OXFORD

Human Molecular Genetics, 2017, Vol. 26, No. 6 1133–1145

doi: 10.1093/hmg/ddx022
Advance Access Publication Date: 1 February 2017
Original Article

ORIGINAL ARTICLE

C9ORF72 hexanucleotide repeat exerts toxicity in a stable, inducible motor neuronal cell model, which is rescued by partial depletion of Pten

Matthew J. Stopford[1], Adrian Higginbottom[1], Guillaume M. Hautbergue[1], Johnathan Cooper-Knock[1], Padraig J. Mulcahy[1], Kurt J. De Vos[1], Alan E. Renton[2], Hannah Pliner[2], Andrea Calvo[3], Adriano Chio[3], Bryan J. Traynor[2], Mimoun Azzouz[1], Paul R. Heath[1], ITALSGEN Consortium, NeuroX Consortium, Janine Kirby[1] and Pamela J. Shaw[1,*]

[1]Department of Neuroscience, Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK, [2]Neuromuscular Diseases Research Section, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA and [3]Department of Neuroscience, University of Turin, Turin, Italy

*To whom correspondence should be addressed at: Tel: +44 1142222295; Fax: +44 1142222290; Email: pamela.shaw@sheffield.ac.uk

Abstract

Amyotrophic lateral sclerosis (ALS) is a devastating and incurable neurodegenerative disease, characterised by progressive failure of the neuromuscular system. A (G4C2)n repeat expansion in *C9ORF72* is the most common genetic cause of ALS and frontotemporal dementia (FTD). To date, the balance of evidence indicates that the (G4C2)n repeat causes toxicity and neurodegeneration via a gain-of-toxic function mechanism; either through direct RNA toxicity or through the production of toxic aggregating dipeptide repeat proteins. Here, we have generated a stable and isogenic motor neuronal NSC34 cell model with inducible expression of a (G4C2)102 repeat, to investigate the gain-of-toxic function mechanisms. The expression of the (G4C2)102 repeat produces RNA foci and also undergoes RAN translation. In addition, the expression of the (G4C2)102 repeat shows cellular toxicity. Through comparison of transcriptomic data from the cellular model with laser-captured spinal motor neurons from C9ORF72-ALS cases, we also demonstrate that the PI3K/Akt cell survival signalling pathway is dysregulated in both systems. Furthermore, partial knockdown of Pten rescues the toxicity observed in the NSC34 (G4C2)102 cellular gain-of-toxic function model of C9ORF72-ALS. Our data indicate that PTEN may provide a potential therapeutic target to ameliorate toxic effects of the (G4C2)n repeat.

Introduction

A (G4C2)n repeat expansion in a non-coding region of the C9ORF72 gene has been established as the most common identified genetic cause of amyotrophic lateral sclerosis (ALS) as well as frontotemporal dementia (FTD) (1,2). Expansions of >30 repeats are considered pathogenic (3,4), but expansions of 200–

Received: November 13, 2016. **Revised:** January 2, 2017. **Accepted:** January 10, 2017

5000 repeats are more commonly detected in ALS patients (5). The mechanism(s) by which the repeat causes neuronal death in the motor cortex, brainstem and spinal cord in ALS and/or neuronal death in the frontal and temporal lobes of the brain in FTD are currently being elucidated, with three hypotheses proposed which are not mutually exclusive: (1) Haploinsufficiency of C9ORF72; (2) RNA toxicity; and (3) Dipeptide repeat protein (DPR) toxicity.

Various reports demonstrate *C9ORF72* mRNA is reduced in post-mortem CNS tissue, lymphoblast cells and iPSC-derived neurons of patients containing the (G4C2)n repeat expansion (1,6–10) and that C9ORF72 protein is also reduced in the frontal cortex of patients with the repeat expansion (10,11), suggesting C9ORF72 haploinsufficiency as a potential pathogenic mechanism. In addition, knockdown or deletion of C9ORF72 orthologues in zebrafish and *Caenorhabditis Elegans*, respectively, caused motor dysfunction (7,12). C9ORF72 protein regulates autophagy induction and autophagic flux, and both C9ORF72-ALS/FTD derived neurons and mice lacking functional C9ORF72 protein have impaired autophagy (13–15). However, mice lacking functional C9ORF72 do not show signs of neurodegeneration or reduced motor function, although they do display autoimmunity and other immune system dysregulation (13,16–19). Equally, when C9ORF72 knockdown (either partial or complete) was confined to the CNS in adult mice, there were no behavioural or motor defects, nor any signs of neuropathology associated with ALS and FTD (20,21).

The (G4C2)n repeat expansion is also suggested to exert a toxic gain-of-function via either direct RNA toxicity and/or DPR proteins. The (G4C2)n repeat is transcribed in both sense and antisense directions, and forms sense (G4C2)n and antisense (C4G2)n RNA foci, respectively. These RNA foci are present in C9ORF72-ALS/FTD CNS tissue and neuronal cells (1,22), and may potentially sequester essential RNA binding proteins leading to dysregulated RNA processing, as seen in other repeat expansion disorders (23). Several RNA binding proteins bind (G4C2)n and/or (C4G2)n RNA *in vitro*, and also colocalise with RNA foci in patient CNS tissue (9,24–29). In addition, transcriptomic analysis of C9ORF72-ALS/FTD patient CNS tissue showed dysregulation in RNA splicing and processing (30,31).

