

Antisense transcripts of the expanded *C9ORF72* hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS

Tania F. Gendron · Kevin F. Bieniek · Yong-Jie Zhang · Karen Jansen-West · Peter E. A. Ash · Thomas Caulfield · Lillian Daugherty · Judith H. Dunmore · Monica Castanedes-Casey · Jeannie Chew · Danielle M. Cosio · Marka van Blitterswijk · Wing C. Lee · Rosa Rademakers · Kevin B. Boylan · Dennis W. Dickson · Leonard Petrucelli

Received: 28 August 2013 / Revised: 29 September 2013 / Accepted: 1 October 2013 / Published online: 16 October 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) are devastating neurodegenerative disorders with clinical, genetic, and neuropathological overlap. A hexanucleotide (GGGGCC) repeat expansion in a non-coding region of *C9ORF72* is the major genetic cause of both diseases. The mechanisms by which this repeat expansion causes “c9FTD/ALS” are not definitively known, but RNA-mediated toxicity is a likely culprit. RNA transcripts of the expanded GGGGCC repeat form nuclear foci in c9FTD/ALS, and also undergo repeat-associated non-ATG

(RAN) translation resulting in the production of three aggregation-prone proteins. The goal of this study was to examine whether antisense transcripts resulting from bidirectional transcription of the expanded repeat behave in a similar manner. We show that ectopic expression of (CCCCGG)₆₆ in cultured cells results in foci formation. Using novel polyclonal antibodies for the detection of possible (CCCCGG)_{exp} RAN proteins [poly(PR), poly(GP) and poly(PA)], we validated that (CCCCGG)₆₆ is also subject to RAN translation in transfected cells. Of importance, foci composed of antisense transcripts are observed in the frontal cortex, spinal cord and cerebellum of c9FTD/ALS cases, and neuronal inclusions of poly(PR), poly(GP) and poly(PA) are present in various brain tissues in c9FTD/ALS, but not in other neurodegenerative diseases, including CAG repeat disorders. Of note, RNA foci and poly(GP) inclusions infrequently co-occur in the same cell, suggesting these events represent two distinct ways in which the *C9ORF72* repeat expansion may evoke neurotoxic effects. These findings provide mechanistic insight into the pathogenesis of c9FTD/ALS, and have significant implications for therapeutic strategies.

Electronic supplementary material The online version of this article (doi:10.1007/s00401-013-1192-8) contains supplementary material, which is available to authorized users.

Tania Gendron and Kevin Bieniek are co-first authors.

T. F. Gendron · K. F. Bieniek · Y.-J. Zhang · K. Jansen-West · T. Caulfield · L. Daugherty · J. H. Dunmore · M. Castanedes-Casey · J. Chew · D. M. Cosio · M. van Blitterswijk · W. C. Lee · R. Rademakers · D. W. Dickson (✉) · L. Petrucelli (✉)
Department of Neuroscience, Mayo Clinic Florida, Jacksonville, FL 32224, USA
e-mail: dickson.dennis@mayo.edu

L. Petrucelli
e-mail: petrucelli.leonard@mayo.edu

K. F. Bieniek · J. Chew
Mayo Graduate School, Mayo Clinic College of Medicine, Rochester, MN 55905, USA

P. E. A. Ash
Department of Pharmacology, Boston University School of Medicine, Boston, MA 02118, USA

K. B. Boylan
Department of Neurology, Mayo Clinic Florida, Jacksonville, FL 32224, USA

Keywords Amyotrophic lateral sclerosis · Bidirectional transcription · *C9ORF72* · Expanded repeat · Frontotemporal dementia · Repeat-associated non-ATG translation · RNA foci

Introduction

An expanded GGGGCC hexanucleotide repeat within a non-coding region of the *C9ORF72* gene is the most common genetic cause of frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), two devastating multisystem neurodegenerative disorders with significant

genetic, neuropathological, and clinical overlap [11, 30]. FTD, a common form of early-onset dementia, is characterized clinically by abnormalities in behavior and language, whereas ALS is characterized by upper and lower motor neuron signs, which include weakness and muscle atrophy. FTD-like cognitive and behavioral impairments are also present in up to 50 % of ALS patients [16, 19, 27], and as many as half of FTD patients develop motor neuron dysfunction (MND) reminiscent of ALS [19].

The mechanisms by which the *C9ORF72* hexanucleotide repeat expansion causes “c9FTD/ALS” are not definitively known, but at least three pathogenic pathways may be at play. Several groups have shown that mRNA levels of certain *C9ORF72* variants are decreased in c9FTD/ALS [11, 15, 25, 30], suggesting loss of *C9ORF72* function as a potential neurotoxic mechanism. In addition, RNA transcripts containing the expanded repeat may cause neurodegeneration by two means: through their accumulation into discrete structures in the nucleus, termed RNA foci, and by serving as a template for the synthesis of aberrantly expressed, aggregation-prone proteins by repeat-associated non-ATG (RAN) translation. Indeed, (GGGGCC)_{exp} RNA foci are observed in c9FTD/ALS [11], and the accumulation of neuronal inclusions composed of “c9RAN proteins” is a pathological hallmark of c9FTD/ALS [3, 25].

RNA foci are thought to cause cellular toxicity by sequestering specific RNA-binding proteins in a sequence-dependent manner, consequently disrupting their function [4]. In myotonic dystrophy type 1 (DM1), for instance, RNA foci formed of CAG·CTG repeat transcripts bind and inactivate the splicing factor muscleblind-like 1 protein (MBNL1) [23, 32, 33]. This sequestration of MBNL1 results in the missplicing of a subset of pre-mRNA targets that account for some of the characteristic features of disease [10, 20]. Transcripts of (GGGGCC)_{exp} that accumulate as nuclear foci in c9FTD/ALS may similarly sequester RNA-binding proteins and cause the misregulation of crucial downstream RNA targets leading to cellular dysfunction [24, 29, 38].

In addition to foci formation, transcripts of expanded repeats may be susceptible to RAN translation, an unconventional mode of translation that occurs across expanded repeat tracts despite the absence of an initiating codon. First described by Ranum and colleagues for expanded trinucleotide CAG·CTG repeats [39], transcripts of expanded CGG repeats [34], and the aforementioned GGGGCC repeats [3, 25] are now known to be RAN translated. Because RAN translation can occur in all possible reading frames, various products can be synthesized from a given transcript. Recently, we, as well as the Edbauer group and colleagues, independently reported that RAN translation of (GGGGCC)_{exp} RNA in c9FTD/ALS results in the production of poly(GP), poly(GA), and poly(GR) proteins [3, 25]. The presence of neuronal inclusions composed of these

c9RAN proteins throughout the central nervous system is now considered pathognomonic of c9FTD/ALS [3, 25].

For many microsatellite expansion disorders, the expanded repeat is bidirectionally transcribed [4]. Detectable expression of both sense and antisense transcripts containing the hexanucleotide repeat in *C9ORF72* patients indicates that bidirectional transcription also occurs in c9FTD/ALS [25]. Consequently, not only do (GGGGCC)_{exp} transcripts form foci and undergo RAN translation, so too may (CCCCGG)_{exp} transcripts resulting from bidirectional transcription of the *C9ORF72* expanded repeat. The goal of this study was thus to examine whether (CCCCGG)_{exp}-derived RNA foci and c9RAN translation proteins are present in c9FTD/ALS.

Materials and methods

Secondary structure prediction and model building of CCCCCG repeats

Folding of the RNA sequence comprising CCCCCG repeats of 10 (60 bases), 50 (300 bases), and 200 (1,200 bases) was carried out as previously described to determine secondary structure motifs [3]. The secondary structure prediction and modeling was built by examining the output from several RNA prediction packages: MFOLD, Sfold, Vienna RNA Package (RNAfold) [12, 17, 18, 21, 22, 31, 40, 41]. MFOLD utilizes a minimum free energy RNA structure prediction algorithm, Sfold utilizes statistical sampling of all possible structures, while Vienna Package has several options, including an minimum free energy calculation. Equivalent structures were given from each package for the ten repeat cases. MFOLD was used as the primary secondary structure prediction package for consistency across all models. Free energies were calculated for each secondary structure prediction at the level of decomposition (base pairing) and global energy [36]. MFOLD verifies each secondary structure prediction generated for valid structure [36]. MFOLD parameters include the following: (1) linear RNA sequence; (2) zero constraints, forces, or prohibitions on all bases allowing maximal sampling; (3) folding temperature of 37 °C; (4) physiological ionic conditions; (5) structure draw mode: untangle with loop fix; and (6) the remainder of MFOLD settings were set to default.

Generation of antibodies

For each of the peptide antigens (C-Ahx-(PR)₈-amide, C-Ahx-(GP)₈-amide and C-Ahx-(PA)₈-amide), two rabbits were immunized. Pre-immune serum from each rabbit was tested against peptide antigens and tissue from c9FTD/ALS

cases by Western blot and immunohistochemistry, respectively, and confirmed negative. Antiserum was used directly or affinity purified before use.

Meso Scale Discovery immunoassays

Peptides diluted in Tris-buffered saline (TBS) were added to duplicate wells (35 μ l/well) of a 96-well MSD assay plate at final concentrations 0.1 μ g/well. Following overnight incubation at 4 °C, wells were washed with TBS containing 0.2 % Tween 20 (TBSTw), and blocked with TBSTw+3 % non-fat milk. Antibody solution (25 μ l/well) containing the indicated anti-PR, anti-GP, anti-PA, anti-GA or anti-GR antibodies (1:1,000) and SULFO-TAG™-rabbit secondary antibody (0.5 μ g/ml, in blocking buffer) was added. Following a 2-h incubation and final washes, antibody binding to immobilized peptides was evaluated by adding MSD Read Buffer and measuring light emission at 620 nm upon electrochemical stimulation using the MSD Sector Imager 2400.

Western blot analysis and immunofluorescence staining for antibody characterization

HEK293T cells were transfected with Lipofectamine™ 2000 with pEGFP-C1 vector only, or pEGFP-C1 (Clontech) plasmids into which oligonucleotides of five repeats of PR, GP, PA, GA or GR were inserted. For Western blotting, cell lysates collected 2 days post-transfection were resolved by 10 % Tris–Glycine SDS-PAGE (Invitrogen) and transferred to nitrocellulose membranes for probing with anti-PR, anti-GP, anti-PA, anti-GA or anti-GR, or with anti-GFP (Abcam, 1:2,000). For immunostaining, coverslips were fixed, permeabilized and blocked, then probed with the indicated antibodies followed by anti-rabbit-AF594 (Alexa Fluor) and Hoechst [3].

