© 2019. Published by The Company of Biologists Ltd | Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

RESEARCH ARTICLE

Nuclear RNA foci from *C9ORF72* expansion mutation form paraspeckle-like bodies

Ana Bajc Česnik[1,2,‡], Simona Darovic[1,3,‡], Sonja Prpar Mihevc[1,‡], Maja Štalekar[1,3], Mirjana Malnar[1,2], Helena Motaln[1], Youn-Bok Lee[4], Julija Mazej[1,5], Jure Pohleven[1,*], Markus Grosch[6], Miha Modic[6], Marko Fonovič[7], Boris Turk[7], Micha Drukker[6], Christopher E. Shaw[4] and Boris Rogelj[1,3,5,§]

ABSTRACT

The GGGGCC (G_4C_2) repeat expansion mutation in the *C9ORF72* gene is the most common genetic cause of frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS). Transcription of the repeat and formation of nuclear RNA foci, which sequester specific RNA-binding proteins, is one of the possible pathological mechanisms. Here, we show that $(G_4C_2)_n$ repeat RNA predominantly associates with essential paraspeckle proteins SFPQ, NONO, RBM14, FUS and hnRNPH and colocalizes with known paraspeckle-associated RNA *hLinc-p21*. As formation of paraspeckles in motor neurons has been associated with early phases of ALS, we investigated the extent of similarity between paraspeckles and $(G_4C_2)_n$ RNA foci. Overexpression of $(G_4C_2)_{72}$ RNA results in their increased number and colocalization with SFPQ-stained nuclear bodies. These paraspeckle-like $(G_4C_2)_{72}$ RNA foci form independently of the known paraspeckle scaffold, the long non-coding RNA *NEAT1*. Moreover, the knockdown of SFPQ protein in *C9ORF72* expansion mutation-positive fibroblasts significantly reduces the number of $(G_4C_2)_n$ RNA foci. In conclusion, $(G_4C_2)_n$ RNA foci have characteristics of paraspeckles, which suggests that both RNA foci and paraspeckles play roles in FTD and ALS, and implies approaches for regulation of their formation.

KEY WORDS: Paraspeckles, *C9ORF72*, SFPQ, RNA foci, *NEAT1*

INTRODUCTION

The $(G_4C_2)_n$ hexanucleotide repeat expansion mutation within the first intron of *C9ORF72* is the most common known pathogenic mutation associated with amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Intronic $(G_4C_2)_n$ repeat expansion was proposed to cause toxicity and neurodegeneration via three pathogenic mechanisms (Vatovec et al., 2014). The first one refers to haploinsufficiency, arising from the formation of complex secondary DNA structures in the mutant allele, which may result in reduced levels of *C9ORF72* transcript and protein (DeJesus-Hernandez et al., 2011; Kovanda et al., 2015; Renton et al., 2011; Shi et al., 2018; Šket et al., 2015; Waite et al., 2014; Xiao et al., 2015). The second refers to RNA toxicity, whereby the transcripts containing $(G_4C_2)_n$ repeats can form nuclear RNA foci, known to bind and sequester RNA-binding proteins, thus consequently affecting post-transcriptional processing (Almeida et al., 2013; DeJesus-Hernandez et al., 2011; Donnelly et al., 2013; Gendron et al., 2013; Haeusler et al., 2014; Lee et al., 2013; Mizielinska et al., 2013; Sareen et al., 2013; Swinnen et al., 2018; Xu et al., 2013; Zu et al., 2013). The last proposed mechanism refers to toxic di-peptide repeat (DPR) polypeptides, arising from repeat-associated non-ATG (RAN) translation of $(G_4C_2)_n$ hexanucleotide transcript, which form p62 (also known as SQSTM1)-positive and TDP-43 (encoded by *TARDBP*)-negative neuronal inclusions (Al-Sarraj et al., 2011; Ash et al., 2013; Gendron et al., 2013; Moens et al., 2018; Mori et al., 2013a,c; Saberi et al., 2018; Troakes et al., 2012; Zu et al., 2013).

RNA toxicity has been associated with other intronic repeat expansion disorders as well, including myotonic dystrophy, fragile X tremor ataxia syndrome and several spinocerebellar ataxias (Galloway and Nelson, 2009; Lee and Cooper, 2009; Orr, 2012; Todd and Paulson, 2010). RNA pulldown studies have reported on proteins that bind to $(G_4C_2)_n$ RNA *in vitro* and colocalize with nuclear RNA foci in transfected cells and mutated *C9ORF72* postmortem brain tissue (Almeida et al., 2013; Donnelly et al., 2013; Haeusler et al., 2014; Lee et al., 2013; Mori et al., 2013b; Rossi et al., 2015; Sareen et al., 2013; Xu et al., 2013; Zhang et al., 2015). Core paraspeckle proteins hnRNPH, SFPQ and NONO were listed in some of these studies (Vatovec et al., 2014).

By definition, paraspeckles are nuclear ribonuclear bodies in which one of the essential paraspeckle proteins colocalizes with the longer non-coding RNA (lncRNA) *NEAT1* (Clemson et al., 2009; Fox et al., 2018; Mao et al., 2011; Sasaki et al., 2009). *NEAT1* lncRNA is transcribed in two variants: short isoform *NEAT1_1* and long isoform *NEAT1_2*. Although both isoforms are components of paraspeckles, only *NEAT1_2* was shown to be essential for paraspeckle formation (Naganuma et al., 2012). Besides *NEAT1_2* scaffold, paraspeckles contain ∼40 proteins, with SFPQ, NONO, RBM14, hnRNPH, hnRNPK, FUS, DAZAP1 and SMARCA4 proving essential for their formation and structural integrity (Fox et al., 2018). The function of paraspeckles is a subject of intense research. They are implicated in modulating post-transcriptional processes in cells, through sequestration of RNA-binding proteins (RBPs), mRNAs and microRNAs, as well as their nuclear retention and subcellular/subnuclear compartmentalization (Anantharaman et al., 2016; Chen and Carmichael, 2009; Chen et al., 2008; Jiang et al., 2017;

[1]Department of Biotechnology, Jozef Stefan Institute, Ljubljana 1000, Slovenia. [2]Graduate School of Biomedicine, Faculty of Medicine, University of Ljubljana, Ljubljana 1000, Slovenia. [3]Biomedical Research Institute BRIS, Ljubljana 1000, Slovenia. [4]Department of Basic and Clinical Neuroscience, Institute of Psychiatry, King's College London, London SE5 9RT, UK. [5]Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana 1000, Slovenia. [6]Helmholtz Center Munich, Institute of Stem Cell Research, German Research Center for Environmental Health, Neuherberg 85764, Germany. [7]Department of Biochemistry and Molecular and Structural Biology, Jozef Stefan Institute, Ljubljana 1000, Slovenia.
*Present address: InnoRenew CoE, Izola 6310, Slovenia.
‡These authors contributed equally to this work

§Author for correspondence (boris.rogelj@ijs.si)

S.D., 0000-0002-4814-2249; S.P.M., 0000-0002-3109-1326; M.Š., 0000-0003-3114-6494; J.M., 0000-0002-5242-8303; J.P., 0000-0002-5911-8439; B.R., 0000-0003-3898-1943

Received 24 August 2018; Accepted 31 January 2019

Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

Prasanth et al., 2005). Paraspeckles have also been shown to associate and possibly retain transcripts harbouring inverted repeat *Alu* (IR*Alu*) elements or AG-rich regions (Chen and Carmichael, 2009; Chen et al., 2008; Prasanth et al., 2005; West et al., 2016; Zhang and Carmichael, 2001). More than 300 human genes contain IR*Alu* elements, which form long intramolecular RNA duplexes that are highly adenosine-to-inosine edited (Chen et al., 2008; Levanon et al., 2004). Among them is the human *hLinc-p21* (also known as *TP53COR1*) RNA, the IR*Alu* element of which is responsible for its nuclear localization and its colocalization with the paraspeckles (Chillon and Pyle, 2016).

Paraspeckle formation was observed in the early stage of ALS pathogenesis (Nishimoto et al., 2013; Shelkovnikova et al., 2018). Importantly, several key molecular players involved in the pathogenesis of ALS and FTD, such as TDP-43, FUS, TAF15, EWSR1, SS18L1 and hnRNPA1, were revealed as components of paraspeckles (Fox et al., 2018; Naganuma et al., 2012; Nishimoto et al., 2013; Shelkovnikova et al., 2014). Likewise, cytoplasmic SFPQ, an essential component of paraspeckles, was shown to be important for axonal maturation and connectivity of motor neurons, and with its mutations afflicting cytoplasmic localization associated with ALS (Thomas-Jinu et al., 2017). Furthermore, nuclear interaction between FUS and SFPQ is affected by disease mutations, and silencing of either one of the proteins induces FTD-like phenotypes in mice (Ishigaki et al., 2017). Recently, a study showed that intron retention and nuclear loss of SFPQ also associate with ALS (Luisier et al., 2018).

Here, we show that $(G_4C_2)_n$ RNA foci arising from the hexanucleotide repeat mutation in *C9ORF72* resemble paraspeckles in distinct characteristics, especially in ones previously defined to denote the structure of paraspeckles. The $(G_4C_2)_n$ RNA foci were shown by us to predominantly interact and colocalize with paraspeckle proteins SFPQ, NONO, RBM14, PSPC1, hnRNPH and FUS. Similar to paraspeckles, they were shown to sequester IR*Alu* containing *hLinc-p21* RNA, but in a *NEAT1*-independent manner. Moreover, the essentiality of SFPQ for paraspeckle formation was also mirrored in the reduction of $(G_4C_2)_n$ RNA foci upon SFPQ knockdown in *C9ORF72* mutation-positive fibroblasts. Revealing these paraspeckle-like characteristics of $(G_4C_2)_n$ RNA foci broadens our understanding of the molecular processes underlying ALS and FTD.

RESULTS

$(G_4C_2)_{72}$ foci colocalize with paraspeckle proteins in transfected HEK293T cells

To investigate whether RNA foci colocalize with paraspeckle proteins in a paraspeckle-like manner, we studied the interaction of paraspeckle proteins with $(G_4C_2)_{72}$ RNA foci in HEK293T cells transfected with a plasmid harbouring $(G_4C_2)_{72}$ repeats. Using fluorescent *in situ* hybridization (FISH) coupled to immunocytochemistry, we observed colocalization of $(G_4C_2)_{72}$ RNA foci with the essential paraspeckle proteins SFPQ, NONO, RBM14, hnRNPH, FUS and non-essential PSPC1 in transfected HEK293T cells (Fig. 1A–G).