The (G4C2)n repeat also undergoes unconventional repeat associated non-ATG (RAN) translation in all reading frames in both the sense and antisense directions, forming aggregation-prone DPR proteins (32,33). The five species of DPR proteins (poly-GA, poly-GR, poly-GP, poly-AP and poly-PR) produced form insoluble inclusions in C9ORF72-ALS/FTD patient CNS tissue (32–35). DPR proteins are toxic in cultured cells and cause neurodegeneration in *Drosophila* models (35–42). The arginine-rich DPR proteins (poly-GR and poly-PR) appear particularly toxic, localise to the nucleolus, disrupt ribosomal RNA biogenesis and cause cell death (36–39,41). Also, in two elegant studies using *Drosophila* models, the toxicity of (G4C2)n repeats was dependent on the production of DPR, and not (G4C2)n RNA foci (41,43). However, while these studies suggest that the DPR are the likely major toxic insult derived from the sense (G4C2)n RNA, antisense (C4G2)n RNA foci, but not sense (G4C2)n RNA foci, correlate with TDP-43 proteinopathy in motor neurons from C9ORF72-ALS patients (44). Furthermore in C9ORF72-ALS patients, DPR load is much lower in spinal motor neurons compared with other unaffected regions of CNS, and TDP-43 inclusions rarely co-localise with DPR suggesting that they may not be the primary toxic insult in motor neuron degeneration (45–48).

Several recently generated mouse models of C9ORF72-ALS have produced fairly variable results. In one study, a (G4C2)66 construct was delivered to the CNS of the mice and resulted in RNA foci, DPR, and TDP-43 pathology, as well as behavioural and motor defects (49). In two other studies, C9ORF72 BAC transgenic mouse models were generated that contain the (G4C2)n repeat expansion within either part of or all of the C9ORF72 gene and display both the RNA foci and DPRs, yet surprisingly did not develop signs of neurodegeneration or ALS/FTD phenotypes (50,51). Another group described similar findings in different C9ORF72 BAC lines, with the addition of a cognitive phenotype (52). However, in a fourth C9ORF72 BAC mouse model, there was TDP-43 pathology, motor neuron degeneration and a neurodegenerative phenotype, including weakness, weight loss, breathing problems and decreased survival, as well as anxiety-like behaviour (53). While these studies are inconsistent in their findings, they demonstrate that (G4C2)n repeat length and expression level, as well as other contributing genetic factors may contribute to ALS pathogenesis.

In further support of a gain-of-toxic function, ASOs that target sense *C9ORF72* transcripts not only reduce (G4C2)n RNA foci number, they also ameliorate transcriptomic changes and reduce toxicity in iPSC-derived neuronal cells from C9ORF72-ALS/FTD patients (9,20,27). The balance of evidence is emerging that a gain-of-toxic function is more likely than C9ORF72 haploinsufficiency to provide the major toxic insult that drives C9ORF72-ALS and/or FTD. However, the relative contributions of the sense and antisense RNA, and each of the DPR species in the neuronal injury of C9ORF72-ALS and FTD have not been fully established. In addition, a loss of C9ORF72 function may still exacerbate the primary toxic insult, contributing to ALS/FTD pathogenesis. Therefore, we aimed to generate a gain-of-toxic function model to identify potential therapeutic targets for C9ORF72-ALS. We generated a stable motor neuron-like cell model with inducible (G4C2)n expression, which allowed identification of the early biochemical changes associated with the expression of the (G4C2)n repeat expansion. Our results showed that (G4C2)102 constructs produced RNA foci, underwent RAN translation and caused toxicity in NSC34 cells. Transcriptomic analysis of the NSC34 cell model identified dysregulation in the PI3K/Akt signalling pathway, which was validated in motor neurons from C9ORF72-ALS patients. Further, we showed that partial knockdown of *Pten*, the normal function of which is to negatively regulate the PI3kinase/Akt cell survival pathway, provided a rescue effect in the NSC34 cell model, suggesting that PTEN could represent a potential therapeutic target for C9ORF72-ALS.

Results

Generation of stable NSC34 cell models with tetracycline-inducible (G4C2)n expression

The primary aim of this work was to establish the gain-of-function effects from the (G4C2)n repeat expansion independent of the C9ORF72 gene context, and thereby identify potential therapeutic targets for C9ORF72-ALS. To address this, we generated stable, isogenic motor neuronal NSC34 cell models with tetracycline-inducible expression of 0 (sham) or 102 (G4C2)n repeats. First, (G4C2)10 repeat constructs were synthesised, and subsequently used to generate interrupted (G4C2)n repeat constructs containing 102 repeats (Fig. 1A). These (G4C2)102 constructs were then sub-cloned into pcDNA5/FRT/TO plasmids (Invitrogen) that contain an FRT site and a CMV/tetracycline operator hybrid promoter (Fig. 1B). The pcDNA5/FRT/TO-(G4C2)102 vectors were then stably integrated into Flp-In™ T-REx™ NSC34 cells (that were generated in house). In addition, empty

