Supporting Information

Zu et al. 10.1073/pnas.1315438110

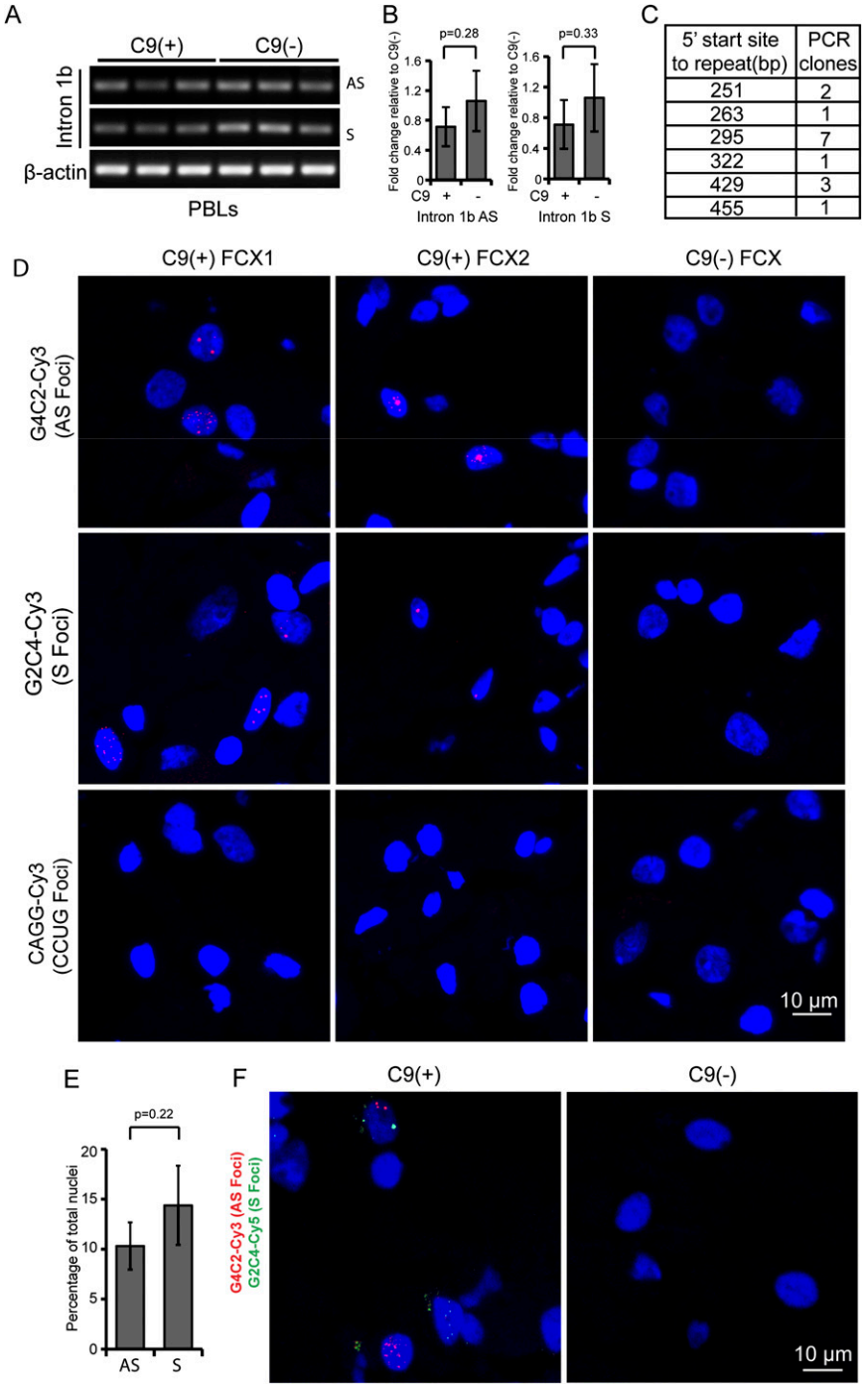

Fig. S1. (*A*) Strand-specific RT-PCR detection of intron 1b S and antisense (AS) transcripts of peripheral blood leukocytes (PBLs) of three *chromosome 9 ORF 72* (*C9ORF72*) expansion-positive [C9(+)] and three C9(–) controls. (*B*) Quantitative RT-PCR of intron 1b AS and S transcripts in C9(+) ($n = 3$) and C9(–) ($n = 3$) controls. (*C*) Summary of 5′ RACE products. (*D*) FISH staining of frontal cortex (FCX) from two representative C9(+) cases showing both sense (G_2C_4) and antisense (G_2C_4) RNA foci, detected by G_4C_2-Cy3 and G_2C_4-Cy3 probes, respectively. No similar staining was found in C9(–) FCX or using a CAGG-Cy3 control probe. (*E* and *F*) Double-labeling of C9(+) FCX shows sense (G_2C_4-Cy5 probe, green) and antisense (C_4C_2-Cy3 probe, red) RNA foci are commonly found in different cells, and occasionally sense and antisense foci are found in the same cell (example of cells with sense, antisense, and both types of foci is shown as an example). No similar staining is found in C9(–) FCX tissue.

