GFP-(GR)$_{100}$, eIF3η and TIA-1 or Ataxin 2 in HEK293T cells expressing GFP-(GR)$_{100}$. Scale bars, 5 μm. (B) Triple-immunofluorescence staining for GFP, TDP-43-Myc and TIA-1 in HEK293T cells expressing GFP and Myc-tagged TDP-43 species. Scale bars, 5 μm. (C) Western blot confirming expression of GFP, GFP-(GR)$_{100}$ and TDP-43-Myc in HEK293T cells co-transfected with either GFP or GFP-(GR)$_{100}$ and various
Myc-tagged TDP-43 species. GAPDH was used to control for protein loading. * Indicates non-specific bands. (D) Densitometric analysis of Myc-tagged TDP-43 species in HEK293T cells co-expressing various Myc-tagged TDP-43 constructs with either GFP or GFP-(GR)_{100} (n = 3 independent experiments). (E) Triple-immunofluorescence staining for poly(GA), TDP-43 and TIA-1 in HEK293T cells expressing GFP-(GA)_{100} and Myc-tagged TDP-43 constructs [including wild-type (WT) and nuclear localization signal mutant (NLSm)] (n = 3 independent experiments). Scale bars, 5 μm. (F) Triple-immunofluorescence staining for GFP-TDP-43, mCherry-(GR)_{100}, and TIA-1 in HEK293T cells co-transfected with mCherry-(GR)_{100} and GFP-tagged TDP-43 species [including WT, NLSm, and the C-terminal fragment (CTF)]. Scale bars, 5 μm. Data shown as the mean ± SEM. **P = 0.0016, ****P < 0.0001, NS (left to right) P = 0.0590, and P = 0.0726, two-way ANOVA, Tukey’s post hoc analysis.
**Fig. S3. TDP-43 and DPR burden in mouse and human.** (A) Representative images of immunohistochemical analysis of poly(GR) in the cortex 3-month-old GFP or GFP-(GR)$_{200}$ mice (diffuse labeling noted by black arrows, aggregates indicated by black arrowheads). Scale bar, 20 µm. (B) Representative images of immunohistochemical analysis of TDP-43 in the cortex of 2-week-old GFP mice. Scale bar, 20 µm. (C) Quantification of the percentage of co-localization between poly(GR) and pTDP-43 in either non-transduced (NT) or transduced cells with diffuse or aggregated poly(GR) (n =6). (D) Triple-immunofluorescence staining for poly(GR), eIF3η, and TDP-43 in the hippocampus of c9FTD/ALS patients (see Table S3 for patient information). Scale bars, 5 µm. (E) Double-immunofluorescence staining for poly(GA) and TDP-43 in the hippocampus of c9FTD/ALS patients (see Table S3 for patient information). Scale bars, 5 µm. (F) Quantification of the percentage of co-localization between TDP-43 and either poly(GA) or poly(GR) in c9FTD/ALS patient tissue. Data shown as the mean ± SEM. In (B), **** $P < 0.0001$, one-way ANOVA, Tukey’s post hoc analysis. In (F), ** $P = 0.0064$, two-tailed unpaired t test.
Fig. S4. Poly(GR) aggregates sequester nuclear pore POM121 protein in vivo. Double-immunofluorescence staining for poly(GR) and POM121 in the cortex of 2-week-old GFP-(GR)₂₀₀ mice (n = 6). Scale bars, 5 μm.
Fig. S5. c9ASO reduces G₄C₂ repeat–containing RNA and sense DPR protein burden. (A) qRT-PCR analysis of exogenous G₄C₂ repeat RNA transcripts in (G₄C₂)₁₄₉ mice using human-specific primers 5’ of the repeat targeting sequence present in the AAV vector. Exogenous G₄C₂ repeat mRNA expression was normalized to the geometric mean of the endogenous controls Gapdh and Rplp0 (PBS, n = 18; c9ASO, n = 12). (B) Percentage of cells containing sense RNA foci in the motor cortex of PBS (n = 19) or c9ASO-treated (n = 12) (G₄C₂)₁₄₉ mice. (C) Poly(GA), poly(GP) and poly(GR) concentrations in brain lysates from PBS (n = 14) or c9ASO-treated (n = 12) (G₄C₂)₁₄₉ mice were measured by immunoassay. Data presented as the mean ± SEM. In (A), ** P = 0.0033, unpaired two-tailed t-test. In (B), ** P = 0.0066, unpaired two-tailed t-test. In (C), ** (left to right) P = 0.0045, P = 0.0089 and **** P < 0.0001, unpaired two-tailed t-test.
Table S1. cDNA sequence of the GFP-(GR)$_{200}$ and mCherry-(GR)$_{100}$ plasmids.