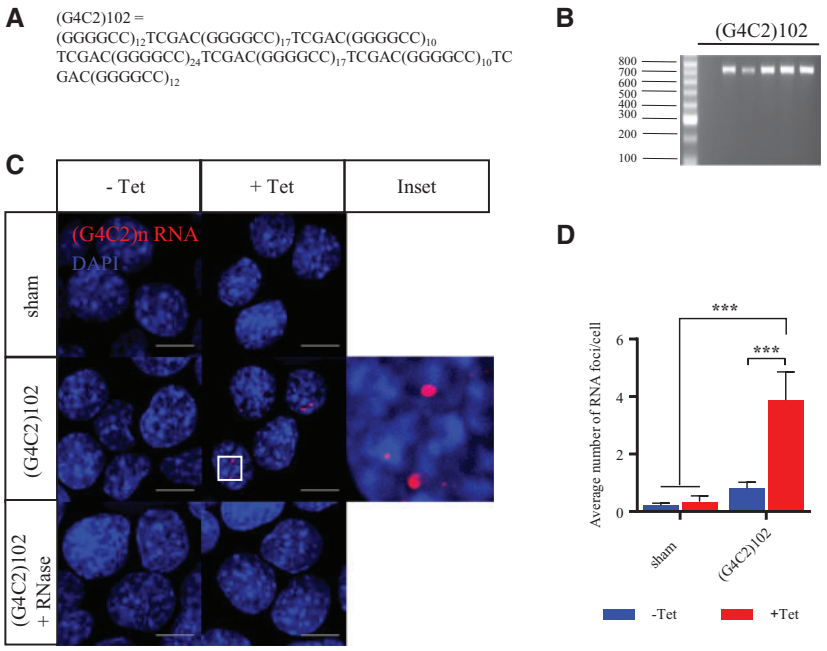

Figure 1 NSC34 (G4C2)102 cells have tetracycline inducible (G4C2)102 RNA expression, which forms RNA foci. **(A)** Sequence of the (G4C2)102 construct. **(B)** Gel electrophoresis sizing the (G4C2)102 repeat constructs in the pcDNA5/FRT/TO-(G4C2)102 plasmids. **(C)** NSC34 sham and NSC34 (G4C2)102 cells were cultured for 3 days with (or without) 0.5 μg/ml tetracycline. NSC34 (G4C2)102 cells were additionally treated with RNAse A as a control. Cells were stained with a locked nucleic acid (C4G2)3 sense probe (Red) and Dapi (Blue). Foci magnified 5× inset. Scale bar = 10 μm. **(D)** Average number of RNA foci per cell (***$p < 0.001$; two-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; $n = 3$).

pcDNA5/FRT/TO vectors were integrated into Flp-In™ T-REx™ NSC34 cells, generating NSC34 sham as a control *in vitro* model.

The interrupted (G4C2)n repeat RNA forms characteristic RNA foci comparable to those detectable in the CNS of C9ORF72-ALS cases. Therefore, RNA fluorescence *in situ* hybridisation (FISH) was used to detect expression of the (G4C2)n RNA in the NSC34 (G4C2)102 cells. The number of sense (G4C2)n RNA foci increased when the cells were induced with tetracycline (Fig. 1C and D). In addition, most RNA foci were nuclear, but occasional cytoplasmic (G4C2)n RNA foci were also detected (Fig. 1C). RNA foci were not detected in NSC34 sham cells (Fig. 1C). In addition, RNAse A treatment ablated foci in the NSC34 (G4C2)102 cells (Fig. 1C).

The NSC34 (G4C2)102 cells were also examined for antisense (C4G2)n RNA transcripts, using an antisense RNA FISH probe. As expected, no antisense (C4G2)n RNA foci were detected in the NSC34 (G4C2)102 cells (Fig. 2), because the (G4C2)n repeat is independent of the genomic C9ORF72 gene and any associated promoter(s) that drive transcription in the antisense direction. To demonstrate that we can detect antisense (C4G2)n RNA foci, we transfected antisense (C4G2)102 constructs into HEK293 cells, as a positive control (Fig. 2).

In addition, immunoblotting was performed to detect translation products of the (G4C2)102 RNA. Translation of the (G4C2)102 was predicted to change frame between DPR motifs at each TCGAC interruption, and, therefore, translation products would contain a mixture of poly-GA, poly-GR, and poly-GP stretches (Fig. 3A). NSC34 (G4C2)102 cells were immunoblotted using anti-poly-GA, anti-poly-GR, anti-poly-GP, anti-poly-AP and anti-poly-PR (Fig. 3B–F). HEK293 cells were transfected with (GA)68, (GR)100, (AP)100 and (PR)100 constructs, and used as positive controls for the respective antibodies (Fig. 3). There was no available positive control for the anti-poly-GP antibody. The anti-poly-GA, anti-poly-GR and anti-poly-GP antibodies all

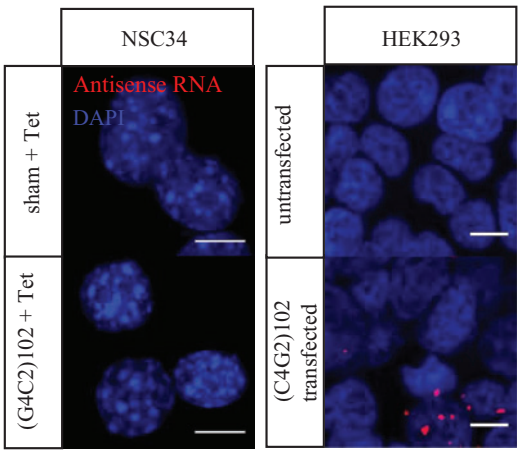

Figure 2 NSC34 (G4C2)n cells do not transcribe (G4C2)n in the antisense direction. NSC34 sham and NSC34 (G4C2)102 cells were cultured for 3 days with 0.5 μg/ml tetracycline. Cells were stained with a locked nucleic acid (G4C2)3 antisense probe (Red) and Dapi (Blue). HEK293 cells transfected with an antisense (C4G2)102 plasmid contain (C4G2)102 foci and serve as a positive control. Scale bar = 10 μm.