Cloning of CCCC GG expression vectors

To generate the antisense (CCCCGG)₂ and (CCCCGG)₆₆ expression vectors, we first generated sense (GGGGCC)₂ and (GGGGCC)₆₆ expression vectors. Toward this end, genomic DNA from muscle or spleen from a *C9ORF72* expanded repeat carrier was used as a template in a nested PCR strategy using ThermalAce DNA Polymerase (Invitrogen) to amplify the (GGGGCC)_n repeat region, including 113 bp of 5' and 99 bp of 3' flanking sequence. The upstream primer used was 5'-AAGGAAGCTTAG-TACTCGCTGAGGGTGAAC-3'; downstream primers used were 5'-GCTTGGATCCCCACTCGCCACCGC-CTG-3' and 5'-GTCAGAGAAATGAGAGGGAAAG-3'. The PCR products were cloned into the pAG3 expression vector (kindly provided by Dr. T. Golde, University of

Florida) using restriction sites *HindIII* and *BamHI*. The pAG3 expression vector has a pcDNA3.0 backbone and a CMV-enhanced chicken B-actin promoter. The clones containing (GGGGCC)₂ or (GGGGCC)₆₆ were screened by colony PCR, and further verified by hairpin sequence analysis. The plasmids of (GGGGCC)₂ or (GGGGCC)₆₆ were digested using *HindIII*, and then the overhangs were filled-in using DNA Polymerase I (Klenow) Fragment. After purification, the fragments were digested with *BamHI*. The inserts with (GGGGCC)₂ or (GGGGCC)₆₆, including the 5' and 3' flanking sequences, were re-cloned in the antisense orientation into a pAG3 expression vector using *BamHI* and *EcoRV* to generate the (CCCCGG)₂ and (CCCCGG)₆₆ expression constructs. The sequence of antisense vectors was verified by hairpin sequence analysis. Note that the DNA sequence of the CCCCCG repeat is provided in Online Resource 1, which also highlights the regions included in the generation of pAG3-(CCCCGG)₂ and pAG3-(CCCCGG)₆₆ expression vectors.

RNA fluorescence in situ hybridization of cultured cells expressing (CCCCGG)_n expression vectors

Evaluation of foci formation in HeLa cells transfected with (CCCCGG)₂ or (CCCCGG)₆₆ expression vectors was carried out by RNA fluorescence in situ hybridization (FISH). In brief, cells grown on glass coverslips were transfected with 0.5 μ g of the (CCCCGG)₂ or (CCCCGG)₆₆ constructs. After 36 h, cells were fixed and permeabilized in 4 % paraformaldehyde + 20 % acetic acid + 2 mM ribonucleoside vanadyl complex (Sigma) for 10 min at room temperature. Cells were then washed with phosphate buffered saline treated with diethylpyrocarbonate (DEPC–PBS), and hybridized with denatured Cy3-conjugated (GGGGCC)₄ probe (2 ng/ μ l) in hybridization buffer (50 % formamide, 10 % dextran sulfate, 0.1 mg/ml yeast tRNA, 2X saline–sodium citrate buffer (SSC), 50 mM sodium phosphate buffer) overnight at 37 °C. After washing once with 40 % formamide/1XSSC for 30 min at 37 °C, and twice with DEPC–PBS for 5 min at room temperature, nuclei were counterstained with Hoechst 33258 (1 μ g/ml, Invitrogen) prior to mounting coverslips. Images were obtained on a Zeiss LSM 510 META confocal microscope.

Western blot analysis of cultured cells expressing (CCCCGG)_n expression vectors

To determine whether ectopic expression of expanded CCCC GG transcripts leads to RAN translation, HEK293T cells were transfected with 5 μ g of the (CCCCGG)₂ or (CCCCGG)₆₆ constructs. After 36 h, cells were harvested and washed with ice-cold PBS (pH 7.4), then cell pellets were lysed in buffer (50 mM Tris–HCl, pH 7.4, 300 mM NaCl,

Table 1 *C9ORF72* antisense transcript RAN translation cohort

Case #	<i>C9ORF72</i> mutation yes/no	Pathological diagnosis	Age at death	Sex	Repeat size		Hippocampal burden			Cerebellar burden		
					Frontal cortex	Cerebellum	PA	PR	GP	PA	PR	GP
1	Y	FTLD-TDP	66	M	16.0	12.5	±	+	+++	±	+	+++
2	Y	FTLD-TDP	71	M	25.2	11.4	±	±	++	±	±	++
3	Y	FTLD-TDP	86	M	50.0	17.5	+	±	+++	±	±	+++
4	Y	FTLD-MND	68	M	25.6	12.8	±	±	++	±	±	+++
5	Y	FTLD-MND	70	M	38.4	10.1	+	+	+++	+	+	+++
6	Y	FTLD-MND	61	F	35.6	13.7	+	+	+++	±	±	+++
7	Y	ALS	53	M	46.8	14.9	+	+	+++	+	+	+++
8	Y	ALS	49	F	23.8	10.0	+	+	+++	±	±	+++
9	Y	ALS	41	F	22.3	10.9	+	+	+++	+	+	+++
10	N	FTLD-TDP	88	F	–	–	–	–	–	–	–	–
11	N	FTLD-TDP	65	M	–	–	–	–	–	–	–	–
12	N	FTLD-TDP	83	F	–	–	–	–	–	–	–	–
13	N	FTLD-MND	64	M	–	–	–	–	–	–	–	–
14	N	FTLD-MND	72	F	–	–	–	–	–	–	–	–
15	N	FTLD-MND	66	F	–	–	–	–	–	–	–	–
16	N	ALS	60	M	–	–	–	–	–	–	–	–
17	N	ALS	61	F	–	–	–	–	–	–	–	–
18	N	ALS	53	F	–	–	–	–	–	–	–	–
19	N	HD	80	F	–	–	–	–	–	–	–	–
20	N	HD	70	F	–	–	–	–	–	–	–	–
21	N	Kennedy's	80	M	–	–	–	–	–	–	–	–
22	N	SCA3	53	M	–	–	–	–	–	–	–	–

– absent, ± sparse (<5 inclusions), + mild, ++ moderate, +++ severe

1 % Triton-X-100, 5 mM EDTA, 2 % sodium dodecyl sulfate (SDS), plus phenylmethylsulfonyl fluoride (PMSF) and both a protease and phosphatase inhibitor mixture). After centrifugation at 16,000g for 20 min at 4 °C, the supernatant was collected and protein concentration determined by BCA assay. For Western blot analysis, samples were prepared in Laemmli's buffer, heated for 5 min at 95 °C, and equal amounts of protein (30 µg) were loaded into Novex® 10–20 % Tricine gels. After transfer, blots were blocked with 5 % non-fat dry milk in TBST for 1 h, and then incubated with the purified anti-PR, anti-GP or anti-PA (1:1,000), or mouse monoclonal GAPDH antibody (1:10,000, Biodesign) overnight at 4 °C. Membranes were washed then incubated with anti-species horseradish peroxidase-linked secondary antibodies (1:5,000; Jackson ImmunoResearch) for 1 h. Protein expression was visualized by enhanced chemiluminescence treatment and exposure to film.

Human case material

All cases examined in this study were selected from a series of autopsied brains submitted to the neuropathology laboratory at Mayo Clinic in Jacksonville. The

sources of this case material include the Mayo Clinic Florida ALS Center ($n = 11$), referral to the Parkinson disease brain bank ($n = 7$), the State of Florida Alzheimer's Disease Initiative ($n = 2$), Florida Alzheimer's Disease Research Center ($n = 1$) and CurePSP/Society of Progressive Supranuclear Palsy brain bank ($n = 1$). The presence or absence of *C9ORF72* repeat expansion was determined using frozen cerebellar tissue from the right hemibrain and a previously described repeat-primed polymerase chain reaction (PCR) method [11]. In addition, repeat length was estimated using Southern blotting techniques as previously described [35] (Table 1). In brief, 7–10 µg of high-quality genomic DNA extracted from frozen frontal cortex and cerebellum was digested with *XbaI*, and electrophoresed in a 0.8 % agarose gel. DNA was then transferred to a positively charged nylon membrane (Roche), cross-linked, and subsequently hybridized with a DIG-labeled probe. Expansions were visualized with anti-DIG antibody (Roche) and CDP-star substrate (Roche) on X-ray film after multiple exposures. The most abundant expansion sizes were estimated using AlphaEaseFC (Alpha Innotech) based on their position relative to DNA molecular weight markers.

Fluorescence in situ hybridization and immunofluorescence staining of human tissue

Formalin-fixed, paraffin-embedded frontal cortex, spinal cord, and cerebellum sections were cut at a 5- μ m thickness and mounted on glass slides, then subjected to RNA FISH followed by immunofluorescence staining. For FISH, slides were deparaffinized and rehydrated, incubated with pepsin (4 mg/ml in 0.9 % NaCl, pH 1.5) for 20 min at 37 °C, rinsed in water, then immersed in ice-cold 20 % acetic acid for 90 s, prior to dehydration. A Cy3-tagged (GGGGCC)₄ probe (IDT) which hybridizes to the expanded CCCC GG repeat was applied to the tissue, which was then sealed under a coverslip. Prior to use, the probe was diluted to 5 ng/ μ l in hybridization buffer (10 % dextran sulfate, 50 % formamide, 2XSSC, 50 mM sodium phosphate buffer, 10 ng/ml tRNA pH 7.0), then heated at 80 °C for 10 min, placed on ice for 5 min, then heated at 37 °C for 10 min. Alternatively, to detect foci formed of (GGGGCC)_{exp} RNA, a TYE563-labeled LNA probe (5'TYE563-CCCCGGCCCCGGCCCC-3'TYE563; Exiqon, Inc) was applied to the tissue. This probe was diluted to 0.4 ng/ml in hybridization buffer (10 % dextran sulfate, 50 % formamide, 2XSSC, 50 mM sodium phosphate buffer, 10 ng/ml tRNA, pH 7.0), then heated at 80 °C for 75 s, according the manufacturer's instructions. Following a 2-day hybridization at 37 or 55 °C for the antisense and sense probe, respectively, coverslips were removed and slides were washed: once in 2X SSC, three times in 50 % formamide/2X SSC at 37 °C, and three times in 1X SSC at 37 °C. Slides were subsequently subjected to immunofluorescence staining: slides were blocked with DAKO Serum-Free Protein Block and then incubated with anti-GP (1:3,000), choline acetyltransferase antibody (ChAT; 1:200; Chemicon AB144P), microtubule-associated protein 2 antibody (MAP2; 1:750, clone AP-20) or glial fibrillary acidic protein antibody (GFAP; 1:500, clone GA-5) overnight at 4 °C. The following day, slides were incubated with an Alexa Fluor 488-conjugated secondary antibody (1:500, Molecular Probes) for 1.5 h at room temperature. Slides were then treated with a solution of Sudan Black for 2 min to block auto-fluorescence, and coverslipped using Vectashield-DAPI mounting medium (Vector Laboratories).