Paraspeckle proteins bind to $(G_4C_2)_{48}$ RNA *in vitro*

Owing to the role of SFPQ and hnRNPH in paraspeckle formation, we hypothesized on the existence of structural and functional similarity between $(G_4C_2)_n$ RNA foci and paraspeckles in terms of the recruitment of their components. To identify paraspeckle-associated proteins that additionally bind to $(G_4C_2)_n$ RNA, we performed an RNA pulldown from rat cortical and cerebellar nuclear fractions using *in vitro*-transcribed 48-repeat G_4C_2 RNA, coupled to the S1 aptamer [$(G_4C_2)_{48}$-S1]. As controls, we used the RNA fragment consisting of 369 bp of DsRed sequence coupled to S1 aptamer (DsRed-S1) with an equivalent length to $(G_4C_2)_{48}$-S1 and the S1 aptamer-only RNA. Silver staining of SDS-PAGE gels showed defined bands (Fig. 1H). Using mass spectrometry to identify the differentially bound protein bands from $(G_4C_2)_{48}$-S1 and control RNA pulldown reactions, SFPQ, hnRNPH, NPM1 and EF1α2 (also known as EEF1A2) were revealed as predominant binders of $(G_4C_2)_{48}$-S1 RNA. Immunoblot analyses with rat cerebellar nuclear extracts confirmed $(G_4C_2)_{48}$ RNA binding to SFPQ, NPM, EF1α2 and hnRNPH (Fig. S1). Further immunoblot analyses were performed with mouse nuclear brain lysates, aiming at cross-species validation of our results, and rescuing for rat sample availability that did not suffice for these additional validations. With these, $(G_4C_2)_{48}$ RNA binding to paraspeckle proteins SFPQ, hnRNPH, NONO, RBM14, PSPC1 and FUS was demonstrated (Fig. 1I; Fig. S2). Of note, although NPM1 and EF1α2 were validated to interact with $(G_4C_2)_{48}$ RNA pulldowns, they failed to colocalize with $(G_4C_2)_n$ RNA foci in cells *in vitro* (Figs S2 and S3).

Paraspeckle proteins colocalize with the RNA foci in *C9ORF72* mutation-positive patient-derived fibroblasts and FTD brain tissues

To avoid misinterpretation of the results due to cell transfection and overexpression of $(G_4C_2)_{72}$ RNA, the pathological relevance of the paraspeckle proteins binding to $(G_4C_2)_n$ RNA foci was assessed in *C9ORF72* mutation-positive patient-derived fibroblasts. There, $(G_4C_2)_n$ RNA foci also colocalized with the essential paraspeckle proteins SFPQ, NONO, RBM14, hnRNPH and FUS (Fig. 2A–F).

Quantitative analysis of FISH coupled to immunostaining revealed ∼50% colocalization of $(G_4C_2)_n$ RNA foci with the paraspeckle proteins (Fig. 2G), indicating the paraspeckle-like activity of endogenous $(G_4C_2)_n$ RNA foci. This was further supported by colocalization of these RNA foci with SFPQ and NONO in cerebellar sections from *C9ORF72* mutation-positive cases (Fig. 3).

$(G_4C_2)_{72}$ RNA foci sequester SFPQ in a *NEAT1*-independent manner

In our colocalization experiments, we observed an increased number of SFPQ-stained nuclear bodies in cells overexpressing $(G_4C_2)_{72}$ RNA. Depending on the nature of $(G_4C_2)_{72}$ RNA interaction with *NEAT1*, this could either mean that there is an overall increase in paraspeckles or that the $(G_4C_2)_{72}$ RNA forms an independent scaffold for the formation of paraspeckle-like structures. Quantification of SFPQ-stained nuclear bodies in HEK293T cells overexpressing $(G_4C_2)_{72}$ repeats revealed that 99% of mock transfected cells contained less than ten SFPQ-stained nuclear bodies per cell, whereas 46% of the cells with $(G_4C_2)_{72}$ RNA nuclear foci contained ten or more SFPQ-stained nuclear bodies per cell. The average number of SFPQ-stained nuclear bodies per cell increased from 2.3±0.2 for mock transfected cells to 9.8±0.4 for cells expressing $(G_4C_2)_{72}$ repeats (Fig. 4A,B), suggesting that $(G_4C_2)_{72}$ RNA foci can modify nuclear compartmentalization of SFPQ.

As *NEAT1_2* RNA represents the scaffold for paraspeckle assembly sufficient to maintain the integrity of paraspeckles (Clemson et al., 2009; Mao et al., 2011; Naganuma et al., 2012; Sasaki et al., 2009), we examined whether this long non-coding RNA colocalizes with $(G_4C_2)_{72}$ RNA foci. We performed double-FISH staining to detect the long *NEAT1_2* isoform and $(G_4C_2)_{72}$ RNA foci in HEK293T cells transfected with a plasmid expressing $(G_4C_2)_{72}$ repeats, followed by immunocytochemistry with SFPQ-specific antibody (Fig. 4A). We observed that 13.8±1.6% of

Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

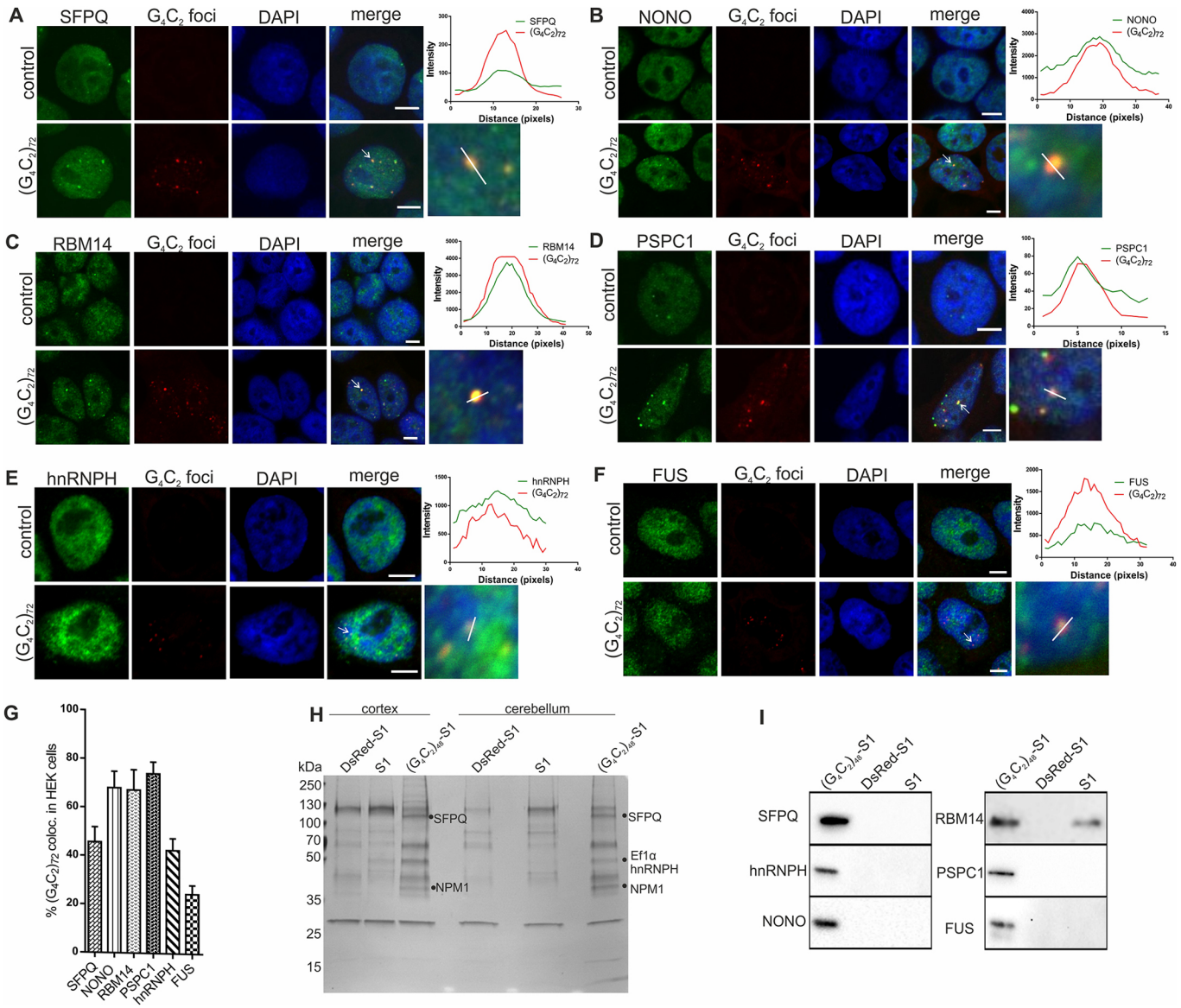

Fig. 1. Paraspeckle proteins colocalize with $(G_4C_2)_{72}$ nuclear foci in HEK293T cells and bind $(G_4C_2)_{48}$ RNA *in vitro*. (A–F) HEK293T cells transfected with an empty vector or a plasmid expressing $(G_4C_2)_{72}$ repeats and probed for $(G_4C_2)_{72}$ and SFPQ (A), NONO (B), RMB14 (C), PSPC1 (D), hnRNPH (E) or FUS (F), with quantification of colocalization in the close-up. Arrows indicate foci used for fluorescence analysis, performed with ImageJ software (https://imagej.nih.gov/ij/). Scale bars: 5 µm. (G) Quantification of $(G_4C_2)_{72}$ RNA foci colocalization with paraspeckle proteins SFPQ (45.7±6.13%), NONO (68.1±6.76%), RBM14 (67.3 ±8.36%), PSPC1 (74.1±4.88%), hnRNPH (42.6±4.91%) and FUS (24.7±3.64%). Three experiments were performed and a minimum of 50 cells per experiment were counted; data are presented as mean±s.e.m. (H) Gel electrophoresis of eluted proteins obtained by RNA pulldown from rat brain nuclear fractions using immobilized $(G_4C_2)_{48}$-S1. Differential bands eluted from $(G_4C_2)_{48}$-S1 and control RNAs [RNA fragment consisting of the first 369 bp of DsRed sequence coupled to S1 aptamer (DsRed-S1) and the S1 aptamer-only RNA] that were analysed by mass spectrometry are indicated by dots. (I) Paraspeckle proteins from mouse brain nuclear lysates SFPQ, RBM14, hnRNPH, PSPC1, NONO and FUS specifically co-precipitated with $(G_4C_2)_{48}$-S1 RNA. RNA fragment consisting of the first 369 bp of DsRed sequence coupled to S1 aptamer (DsRed-S1) with an equivalent length to $(G_4C_2)_{48}$-S1 and the S1 aptamer-only RNA were used as controls in the RNA pulldown experiment.

$(G_4C_2)_{72}$ RNA foci colocalized with *NEAT1_2*, suggesting some overlap of the RNAs. However, immunocytochemistry against SFPQ performed on those double-FISH-stained cells revealed that, although the majority of $(G_4C_2)_{72}$ RNA foci did not colocalize with *NEAT1_2* lncRNA, they still colocalized with SFPQ in much higher number (61.8±4.7%), which implied *NEAT1*-independent formation of $(G_4C_2)_{72}$ RNA (Fig. 4C). To further substantiate this claim, we established *NEAT1* knockdown HEK293T cells holding a 1.1 kbp deletion around the *NEAT1* transcription start site (a list of RNAs and primers for *NEAT1* knockdown is available in Table S1). This deletion reduced the expression level of *NEAT1* to 2.1±0.9% and expression level of *NEAT1_2* to 1.7±0.6%, compared to the expression levels in wild-type HEK293T cells (Fig. S4A). Consequently, the paraspeckle formation was strongly reduced in this cell line (Fig. S4B). Furthermore, there were no changes observed in occurrence of $(G_4C_2)_{72}$ RNA foci between nuclei of $(G_4C_2)_{72}$ RNA-transfected *NEAT1* knockdown HEK293T cells and wild-type HEK293T cells (Fig. 4A,D,E). However, colocalization of SFPQ with $(G_4C_2)_{72}$ RNA foci in these *NEAT1* knockdown HEK293T cells was still observed, confirming that the localization of the paraspeckle proteins in $(G_4C_2)_{72}$ RNA foci is a *NEAT1*-independent event (Fig. 4D).

Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

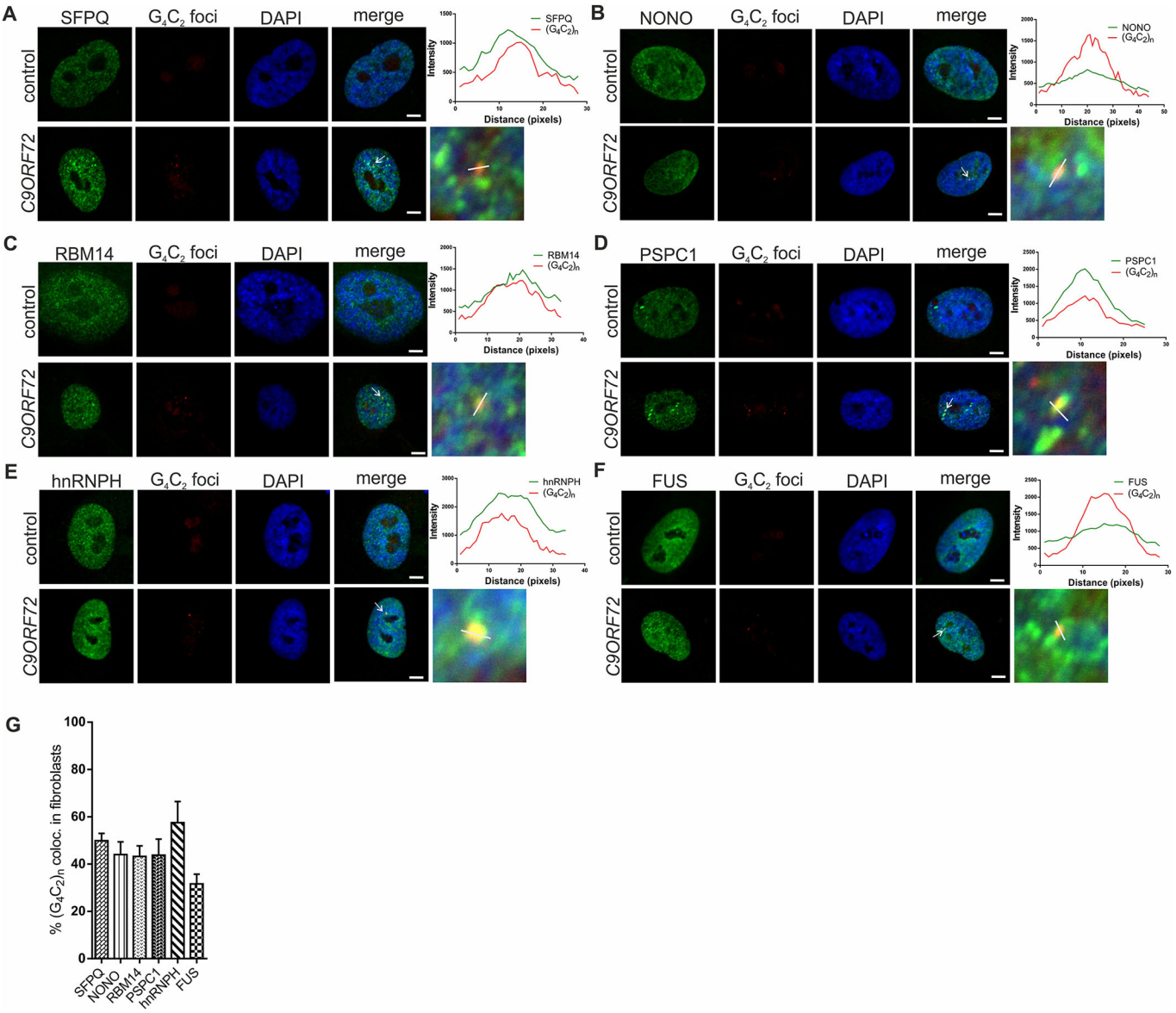

Fig. 2. Paraspeckle proteins colocalize with $(G_4C_2)_n$ nuclear foci in *C9ORF72* mutation-positive patient-derived fibroblasts. (A–F) Control and *C9ORF72* mutation-positive patient-derived fibroblasts probed for $(G_4C_2)_n$ and SFPQ (A), NONO (B), RMB14 (C), PSPC1 (D), hnRNPH (E) or FUS (F), with quantification of colocalization in the close-up. Arrows indicate foci used for fluorescence analysis, performed with ImageJ software. Scale bars: 5 µm. (G) Quantification of RNA foci colocalization with paraspeckle proteins SFPQ (49.9±3.05%), NONO (44.0±5.39%), RBM14 (43.2±4.47%), PSPC1 (43.8±6.79%), hnRNPH (57.5±8.7%) and FUS (31.6±4.11%). Three experiments were performed and a minimum of 50 cells per experiment were counted; data are presented as mean±s.e.m.

$(G_4C_2)_{72}$ RNA foci colocalize with paraspeckle-associated *hLincRNA-p21*

Following our previous observations, we hypothesized that $(G_4C_2)_{72}$ RNA foci could exhibit RNA-binding behaviour similar to *NEAT1* paraspeckles. Double-FISH staining for *hLincRNA-p21* and $(G_4C_2)_{72}$ RNA in $(G_4C_2)_{72}$-transfected HEK293T cells confirmed their colocalization, thereby evidencing paraspeckle-like RNA-binding function of $(G_4C_2)_{72}$ RNA foci (Fig. 4E).

Knockdown of SFPQ reduces the number of $(G_4C_2)_n$ foci in *C9ORF72* mutation-positive patient-derived fibroblasts

Reduction of SFPQ expression has been shown to reduce paraspeckle formation in HeLa cells (Sasaki et al., 2009). To further explore the similarity between paraspeckles and $(G_4C_2)_n$ RNA foci, we used lentiviral particles with short hairpin RNA (shRNA) to knock down SFPQ expression in *C9ORF72* mutation-positive patient-derived fibroblasts (Fig. 5A,B; Fig. S5).

Quantification of $(G_4C_2)_n$ RNA foci in fibroblasts transduced with SFPQ shRNA in comparison with scrambled shRNA, revealed a reduced number of $(G_4C_2)_n$ RNA foci in those with silenced SFPQ (Fig. 5C,D). The average number of $(G_4C_2)_n$ RNA foci per cell dropped from 5.8 for fibroblasts transduced with scrambled shRNA to 1.8 for those transduced with SFPQ shRNA. By showing that SFPQ affects the formation of $(G_4C_2)_n$ RNA foci, these results additionally substantiate the similarity between $(G_4C_2)_n$ RNA foci and paraspeckles.

DISCUSSION

In this work, we show similarities between $(G_4C_2)_n$ RNA foci and paraspeckles in four different aspects, denoting both their structure

Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

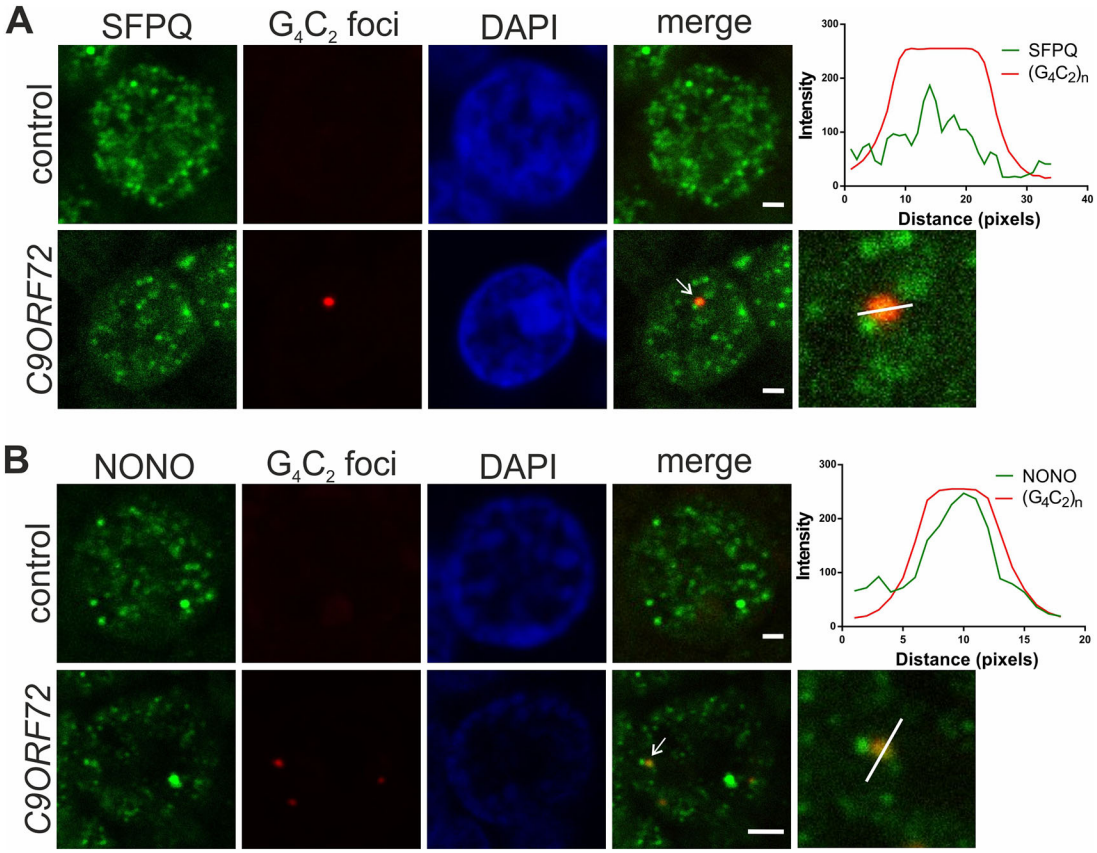

Fig. 3. Core paraspeckle proteins SFPQ and NONO colocalize with $(G_4C_2)_n$ nuclear foci in *C9ORF72* mutation-positive cerebellum. (A,B) Control or *C9ORF72* mutant cerebellar sections probed for $(G_4C_2)_n$ RNA foci and SFPQ (A) or NONO (B). The brain area between the granular and molecular layer of the cerebellum was imaged and RNA foci were found most frequently in neurons adjacent to Purkinje cells. Scale bars: 1 µm.

and function. First, we show that $(G_4C_2)_n$ RNA foci, characteristic for *C9ORF72*-associated pathology, colocalize with the essential paraspeckle proteins SFPQ, NONO, RBM14, hnRNPH and FUS, thus resembling the structure of paraspeckles. The presumed paraspeckle function is sequestration of RBPs or RNA molecules. Hence, upregulating the amount of *NEAT1_2* results in larger paraspeckles and consequent sequestration of more paraspeckle-associated proteins from nucleoplasm, which leads to their reduced availability (Fox et al., 2018; Hirose et al., 2014; Imamura et al., 2014). A similar phenomenon might be underlying $(G_4C_2)_n$ RNA foci formation in *C9ORF72*-associated ALS/FTD pathology.