detected proteins at 24 and 27 kDa in the NSC34 (G4C2)102 cells treated with tetracycline (Fig. 3B–D), and these proteins were more abundant in the tetracycline-treated NSC34 (G4C2)102 compared with untreated NSC34 (G4C2)102 cells (Fig. 3B–D). In addition, these proteins were not detected in NSC34 sham, indicating the (G4C2)102 RNA was translated to produce proteins containing all three sense DPR motifs. Also, we concluded the (G4C2)102 RNA was translated via RAN translation as there are no ATG start sites preceding the (G4C2)102 construct in any frame (Fig. 3A). Further, poly-AP and poly-PR are specifically

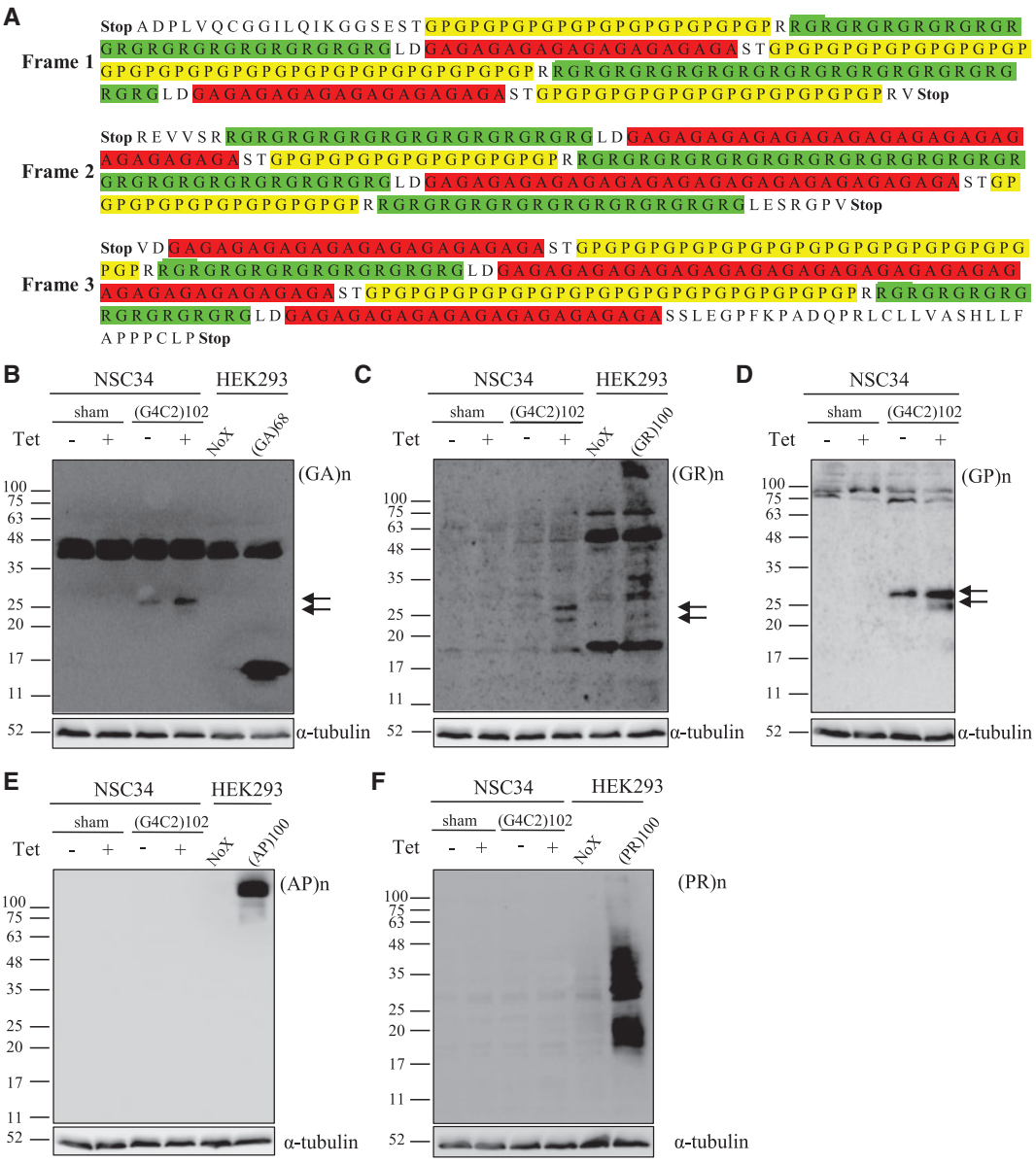

Figure 3 The (G4C2)102 construct undergoes RAN translation in the NSC34 (G4C2)102 (G4C2)102 cells. (**A**) Schematic of translation products (including Stop codons) from interrupted (G4C2)102 RNA construct in all three reading frames. GA, GP and GR repeats are highlighted in red, yellow and green, respectively. (**B–F**) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days with (or without) 0.5 μg/ml tetracycline. HEK293 cells were transfected with (GA)68, (GR)100, (AP)100 and (PR)100 constructs to serve as a positive control for the respective DPR antibody. Cells were lysed and immunoblotted with anti-poly-GA (B), anti-poly-GR (C), anti-poly-GP (D), anti-poly-AP (E) and anti-poly-PR (F). Arrows indicate the RAN translation products in the NSC34 (G4C2)102 cells. α-Tubulin was used as a loading control for blots. Molecular weight markers are indicated (kDa).