Imaging was performed using a Zeiss Axio Imager Z1 microscope to visualize foci and poly(GP) inclusions. Although it is difficult to distinguish binding of the probe to an individual RNA transcript given that the fluorescence resulting thereof is not sufficiently strong to yield a signal above background, the FISH method employed is well suited to examine nuclear foci. The presence of multiple labeled transcripts in a compact location, as is the case for foci, results in a highly distinct, punctate fluorescent signal much brighter than background. Note that, to validate

specificity of the probes targeting GGGGCC and CCCCGG RNA, FISH was carried out on frontal cortex and cerebellar tissue from ALS, frontotemporal lobar degeneration (FTLD), and FTLD-MND patients with normal *C9ORF72* repeat length. In addition, a (CAGG)₆ probe targeting the CCTG repeat was tested and shown also to be negative.

To determine what percentage of affected cells (i.e., cells having either foci or a poly(GP) inclusion) have *both* nuclear foci and poly(GP) pathology, we analyzed frontal cortex and cerebellar sections of four c9FTD/ALS cases co-stained for either sense or antisense foci and poly(GP) inclusions. For each section, we examined 25 cells with inclusions by fluorescence microscopy, and then changed the excitation filter to determine whether they also had foci; we then examined 25 cells with foci and determined whether they had inclusions. In this manner, a total of 50 cells were examined per section. For each group (Group 1—frontal cortex probed for antisense foci and poly(GP) inclusions; Group 2—cerebellum probed for antisense foci and poly(GP) inclusions; Group 3—frontal cortex probed for sense foci and poly(GP) inclusions; Group 4—cerebellum probed for sense foci and poly(GP) inclusions), the percentage of cells having both foci and inclusions was calculated and compared by two-way ANOVA. In addition, we compared the frequency of sense and antisense foci in the frontal cortex of the four examined cases. For each section, the number of cells with foci and the number of total cells were counted in 12 randomly selected, non-overlapping fields from layers I–III. The average percentage of cells with antisense foci among the four cases was compared to the average percentage of cells with sense foci by paired, two-tailed *t* test.

Immunohistochemistry

For immunohistochemical analysis, 9 cases exhibiting an expanded *C9ORF72* repeat (3 FTLD-TDP, 3 FTLD-MND, 3 ALS) were matched with 13 negative control cases that lacked the expanded repeat, including 3 FTLD-TDP, 3 FTLD-MND, 3 ALS, and 4 cases with CAG repeat disorders, including 2 cases with Huntington's disease, 1 case with Kennedy's disease, and 1 case with spinocerebellar ataxia type 3 (Table 1). For the 9 c9FTD/ALS cases and the 9 FTLD/ALS matched controls, immunohistochemistry was performed on formalin-fixed paraffin-embedded tissue from the motor cortex, hippocampus (including temporal cortex), basal forebrain (including amygdala), thalamus (at the level of the subthalamic nucleus), medulla, and cerebellum. For the four trinucleotide repeat disorders, sections from the basal forebrain (2 Huntington's disease cases), medulla (Kennedy's disease case) or pons (spinocerebellar ataxia type 3 case) were used. Five-micron-thick tissue sections were cut from paraffin blocks, mounted on charged

glass slides, and allowed to dry overnight at 65 °C. The following day, slides were deparaffinized and rehydrated in serial washes in xylene and alcohol before steaming the slides for 30 min in 1X Tris–EDTA (pH 9) buffer solution. Immunohistochemistry was performed using the Dako Autostainer and the Dako EnVision™ + Rabbit (DAB) kits. Immunostaining was performed with the following antibodies: anti-PA (Rb8604, 1:2,500), anti-GP (Rb7379, 1:10,000), and anti-PR (Rb8736, affinity purified, 1:100). Following immunohistochemistry, slides were counterstained with Lerner's hematoxylin, dehydrated, and coverslipped. All imaging was conducted using the Zeiss Axio Imager Z1 microscope.

Results

Characterization of antibodies for the detection of c9RAN proteins produced by RAN translation of (CCCCGG)_{exp} RNA

The RNA structure of expanded repeats is believed to influence their susceptibility to RAN translation [9, 39]. We had previously reported that (GGGGCC)_{exp} RNA is predicted to form imperfect hairpins and is RAN translated [3]. Using methodologies for generating secondary structure predictions, we thus sought to determine whether (CCCCGG)_{exp} RNA is predicted to have a similar structure and also be prone to RAN translation. It should be noted that, while the reverse complement of the 5'-GGGGCC-3' repeat is 5'-GGCCCC-3', there are 4 Gs downstream of the last sense GGGGCC repeat, thus making the first antisense repeat CCCCCG (Fig. 1, Online Resource 1). The major RNA secondary structure prediction for 10 antisense CCCCCG repeats or 10 sense GGGGCC repeats is stable imperfect hairpins, but the GGGGCC repeat structure has a lower composite global energy ($\Delta G = -40.8$ kcal/mol), and therefore greater stability, than the antisense CCCCCG repeats



Fig. 1 Schematic representation of the possible proteins generated by RAN translation of expanded GGGGCC and CCCCCG repeats in all possible reading frames

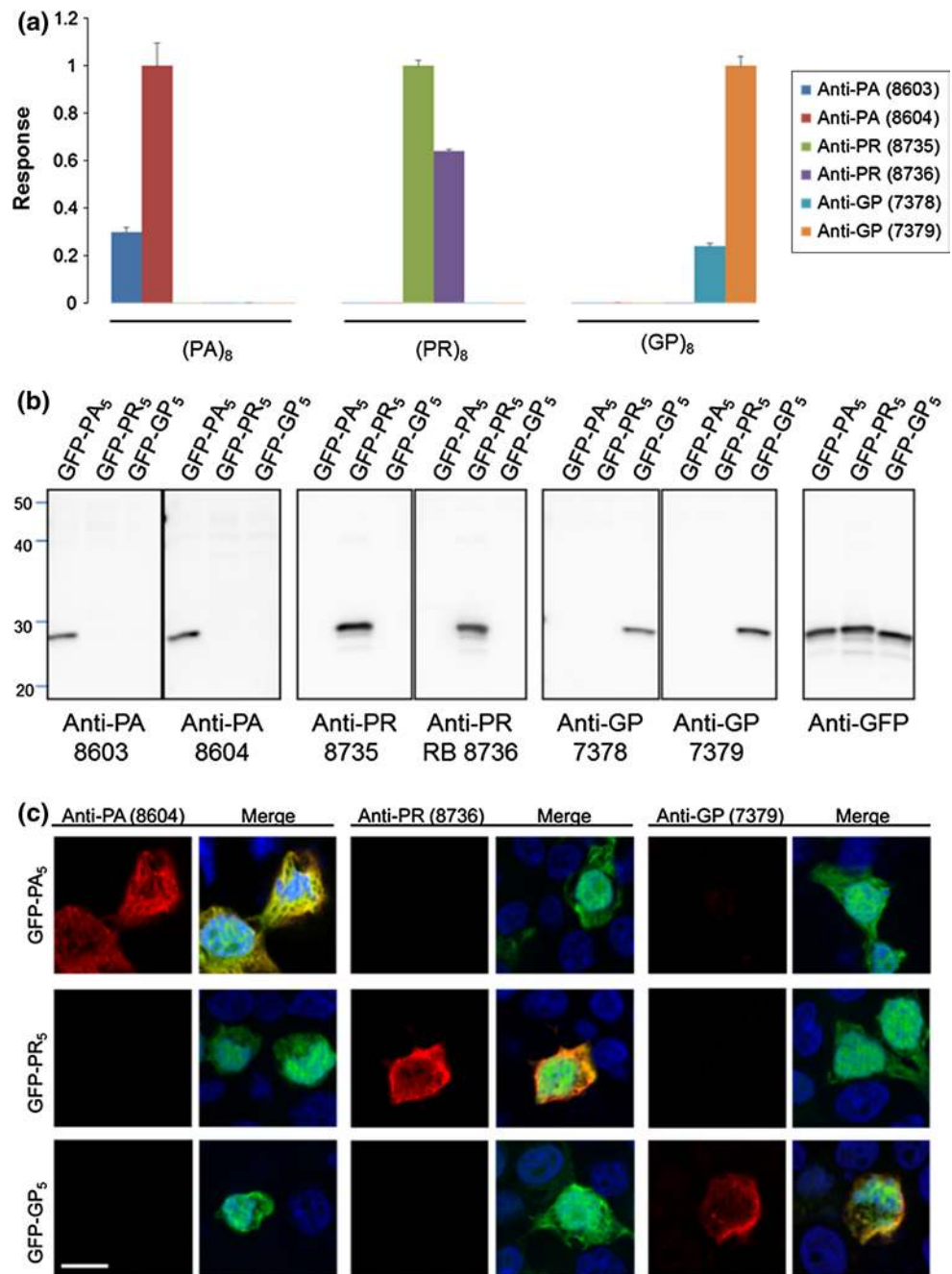
($\Delta G = -35.3$ kcal/mol) (Online Resource 2). Nonetheless, both GGGGCC and CCCCCG repeats become successively more stable as repeat length increases (Online Resource 2b). Given these observations, the CCCCCG repeat, like its GGGGCC counterpart, may be RAN translated. To determine whether this is indeed the case, two independent rabbit polyclonal antibodies were generated against each of the products that could be produced from RAN translation of (CCCCGG)_{exp} transcripts in the three alternate reading frames: poly(PR), poly(PG), and poly(PA) (Fig. 1). Given that poly(PG) [aka poly(GP)] proteins can be translated from both sense and antisense transcripts of the *C9ORF72* repeat expansion (Fig. 1), “GP” is henceforth used when referring to these proteins.