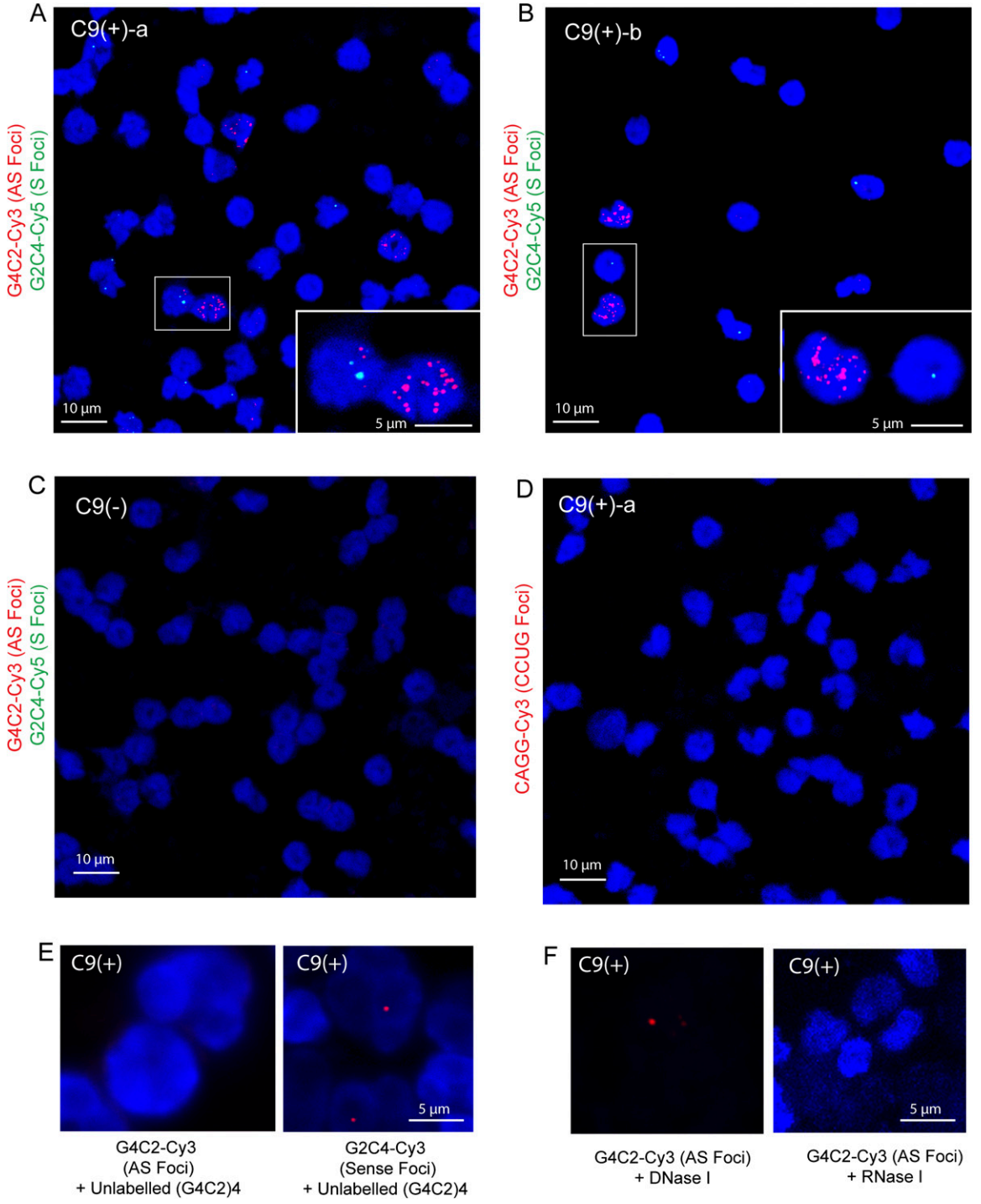

Fig. S2. FISH staining of PBLs showing the accumulation of antisense G_2C_4 and sense G_4C_2 RNA foci in C9(+) but not C9(−) cells. (*A*–*C*) Double-labeling of sense (G_2C_4-Cy5 probe, green) and antisense RNA (C_4C_2-Cy3 probe, red) RNA foci in cells from two C9(+) patients [C9(+)a and C9(+)b] but not C9(−) PBLs. (*D*) No similar staining was seen in C9(+) samples probed with a CAGG-Cy3 control probe. (*E*) Antisense foci specificity assay showing excess unlabeled $(G_4C_2)_4$ oligo blocks labeling of G_4C_2-Cy3 antisense but not G_2C_4-Cy3 labeled sense foci. (*F*) Additional controls for antisense RNA foci showing expected DNase I resistance and RNase I sensitivity.