Second, we show that, akin to paraspeckles, $(G_4C_2)_{72}$ RNA foci colocalize with long non-coding RNAs, such as IR*Alu* repeat RNA *hLincRNA-p21*, indicating possible disruption of normal RNA localization and export due to its nuclear retention, and pathological modulation of RBP levels. In support of our results, both cells transiently expressing $(G_4C_2)_{72}$ repeats and cortical neurons established via stem cell state from *C9ORF72* mutation-positive patient-derived fibroblasts, exhibited nuclear retention of various mRNAs (Freibaum et al., 2015; Rossi et al., 2015). Furthermore, poly(A)-binding protein PABPC1 localizes from the cytoplasm to the nuclear $(G_4C_2)_{72}$ RNA foci in transiently transfected cells, which is indicative of nuclear retention of poly(A) mRNAs (Rossi et al., 2015). Hence, the formation of the paraspeckle-like structures may also be one of the mechanisms leading to nuclear retention of mRNAs and mRNA export disruption. Whether or not the nuclear accumulation of RNAs represents a critical step in neurodegenerative disease development still needs to be elucidated. *NEAT1* upregulation observed in the brains of FTD patients (Tollervey et al., 2011), and the possibility that $(G_4C_2)_n$ RNA forms alternative paraspeckle-like foci, increases the importance of the paraspeckles and paraspeckle-like structures in the pathogenesis of ALS and FTD.

Third, we show that $(G_4C_2)_{72}$ RNA foci form paraspeckle-like bodies in a *NEAT1*-independent manner. *NEAT1_2* is defined as an RNA scaffold component of paraspeckles (Chen and Carmichael, 2009; Clemson et al., 2009; Sasaki et al., 2009; Sunwoo et al., 2009), yet our findings suggest that $(G_4C_2)_n$ RNA repeats could replace *NEAT1* RNA as a scaffold in paraspeckle-like structures. Likewise, phosphorothioate-modified antisense oligonucleotides have been revealed to serve as scaffolds for paraspeckle protein assembly in the absence of *NEAT1* RNA (Shen et al., 2014). Together with our results, this implies the heterogeneity of RNA foci in the cells that, although differently structured, may nonetheless function in a paraspeckle-like manner. As mentioned, $(G_4C_2)_n$ RNA foci colocalize with paraspeckle proteins SFPQ, NONO, RBM14, PSPC1, hnRNPH and FUS, as demonstrated by mapping $(G_4C_2)_n$-protein interactions by quantitative proteomics and microscopy studies of colocalization. Similar to paraspeckles, they also sequester *hLinc-p21* lncRNA, but in a *NEAT1_2*-independent manner. This altogether thus calls for a broadening of the paraspeckle definition, originally stating a paraspeckle to be a nuclear body in which one of the essential paraspeckle proteins colocalizes with *NEAT1* RNA (Fox et al., 2018), to also include other RNA scaffolds amenable of binding the paraspeckle proteins and RNAs.

Finally, we show that, besides in paraspeckles, SFPQ also plays an important role in the formation of $(G_4C_2)_n$ RNA foci, which

Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

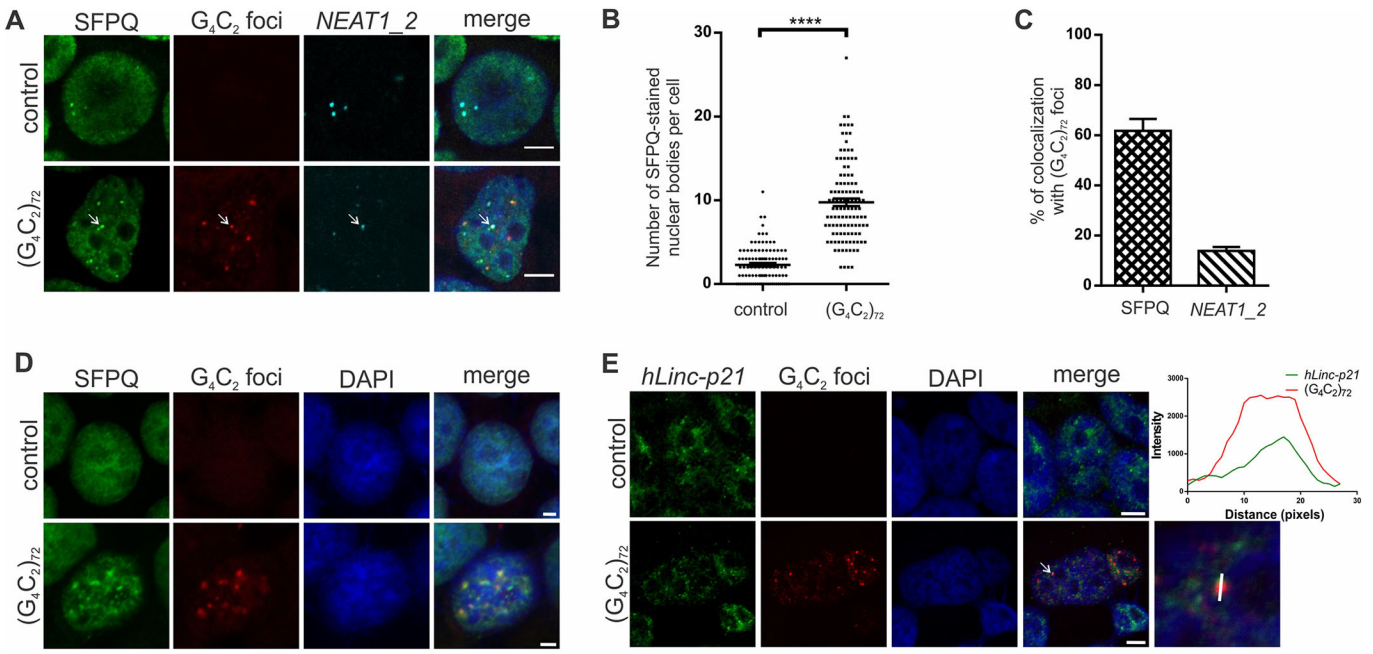

Fig. 4. $(G_4C_2)_{72}$ nuclear foci form paraspeckle-like structures independent of *NEAT1_2* and colocalize with paraspeckle-associated RNA *hLincRNA-p21*. (A) HEK293T cells transfected with an empty vector or a plasmid expressing $(G_4C_2)_{72}$ repeats and probed for *NEAT1_2*, $(G_4C_2)_{72}$ and SFPQ. Arrows indicate foci colocalizing with *NEAT1_2* and SFPQ. Scale bars: 5 µm. (B) The number of SFPQ-stained nuclear bodies in HEK293T cells transfected with an empty vector or a plasmid expressing $(G_4C_2)_{72}$ repeats (Student's *t*-test, ****P<0.0001). Three experiments were performed and a minimum of 30 cells were counted per transfection; data are presented as mean±s.e.m. (C) Percentage of $(G_4C_2)_{72}$ RNA foci colocalizing with SFPQ and *NEAT1_2*. Three experiments were performed and a minimum of 100 foci were counted per transfection; data are presented as mean±s.e.m. (D) *NEAT1* knockdown HEK293T cells transfected with an empty vector or a plasmid expressing $(G_4C_2)_{72}$ repeats and probed for $(G_4C_2)_{72}$ and SFPQ. Scale bars: 2 µm. (E) HEK293T cells transfected with an empty vector or a plasmid expressing $(G_4C_2)_{72}$ repeats and probed for *hLincRNA-p21* and $(G_4C_2)_{72}$, with quantification of colocalization in the close-up. Arrow indicates the focus used for fluorescence analysis, performed with ImageJ software. Scale bars: 5 µm.

makes SFPQ an interesting target for further research in terms of disease mechanisms and therapeutics development. SFPQ is one of the essential proteins responsible for paraspeckle integrity and, together with NONO, ensures the stability of the *NEAT1_2* scaffold by binding along its length (Fox et al., 2018; Sasaki et al., 2009). Reduced levels of functional SPFQ in the motor axons have been linked to ALS pathology (Thomas-Jinu et al., 2017). We observed colocalization of $(G_4C_2)_n$ RNA foci with SFPQ that suggested a structural role of SFPQ also in the formation of $(G_4C_2)_n$ RNA foci. Accordingly, SFPQ knockdown in *C9ORF72* mutation-positive patient-derived fibroblasts resulted in a lower number of nuclear $(G_4C_2)_n$ RNA foci. There, $(G_4C_2)_n$ repeat RNA may have been potentially released to the cytoplasm and made available for RAN translation. In this respect, a protective function could be assigned to $(G_4C_2)_n$ RNA foci, ensured by bound paraspeckle proteins. Indeed, in recent studies investigating RNA toxicity, diffuse RNA repeats and $(G_4C_2)_n$ RNA foci present in cytoplasm were shown to cause axonal abnormalities in a zebrafish model (Swinnen et al., 2018), whereas in fly models exhibiting numerous cytoplasmic or nuclear foci, no toxic effect was noted (Moens et al., 2018). Accordingly, Shi and co-workers claimed that toxicity in ALS/FTD human induced motor neurons arises from a synergistic effect of *C9ORF72* haploinsufficiency and DPR accumulation (Shi et al., 2018). On the other hand, due to similarities between $(G_4C_2)_n$ RNA foci and paraspeckles, possible RNA toxicity may be assigned to paraspeckles as well. Indeed, the increase in the number of paraspeckles has already been observed in the early phase of motor neuron degeneration in ALS (Nishimoto et al., 2013). Further ahead, the toxic or protective status of $(G_4C_2)_n$ RNA foci could guide the development of therapeutics based on modulating levels of associated proteins. This could result in sequestering RNA repeats into protective foci or reducing the formation of toxic RNA foci.

MATERIALS AND METHODS

Ethics statement

Human postmortem brain sections were provided by the Medical Research Council (MRC) London Neurodegenerative Diseases Brain Bank (Institute of Psychiatry, King's College London) and were collected and distributed in accordance with local and national research ethics committee approval. Rat and mouse brains were isolated by approval by the Veterinary Administration of the Ministry of Agriculture and the Environment, Slovenia.

Antibodies

The following commercial antibodies were used: NPM1-specific mouse monoclonal antibody [sc-56622, Santa Cruz Biotechnologies; immunocytochemistry (ICC), 1:50], EF1α2-specific rabbit polyclonal antibody [sc-68481, Santa Cruz Biotechnologies; western blotting (WB), 1:100], NONO-specific mouse monoclonal antibody (sc-166702, Santa Cruz Biotechnologies; WB, 1:250), NONO-specific rabbit polyclonal antibody (ab70335, Abcam; ICC, 1:1000), SFPQ-specific rabbit polyclonal antibody (sc-28730, Santa Cruz Biotechnologies; ICC, 1:1000), SFPQ-specific rabbit monoclonal antibody (ab177149, Abcam; ICC, 1:2000), RBM14-specific rabbit polyclonal antibody (NBP1-84416, Novus Biologicals; ICC, 1:250), PSPC1-specific mouse monoclonal antibody (SAB4200503, Sigma-Aldrich; ICC, 1:500), PSPC1-specific rabbit polyclonal antibody (sc-84577, Santa Cruz Biotechnologies; ICC, 1:500), FUS-specific rabbit polyclonal antibody (NB100-565, Novus Biologicals; WB, 1:2000; ICC, 1:500), hnRNPH-specific rabbit polyclonal antibody (NB100-385, Novus Biologicals; ICC, 1:200) and hnRNPH-specific rabbit polyclonal antibody (ab10374, Abcam; WB, 1:500).