Specificity of anti-PR, anti-GP and anti-PA antibodies toward their respective antigen was evaluated by quantitative electrochemiluminescent immunoassays. Measurement of antibody binding to immobilized (PR)₈, (GP)₈ or (PA)₈ peptides showed that all antibodies bound only their respective antigen (Fig. 2a). Consistent with these findings, anti-PR, anti-GP and anti-PA antibodies respectively detected exogenously expressed enhanced GFP-tagged (PR)₅, (GP)₅ or (PA)₅ in HEK293T cells, as assessed by Western blot of cell lysates (Fig. 2b), and immunofluorescence staining of cells (Fig. 2c). Anti-PR, -GP and -PA antibodies were additionally tested against poly(GA) and poly(GR) peptides, which represent c9RAN proteins produced from the sense transcript (Fig. 1). All antibodies but one detected only their respective antigen, as assessed by immunoassay, Western blot and immunofluorescence staining of cultured cells (Online Resource 3). Anti-PR (7378), however, did show modest cross-reactivity with the GR peptide by immunoassay, but such cross-reactivity was not observed by Western blot or immunofluorescence staining (Online Resource 3).

Exogenous (CCCCGG)_{exp} transcripts are subject to RAN translation in cultured cells

Having confirmed specificity of these novel antibodies toward potential (CCCCGG)_n c9RAN proteins, they were then used to evaluate whether poly(PR), poly(GP) or poly(PA) proteins can indeed be RAN translated from expanded CCCCCG repeats. To address this question, a (CCCCGG)₆₆ expression vector having no initiating ATG start codon upstream of the repeats was generated for the transfection of cultured cells. As shown in Fig. 3a, expression of (CCCCGG)₆₆, but not of non-expanded (CCCCGG)₂, led to the formation of nuclear foci, as assessed by RNA FISH using a Cy3-tagged (GGGGCC)₄ probe targeting the CCCCCG repeat. Ectopic expression of (CCCCGG)₆₆ also resulted in the synthesis of poly(GP)

Fig. 2 Antibody characterization for c9RAN proteins. **a** The immunoreactivity of antibodies to c9RAN proteins towards (PA)₈, (PR)₈ and (GP)₈ peptides was measured by adsorbing peptides onto carbon electrodes in 96-well MSD plates, and co-incubating wells with anti-PA, anti-PR or anti-GP antibodies, and a SULFO-tagged anti-rabbit secondary antibody. Antibody binding to respective peptides was quantified by measuring the intensity of emitted light upon electrochemical stimulation of the plate using the MSD Sector Imager 2400. For each pair of antibodies, binding responses were normalized to the signal of the antibody showing the highest binding to its respective antigen. *Error bars* indicate standard deviations from duplicate wells. **b** Western blot analysis of lysates from HEK293T cells transfected to express enhanced GFP-tagged (PA)₅, (PR)₅ or (GP)₅. Blots were probed with the indicated antibodies. **c** Immunofluorescence staining of HEK293T cells transfected to express the indicated enhanced GFP (*green*)-tagged peptides using anti-PA, anti-PR or anti-GP (*red*) antibodies. Nuclei are stained with Hoechst (*blue*). *Scale bar* 10 μm. Note that similar studies were conducted to test potential cross-reactivity of antibodies to poly(GA) and poly(GR), as shown in Online Resource 3



and poly(PR) proteins, but not poly(PA), as assessed by Western blot of cell lysates (Fig. 3b). These c9RAN proteins were not observed in cells expressing (CCCCGG)₂ suggesting that their formation is repeat length-dependent (Fig. 3b).

Nuclear foci of (CCCCGG)_{exp} transcripts are present in brain and spinal cord of c9FTD/ALS patients

It was previously shown that transcripts of the expanded GGGGCC repeat accumulate as nuclear RNA foci in the

frontal cortex and spinal cord of c9FTD/ALS patients [11]. To determine whether expression of antisense transcripts of (CCCCGG)_{exp} similarly result in foci formation in human tissue, FISH was performed on paraffin-embedded frontal cortex and spinal cord sections from c9ALS, c9FTLD or c9FTLD-MND patients using the Cy3-(GGGGCC)₄ probe, the specificity of which was verified using sections from affected patients with normal *C9ORF72* repeat lengths. Cells harboring (CCCCGG)_{exp} nuclear foci were observed throughout all layers of the frontal cortex, as well as in the white matter, of c9ALS,

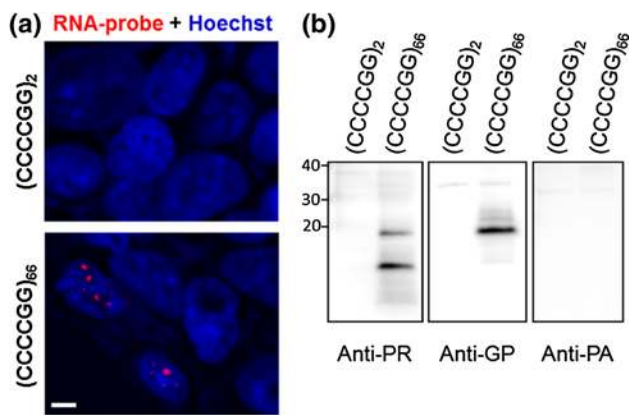


Fig. 3 Expression of expanded CCCC GG repeats in cultured cells leads to foci formation and expression of c9RAN proteins. **a** HEK293T cells transfected to express (CCCCGG)₂ or (CCCCGG)₆₆ were subjected to RNA fluorescence in situ hybridization using a probe against CCCC GG repeat transcripts. Note the foci (red) in Hoechst-stained nuclei (blue) of (CCCCGG)₆₆-expressing cells, but not (CCCCGG)₂-expressing cells. Scale bar 5 μ m. **b** Western blot analysis of lysates from (CCCCGG)_n-expressing cells shows that poly(PR) and poly(GP) proteins, but not poly(PA) proteins, are expressed in cells transfected with (CCCCGG)₆₆. No c9RAN protein was detected in control cells expressing non-expanded (CCCCGG)₂

c9FTLD and c9FTLD-MND patients, with the number of foci in affected cells ranging from few (e.g., 1–2) to many (>10), and with the size of foci being variable (Fig. 4a). Likewise, (CCCCGG)_{exp} nuclear foci were present in cells of the spinal cord, including ChAT-immunopositive motor neurons (Fig. 4b).

Compared to FTD and ALS cases lacking the *C9ORF72* expansion, c9FTD/ALS is marked by abnormal cerebellar pathology, such as inclusions composed of (GGGGCC)_{exp} RAN proteins [3, 25], as well as inclusions negative for TDP-43, but immunopositive for p62, ubiquitin, and various ubiquitin-binding proteins [1, 28]. Therefore, we evaluated whether (CCCCGG)_{exp} RNA foci, as well as (GGGGCC)_{exp} RNA foci, are present in the cerebellum of *C9ORF72* expanded repeat carriers. Multiple cells with RNA foci composed of sense or antisense transcripts were detected, and these were most often observed in proximity to the Purkinje cell layer separating the granular and molecular layers (Fig. 4c). Foci were also occasionally present in cells of the molecular layer, deep within the granular layer, in Purkinje cells, and in cells within the white matter (Fig. 4c). Post-FISH immunofluorescent staining with MAP2, a neuron marker, and GFAP, an astrocyte marker, confirmed that nuclear foci were present in both cell populations (Fig. 5). These findings indicate that, in addition to the frontal cortex and spinal cord, RNA foci resulting from the hexanucleotide repeat expansion in *C9ORF72* accumulate in the cerebellum.

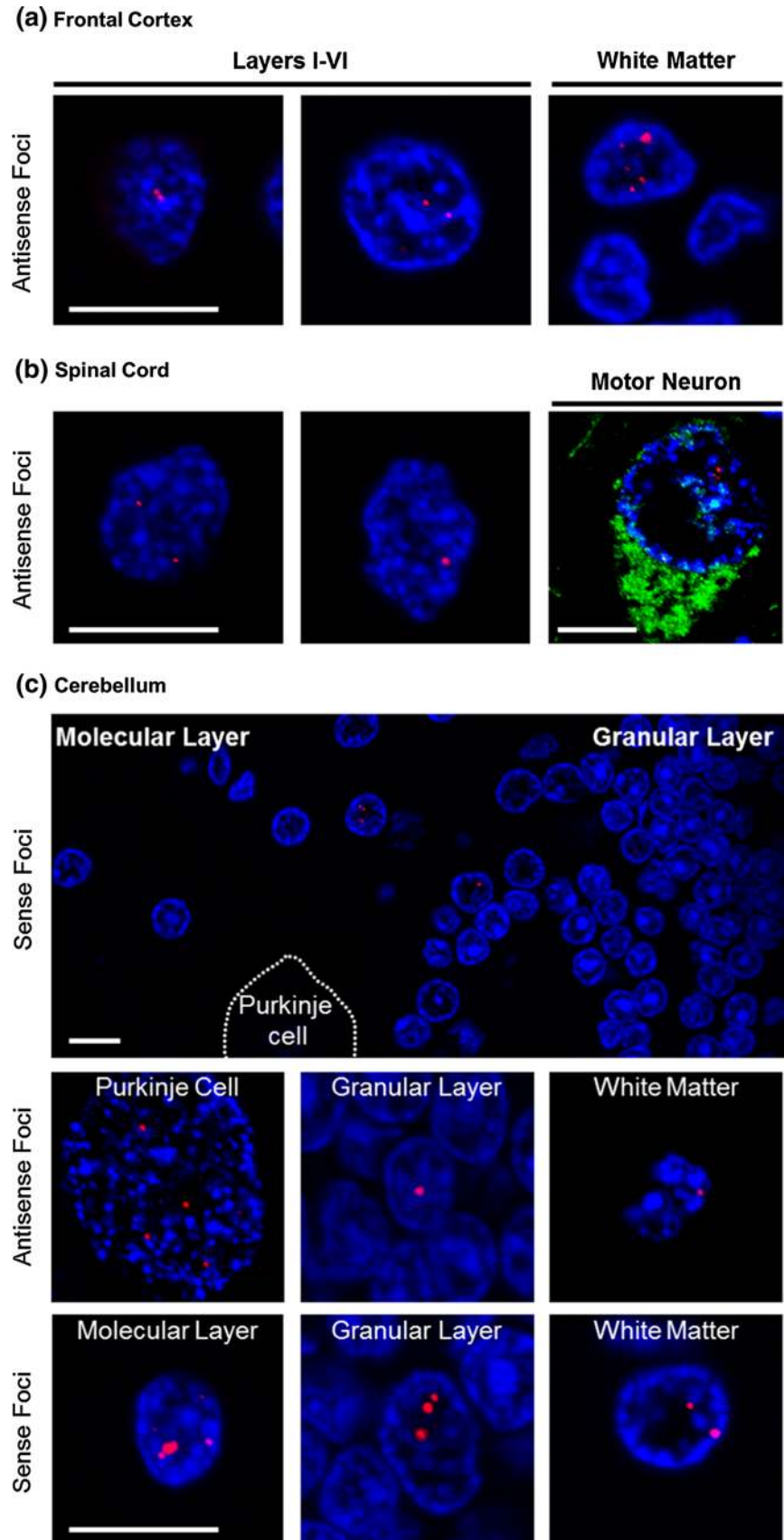
Inclusions of (CCCCGG)_{exp} c9RAN proteins are present in c9FTD/ALS