G_4C_2 (sense strand)

Sense frame 1 (GP_S)
GPGPGPGPGPGPGPGPGP(GP)$_{EXP}$GPGPGRGRGGPGGGPGAGLRLRCLRPRRRRRRRRWRVGE

Sense frame 2 (GR_S)
*RLTRRKQGGKQPQPVASSGTQESRARGRGRGRGRGRGRGRGRGR(GR)$_{EXP}$GRGRGVVGAGPGAG
PGRGCGCGACARGGGGAGGGEWVSEEAASWRVAVWGSAAGKRRG*

Sense frame 3 (GA_S)
*QALELRSRALGAGAGAGAGAGAGAGAGAGA(GA)$_{EXP}$GAGAWSGRARGRARGGAAVAVPAPAAAE
AQAVASG*

G_2C_4 (antisense strand)

Antisense frame 1 (PA_{AS})
*GEPPLLPAPLPGSRTPNSHPPGCRLLTHPLATACASAAAGAGTATAAPPRARPRARPDHAPAPA
PAPAPAPAPAPAPA(PA)$_{EXP}$PAPAPSARLLSSRACYRLRLFPSLFSSG*

Antisense frame 2 (PR_{AS})
MQAIPPVARGESPTPSFGQRNERESKNASSSEESPRFYPRLFPAAEPQTATRQDAASSLTHSPPP
APPPPRAQAPQPQPRPGPAPGPAPTTPRPRPRPRPRPRPRPRPR(PR)$_{EXP}$PRPRPLARDS*

Antisense frame 3 (GP_{AS})
MRGKVKMRRALRRAPASTRASSRQPNPKQPPARMPPPHSPTRHRLRLRRRGRRHRNRSPAPGPPP
GPPRPRPGPGPGPGPGPGPGPGP(GP)$_{EXP}$GPGP*

Fig. S3. Putative protein products in sense and antisense directions for all six reading frames. Repeat motifs are highlighted in red. Peptides used to generate polyclonal antibodies to the various repeat motifs or unique C-terminal flanking sequences are underlined. Possible upstream AUG start sites (M) are shown in red for antisense frame 2 and frame 3. Stop codons are indicated by an asterisk.