Journal of Cell Science (2019) 132, jcs224303. doi:10.1242/jcs.224303

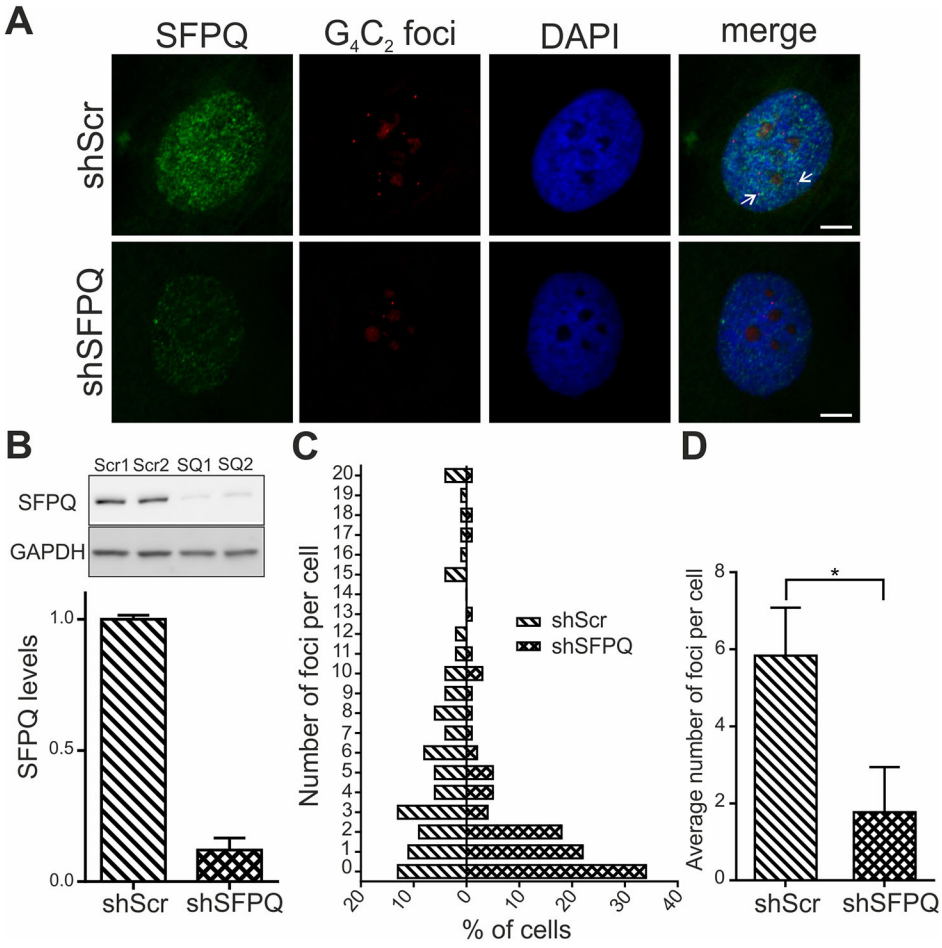

Fig. 5. Knockdown of the essential paraspeckle protein SFPQ reduces the number of $(G_4C_2)_n$ foci in *C9ORF72* mutation-positive patient-derived fibroblasts. (A) *C9ORF72* mutation-positive fibroblasts transduced with lentiviral particles with scrambled or SFPQ shRNA probed for $(G_4C_2)_n$ and SFPQ. Overlapping $(G_4C_2)_n$ foci with SFPQ are indicated by arrows. Scale bars: 5 µm. (B) Quantification of SFPQ knockdown in *C9ORF72* mutation-positive fibroblasts transduced with lentiviral particles with scrambled or SFPQ shRNA. (C) Foci distribution in *C9ORF72* mutation-positive fibroblasts transduced with lentiviral particles with scrambled or SFPQ shRNA. A sum of three experiments is shown. (D) Quantification of $(G_4C_2)_n$ foci in *C9ORF72* mutation-positive fibroblasts transduced with lentiviral particles with scrambled or SFPQ shRNA. Three experiments were performed and a minimum of 60 cells per experiment were counted (Student's *t*-test, $^*P<0.05$). Data are presented as mean±s.e.m.

Plasmid DNA

Vector pcDNA3.2/GW/D-TOPO containing $(G_4C_2)_{72}$, vector pcDNA3 containing S1 aptamer, $(G_4C_2)_{48}$ with S1 aptamer at the 3′ end or DsRed [1–369 nucleotides (nt)] with S1 aptamer at the 3′ end have been described previously (Lee et al., 2013). pGEMT plasmid with *hLincRNA-p21* probe sequence was a gift from Dr Anna M. Pyle (Yale University, New Haven, CT) (Chillon and Pyle, 2016). pSpCas9(BB)-2A-GFP vector was Addgene #48138, deposited by Dr Feng Zhang (Ran et al., 2013). We thank Dr Don W. Cleveland (Ludwig Institute for Cancer Research, La Jolla, CA) for providing us with plasmids pMD2.G (Addgene #12259, deposited by Didier Trono) and psPAX2 (Addgene #12260, deposited by Didier Trono). Plasmid pLKO.1 scramble shRNA was Addgene #1864 (deposited by David Sabatini) (Sarbassov et al., 2005).

Cell culture and transfection

HEK293T cells obtained from ATTC were maintained in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% fetal bovine serum (Gibco), 2 mM GlutaMAX (Gibco) and 100 U/ml penicillin-streptomycin (Gibco). Cells were transfected using Lipofectamine 2000 (Life Technologies) according to the manufacturer's protocol. FISH with immunofluorescence were performed 1 day after transfection. *C9ORF72* mutation-positive fibroblasts were a kind gift from Dr Don W. Cleveland (Lagier-Tourenne et al., 2013). Control and *C9ORF72* mutation-positive fibroblasts were maintained in DMEM (Gibco) supplemented with 20% fetal bovine serum (Gibco), 2 mM GlutaMAX (Gibco) and 100 U/ml penicillin-streptomycin (Gibco). For colocalization experiments, they were plated on poly-L-lysine-coated glass coverslips. All cell lines were tested and confirmed free of mycoplasma.

For SFPQ knockdown experiments, HEK293T lentivirus production cells (a kind gift from Dr Don W. Cleveland) were plated on a 6 cm plate to reach 70% confluence 24 h later, when they were co-transfected with pMD2.G (Addgene #12259), psPAX2 (Addgene #12260) and pLKO.2 shSFPQ (Sigma-Aldrich, NM_005066.x-977s1c1; Table S2) or pLKO.1 shScramble (Addgene #1864; Table S2) in 1:2:3 ratios, using Polyjet transfection reagent (SignaGen Laboratories, SL100688), according to the manufacturer's instructions. After 6 h, the growth medium was replaced with 4 ml fibroblast growth medium. After 48 h, the supernatant was collected from HEK293T lentivirus-producing cells, filtered through a 0.45 µm cellulose acetate membrane and added to the same volume of fresh fibroblast medium. Then, 700 µl virus mixture per well was added to the fibroblasts, plated a day before on glass coverslips onto a 24-well plate. After 24 h, the medium was changed, and the fibroblasts were incubated for an additional 65 h. Then, they were either collected for WB analyses or fixed for FISH and immunofluorescence.

FISH with immunofluorescence

5TYE563-labelled G_4C_2 (CCCCGGCCCCGGCCCC) LNA probe was synthesized by Exiqon. Quasar670-labelled and Quasar570-labelled *NEAT1_2* Stellaris FISH probes and Quasar670-labelled Stellaris *NEAT1* FISH probe were purchased from Biosearch Technologies. Digoxigenin-labelled *hLincRNA-p21* probe was prepared as described previously (kind gift from Dr Anna M. Pyle) (Chillon and Pyle, 2016; West et al., 2016). Cells were fixed in 4% paraformaldehyde in PBS for 15 min and permeabilized with 0.1% Triton X-100 in PBS for 5 min. The coverslips were incubated in pre-hybridization solution [40% formamide, 2× saline sodium citrate (SSC)] for 15 min, followed by overnight hybridization with 2 µM G_4C_2 probe diluted in hybridization buffer [40% formamide, 1 mg/ml transfer RNA (tRNA), 10% dextran sulphate, 2× SSC] at 60°C. The next day, coverslips were washed with 0.1% Tween 20 in 2× SSC for 5 min at room temperature, followed by three washes with 0.1× SSC at 60°C, 10 min each. Coverslips were then used for detection of *NEAT1_2* and immunofluorescence, or detection of *hLincRNAp21*. Quasar670-labelled *NEAT1_2* probe was diluted in hybridization buffer at a concentration of

2 ng/µl and coverslips were incubated at 37°C for 5 h. Afterwards, coverslips were briefly washed with 2× SSC and blocked in 3% bovine serum albumin (Sigma-Aldrich) in PBS for 30 min. After incubation with primary antibodies, coverslips were incubated with secondary anti-rabbit Alexa Fluor 488 or anti-mouse Alexa Fluor 488 antibodies (Life Technologies). For detection of *hLincRNA-p21*, the probe was diluted in hybridization buffer at a concentration of 20 ng/µl and coverslips were incubated overnight at 60°C. The following day, coverslips were washed with 0.1% Tween 20 in 2× SSC for 5 min at room temperature, followed by three washes with 0.1× SSC at 60°C for 10 min each and then incubated with a 1:200 dilution of Alexa Fluor 488-conjugated anti-digoxigenin antibody (Jackson ImmunoResearch) in 1% bovine serum albumin in PBS for 1 h. Nuclei were stained with 4′,6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich) and coverslips mounted using FluorSave Reagent (Millipore). Zeiss LSM 710 or Leica confocal SP systems were used for imaging.

For simultaneous detection of *NEAT1* and *NEAT1_2*, cells were fixed in 4% paraformaldehyde in PBS for 15 min at room temperature and permeabilized in 70% ethanol at 4°C for at least 12 h. The next day, cells were washed twice with PBS for 5 min each and incubated for 5 min in 2× SSC buffer with 10% formamide. Then, hybridization was performed for 5 h at 37°C using 2 ng/µl Quasar670-labelled *NEAT1* and 0.5 ng/µl Quasar570-labelled *NEAT1_2* probe diluted in buffer containing 2× SSC, 10% formamide, 50 µg/µl tRNA, 10% dextrane sulphate, 2 mg/ml BSA and 10 mM vanadyl-ribonucleoside complex. Afterwards, cells were washed twice with pre-warmed 2× SCC with 10% formamide for 30 min each at 37°C and then twice with PBS for 30 min each at room temperature. Cells were mounted with ProLong Gold Antifade Reagent containing DAPI (Thermo Fisher Scientific). Imaging was performed on an AppliedPrecision DeltaVision RT wide-field microscope.

FISH with immunofluorescence for human tissue

Human cerebellar sections were provided as 10% formalin-fixed and paraffin-embedded blocks. Paraffin was removed with xylene, and sections were rehydrated in a series of ethanol dilutions (100%, 95%, 70%) for 3 min per step. Afterwards, sections were incubated in 0.3% Sudan Black in 70% ethanol for 5 min and washed with water for 5 min. Antigen retrieval was achieved with 20 mg/ml proteinase-K (Qiagen) diluted in TBS (pH 7.4) at 37°C for 20 min. Afterwards, slides were treated with ice-cold 20% acetic acid in TBS (pH 7.4) for 2 min and incubated in pre-hybridization buffer for 15 min. Hybridization was performed with 2 µM G_4C_2 probe in hybridization buffer overnight at 60°C. Sections were washed once with 0.1% Tween 20 in 2× SSC at room temperature for 5 min and three times with 0.1× SSC at 60°C for 15 min. Afterwards, sections were washed with PBS for 15 min and blocked in 20% donkey serum for 1 h. Incubation with primary antibodies was carried out overnight at 4°C, followed by incubation with secondary anti-rabbit Alexa Fluor 488 antibodies (Life Technologies). DAPI (Sigma-Aldrich) was used for counterstaining. Leica confocal SP or Zeiss LSM 710 systems were used for high-resolution imaging.