Next, the presence of c9RAN proteins synthesized from the antisense repeat was analyzed in human brain tissue. Immunohistochemistry with the anti-PA and anti-PR antibodies revealed sparse neuronal cytoplasmic inclusions in all sections examined from the nine c9FTD/ALS cases. Lesion burden was greatest in the hippocampus (dentate fascia and hippocampus proper), lesser in the motor cortex, temporal cortex, amygdala, as well as thalamus, and lowest in the cerebellum and medulla (Fig. 6). Many of the lesions were of the characteristic “star-shaped” morphology. Immunohistochemistry with the anti-GP antibody revealed abundant pathology in all regions that paralleled the pathologic burden and distribution previously described with C9RANT antibodies [3]. Semi-quantitative analysis of anti-GP, anti-PA or anti-PR immunoreactive-inclusion burden in the hippocampus and cerebellum is shown in Table 1. Neuronal c9RAN protein inclusions were not seen in the nine FTD/ALS matched control cases, or in the four trinucleotide repeat disorder cases.

RNA foci and poly(GP)-inclusions infrequently coexist in the same cell

The findings above, together with previous studies [2, 3, 11, 25], establish that (GGGGCC)_{exp} and (CCCCGG)_{exp} transcripts form nuclear RNA foci and undergo RAN translation; nonetheless, the relationship between these two events remains unclear. The presence of foci could be indicative of high levels of (GGGGCC)_{exp} and (CCCCGG)_{exp} transcripts, thus increasing the possibility of RAN translation. Conversely, the formation of nuclear foci may sequester these transcripts away from translational machinery, thus decreasing the likelihood of RAN translation. Given that both foci formation and RAN translation are potentially pathogenic, we examined the relationship between foci and inclusions composed of poly(GP), which is RAN translated from both sense and antisense transcripts. To this end, sections of cerebellar and frontal cortex from c9ALS, c9FTLD and c9FTLD-MND patients were subjected to FISH for the detection of (CCCCGG)_{exp} or (GGGGCC)_{exp} RNA foci, followed by immunofluorescence staining for poly(GP). Remarkably, while foci and anti-(GP)-immunoreactive inclusions were occasionally present in the same cell, the majority of affected cells had only RNA foci or only poly(GP) inclusions (Fig. 7a, b). To determine what percentage of cells have both foci and poly(GP) inclusions in the frontal cortex and cerebellum, quantitative analysis was undertaken on four cases co-stained for poly(GP) inclusions and either sense or antisense foci. For each section, we examined 25 cells with inclusions and determined

Fig. 4 *C9ORF72* hexanucleotide repeat transcripts form nuclear RNA foci in frontal cortex, spinal cord and cerebellum in *c9FTD/ALS*. Fluorescence in situ hybridization (FISH) of *c9FTD/ALS* frontal cortex (a) and *c9ALS* spinal cord (b) tissue using a probe against the CCCCCG repeat transcripts shows RNA foci (red) in the nucleus (stained with DAPI, blue) of cells. In (b), note that foci are observed in motor neurons that stain positively for ChAT. c Cerebellar sections of *c9FTD/ALS* cases were subjected to FISH using a probe against CCCCCG repeat transcripts or GGGGCC repeat transcripts. In most instances, foci-bearing cells within the cerebellum were found in proximity to the Purkinje cell layer separating the molecular and granular layers. However, RNA foci were also observed in cells of the molecular layer, deep within the granular layer, in Purkinje cells, and in cells within the white matter. Scale bars 10 μm



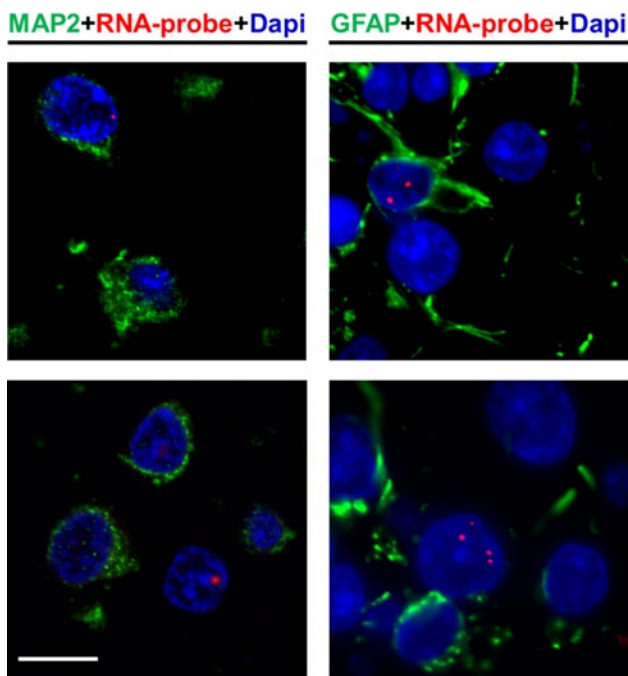


Fig. 5 Nuclear RNA foci are present in both neurons and glia. Fluorescence in situ hybridization of c9FTD/ALS cerebellar tissue using a probe against the GGGGCC repeat was followed by immunofluorescence staining with the neuronal marker, MAP2, or the astrocytic marker, GFAP. Note that nuclear RNA foci are present in MAP2-positive and MAP2-negative cells, as well as GFAP-positive and GFAP-negative cells. Scale bar 10 μ m

whether they had foci, and examined an additional 25 cells with foci and determined whether they had inclusions. Shown in Fig. 7c is the percentage of affected cells that had both foci and inclusions. Note that the brain region sampled (frontal cortex vs. cerebellum) and foci type (antisense vs. sense) both significantly affect the percentage of cells having both foci and inclusions, as assessed by two-way ANOVA (Fig. 7c). The apparent increase in the percentage of cells having both inclusions and sense foci, compared to the percentage of cells having both inclusions and antisense foci, is not likely caused by increased frequency of sense foci. In the frontal cortex, where the difference between sense and antisense foci co-occurring with poly(GP)-inclusions is most pronounced, there was no difference in the percentage of cells with sense or antisense foci in the cases examined (Fig. 7d).

Discussion

The discovery of new neuropathologic features specific to c9FTD/ALS, namely the formation of RNA foci and the production of c9RAN proteins resulting from the synthesis of antisense transcripts of the expanded *C9ORF72*