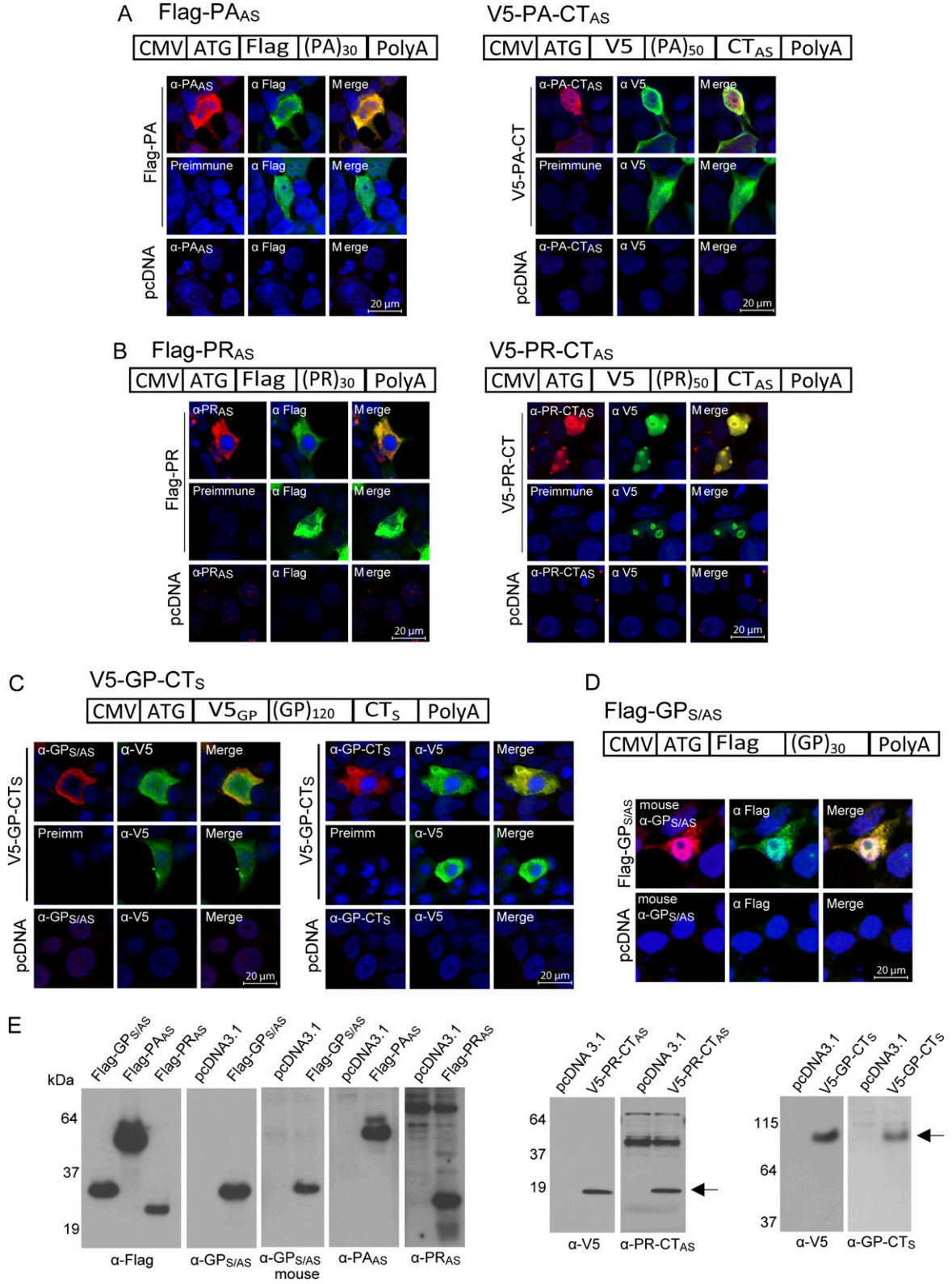

Fig. S4. Validation of dual antibodies to detect putative polyPA, polyPR, and polyGP proteins by immunofluorescence (IF) and protein blot. (*A–D*, *Upper*) Schematic diagrams of constructs expressing ATG-initiated N-terminal epitope-tagged (V5 or Flag) repeat proteins with or without endogenous C-terminal sequences. (*A–D*, *Lower*) Colocalization of α-Flag or α-V5 (green) staining in transfected HEK293T cells with staining (red) using the following newly developed antibodies: (*A*) α-PA$_A$S or α-PA-CT$_{AS}$; (*B*) α-PR$_{AS}$ or α-PR-CT$_{AS}$; (*C*) rabbit α-GP$_{S/AS}$ or α-GP-CT$_S$; (*D*) mouse α-GP$_{S/AS}$. Similar staining was not seen in preimmune or pcDNA3.1 empty vector controls. (*E*) Corresponding immunoblots showing six of the seven antibodies tested also detect recombinant proteins by Western blot.

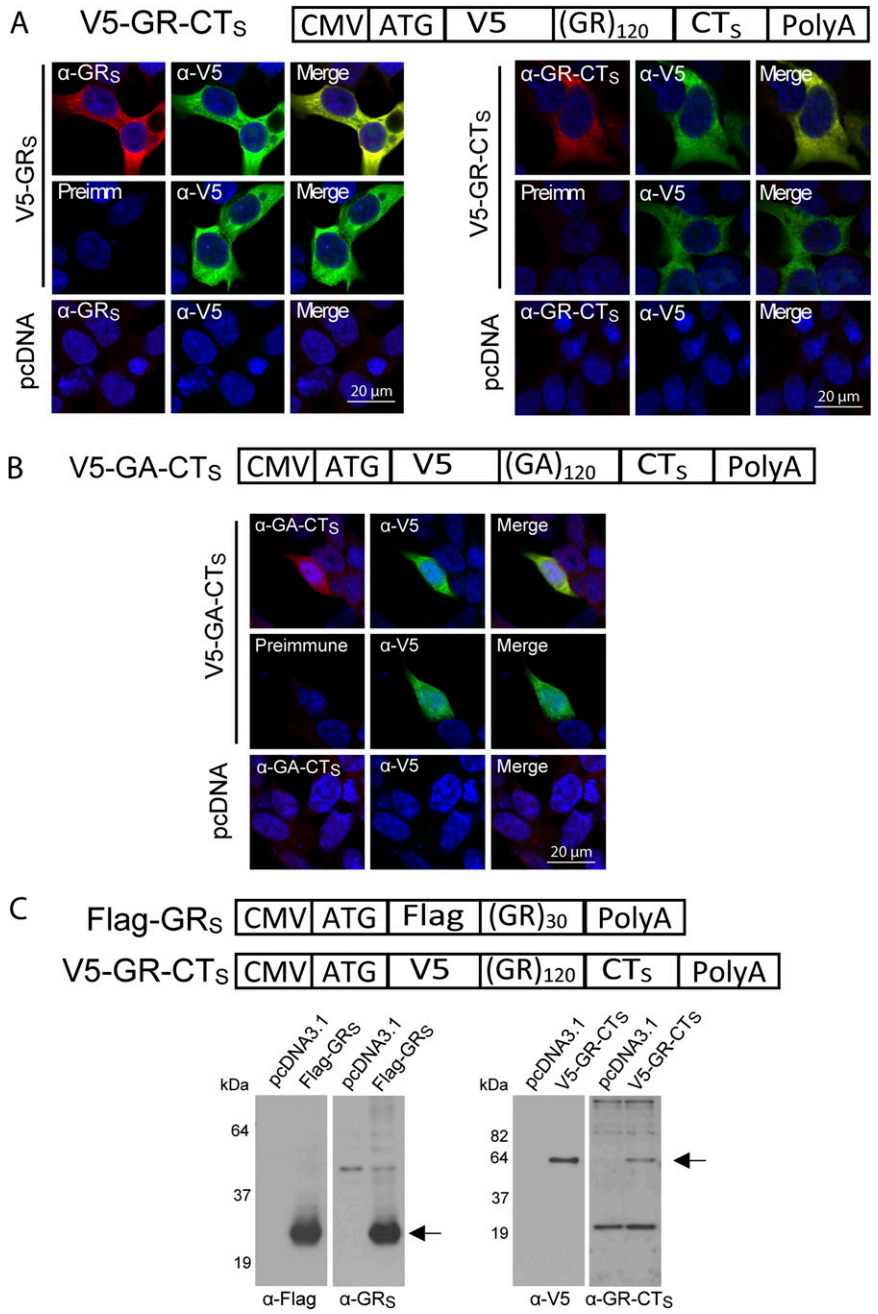

Fig. S5. Validation of additional sense repeat and C-terminal polyclonal antibodies. (*A* and *B*, *Upper*) Schematic diagrams of constructs expressing ATG-initiated N-terminal V5-epitope tagged GR or GA repeat proteins with endogenous C-terminal sequences in the sense (s) direction (V5-GR-CT_S and V5-GA-CT_S). (*A* and *B*, *Lower*) Colocalization of α-V5 (green) staining in transfected HEK293T cells with α-GR_S, α-GR-CT_S, and α-GA-CT_S, respectively. Similar staining was not seen in corresponding preimmune or pcDNA3.1 empty vector controls. (*C*) Immunoblots showing novel α-GR_S and α-GR-CT_S antibodies detect recombinant protein (indicated by arrows) in Flag-GR_S transfected cells by protein blot.