### GFP-(GR)$_{200}$ plasmid

| ATGGTGAGCACAAGGCGAGGAGGTGGTCTCACCAGGGGTTGCGCCCATCCTGTCGAGC |
| TGGACGCCAGACGATACACGCACACGAGGTGGTCGCCGAGGGCGAGGGCGAGAGG |
| ATGCCCCAGCTACGCAACCCATCTTCTTCTGAAAGATGCGAGTACGCACCACAG |
| TGGACACGCCACCCTTGAGAAAGGACGAACGGCAATTACAAGACCACC |
| GCGGAGGTGGTACGAGTGGCCAGAACACCTGTGTAACCCAGAGCTCAGTGAGG |
| TCAGCTTCAAGGAGGACGCAATCCCTGAGGAGGCAAAGTCGAGTAGAATAA |
| CAGCCCCAAGCTCTATATATCATGCCGAGACACGACGATCACAGTGAAC |
| TTCAAGACCGCACAACATCGAGACGCACGATCGATCGTCCGACACCACC |
| AGCAGAAGCCACCATCGGCCAGCCGAGCCTTGCTGCTGCCAGACACCAC |
| GAGCAGCAGCTCCGAGCCACAAGGACGAGTAGACGTGAGTGTCGAGG |

The (GR)$_{200}$ sequence is highlighted in yellow.

### mCherry-(GR)$_{100}$ plasmid

| ATGGTGAGCACAAGGCGAGGAGGTGGTCTCACCAGGGGTTGCGCCCATCCTGTCGAGC |
| TGGACGCCAGACGATACACGCACACGAGGTGGTCGCCGAGGGCGAGGGCGAGAGG |
| ATGCCCCAGCTACGCAACCCATCTTCTTCTGAAAGATGCGAGTACGCACCACAG |
| TGGACACGCCACCCTTGAGAAAGGACGAACGGCAATTACAAGACCACC |
| GCGGAGGTGGTACGAGTGGCCAGAACACCTGTGTAACCCAGAGCTCAGTGAGG |
| TCAGCTTCAAGGAGGACGCAATCCCTGAGGAGGCAAAGTCGAGTAGAATAA |
| CAGCCCCAAGCTCTATATATCATGCCGAGACACGACGATCACAGTGAAC |
| TTCAAGACCGCACAACATCGAGACGCACGATCGATCGTCCGACACCACC |
| AGCAGAAGCCACCATCGGCCAGCCGAGCCTTGCTGCTGCCAGACACCAC |
| GAGCAGCAGCTCCGAGCCACAAGGACGAGTAGACGTGAGTGTCGAGG |

mCherry-(GR)$_{100}$ is highlighted in yellow.
The (GR)\textsubscript{100} sequence is highlighted in yellow.
Table S2. Primary antibodies for Western blot, immunohistochemistry, and immunofluorescence staining.

<table>
<thead>
<tr>
<th>Western blot</th>
<th>Antibody</th>
<th>Species</th>
<th>Dilution</th>
<th>Number</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>anti-GR</td>
<td>rabbit</td>
<td>1:2000</td>
<td>Rb7810&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Life Technologies</td>
</tr>
<tr>
<td></td>
<td>anti-GFP</td>
<td>rabbit</td>
<td>1:4000</td>
<td>A-6455</td>
<td>Invitrogen</td>
</tr>
<tr>
<td></td>
<td>anti-Myc</td>
<td>mouse</td>
<td>1:1000</td>
<td>MA1-980</td>
<td></td>
</tr>
<tr>
<td></td>
<td>anti-GAPDH</td>
<td>mouse</td>
<td>1:5000</td>
<td>H86504M</td>
<td>Meridian Life Science</td>
</tr>
</tbody>
</table>

**Immunohistochemistry**

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Dilution</th>
<th>Number</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-GA</td>
<td>rabbit</td>
<td>1:5000</td>
<td>Rb9880&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>anti-GP</td>
<td>rabbit</td>
<td>1:10000</td>
<td>Rb5823&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>anti-GR</td>
<td>rabbit</td>
<td>1:2500</td>
<td>Rb7810&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>anti-pTDP-43</td>
<td>rabbit</td>
<td>1:1000</td>
<td>Rb3655&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>anti-TDP-43</td>
<td>rabbit</td>
<td>1:1000</td>
<td>MC2079&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>anti-GR</td>
<td>rat</td>
<td>1:250</td>
<td>MABN778</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>anti-Ataxin 2</td>
<td>rabbit</td>
<td>1:500</td>
<td>21776-1-AP</td>
<td>Proteintech</td>
</tr>
</tbody>
</table>