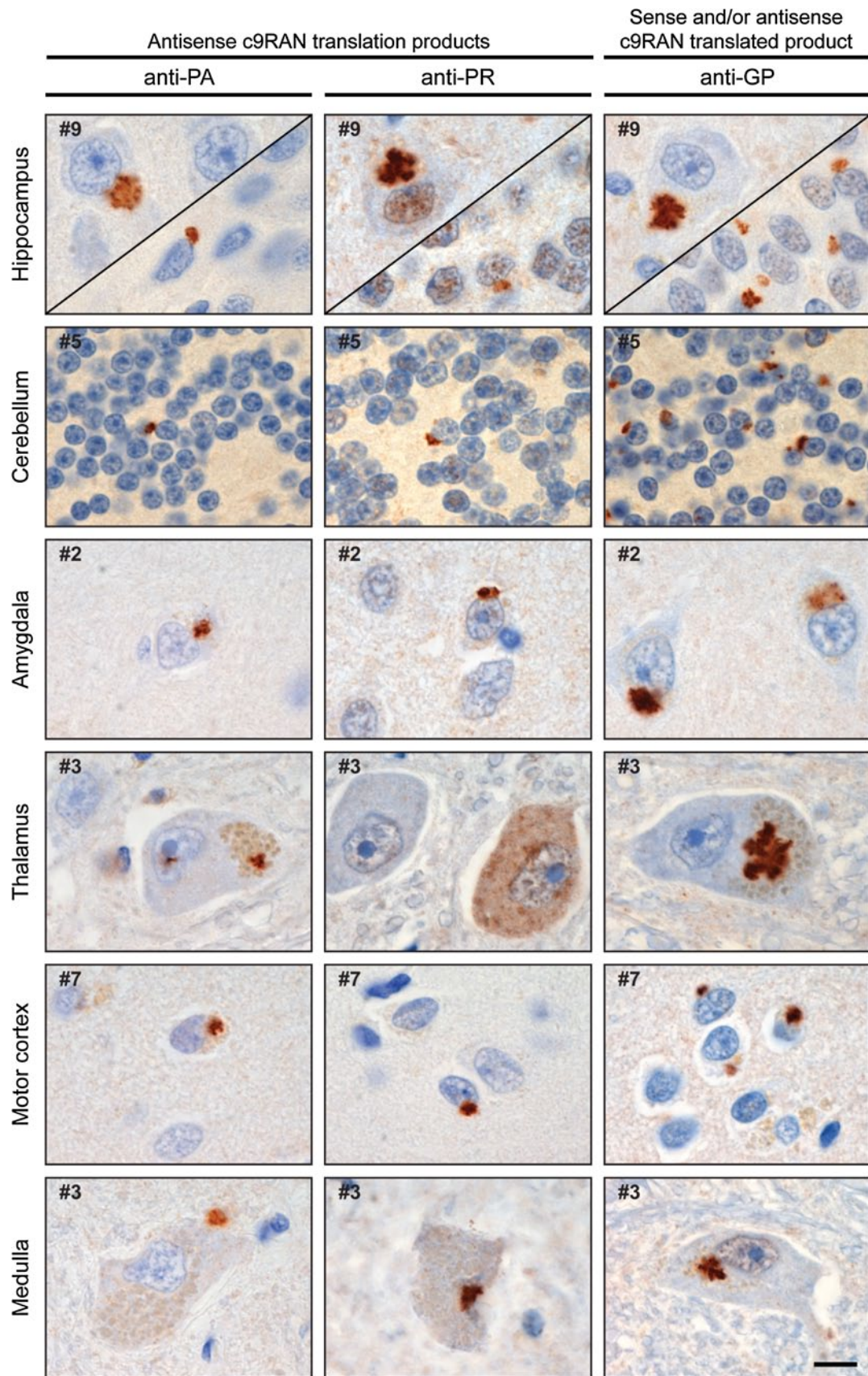
Fig. 6 Antisense c9RAN proteins in human post-mortem tissue. Immunohistochemistry reveals poly(PA)- [left column], poly(PR)- [middle column], and poly(GP)- [right column] reactive lesions throughout the central nervous system, including the hippocampus (endplate-CA3 on the top left, dentate fascia on the bottom right), cerebellum, amygdala, thalamus, motor cortex (layers 2–3), and medulla (inferior olivary nucleus). Lesions are often neuronal cytoplasmic inclusions (NCI) with a star-shaped morphology, but can also appear as dense NCI, small neuronal intranuclear inclusions, or diffuse neuronal “pre-inclusions”. Anti-GP, which detects poly(GP) proteins that can be made from both sense and antisense transcripts of the *C9ORF72* expanded repeat, reveal greater pathologic burden compared to the anti-PA and PR antibodies. Case numbers correspond to c9FTD/ALS cases in Table 1. Scale bar 10 μ m

repeat, provides additional insight into the pathobiology of c9FTD/ALS. Through the production of both sense and antisense expanded repeat RNA and five distinct c9RAN proteins, the repeat expansion leads to the production of seven potentially toxic biomolecules.

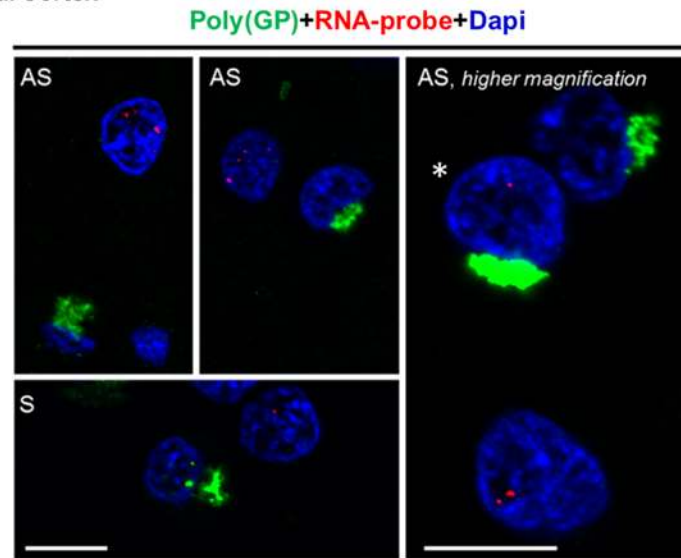
To determine whether antisense transcripts of the expanded *C9ORF72* repeat are RAN translated, we generated novel rabbit polyclonal antibodies for the detection of poly(GP), poly(PR) and poly(PA) proteins. Examination of RAN translation in cultured cell models showed that poly(GP) and poly(PR) proteins were synthesized in (CCCCGG)₆₆-expressing cells, but not in cells expressing only two CCCCCG repeats. That poly(PA) proteins were not detected in (CCCCGG)₆₆-expressing cells may be because a crucial upstream sequence necessary for translation of poly(PA) peptides is missing from the (CCCCGG)₆₆ expression vector. Alternatively, it may be due to the fact that RAN translation is repeat length-dependent, with different reading frames having different length thresholds [9, 39]. Nonetheless, our immunohistochemical analysis of c9RAN proteins in human tissue indicates either this sequence is present, or the repeat length threshold is met, in c9FTD/ALS patients.

Because poly(GP) proteins can be synthesized from sense and antisense transcripts of the expanded *C9ORF72* repeat, their exact origin is not definitely known. Yet, the presence of poly(PA) and poly(PR) neuronal inclusions in post-mortem c9FTD/ALS brain tissue is indicative of RAN translation of the antisense transcript. These inclusions are specific to c9FTD/ALS cases, not being found in matched FTD/ALS controls lacking the *C9ORF72* expanded repeat, or in other repeat disorders. We did note a difference between poly(PA) and poly(PR) pathology in comparison to poly(GP) pathology, the latter being markedly more frequent. For example, poly(GP) pathology is extensive in granule cells and primary neurons of the cerebellum, but cerebellar poly(PA) and poly(PR) inclusions are sparse in these same populations.

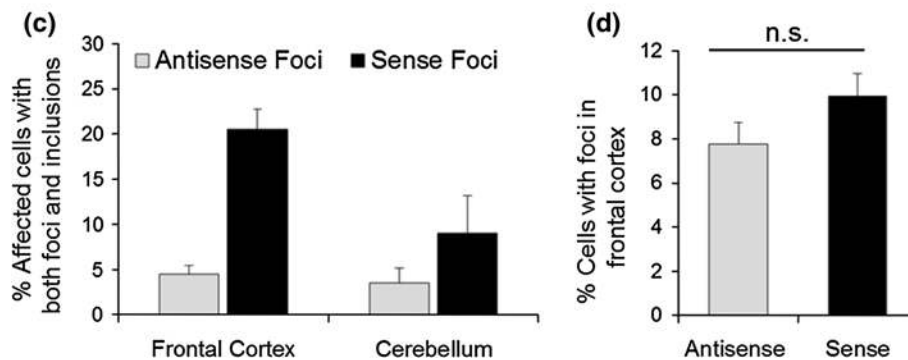
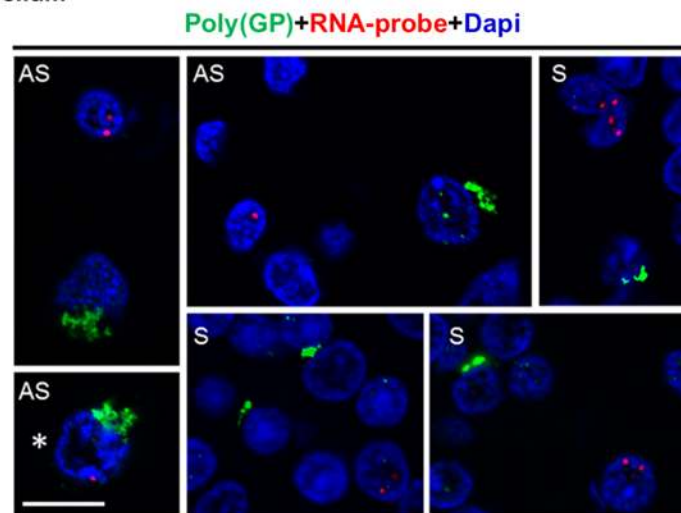
While we cannot rule-out the possibility that poly(GP) inclusions appear more abundant because of differences



(a) Frontal Cortex



(b) Cerebellum



in antibody affinities, their increased frequency may be due to the fact that poly(GP) proteins are synthesized from both sense and antisense transcripts. In addition, the antisense transcript may be less efficiently translated than the sense transcript. CCCC GG repeats are predicted to fold

into an imperfect hairpin akin to what we have previously shown for GGGGCC repeats [3]. However, compared to the GGGGCC repeat, which has an optimal organization to maximize base pairing, the CCCC GG repeat has fewer base pairs and more frequent sets of four unpaired

◀ **Fig. 7** Nuclear RNA foci and poly(GP) inclusions are seldom observed in the same cell. Fluorescence in situ hybridization of frontal cortex and cerebellar tissue of c9FTD/ALS cases using a probe against sense and antisense *C9ORF72* hexanucleotide repeat transcripts was followed by immunofluorescence staining to detect poly(GP) inclusions, which may result from RAN translation of both sense and antisense transcripts. Though infrequent, both foci and poly(GP) inclusions can co-occur in the same cell (cells indicated by an asterisk in **a** and **b**). Scale bars 10 μ m. AS antisense foci, S sense foci. **c** To determine the percentage of cells having both foci and poly(GP) inclusions in the frontal cortex and cerebellum, quantitative analysis was undertaken on four cases co-stained for poly(GP) inclusions and either sense or antisense foci. For each section, we examined 25 cells with inclusions and determined whether they had foci, and examined 25 cells with foci and determined whether they had inclusions, for a total of 50 cells per section. Data are presented as mean \pm SEM, $n = 4$. The brain region sampled (frontal cortex vs. cerebellum, $P = 0.0292$) and foci type (antisense vs. sense, $P = 0.0011$) both significantly affect the percentage of cells having foci and inclusions, as assessed by two-way ANOVA. To determine the percentage of cells with antisense and sense foci in the frontal cortex, the number of cells with foci and the total number of cells were counted in 12 randomly selected, non-overlapping fields for each case. Data are presented as mean \pm SEM, $n = 4$. No significant difference between the percentage of cells with sense or antisense foci was detected, as assessed by paired two-tailed t test ($P = 0.1916$) (**d**). n.s. not significant

C nucleotides within the stem, and thus relies more on base stacking effects. While the stability of both CCCCCG and GGGGCC repeat transcripts increases with increasing repeat length, the CCCCCG repeat is predicted to be less stable than its GGGGCC counterpart, which may decrease its susceptibility to RAN translation. It should be mentioned that two previous studies report that r(GGGGCC) repeats form an intramolecular G-quadruplex structure [13, 29]. As has been suggested for other quadruplexes [8] and r(CG) repeats, the r(GGGGCC) repeat, and perhaps the r(CCCCCG) repeat, may adopt two conformations that are in equilibrium: an extended hairpin structure and a quadruplex.