Immunoblotting

Proteins were separated by reducing SDS-PAGE and transferred onto nitrocellulose membrane using wet transfer at 200 mA for 90 min. Membranes were blocked in 5% skim milk in TBS with 0.05% Tween 20. Blocking was carried out at room temperature for 1 h. Primary antibodies diluted in blocking solution were incubated for 1–4 h at room temperature. Following three washes with TBS-Tween 20, the membranes were incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies [anti-rabbit-423 HRP (1:10,000, Jackson ImmunoResearch) or anti-mouse HRP (1:5000, Millipore)], washed and incubated with chemiluminescent reagent (Roche).

RNA pulldown

Rat and mouse brain tissue nuclear extracts were prepared as described previously (Lee et al., 2013). pcDNA3$(G_4C_2)_{48}$-S1 and controls pcDNA3 S1 and pcDNA3 DsRed (1–369 nt)-S1 plasmids were linearized after S1 aptamer at the *Xba*I site and purified with phenol/chloroform extraction. *In vitro* transcription was performed using the T7 promoter on pcDNA3 vector with TranscriptAid T7 High Yield Transcription Kit (Fermentas). Single-strand binding protein (Sigma-Aldrich) at a final concentration of 7.5 µg per 1 µg DNA was added to facilitate transcription of hexanucleotide repeats. RNA pulldown was performed as previously described (Butter et al., 2009) with some modifications. S1-tagged RNAs were incubated with streptavidin magnetic beads (Promega) in RNA-binding buffer [50 mM HEPES (pH 7.4), 100 mM KCl, 10 mM $MgCl_2$, 0.5% IGEPAL CA-630] for 40 min at 4°C. Beads with bound RNA were washed three times with RNA-binding buffer and incubated with 3 mg rat or mouse brain extracts, 20 U RiboLock RNase inhibitor (Fermentas) and 50 µg yeast tRNA (Sigma-Aldrich) for 4 h at 4°C. Afterwards, they were washed five times with RNA-binding buffer and eluted with 15 U RNAseI in RNA-binding buffer, before 2× SDS loading buffer with 200 mM dithiothreitol (DTT) was added to the eluates.

Sample preparation for mass spectrometry analyses

Samples were separated on a 12.5% pre-cast SDS-PAGE gel (Lonza) and visualized by silver staining (Gharahdaghi et al., 1999). Protein bands were excised from the gel, destained and subjected to reduction with 10 mM DTT in 25 mM ammonium bicarbonate, followed by alkylation with 55 mM iodoacetamide in the same buffer. Then, they were washed twice with 25 mM ammonium bicarbonate, dried in a SpeedVac and rehydrated in 25 mM ammonium bicarbonate containing 1 µg porcine sequence-grade modified trypsin (Promega), prior to overnight digestion at 37°C. Digested peptides were extracted from the gel with 50% acetonitrile solution containing 5% formic acid and concentrated to 15 µl and analysed with a LC/MSD Trap XCT Ultra mass spectrometer coupled to a Series 1200 liquid chromatography unit (Agilent Technologies). Peptides were loaded on an HPLC Chip with integrated 40 nl trap column and C_{18} separation column (150 mm×75 µm) (ProtID Chip-150). Elution was performed with a 41-min acetonitrile gradient from 3% to 50% in a 0.1% solution of formic acid, with a flow rate of 350 nl/min. The five most intense precursor ions in each full scan were selected for collision-induced dissociation (CID) fragmentation. Dynamic exclusion was set at a repeat count of 2, with an exclusion duration of 30 s. Database searches were performed against the NCBInr database using the Spectrum Mill database search software. Carbamidomethylation of cysteines was set as fixed and oxidation of methionines as dynamic modification.

Generation of *NEAT1* knockdown HEK293T cells

NEAT1 knockdown cells were generated by cutting out 1.1 kb around the transcription start site of *NEAT1* using the CRISPR/Cas9 protocol as described previously (Ran et al., 2013). Briefly, forward and reverse guide RNAs (gRNAs) (Table S1) with *Bbs*I restriction site overhangs were designed, phosphorylated, annealed and cloned into *Bbs*I (NEB)-digested pSpCas9(BB)-2A-GFP vector [Addgene #48138 (Ran et al., 2013)] using T4 DNA Ligase (NEB). HEK293T cells were transfected with 500 ng of both gRNA-Cas9-2A-GFP plasmids using Lipofectamine 2000 (Life Technologies), according to the manufacturer's protocol. Single clones were picked, expanded and screened for successful knockout using the primers listed in Table S1.

RNA extraction and quantitative RT-PCR

RNA extraction was performed with the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol and including treatment with DNaseI. Reverse transcription was performed with SuperScriptIII Reverse Transcriptase (Life Technologies) using 500 ng RNA. Quantitative PCR was performed with the Taqman Gene Expression Master Mix (Life Technologies), using Taqman primers for *NEAT1* (Hs01008264_s1; Life Technologies) and *NEAT1_2* (Hs03924655_s1; Life Technologies).

Statistical analyses

All experiments were performed in duplicate and independently repeated at least three times unless otherwise stated. Statistical analyses of the data were performed by Student's *t*-test. $P<0.05$ was considered significant. Data were expressed as means±s.e.m.

Acknowledgements

pGEMT plasmid with hLincRNA-p21 probe sequence was a gift from Dr Anna M. Pyle (Yale University, New Haven, CT). We thank Dr Don W. Cleveland (Ludwig Institute for Cancer Research, La Jolla, CA) for providing us with plasmids pMD2.G and psPAX2, HEK293T lentivirus production cells and primary fibroblasts.

Competing interests
The authors declare no competing or financial interests.

Author contributions
Conceptualization: S.D., Y.-B.L., M. Modic, C.E.S., B.R.; Methodology: A.B.C., S.D., M.S., B.T., B.R.; Validation: A.B.C., S.D., S.P.M., M.S., M. Malnar, H.M., B.R.; Formal analysis: A.B.C., S.D., S.P.M., M.S., M. Malnar, Y.-B.L., J.M., M.G., M. Modic, M.F., B.R.; Investigation: A.B.C., S.D., S.P.M., M.S., M. Malnar, Y.-B.L., J.M., J.P., M.G., M. Modic, M.D., B.R.; Resources: B.T., M.D., C.E.S., B.R.; Data curation: J.P., M.F., B.T.; Writing - original draft: A.B.C., S.D., S.P.M., M.S., M. Malnar, H.M., Y.-B.L., J.P., M. Modic, M.F., M.D., C.E.S., B.R.; Writing - review & editing: A.B.C., S.D., S.P.M., H.M., B.R.; Visualization: A.B.C., S.D., M.S., M. Malnar, H.M., M.G., M. Modic, B.R.; Supervision: Y.-B.L., B.T., M.D., C.E.S., B.R.; Project administration: B.R.; Funding acquisition: B.T., M.D., C.E.S., B.R.

Funding
This work was supported by Javna Agencija za Raziskovalno Dejavnost RS (the Slovenian Research Agency) [P4-0127, J3-9263, J3-8201, J3-6789, J3-5502, J7-5460, P1-0140]; Alzheimer's Research UK; the National Institute of Health Research Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust; and King's College London, in partnership with King's College Hospital.

Supplementary information
Supplementary information available online at
http://jcs.biologists.org/lookup/doi/10.1242/jcs.224303.supplemental

References
Al-Sarraj, S., King, A., Troakes, C., Smith, B., Maekawa, S., Bodi, I., Rogelj, B., Al-Chalabi, A., Hortobágyi, T. and Shaw, C. E. (2011). p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9ORF72-linked FTLD and MND/ALS. *Acta Neuropathol.* **122**, 691-702.

Almeida, S., Gascon, E., Tran, H., Chou, H. J., Gendron, T. F., Degroot, S., Tapper, A. R., Sellier, C., Charlet-Berguerand, N., Karydas, A. et al. (2013). Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathol.* **126**, 385-399.

Anantharaman, A., Jadaliha, M., Tripathi, V., Nakagawa, S., Hirose, T., Jantsch, M. F., Prasanth, S. G. and Prasanth, K. V. (2016). Paraspeckles modulate the intranuclear distribution of paraspeckle-associated Ctn RNA. *Sci. Rep.* **6**, 34043.

Ash, P. E. A., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W.-L., Dejesus-Hernandez, M., van Blitterswijk, M. M., Jansen-West, K., Paul, J. W., III, Rademakers, R. et al. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77**, 639-646.

Butter, F., Scheibe, M., Morl, M. and Mann, M. (2009). Unbiased RNA-protein interaction screen by quantitative proteomics. *Proc. Natl. Acad. Sci. USA* **106**, 10626-10631.

Chen, L.-L. and Carmichael, G. G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol. Cell* **35**, 467-478.

Chen, L.-L., DeCerbo, J. N. and Carmichael, G. G. (2008). Alu element-mediated gene silencing. *EMBO J.* **27**, 1694-1705.

Chillon, I. and Pyle, A. M. (2016). Inverted repeat Alu elements in the human lincRNA-p21 adopt a conserved secondary structure that regulates RNA function. *Nucleic Acids Res.* **44**, 9462-9471.

Clemson, C. M., Hutchinson, J. N., Sara, S. A., Ensminger, A. W., Fox, A. H., Chess, A. and Lawrence, J. B. (2009). An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **33**, 717-726.

DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. C. A., Flynn, H., Adamson, J. et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245-256.

Donnelly, C. J., Zhang, P.-W., Pham, J. T., Haeusler, A. R., Mistry, N. A., Vidensky, S., Daley, E. L., Poth, E. M., Hoover, B., Fines, D. M. et al. (2013). RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* **80**, 415-428.

Fox, A. H., Nakagawa, S., Hirose, T. and Bond, C. S. (2018). Paraspeckles: where long noncoding RNA meets phase separation. *Trends Biochem. Sci.* **43**, 124-135.

Freibaum, B. D., Lu, Y., Lopez-Gonzalez, R., Kim, N. C., Almeida, S., Lee, K.-H., Badders, N., Valentine, M., Miller, B. L., Wong, P. C. et al. (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129-133.

Galloway, J. N. and Nelson, D. L. (2009). Evidence for RNA-mediated toxicity in the fragile X-associated tremor/ataxia syndrome. *Future Neurol.* **4**, 785.

Gendron, T. F., Bieniek, K. F., Zhang, Y.-J., Jansen-West, K., Ash, P. E. A., Caulfield, T., Daughrity, L., Dunmore, J. H., Castanedes-Casey, M., Chew, J. et al. (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol.* **126**, 829-844.

Gharahdaghi, F., Weinberg, C. R., Meagher, D. A., Imai, B. S. and Mische, S. M. (1999). Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis* **20**, 601-605.

Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. A. J., Shaw, P. G., Kim, M.-S., Maragakis, N. J., Troncoso, J. C., Pandey, A., Sattler, R. et al. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **507**, 195-200.

Hirose, T., Virnicchi, G., Tanigawa, A., Naganuma, T., Li, R., Kimura, H., Yokoi, T., Nakagawa, S., Bénard, M., Fox, A. H. et al. (2014). NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol. Biol. Cell* **25**, 169-183.

Imamura, K., Imamachi, N., Akizuki, G., Kumakura, M., Kawaguchi, A., Nagata, K., Kato, A., Kawaguchi, Y., Sato, H., Yoneda, M. et al. (2014). Long noncoding RNA NEAT1-dependent SFPQ relocation from promoter region to paraspeckle mediates IL8 expression upon immune stimuli. *Mol. Cell* **53**, 393-406.