**Immunofluorescence**

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Dilution</th>
<th>Number</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-Myc</td>
<td>mouse</td>
<td>1:1000</td>
<td>MA1-980</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>anti-TIA-1</td>
<td>rabbit</td>
<td>1:2000</td>
<td>ab40693</td>
<td>Abcam</td>
</tr>
<tr>
<td>anti-eIF3η</td>
<td>goat</td>
<td>1:200</td>
<td>sc-16377</td>
<td>Santa Cruz Biotechnology</td>
</tr>
<tr>
<td>anti-GR</td>
<td>rabbit</td>
<td>1:2000</td>
<td>Rb7810&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Cosmo Bio</td>
</tr>
<tr>
<td>anti-GR</td>
<td>rabbit</td>
<td>1:500</td>
<td>MABN778</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>Anti-pTDP-43</td>
<td>rabbit</td>
<td>1:1000</td>
<td>CAC-TIP-PTD-P02</td>
<td>Cosmo Bio</td>
</tr>
<tr>
<td>anti-Ataxin 2</td>
<td>rabbit</td>
<td>1:500</td>
<td>21776-1-AP</td>
<td>Proteintech</td>
</tr>
<tr>
<td>anti-NPC</td>
<td>mouse</td>
<td>1:100</td>
<td>ab24609</td>
<td>Abcam</td>
</tr>
<tr>
<td>anti-NUP98</td>
<td>rat</td>
<td>1:200</td>
<td>ab50610</td>
<td>Abcam</td>
</tr>
<tr>
<td>anti-POM121</td>
<td>rabbit</td>
<td>1:100</td>
<td>PA5-36498</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>anti-RanGAP1</td>
<td>rabbit</td>
<td>1:100</td>
<td>sc-25630</td>
<td>Santa Cruz Biotechnology</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Dilution</th>
<th>Number</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>anti-GA</td>
<td>mouse</td>
<td>1:500</td>
<td>MABN889</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>anti-GFP</td>
<td>mouse</td>
<td>1:1000</td>
<td>33-2600</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>anti-TDP-43</td>
<td>rabbit</td>
<td>1:1000</td>
<td>MC2079&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Importin α5</td>
<td>rabbit</td>
<td>1:100</td>
<td>18137-1-AP</td>
<td>Proteintech</td>
</tr>
<tr>
<td>KPNA2</td>
<td>rabbit</td>
<td>1:100</td>
<td>0819-1-AP</td>
<td>Proteintech</td>
</tr>
</tbody>
</table>

<sup>a, d</sup> Antibody described in: T. F. Gendron et al., Acta Neuropathol 126, 829-844 (2013).
Table S3. Characteristics of patients with c9FTD/ALS.

<table>
<thead>
<tr>
<th>Case #</th>
<th>Pathological Diagnosis</th>
<th>Gender</th>
<th>Age at Onset</th>
<th>Age at death</th>
<th>Disease Duration</th>
<th>C9orf72 repeat expansion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FTLD/ALS</td>
<td>F</td>
<td>60.4</td>
<td>61.4</td>
<td>1.0</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>FTLD</td>
<td>M</td>
<td>62</td>
<td>73.0</td>
<td>11.0</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>FTLD/ALS</td>
<td>F</td>
<td>52</td>
<td>60.3</td>
<td>8.3</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>FTLD/ALS</td>
<td>M</td>
<td>57</td>
<td>62.2</td>
<td>5.2</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>FTLD</td>
<td>M</td>
<td>68</td>
<td>73.9</td>
<td>5.9</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>FTLD/ALS</td>
<td>M</td>
<td>57</td>
<td>62</td>
<td>6.0</td>
<td>Yes</td>
</tr>
</tbody>
</table>

FTLD, frontotemporal lobar degeneration; ALS, amyotrophic lateral sclerosis
<table>
<thead>
<tr>
<th>Model</th>
<th>Target</th>
<th>Primers</th>
</tr>
</thead>
</table>
| mouse brain   | 5’flanking sequence | 5’-TAGTACTCGCTGAGGGTGAAC-3’  
                     |              | 5’-CTACAGGCTGCCTGTTGTTTC-3’   |
| mouse brain   | Gfap        | 5’-CATGGCCCTTCGGTTCCTA-3’  
                     |              | 5’-CCTGCTTCACCACCTTCTTGTGAT-3’ |
| mouse brain   | Rplp0       | 5’-ACTGGTCTAGGACCCGAGAAG-3’  
                     |              | 5’-CTCCACCTTGCTCCAGTC-3’     |
Data file S1. Raw data for all the quantitative figures where $n < 20$. Provided as a separate Excel file.