Another possible explanation for the higher frequency of poly(GP) inclusions in c9FTD/ALS may be due to ATG-initiated translation. While no sequence has been reported for the antisense transcript, resorting to analysis of the genomic DNA consensus sequence for *C9ORF72* has revealed no ATG codons for sense transcript-derived c9RAN proteins, or for the antisense transcript-derived poly(PA) protein. However, one and three potential ATG start codons were found in the antisense poly(PR) and poly(GP) frames, respectively (Online Resource 1). Nonetheless, whether ATG codons are in fact present in RNA transcripts is not yet known, and our cell culture data provide evidence that poly(PR) and poly(GP) proteins can be translated from (CCCCGG)_{exp} in the absence of an ATG initiation site.

Whether aberrant expression of c9RAN proteins, and the inclusions formed thereof, influence disease

pathogenesis remains to be elucidated. Nevertheless, the formation of abnormal proteinaceous inclusions is associated with neurotoxicity in various neurodegenerative diseases. The inclusions may sequester other proteins causing loss of function, impair/overwhelm protein degradation systems, displace cytoplasmic organelles, and may themselves have neurotoxic properties. It is noteworthy that polyA and polyG proteins synthesized by RAN translation of expanded CAG-CTG repeats accumulate in disease-relevant tissues of patients with spinocerebellar ataxia type 8 and DM1, and that their expression in cultured cells is sufficient to cause apoptotic cell death [26]. In addition, in both DM1 patients and in mice expressing (CUG)_{exp}, polyG aggregates co-localize with caspase-8, an early indicator of polyG-induced apoptosis [26]. In c9FTD/ALS, neuronal inclusions of c9RAN proteins are similarly present in vulnerable areas (e.g., neurons of neocortex and hippocampus), but there is a paucity of such inclusions in certain affected areas, such as the spinal cord [3, 25]. As with other aggregation-prone proteins involved in neurodegeneration (e.g., tau [14]), it remains unclear whether c9RAN inclusions per se are toxic, or whether RAN translation of transcripts from the *C9ORF72* repeat contributes to neurodegeneration through the formation of toxic, soluble oligomers.

As with RAN translation, the consequence of foci formation in c9FTD/ALS is currently under investigation. Taking cues from other repeat expansion disorders, it is anticipated that foci will sequester select RNA-binding proteins, and cause the misregulation of crucial downstream RNA targets. To date, studies have identified several proteins that bind GGGGCC transcripts [24, 29, 38]. The findings herein highlight the necessity to similarly identify proteins bound and sequestered by CCCCCG foci.

Examination of foci formed of sense or antisense transcripts of the expanded *C9ORF72* repeat revealed that, in addition to the frontal cortex and spinal cord, RNA foci accumulate in the cerebellum. Together with other studies, it is now evident that numerous pathological features are present in the cerebellum of c9FTD/ALS patients, including nuclear RNA foci, c9RAN protein inclusions, as well as p62-positive inclusions [1, 3, 6, 25, 26]. Furthermore, FTD cases caused by the *C9ORF72* repeat expansion show atrophy of the parietal lobe and cerebellum, in addition to frontal and temporal lobe atrophy [37]. In fact, c9FTD is characterized by greater cerebellar atrophy than sporadic FTD, as well as FTD caused by mutations in *MAPT* [37]. Consequently, cerebellar atrophy, nuclear RNA foci, and proteinaceous inclusions may be considered characteristic features of c9FTD/ALS. Future studies of c9FTD/ALS should therefore encompass evaluations of the cerebellum which, to date, has been largely neglected.

Of interest, foci and poly(GP) inclusions were seldom observed in the same cell in the frontal cortex and cerebellum of c9FTD/ALS patients. Although it is possible that foci too small to be easily detected are present in cells with poly(GP) inclusions, our findings suggest that only those (GGGGCC)_{exp} or (CCCCGG)_{exp} transcripts that escape being sequestered as foci, and are instead exported to the cytoplasm, become available for RAN translation. That RNA foci are found in both neuronal and glial cells (Figs. 4, 5), while poly(GP) inclusions are neuronal [3], coupled with the fact that RNA foci in the cerebellum are found in greatest abundance in cells in proximity to the Purkinje cell layer, whereas poly(GP) inclusions are widely expressed throughout the molecular and granule layers, does provide a basis for the lack of co-occurrence of these two features in a given cell. While it remains to be determined whether other c9RAN proteins are more frequently found in the same cells as foci, and whether both sense and antisense foci are present within the same cell, these findings support the notion that foci and inclusions represent two distinct pathogenic mechanisms for *C9ORF72* repeat expansions.

Since the 2011 discovery that the expanded hexanucleotide repeat in *C9ORF72* causes chromosome 9p-linked FTD and ALS [11, 30], several neuropathological features unique to c9FTD/ALS have been identified [1, 3, 5, 7, 11, 24, 25, 28]. The findings from the present study expand this list and highlight the need to broaden our view of potential disease mechanisms to include toxicity potential stemming from the antisense transcript. Going forward, it will be critical to distinguish if and how RNA transcripts and c9RAN proteins contribute to the pathogenesis of disease, and whether the frequency or regional localization of RNA foci and c9RAN inclusions correlate with distinct clinical features. While these questions are being investigated, c9RAN proteins should be explored as a biomarker for c9FTD/ALS, as should treatment strategies aimed at eradicating the putatively toxic sense and antisense transcripts responsible for both foci formation and RAN translation.

Acknowledgments We are grateful to all patients, family members, and caregivers who agreed to brain donation, without which these studies would have been impossible. We also acknowledge expert technical assistance of Linda Rousseau and Virginia Phillips for histology, and Beth Marten, Pamela Desaro, Amelia Johnston and Kristin Staggs-Douberly for brain banking. This work was supported by Mayo Clinic Foundation; National Institutes of Health/National Institute on Aging [R01 AG026251 (LP, RR), P01 AG003949 (DWD), P50 AG016574 (DWD, RR)]; National Institutes of Health/National Institute of Neurological Disorders and Stroke [R21 NS074121 (TFG, KB), R21 NS079807 (YZ), R01 NS080882 (RR), R01 NS063964 (LP); R01 NS077402 (LP), P50 NS072187 (DWD, RR); R21 NS084528 (LP)]; National Institute of Environmental Health Services [R01 ES20395 (LP)]; Amyotrophic Lateral Sclerosis Association (KB, LP); ALS Therapy Alliance (RR).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Al-Sarraj S, King A, Troakes C, Smith B, Maekawa S, Bodi I, Rogelj B, Al-Chalabi A, Hortobagyi T, Shaw CE (2011) p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTL and MND/ALS. *Acta Neuropathol* 122(6):691–702. doi:10.1007/s00401-011-0911-2
- Almeida S, Gascon E, Tran H, Chou HJ, Gendron TF, Degroot S, Tapper AR, Sellier C, Charlet-Berguerand N, Karydas A, Seeley WW, Boxer AL, Petrucelli L, Miller BL, Gao FB (2013) Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathol* 126(3):385–399. doi:10.1007/s00401-013-1149-y
- Ash PE, Bieniek KF, Gendron TF, Caulfield T, Lin WL, DeJesus-Hernandez M, van Blitterswijk MM, Jansen-West K, Paul JW 3rd, Rademakers R, Boylan KB, Dickson DW, Petrucelli L (2013) Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 77(4):639–646. doi:10.1016/j.neuron.2013.02.004
- Belzil VV, Gendron TF, Petrucelli L (2012) RNA-mediated toxicity in neurodegenerative disease. *Mol Cell Neurosci* 56:406–419. doi: 10.1016/j.mcn.2012.12.006
- Bieniek KF, Murray ME, Rutherford NJ, Castanedes-Casey M, DeJesus-Hernandez M, Liesinger AM, Baker MC, Boylan KB, Rademakers R, Dickson DW (2013) Tau pathology in frontotemporal lobar degeneration with C9ORF72 hexanucleotide repeat expansion. *Acta Neuropathol* 125(2):289–302. doi:10.1007/s00401-012-1048-7
- Boxer AL, Mackenzie IR, Boeve BF, Baker M, Seeley WW, Crook R, Feldman H, Hsiung GY, Rutherford N, Laluz V, Whitwell J, Foti D, McDade E, Molano J, Karydas A, Wojtas A, Goldman J, Mirsky J, Sengdy P, Dearmond S, Miller BL, Rademakers R (2011) Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *J Neurol Neurosurg Psychiatry* 82(2):196–203. doi:10.1136/jnnp.2009.204081
- Brettschneider J, Van Deerlin VM, Robinson JL, Kwong L, Lee EB, Ali YO, Safren N, Monteiro MJ, Toledo JB, Elman L, McCluskey L, Irwin DJ, Grossman M, Molina-Porcel L, Lee VM, Trojanowski JQ (2012) Pattern of ubiquilin pathology in ALS and FTL indicates presence of C9ORF72 hexanucleotide expansion. *Acta Neuropathol* 123(6):825–839. doi:10.1007/s00401-012-0970-z
- Bugaut A, Murat P, Balasubramanian S (2012) An RNA hairpin to G-quadruplex conformational transition. *J Am Chem Soc* 134(49):19953–19956. doi:10.1021/ja308665g
- Cleary JD, Ranum LP (2013) Repeat-associated non-ATG (RAN) translation in neurological disease. *Hum Mol Genet* 22(R1):R45–R51. doi:10.1093/hmg/ddt371
- Dansithong W, Paul S, Comai L, Reddy S (2005) MBNL1 is the primary determinant of focus formation and aberrant insulin receptor splicing in DM1. *J Biol Chem* 280(7):5773–5780. doi:10.1074/jbc.M410781200
- DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, Nicholson AM, Finch NA, Flynn H, Adamson J, Kouri N, Wojtas A, Sengdy P, Hsiung GY, Karydas A, Seeley WW, Josephs KA, Coppola G, Geschwind DH, Wszolek ZK, Feldman H, Knopman DS, Petersen RC, Miller BL, Dickson DW, Boylan KB, Graff-Radford NR, Rademakers