translated from the antisense (C4G2)n RNA, and neither anti-poly-AP nor anti-poly-PR detected any proteins specific to the NSC34 (G4C2)102 cells compared with NSC34 sham cells (Fig. 3E and F), supporting an absence of anti-sense (C4G2)n RNA transcription and subsequent translation. Finally, translation products were also detected using the anti-poly-GA, anti-poly-GR and anti-poly-GP antibodies in HEK293 cells containing the same interrupted (G4C2)102 repeat construct (data not shown).

NSC34 (G4C2)102 cells recapitulate certain aspects of ALS pathology

To validate the NSC34 (G4C2)102 cells as a model of C9ORF72-ALS, they were assessed for biochemical changes that are relevant to the pathophysiology of C9ORF72-ALS. First, we and others have previously shown that several RNA binding proteins (RBPs) such as Serine/Arginine-Rich Splicing Factor 1 (SRSF1), Serine/Arginine-Rich Splicing Factor 2 (SRSF2) and Nucleolin (NCL) bind (G4C2)n RNA, and co-localise with (G4C2)n RNA foci in C9ORF72-ALS cells (24,28). Tetracycline-induced NSC34 (G4C2)102 cells were co-stained for sense (G4C2)n RNA foci as well as either SRSF1, SRSF2 or NCL (Fig. 4A–C). We analysed 50 cells for each RBP co-stain. Of the (G4C2)n RNA foci counted, 19.5% co-localised with SRSF1 puncta; 11.9% co-localised with SRSF2 puncta; and 22.6% co-localized with nucleolar NCL (Fig. 4A–C). We have previously shown (G4C2)n RNA foci co-localisation with SRSF2 in cerebellar granule cells and motor neurons in C9ORF72-ALS patients (24). NCL was also