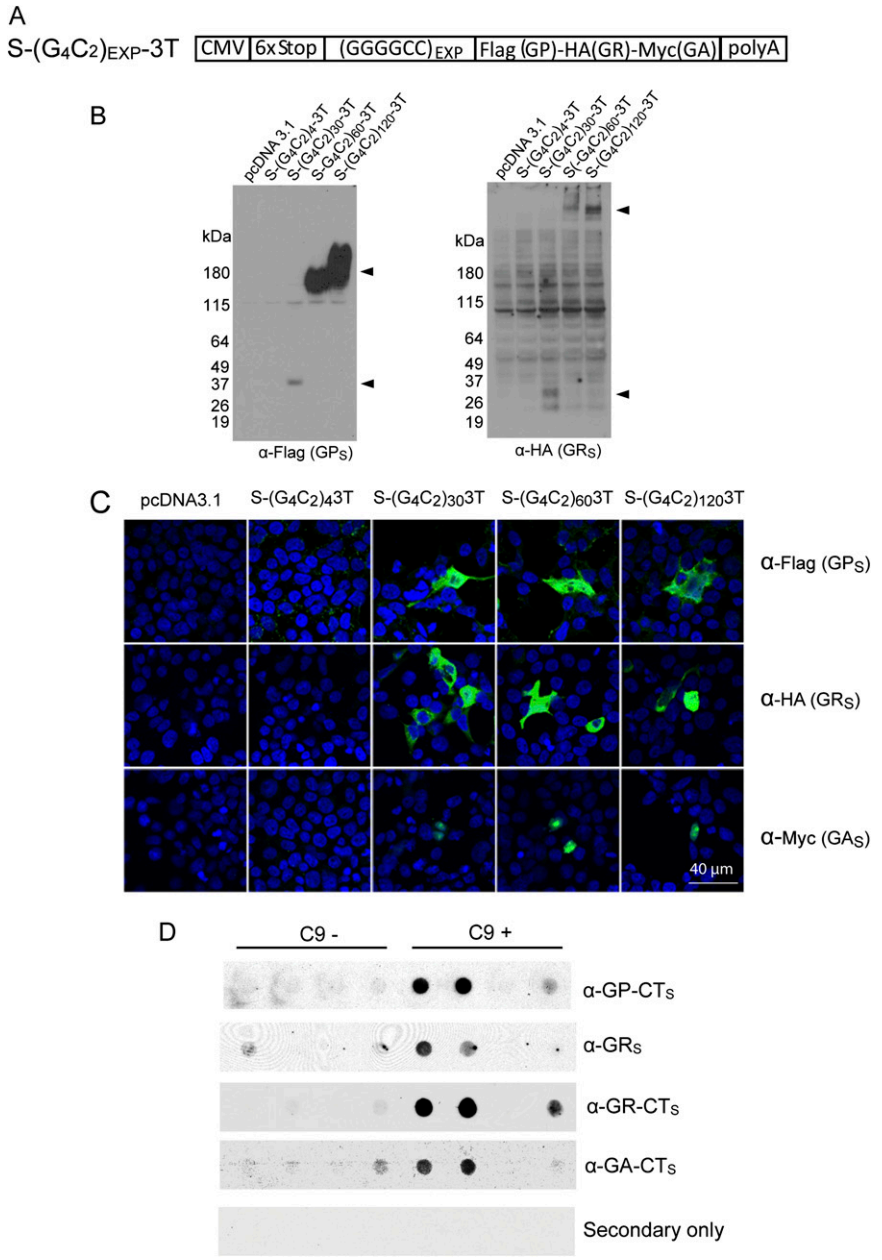

Fig. S6. In vitro and in vivo evidence for repeat-associated non-ATG (RAN) translation of the sense GGGGCC repeat expansion. (*A*) Constructs containing varying GGGGCC repeat lengths with upstream 6xStop cassette and 3′ epitope tags in each reading frame. Immunoblots (*B*) probed with α-Flag or α-HA antibodies detect RAN proteins (indicated by arrowheads). IF staining (*C*) showing RAN translation occurs in all three frames (GP_S, GR_S, GA_S) in cells transfected with constructs containing 30, 60, and 120 repeats. (*D*) Immunoblots of 2% soluble lysates from C9(+) and C9(−)amyotrophic lateral sclerosis (ALS) FCX lysates with antibodies against sense RAN proteins. Dot blots containing frontal cortex lysates from four different C9(−) and four different C9(+) ALS/frontotemporal dementia (FTD) patients were probed with α-GP-CTs, α-GRs, α-GR-CT_S, and α-GA_S antibodies. Results show evidence of RAN protein accumulation with all four antibodies in two of four C9(+) samples and weaker positive staining in one sample with α-GP-CT_S and α-GR-CT_S.