- R (2011) Expanded GGGGCC hexanucleotide repeat in non-coding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72(2):245–256. doi:10.1016/j.neuron.2011.09.011
12. Ding Y, Lawrence CE (2001) Statistical prediction of single-stranded regions in RNA secondary structure and application to predicting effective antisense target sites and beyond. *Nucleic Acids Res* 29(5):1034–1046
 13. Fratta P, Mizielińska S, Nicoll AJ, Zloh M, Fisher EM, Parkinson G, Isaacs AM (2012) C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci Rep* 2:1016. doi:10.1038/srep01016
 14. Gendron TF, Petrucelli L (2009) The role of tau in neurodegeneration. *Mol Neurodegen* 4:13. doi:10.1186/1750-1326-4-13
 15. Gijssels I, Van Langenhove T, van der Zee J, Slegers K, Philijns S, Kleinberger G, Janssens J, Bettens K, Van Cauwenberghe C, Pereson S, Engelborghs S, Sieben A, De Jonghe P, Vandenberghe R, Santens P, De Bleecker J, Maes G, Baumer V, Dillen L, Joris G, Cuijt I, Corsmit E, Elinck E, Van Dongen J, Vermeulen S, Van den Broeck M, Vaerenberg C, Mattheijssens M, Peeters K, Robberecht W, Cras P, Martin JJ, De Deyn PP, Cruts M, Van Broeckhoven C (2012) A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol* 11(1):54–65. doi:10.1016/S1474-4422(11)70261-7
 16. Giordana MT, Ferrero P, Grifoni S, Pellerino A, Naldi A, Montuschi A (2011) Dementia and cognitive impairment in amyotrophic lateral sclerosis: a review. *Neurol Sci* 32(1):9–16. doi:10.1007/s10072-010-0439-6
 17. Hofacker IL (2003) Vienna RNA secondary structure server. *Nucleic Acids Res* 31(13):3429–3431
 18. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P (1994) Fast folding and comparison of rna secondary structures. *Monatsh Chem* 125:167–188
 19. Lomen-Hoerth C, Murphy J, Langmore S, Kramer JH, Olney RK, Miller B (2003) Are amyotrophic lateral sclerosis patients cognitively normal? *Neurology* 60(7):1094–1097
 20. Mankodi A, Takahashi MP, Jiang H, Beck CL, Bowers WJ, Moxley RT, Cannon SC, Thornton CA (2002) Expanded CUG repeats trigger aberrant splicing of CIC-1 chloride channel pre-mRNA and hyperexcitability of skeletal muscle in myotonic dystrophy. *Mol Cell* 10(1):35–44. doi:S1097276502005634
 21. Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH (2004) Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci USA* 101(19):7287–7292. doi:10.1073/pnas.0401799101
 22. Mathews DH, Zuker M (2004) Predictive methods using RNA sequences. In: Baxevanis AD, Ouellette BFF (eds) *Bioinformatics: a practical guide to the analysis of genes and proteins*. Wiley, New York, pp 143–170
 23. Mooers BH, Logue JS, Berglund JA (2005) The structural basis of myotonic dystrophy from the crystal structure of CUG repeats. *Proc Natl Acad Sci USA* 102(46):16626–16631
 24. Mori K, Lammich S, Mackenzie IR, Forne I, Zilow S, Kretschmar H, Edbauer D, Janssens J, Kleinberger G, Cruts M, Hermes J, Neumann M, Van Broeckhoven C, Arzberger T, Haass C (2013) hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol* 125(3):413–423. doi:10.1007/s00401-013-1088-7
 25. Mori K, Weng SM, Arzberger T, May S, Rentzsch K, Kremmer E, Schmid B, Kretschmar HA, Cruts M, Van Broeckhoven C, Haass C, Edbauer D (2013) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science* 339(6125):1335–1338. doi:10.1126/science.1232927
 26. Murray ME, DeJesus-Hernandez M, Rutherford NJ, Baker M, Duara R, Graff-Radford NR, Wszolek ZK, Ferman TJ, Josephs KA, Boylan KB, Rademakers R, Dickson DW (2011) Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. *Acta Neuropathol* 122(6):673–690. doi:10.1007/s00401-011-0907-y
 27. Phukan J, Pender NP, Hardiman O (2007) Cognitive impairment in amyotrophic lateral sclerosis. *Lancet Neurol* 6(11):994–1003
 28. Pikkarainen M, Hartikainen P, Alafuzoff I (2010) Ubiquitinated p62-positive, TDP-43-negative inclusions in cerebellum in frontotemporal lobar degeneration with TAR DNA binding protein 43. *Neuropathology* 30(2):197–199. doi:10.1111/j.1440-1789.2009.01043.x
 29. Reddy K, Zamiri B, Stanley SY, Macgregor RB Jr, Pearson CE (2013) The disease-associated r(GGGGCC)_n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J Biol Chem* 288(14):9860–9866. doi:10.1074/jbc.C113.452532
 30. Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, Schymick JC, Laaksovirta H, van Swieten JC, Myllykangas L, Kalimo H, Paetau A, Abramzon Y, Remes AM, Kaganovich A, Scholz SW, Duckworth J, Ding J, Harmer DW, Hernandez DG, Johnson JO, Mok K, Ryten M, Trabzuni D, Guerreiro RJ, Orrell RW, Neal J, Murray A, Pearson J, Jansen IE, Sondervan D, Seelaar H, Blake D, Young K, Halliwell N, Callister JB, Toulson G, Richardson A, Gerhard A, Snowden J, Mann D, Neary D, Nalls MA, Peuralinna T, Jansson L, Isoviita VM, Kaivorinne AL, Holtta-Vuori M, Ikonen E, Sulkava R, Benatar M, Wu J, Chio A, Restagno G, Borghero G, Sabatelli M, Heckerman D, Rogava E, Zinman L, Rothstein JD, Sendtner M, Drepper C, Eichler EE, Alkan C, Abdullaev Z, Pank SD, Dutra A, Pak E, Hardy J, Singleton A, Williams NM, Heutink P, Pickering-Brown S, Morris HR, Tienari PJ, Traynor BJ (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72(2):257–268. doi:10.1016/j.neuron.2011.09.010
 31. Rouillard JM, Zuker M, Gulari E (2003) OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res* 31(12):3057–3062
 32. Sobczak K, de Mezer M, Michlewski G, Krol J, Krzyzosiak WJ (2003) RNA structure of trinucleotide repeats associated with human neurological diseases. *Nucleic Acids Res* 31(19):5469–5482
 33. Taneja KL, McCurrach M, Schalling M, Housman D, Singer RH (1995) Foci of trinucleotide repeat transcripts in nuclei of myotonic dystrophy cells and tissues. *J Cell Biol* 128(6):995–1002
 34. Todd PK, Oh SY, Krans A, He F, Sellier C, Frazer M, Renoux AJ, Chen KC, Scaglione KM, Basur V, Elenitoba-Johnson K, Vonsattel JP, Louis ED, Sutton MA, Taylor JP, Mills RE, Charlet-Berguerand N, Paulson HL (2013) CGG Repeat-Associated Translation Mediates Neurodegeneration in Fragile X Tremor Ataxia Syndrome. *Neuron* 78(3):440–455. doi:10.1016/j.neuron.2013.03.026
 35. van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, Murray ME, Heckman MG, Diehl NN, Brown PH, Baker MC, Finch NA, Bauer PO, Serrano G, Beach TG, Josephs KA, Knopman DS, Petersen RC, Boeve BF, Graff-Radford NR, Boylan KB, Petrucelli L, Dickson DW, Rademakers R (2013) Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol* 12(10):978–988. doi:10.1016/S1474-4422(13)70210-2
 36. Walter AE, Turner DH, Kim J, Lyttle MH, Muller P, Mathews DH, Zuker M (1994) Coaxial stacking of helices enhances

- binding of oligoribonucleotides and improves predictions of RNA folding. *Proc Natl Acad Sci USA* 91(20):9218–9222
37. Whitwell JL, Weigand SD, Boeve BF, Senjem ML, Gunter JL, DeJesus-Hernandez M, Rutherford NJ, Baker M, Knopman DS, Wszolek ZK, Parisi JE, Dickson DW, Petersen RC, Rademakers R, Jack CR Jr, Josephs KA (2012) Neuroimaging signatures of frontotemporal dementia genetics: C9ORF72, tau, progranulin and sporadics. *Brain* 135(Pt 3):794–806. doi:[10.1093/brain/aws001](https://doi.org/10.1093/brain/aws001)
38. Xu Z, Poidevin M, Li X, Li Y, Shu L, Nelson DL, Li H, Hales CM, Gearing M, Wingo TS, Jin P (2013) Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc Natl Acad Sci USA* 110(19):7778–7783. doi:[10.1073/pnas.1219643110](https://doi.org/10.1073/pnas.1219643110)
39. Zu T, Gibbens B, Doty NS, Gomes-Pereira M, Huguet A, Stone MD, Margolis J, Peterson M, Markowski TW, Ingram MA, Nan Z, Forster C, Low WC, Schoser B, Somia NV, Clark HB, Schmechel S, Bitterman PB, Gourdon G, Swanson MS, Moseley M, Ranum LP (2011) Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci USA* 108(1):260–265. doi:[10.1073/pnas.1013343108](https://doi.org/10.1073/pnas.1013343108)
40. Zuker M, Jacobson AB (1998) Using reliability information to annotate RNA secondary structures. *RNA* 4(6):669–679
41. Zuker M, Mathews DH, Turner DH (1999) Algorithms and thermodynamics for RNA secondary structure prediction: A practical guide. In: Barciszewski J, Clark BFC (eds) *RNA Biochemistry and Biotechnology*. Kluwer Academic Publishers, Boston, pp 11–43