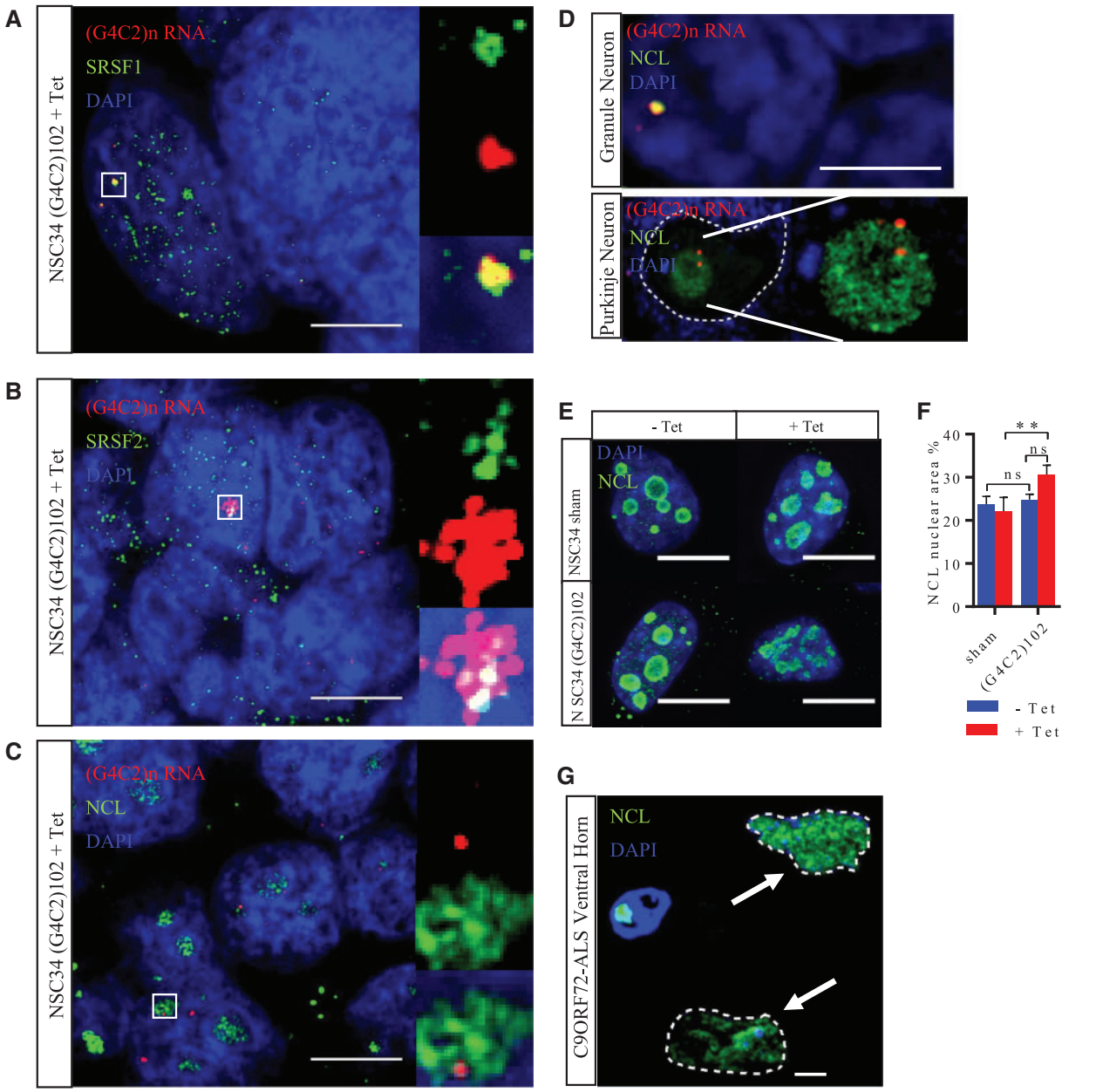

Figure 4 NSC34 (G4C2)102 cells recapitulate pathological features of C9ORF72-ALS. **(A–C, E)** NSC34 (G4C2)102 cells were induced with 0.5 μg/ml tetracycline for 5 days. Cells were then stained with a locked nucleic acid (C4G2)3 sense probe (Red), anit-SRSF1 (A), anti-SRSF2 (B) or anti-NCL (C) (all green), and Dapi (Blue). Scale bar = 10 μm. **(D)** Cerebellar slices from C9ORF72-ALS cases were stained with a locked nucleic acid (C4G2)3 sense probe (Red), anti-NCL (green) and Dapi (Blue). Scale bar = 3 μm. (E) NSC34 sham and NSC34 (G4C2)102 cells were stained for NCL (green) and DAPI (blue). Scale bar = 10 μm. **(F)** Quantification of NCL area as a percentage of nuclear area in NSC34 sham and (G4C2)102 cells (*$p < 0.05$; **$p < 0.01$; two-way ANOVA with Tukey's multiple comparisons post hoc test; $n = 3$). **(G)** Ventral horn from C9ORF72-ALS was stained for NCL and DAPI. Motor neurons (indicated with arrows and nuclei outlined) show disrupted NCL staining, while glial cells do not show disrupted NCL staining. Scale bar = 3 μm.

observed to co-localise with foci in cerebellar granule and Purkinje neurons from C9ORF72-ALS patients (Fig. 4D).

Second, nucleolar stress is implicated in C9ORF72-ALS pathology (28). Here we detected a degree of disruption to the nucleolar morphology, in the NSC34 (G4C2)102 cells (Fig. 4E). The nucleolar area as a percentage of the nucleus was increased by 38.6 ± 9.8% ($p < 0.01$) in the NSC34 (G4C2)102 + tet vs. NSC34 sham cells + tet (Fig. 4E and F). There was no difference in the overall nuclear area between the cell groups (data not shown). This effect was much more apparent in motor neurons from the ventral horn, but not glial cells, from C9ORF72-ALS patients (Fig. 4G). To establish if the wider ALS pathological features of TDP-43 mislocalisation or oxidative stress were also present in this cellular model, immunofluorescence microscopy and dichlorofluorescein (DCF) assays were performed on the NSC34 (G4C2)n cells. However, there was no TDP-43 mislocalisation or aggregation (Supplementary Material, Fig. S1), or increase in oxidative stress were observed in the NSC34 (G4C2)102 cells (Supplementary Material, Fig. S2), compared with NSC34 sham cells, at least under basal/unstressed cell culture conditions.

(G4C2)102 repeat induces cytotoxicity

Next, we assessed whether the (G4C2)102 repeat was toxic in the motor neuronal NSC34 cells. The NSC34 cells were grown for 7 days, and induced with tetracycline for increasing lengths of time. The NSC34 (G4C2)102 showed reduced viability of $29.9 \pm 8.6\%$ ($p < 0.05$) in an MTT assay after 7 days tetracycline induction compared with NSC34 sham cells. In addition, tetracycline exposure did not reduce viability in the NSC34 sham cells (Fig. 5A). A growth curve was also performed on the NSC34 sham and NSC34 (G4C2)102 cells, with or without 0.5 μg/ml tetracycline (Fig. 5B). Cells were cultured for 16 days in total, viable cells were counted every 4 days and 1.5×10^6 cells were reseeded and cultured. There were significantly fewer viable NSC34 (G4C2)102 cells when induced with tetracycline after 12 days compared with NSC34 sham cells ± tetracycline and NSC34 (G4C2)102 cells without tetracycline. Specifically, there were only $66.8 \pm 26.5\%$ ($p < 0.001$) of NSC34 (G4C2)102 + tetracycline at day 12 and $52.4 \pm 11.6\%$ ($p < 0.0001$) at day 16 compared with NSC34 sham + tetracycline (Fig. 5B). In contrast, there were no significant differences in the number of NSC34 sham with or without tetracycline and NSC34 (G4C2)102 cells without tetracycline.