Ishigaki, S., Fujioka, Y., Okada, Y., Riku, Y., Udagawa, T., Honda, D., Yokoi, S., Endo, K., Ikenaka, K., Takagi, S. et al. (2017). Altered Tau isoform ratio caused by loss of FUS and SFPQ function leads to FTLD-like phenotypes. *Cell Rep.* **18**, 1118-1131.

Jiang, L., Shao, C., Wu, Q.-J., Chen, G., Zhou, J., Yang, B., Li, H., Gou, L.-T., Zhang, Y., Wang, Y. et al. (2017). NEAT1 scaffolds RNA-binding proteins and the Microprocessor to globally enhance pri-miRNA processing. *Nat. Struct. Mol. Biol.* **24**, 816-824.

Kovanda, A., Zalar, M., Šket, P., Plavec, J. and Rogelj, B. (2015). Anti-sense DNA d(GGCCCC)n expansions in C9ORF72 form i-motifs and protonated hairpins. *Sci. Rep.* **5**, 17944.

Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H.-R., Jiang, J., Watt, A. T., Chun, S., Katz, M. et al. (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. USA* **110**, E4530-E4539.

Lee, J. E. and Cooper, T. A. (2009). Pathogenic mechanisms of myotonic dystrophy. *Biochem. Soc. Trans.* **37**, 1281-1286.

Lee, Y.-B., Chen, H.-J., Peres, J. N., Gomez-Deza, J., Attig, J., Štalekar, M., Troakes, C., Nishimura, A. L., Scotter, E. L., Vance, C. et al. (2013). Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **5**, 1178-1186.

Levanon, E. Y., Eisenberg, E., Yelin, R., Nemzer, S., Hallegger, M., Shemesh, R., Fligelman, Z. Y., Shoshan, A., Pollock, S. R., Sztybel, D. et al. (2004). Systematic identification of abundant A-to-I editing sites in the human transcriptome. *Nat. Biotechnol.* **22**, 1001-1005.

Luisier, R., Tyzack, G. E., Hall, C. E., Mitchell, J. S., Devine, H., Taha, D. M., Malik, B., Meyer, I., Greensmith, L., Newcombe, J. et al. (2018). Intron retention and nuclear loss of SFPQ are molecular hallmarks of ALS. *Nat. Commun.* **9**, 2010.

Mao, Y. S., Sunwoo, H., Zhang, B. and Spector, D. L. (2011). Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. *Nat. Cell Biol.* **13**, 95-101.

Mizielinska, S., Lashley, T., Norona, F. E., Clayton, E. L., Ridler, C. E., Fratta, P. and Isaacs, A. M. (2013). C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* **126**, 845-857.

Moens, T. G., Mizielinska, S., Niccoli, T., Mitchell, J. S., Thoeng, A., Ridler, C. E., Grönke, S., Esser, J., Heslegrave, A., Zetterberg, H. et al. (2018). Sense and antisense RNA are not toxic in Drosophila models of C9orf72-associated ALS/FTD. *Acta Neuropathol.* **135**, 445-457.

Mori, K., Arzberger, T., Grässer, F. A., Gijselinck, I., May, S., Rentzsch, K., Weng, S.-M., Schludi, M. H., van der Zee, J., Cruts, M. et al. (2013a). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.* **126**, 881-893.

Mori, K., Lammich, S., Mackenzie, I. R. A., Forné, I., Zilow, S., Kretzschmar, H., Edbauer, D., Janssens, J., Kleinberger, G., Cruts, M. et al. (2013b). 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**, 413-423.

Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H. A., Cruts, M., Van Broeckhoven, C. et al. (2013c). The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335-1338.

Naganuma, T., Nakagawa, S., Tanigawa, A., Sasaki, Y. F., Goshima, N. and Hirose, T. (2012). Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* **31**, 4020-4034.

Nishimoto, Y., Nakagawa, S., Hirose, T., Okano, H. J., Takao, M., Shibata, S., Suyama, S., Kuwako, K., Imai, T., Murayama, S. et al. (2013). The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle

formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol. Brain* **6**, 31.
Orr, H. T. (2012). Cell biology of spinocerebellar ataxia. *J. Cell Biol.* **197**, 167-177.
Prasanth, K. V., Prasanth, S. G., Xuan, Z., Hearn, S., Freier, S. M., Bennett, C. F., Zhang, M. Q. and Spector, D. L. (2005). Regulating gene expression through RNA nuclear retention. *Cell* **123**, 249-263.
Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281-2308.
Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., Schymick, J. C., Laaksovirta, H., van Swieten, J. C., Myllykangas, L. et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-268.
Rossi, S., Serrano, A., Gerbino, V., Giorgi, A., Di Francesco, L., Nencini, M., Bozzo, F., Schinina, M. E., Bagni, C., Cestra, G. et al. (2015). Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *J. Cell Sci.* **128**, 1787-1799.
Saberi, S., Stauffer, J. E., Jiang, J., Garcia, S. D., Taylor, A. E., Schulte, D., Ohkubo, T., Schloffman, C. L., Maldonado, M., Baughn, M. et al. (2018). Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. *Acta Neuropathol.* **135**, 459-474.
Sarbassov, D. D., Guertin, D. A., Ali, S. M. and Sabatini, D. M. (2005). Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* **307**, 1098-1101.
Sareen, D., O'Rourke, J. G., Meera, P., Muhammad, A. K., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A. et al. (2013). Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci. Transl. Med.* **5**, 208ra149.
Sasaki, Y. T. F., Ideue, T., Sano, M., Mituyama, T. and Hirose, T. (2009). MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. USA* **106**, 2525-2530.
Shelkovnikova, T. A., Robinson, H. K., Troakes, C., Ninkina, N. and Buchman, V. L. (2014). Compromised paraspeckle formation as a pathogenic factor in FUSopathies. *Hum. Mol. Genet.* **23**, 2298-2312.
Shelkovnikova, T. A., Kukharsky, M. S., An, H., Dimasi, P., Alexeeva, S., Shabir, O., Heath, P. R. and Buchman, V. L. (2018). Protective paraspeckle hyper-assembly downstream of TDP-43 loss of function in amyotrophic lateral sclerosis. *Mol. Neurodegener.* **13**, 30.
Shen, W., Liang, X.-H. and Crooke, S. T. (2014). Phosphorothioate oligonucleotides can displace NEAT1 RNA and form nuclear paraspeckle-like structures. *Nucleic Acids Res.* **42**, 8648-8662.
Shi, Y., Lin, S., Staats, K. A., Li, Y., Chang, W.-H., Hung, S.-T., Hendricks, E., Linares, G. R., Wang, Y., Son, E. Y. et al. (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* **24**, 313-325.
Šket, P., Pohleven, J., Kovanda, A., Štalekar, M., Župunski, V., Zalar, M., Plavec, J. and Rogelj, B. (2015). Characterization of DNA G-quadruplex species forming from C9ORF72 G4C2-expanded repeats associated with amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **36**, 1091-1096.
Sunwoo, H., Dinger, M. E., Wilusz, J. E., Amaral, P. P., Mattick, J. S. and Spector, D. L. (2009). MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res.* **19**, 347-359.
Swinnen, B., Bento-Abreu, A., Gendron, T. F., Boeynaems, S., Bogaert, E., Nuyts, R., Timmers, M., Scheveneels, W., Hersmus, N., Wang, J. et al. (2018). A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol.* **135**, 427-443.
Thomas-Jinu, S., Gordon, P. M., Fielding, T., Taylor, R., Smith, B. N., Snowden, V., Blanc, E., Vance, C., Topp, S., Wong, C.-H. et al. (2017). Non-nuclear pool of splicing factor SFPQ regulates axonal transcripts required for normal motor development. *Neuron* **94**, 322-336.e5.
Todd, P. K. and Paulson, H. L. (2010). RNA-mediated neurodegeneration in repeat expansion disorders. *Ann. Neurol.* **67**, 291-300.
Tollervey, J. R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A. L., Župunski, V. et al. (2011). Characterizing the RNA targets and position-dependent splicing regulation by TDP-43. *Nat. Neurosci.* **14**, 452-458.
Troakes, C., Maekawa, S., Wijesekera, L., Rogelj, B., Siklós, L., Bell, C., Smith, B., Newhouse, S., Vance, C., Johnson, L. et al. (2012). An MND/ALS phenotype associated with C9orf72 repeat expansion: abundant p62-positive, TDP-43-negative inclusions in cerebral cortex, hippocampus and cerebellum but without associated cognitive decline. *Neuropathology* **32**, 505-514.
Vatovec, S., Kovanda, A. and Rogelj, B. (2014). Unconventional features of C9ORF72 expanded repeat in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **35**, 2421.e1-2421.e12.
Waite, A. J., Bäumer, D., East, S., Neal, J., Morris, H. R., Ansorge, O. and Blake, D. J. (2014). Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol. Aging* **35**, 1779.e5-1779.e13.
West, J. A., Mito, M., Kurosaka, S., Takumi, T., Tanegashima, C., Chujo, T., Yanaka, K., Kingston, R. E., Hirose, T., Bond, C. et al. (2016). Structural, super-resolution microscopy analysis of paraspeckle nuclear body organization. *J. Cell Biol.* **214**, 817-830.
Xiao, S., MacNair, L., McGoldrick, P., McKeever, P. M., McLean, J. R., Zhang, M., Keith, J., Zinman, L., Rogaeva, E. and Robertson, J. (2015). Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Ann. Neurol.* **78**, 568-583.
Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D. L., Li, H., Hales, C. M., Gearing, M., Wingo, T. S. et al. (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl. Acad. Sci. USA* **110**, 7778-7783.
Zhang, Z. and Carmichael, G. G. (2001). The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106**, 465-475.
Zhang, K., Donnelly, C. J., Haeusler, A. R., Grima, J. C., Machamer, J. B., Steinwald, P., Daley, E. L., Miller, S. J., Cunningham, K. M., Vidensky, S. et al. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61.
Zu, T., Liu, Y., Banez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B., Falchook, A. E., Subramony, S. H. et al. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc. Natl. Acad. Sci. USA* **110**, E4968-E4977.

Table S1. List of gRNAs and primers for *NEAT1* knockdown HEK293 cells.

NEAT1 gRNA up fwd	CACCGCGAAAGTCACGCGCGCCTCC
NEAT1 gRNA up rev	AAACGGAGGCGCGCGTGACTTTCGC
NEAT1 gRNA down fwd	CACCGCCAGACCTGGACGCTCCACC
NEAT1 gRNA down rev	AAACGGTGGAGCGTCCAGGTCTGGC
NEAT1 det primer fwd	GTTGTCACCCACTAGCTCCT
NEAT1 det primer rev	AAGTCCAAAAGGAGCACTGC

Table S2. shSFPQ and shScramble hairpin sequences.

shScramble	CCTAAGGTTAAGTCGCCCTCGCTCGAGCGAGGGCGACTTAACCTTAGG	Addgene	#1864
shSFPQ	CCGGCGGTTGTTTGTTGGGAATCTACTCGAGTAGATTCCCAACAAACAACCGTTTTT	Sigma Aldrich	NM_005066.x-977s1c1

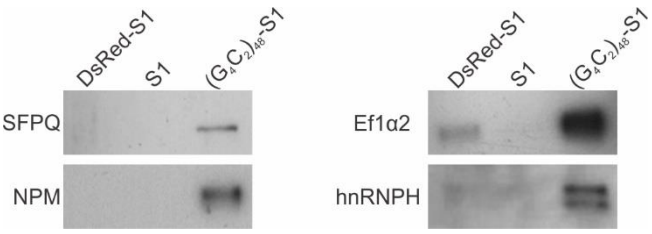

Fig. S1. Immunoblots confirming SFPQ, NPM1, EF1α2 and hnRNPH specifically co-precipitate with $(G4C2)_{48}$-S1 RNA in rat cerebellar nuclear extracts.