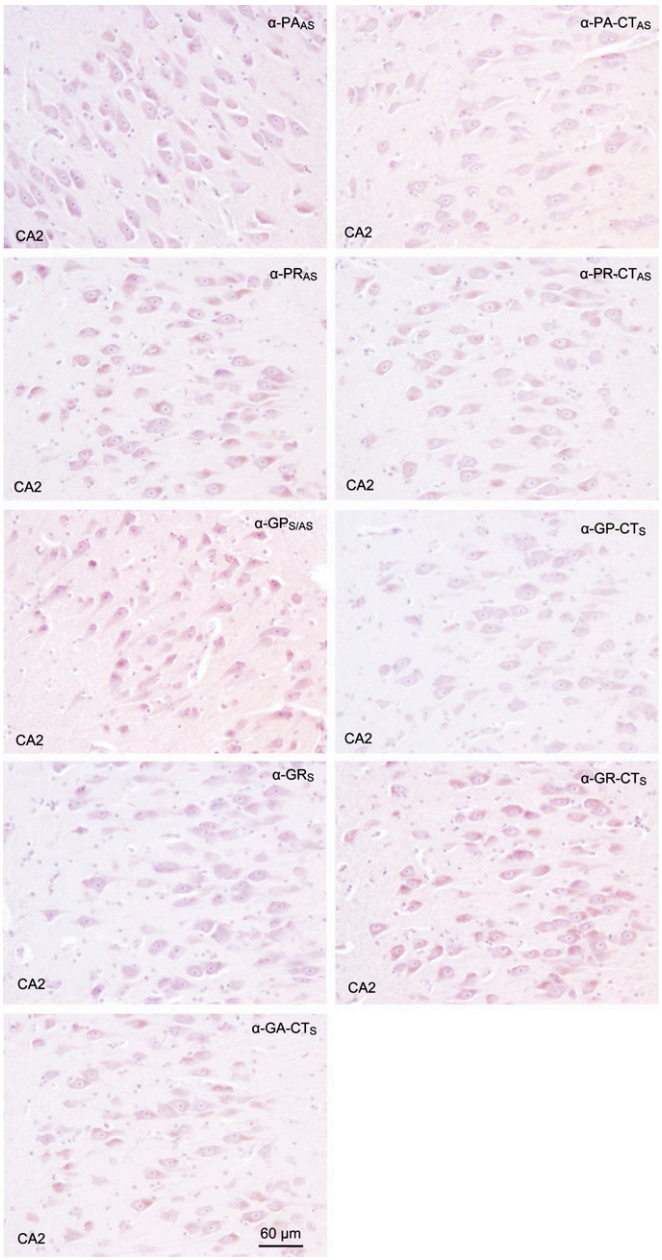

Fig. S7. Examples of negative IHC staining of C9(−) ALS/FTD hippocampal sections from cornu ammonis region 2 (CA2). Antibodies against sense and antisense RAN proteins are indicated.

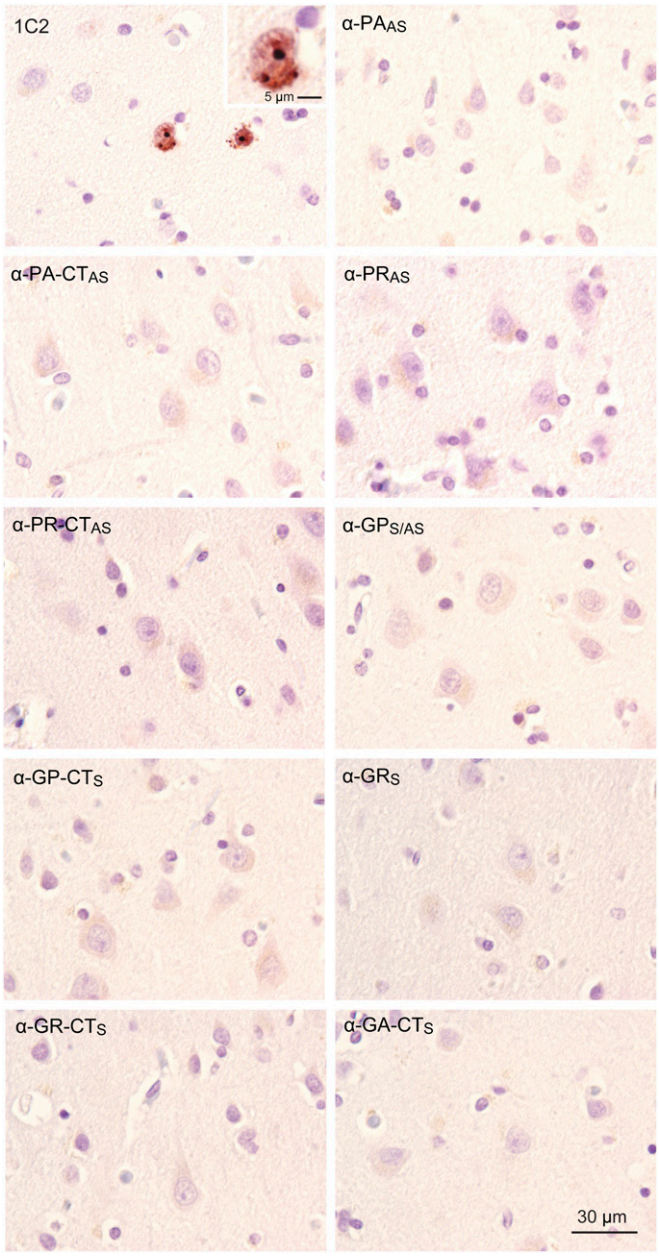

Fig. S8. Antibody validation. Intranuclear inclusions in frontal cortex of Huntington disease were detected by antipolyglutamine antibody (1C2); (*Inset*) higher-power image. No similar staining was found using any of the nine C9 antisense and sense antibodies.