(G4C2)n affects the PI3K/Akt signalling pathway at transcriptomic level

We had previously undertaken transcriptome analysis on laser captured microdissected (LCM) motor neurons from the spinal cord of C9ORF72-ALS patients (30). Within this dataset, *PTEN* (phosphatase and tensin homolog) had the highest fold change of all differentially expressed genes (fold change = +11.3, $p = 0.00001$, Supplementary Material, Table 1). Additionally, previous transcriptome analysis on LCM motor neurons from spinal cord motor neurons of SOD1-ALS patients showed that the PI3K/Akt signalling pathway, negatively regulated by PTEN, was increased in the surviving motor neurons, and may provide a therapeutic target (54). Therefore, we investigated the PI3K/Akt signalling pathway in the Kyoto Encyclopaedia of Genes and Genomes (KEGG), in both the human C9ORF72-ALS LCM motor neurons and the murine NSC34 (G4C2)n cells. Careful statistical analysis revealed that this pathway was significantly dysregulated in LCM motor neurons carrying the (G4C2)n C9ORF72 repeat expansion (rank-product, $p = 0.01$), and the NSC34 cells expressing the (G4C2)102 repeat (rank-product, $p < 0.00001$). Transcripts in the PI3K/Akt signalling pathway (KEGG) are listed in Supplementary Tables 1 and 2 for human motor neurons and murine NSC34 cells, respectively.

PTEN single nucleotide polymorphism (SNP) associated with protection against ALS

We also examined a large genome wide association study (GWAS) using the NeuroX chip (55) to genotype 4890 ALS cases and 5649 normal controls; the NeuroX chip includes genotyping of standard Illumina exome content of approximately 240 000 variants, and additionally, more than 24 000 custom content variants to improve coverage in genes associated with neurological diseases. A list of SNPs found in *PTEN* was identified using BioMart (56). The threshold for significant association with the ALS phenotype was set as a Benjamini–Hochberg-corrected false discovery rate (FDR) < 0.05, which is appropriate for a hypothesis-based rather than a screening-based approach. A single SNP in *PTEN* was associated with risk of ALS (FDR = 0.005).

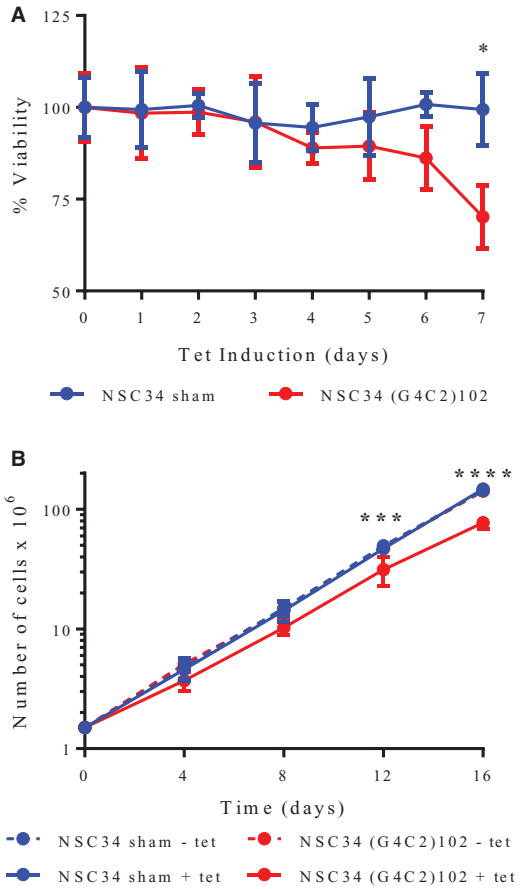

Figure 5 (G4C2)102 repeat is toxic in NSC34 cells. **(A)** NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days and were induced for various lengths of time with 0.5 μg/ml tetracycline. Cell viability was measured using an MTT assay (*$p < 0.05$; two-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; $n = 3$). **(B)** NSC34 sham and NSC34 (G4C2)102 cells were cultured for 16 days with or without 0.5 μg/ml tetracycline. The cells were counted every 4 days, and then 1.5×10^6 cells were reseeded (***$p < 0.001$; ****$p < 0.0001$; two-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; $n = 4$).

The minor (A) allele of rs202004587 was five times more common in controls than in ALS patients. The minor allele is a missense change which has been associated with hereditary neoplasms by clinical testing (http://www.ncbi.nlm.nih.gov/clinvar/RCV000129085/) and is, therefore, thought to be associated with loss of PTEN function (57). Based on our data, the minor allele should, therefore, be protective against ALS. This further validates the transcriptomic findings from our cell model, and C9ORF72-ALS LCM motor neurons, and suggests that PTEN function could be an important therapeutic target for all ALS. It is possible that the SNP we have identified is in linkage disequilibrium with another more functionally significant variant in the PTEN gene; however, since we are relying on a GWAS dataset with incomplete coverage of the PTEN gene, it is not possible to explore this further with the current data-set.

Partial Pten knockdown rescues (G4C2)102-induced toxicity

Up-regulation of Pten is consistent with inhibition of the PI3K/Akt signalling pathway. This led us to hypothesise that reducing the expression of Pten in the C9ORF72 *in vitro* model system