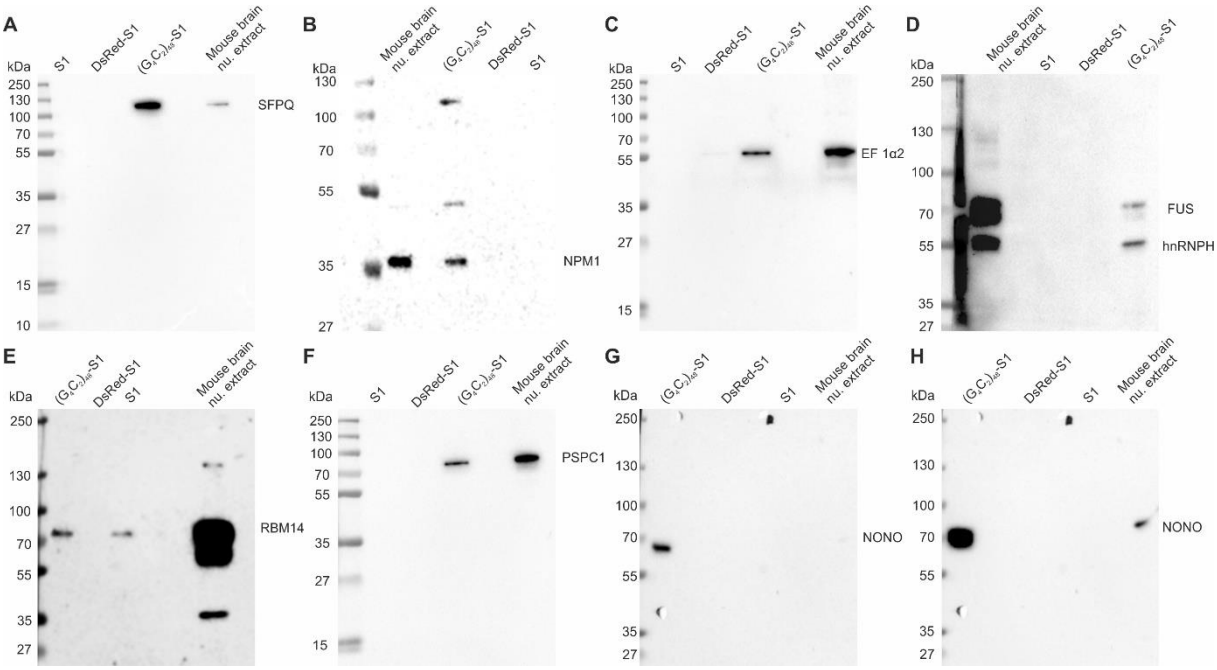

Fig. S2. Full length blots from Fig. 1.

(A, B, C, D, E, F, G, H) Full length immunoblots with positive control confirming that proteins SFPQ, NPM1, EF1α2, hnRNPH, FUS, RBM14, PSPC1 and NONO from mouse brain nuclear extracts specifically co-precipitate with $(G_4C_2)_{48}$-S1 RNA.

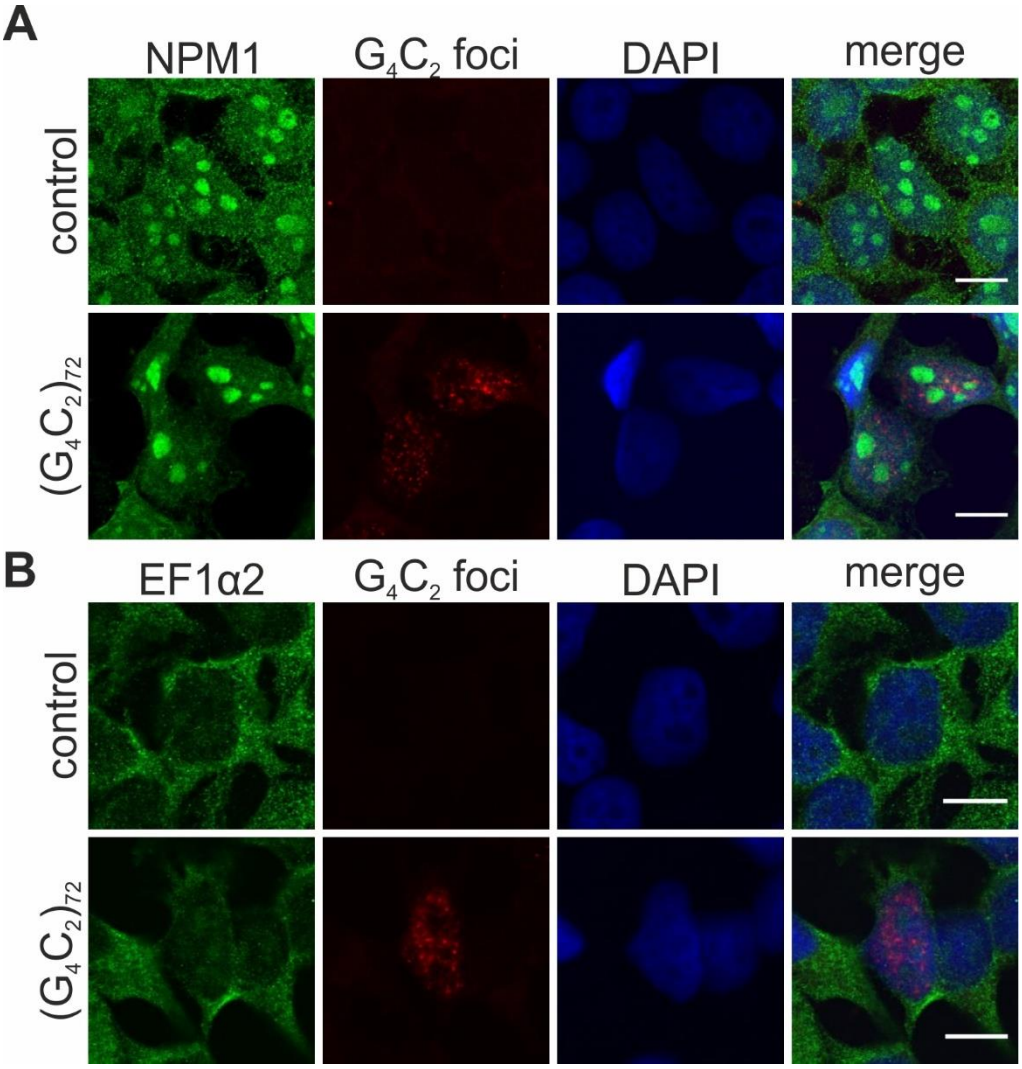

Fig. S3. NPM1 and EF1α2 do not colocalize with $(G_4C_2)_{72}$ nuclear foci.

HEK293T cells transfected with a plasmid expressing $(G_4C_2)_{72}$ repeats and probed for $(G_4C_2)_{72}$ and NPM1 (A) or EF1α2 (B). Scale bars: 10 µm.

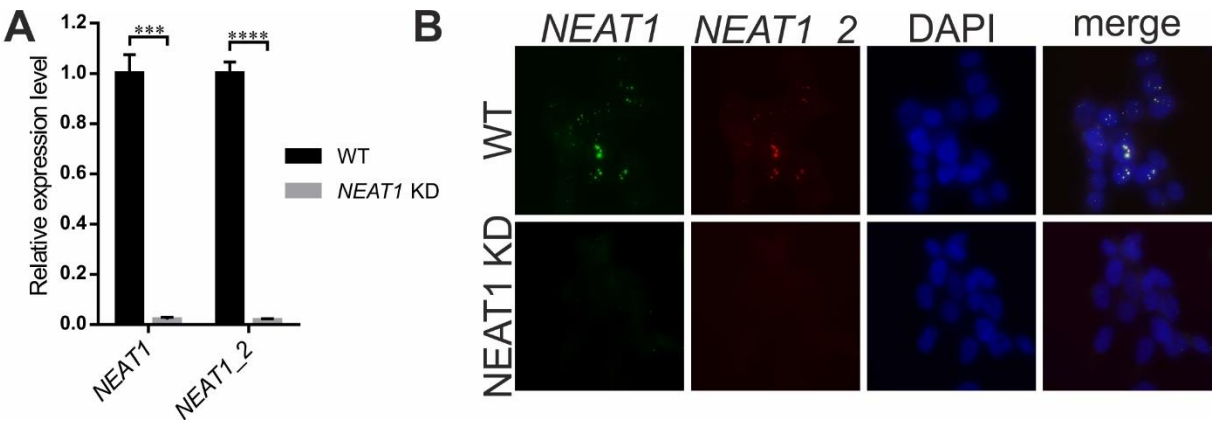

Fig. S4. Characterization of *NEAT1* knockdown HEK293T cells.

(A) Relative expression levels of *NEAT1* and *NEAT1_2* in *NEAT1* knockdown HEK293T cells (*NEAT1* KD) compared to wild-type HEK293T cells (WT) determined by quantitative RT-PCR (t-test, ***p<0.001 for *NEAT1* and ****p<0.0001 for *NEAT1_2*; three experiments were performed; data are presented as mean ± s.e.m). (B) *NEAT1* knockdown HEK293T and wild-type HEK293T cells probed for *NEAT1* and *NEAT1_2*.

(B)

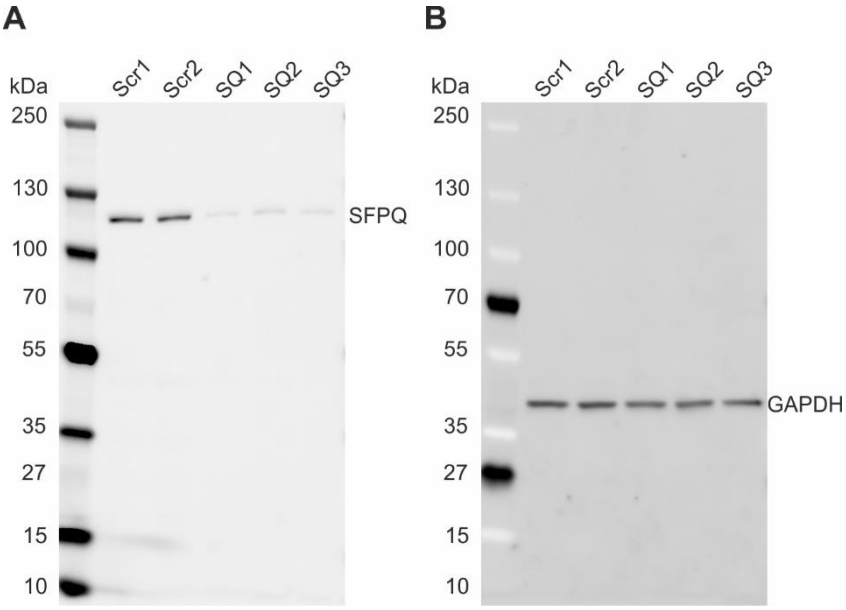

Fig. S5. Full length blot from Fig 5.

(A) Full length blot from Fig. 5 for quantification of SFPQ knockdown in *C9ORF72* mutant fibroblasts transduced with lentiviral particles containing sequences for scrambled shRNA or SFPQ shRNA. (B) GAPDH was used as loading control. Signal was detected using dual-colour imaging system.