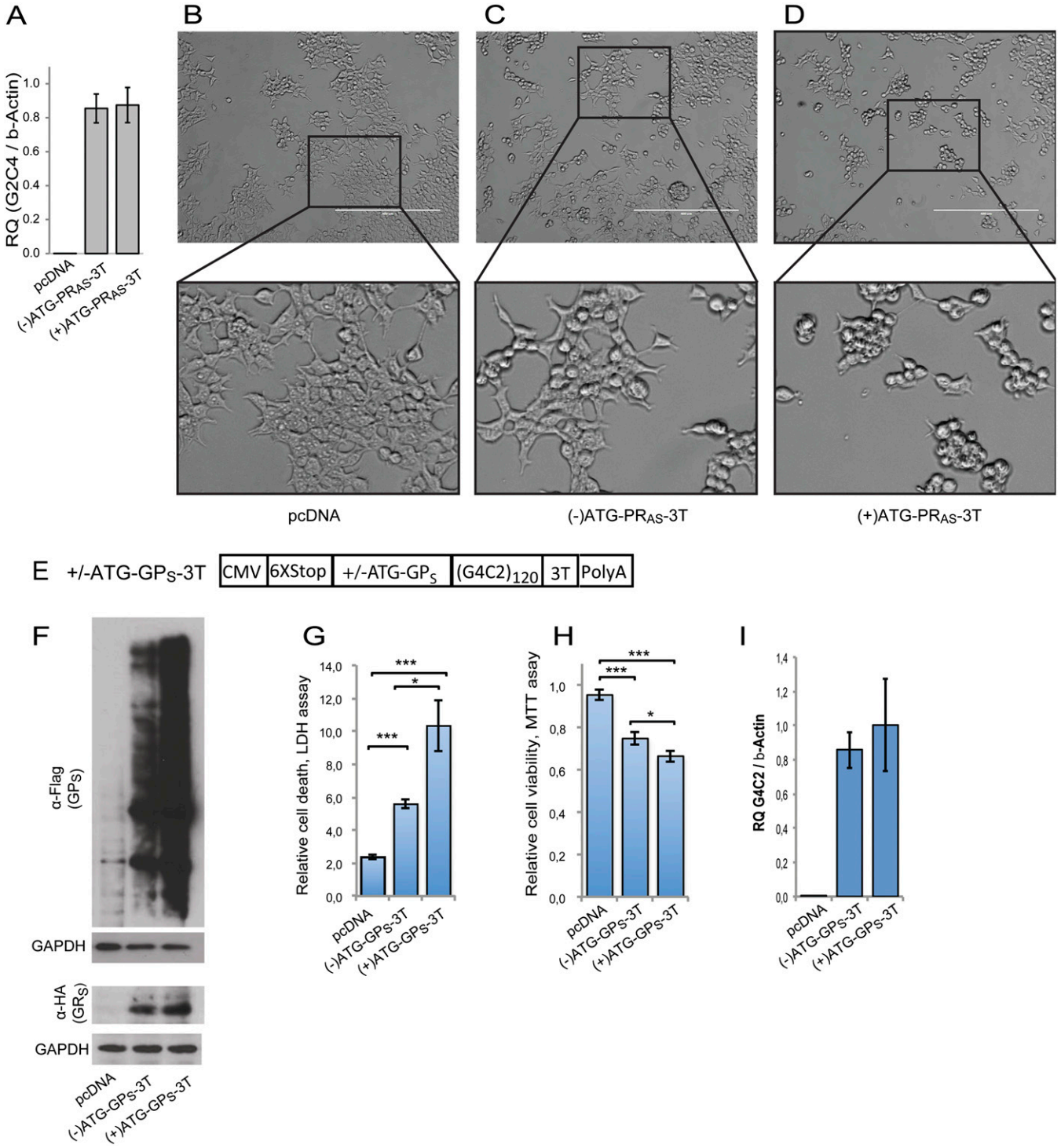

Fig. S9. RAN translation and PR_{AS} and GP_S protein expression affect cell viability. (*A*) Quantitative RT-PCR shows expression of expansion transcripts are similar in HEK293T cells transfected with (–)ATG-PR_{AS}-3T and (+)ATG-PR_{AS}-3T constructs. (*B–D*) Bright-field microscopy images showing changes in cell morphology in cells expressing RNA and RAN proteins from (–)ATG-PR_{AS}-3T constructs compared with empty vector control (pcDNA3.1) and worsening effects in (+)ATG-PR_{AS}-3T cells expressing increased levels of PR protein. (*E–I*) RAN and GP_S toxicity studies. (*E*) Schematic diagram of G_4C_2 expansion constructs (+/–ATG-GP_S-3T) with and without an ATG initiation codon in GP_S frame and 3′ epitope tags. (*F*) Protein blots showing levels of PR and GP in cells transfected with constructs in *E*. (*G*) Lactate dehydrogenase (LDH) (n = 9 independent experiments) and (*H*) methylthiazol tetrazolium (MTT) assays (n = 10 independent experiments) of transfected HEK293T cells. (*I*) Quantitative RT-PCR showing cells transfected with +ATG-GP_S-3T and –ATG-GP_S-3T constructs express comparable levels of RNA. *$P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$.

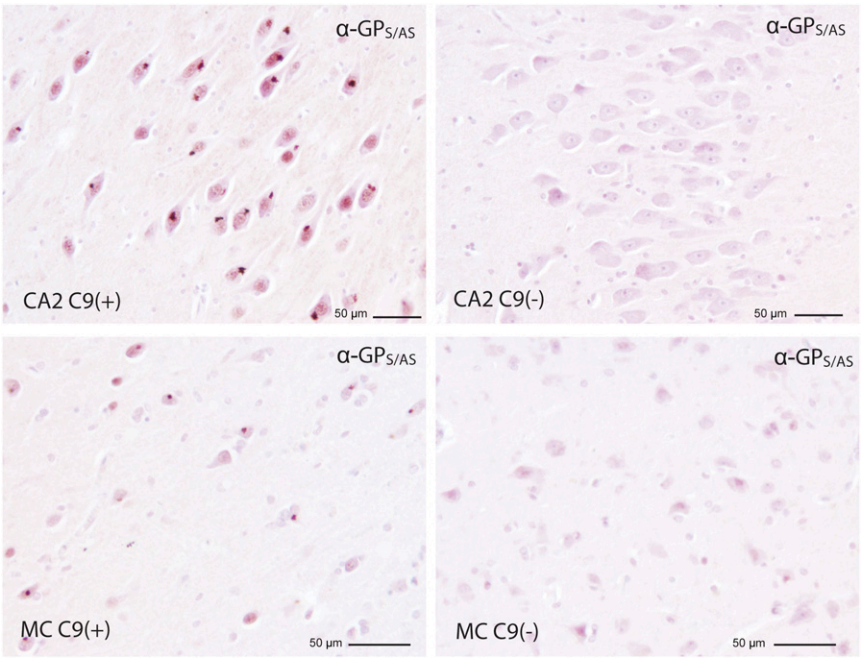

Fig. S10. Clustered GP-RAN protein aggregates in hippocampus and motor cortex. Low-power images of IHC staining with α-$GP_{S/AS}$ show $GP_{S/AS}$ aggregates in ALS/FTD C9(+) cornu ammonis region 2 (CA2) and motor cortex (MC) autopsy samples, but not in ALS/FTD C9(−) control tissue.

Table S1. Antibody validation summary

Strand	Antigen	ID #	Sequence	Species	IB	IHC	IF
AS-G_2C_4	poly(PA)	H3152	H2N-APAPAPAPAPAPAPACKKKK-amide	Rabbit	Y	Y	Y
	PA C-term	H3159	Ac-CYRLRLFPSLFSSG-OH	Rabbit	N	Y	Y
	poly(PR)	H3150	Ac-RPRPRPRPRPRPRPRPRC-amide	Rabbit	Y	Y	Y
	PR C-term	H3162	Ac-CRPRPLARDS-OH	Rabbit	Y	Y	Y
Both strands	poly(GP)	H3154	H2N-GPGPGPGPGPGPGPGPGCKK-amide	Rabbit	Y	Y	Y
	poly(GP)	F3M1	H2N-GPGPGPGPGPGPGPGPGCKK-amide	Mouse	Y	Y	Y
S-G_4C_2	GP C-term	H3157	Ac-CRRRRWRVGE-OH	Rabbit	Y	Y	Y
	poly(GR)	H3148	Ac-RGRGRGRGRGRGRGRGRC-amide	Rabbit	Y	Y	Y
	GR C-term	H3160	Ac-CRVAVWGSAAGKRRG-OH	Rabbit	Y	Y	Y
	GA C-term	H3164	Ac-CSGRARGRARGGA-amide	Rabbit	N	Y	Y

IB, immunoblotting; IF, immunofluoresence; IHC, immunohistochemistry.

Table S2. Primer sequences

Primer name	Primer sequence (5′ to 3′)
ASORF-F	AGTCGCTAGAGGCGAAAGC
ASORF-R	CGAGTGGGTGAGTGAGGAG
LK-ASORF-R	CGACTGGAGCACGAGGACACTGACGAGTGGGTGAGTGAGGAG
LK-ASORF-F	CGACTGGAGCACGAGGACACTGAAGTCGCTAGAGGCGAAAGC
1a-F	GCCCACGTAAAAGATGACGC
1a-R	CCTCCTAAACCCACACCTGC
LK-1a-R	CGACTGGAGCACGAGGACACTGACCTCCTAAACCCACACCTGC
LK-1a-F	CGACTGGAGCACGAGGACACTGAGCCCACGTAAAAGATGACGC
LK	CGACTGGAGCACGAGGACACTGA
5′ GSP1	GCTTTCGCCTCTAGCGACT
5′ GSP2	TCTAGCGACTGGTGGAATTGCCT
3′ GSP1	CTGCGGTTGTTTCCCTCCTT
3′ GSP2	TTTCTTGTTCACCCTCAGCGA
ACTB3	CTGGAACGGTGAAGGTGACA
ACTB4	GGGAGAGGACTGGGCCATT
3xTag-Fw	ACGACATCGATTACAAGGACG
3xTag-RV	ATCAGCTTCTGCTCGCTATG