might rescue the observed (G4C2)n-induced toxicity. NSC34 sham and NSC34 (G4C2)102 were both stably transduced with a lentivirus (LV) expressing either Pten shRNA, or GFP (as a control). Pten was knocked down by 56 ± 7.6% ($p < 0.05$) in the NSC34 sham cells stably transduced with Pten shRNA LV compared with the NSC34 sham cells stably transduced with control GFP LV, and 62.5 ± 22.7% ($p < 0.01$) in the NSC34 (G4C2)102 cells stably transduced with Pten shRNA LV compared with the NSC34 (G4C2)102 cells stably transduced with control GFP LV (Fig. 6A). Additionally, Pten knockdown does not significantly affect the number of RNA foci in the transduced NSC34 (G4C2)102 cells (Supplementary Material, Fig. S4), suggesting that this manipulation does not affect the (G4C2)102 RNA level. The LV transduced NSC34 cells were then grown for 7 days, with or without tetracycline, and viability was assessed. As expected, we detected a significant reduction (21.1 ± 9.1%; $p < 0.01$) in the NSC34 (G4C2)102 GFP LV cell viability when induced with tetracycline (Fig. 6B). However, there was no significant reduction in the NSC34 (G4C2)102 Pten shRNA LV cell viability when induced with tetracycline (Fig. 6B). Therefore, the partial Pten knockdown rescues the NSC34 cells from (G4C2)n-induced toxicity. These results were also replicated using NSC34 sham and (G4C2)102 cells that were transduced with LV expressing scrambled shRNA where the expected cytotoxicity was observed (data not shown). Modulation of PTEN may, therefore, provide a potential therapeutic approach to ameliorate the toxic effects of (G4C2)n-induced toxicity in C9ORF72-ALS.

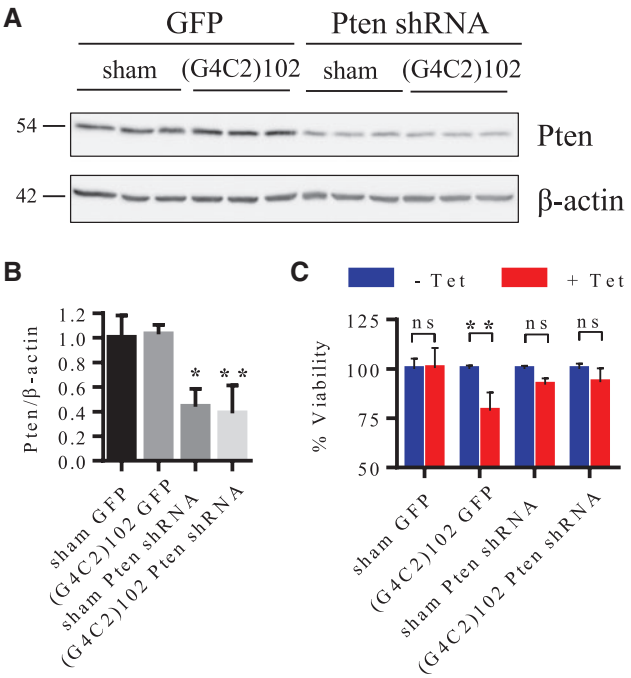

Figure 6 Pten knockdown rescues NSC34 cells from (G4C2)102-induced toxicity. **(A)** Cells were lysed and immunoblotted with anti-Pten. β-Actin was used as a loading control for blots. Molecular weight markers are indicated (kDa). **(B)** Quantification of Pten normalised to β-actin (*$p < 0.05$; **$p < 0.01$; one-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; $n = 3$). **(C)** NSC34 sham GFP, NSC34 (G4C2)102 GFP, NSC34 sham Pten shRNA and NSC34 (G4C2)102 Pten shRNA cells were cultured for 7 days, with (or without) 0.5 μg/ml tetracycline. Cell viability was measured using an MTT assay, and the viability of the tetracycline induced cells was normalised to the non-induced control for each individual cell line (**$p < 0.01$; two-way ANOVA with Sidak's multiple comparisons post hoc test; data shown are mean and SD; $n = 3$).

Discussion

Recently, various studies have attempted to elucidate the toxic function associated with the (G4C2)n repeat expansion in C9ORF72, which is the most common genetic cause of ALS and FTD (1,2). The repeat expansion is thought to either cause neuronal injury through loss of C9ORF72 function (haploinsufficiency), and/or a gain-of-toxic function from the RNA or DPR proteins. Although there is still much to learn, emerging evidence strongly supports a gain-of-toxic function. First, multiple studies show that the (G4C2)n RNA and/or DPR proteins cause toxicity and/or neurodegeneration in various model systems. Second, while studies interrogating C9ORF72 loss-of-function have reported microglial pathology, they have not reported motor neuron degeneration or TDP-43 pathology associated with C9ORF72 knock down in mice (13,16–21). Third, the targeted knockdown of *C9ORF72* transcripts using ASOs rescues toxicity, and associated transcriptomic changes in cellular models from C9ORF72-ALS/FTD patients (9,20,27).

To investigate the early biochemical changes caused by (G4C2)n repeat expression, we generated stable, isogenic motor neuron-like cell lines with inducible expression of interrupted (G4C2)n repeat constructs. The interrupted (G4C2)102 repeat constructs produced (G4C2)n RNA foci, and underwent RAN translation in the NSC34 (G4C2)n cells, recapitulating key pathological features of C9ORF72-ALS/FTD patient cells. Not surprisingly, there were no antisense-derived RNA foci or RAN translation products in the NSC34 (G4C2)n cells, making this model a sense-only (G4C2)n gain-of-function model. We believe that the new *in vitro* model we have generated will prove useful for screening compounds that modulate the number of RNA foci, the burden of RAN translation products, or the cytotoxicity relating to the presence of the *C9ORF72* G4C2 expansion.

However, while the NSC34 (G4C2)102 cells did recapitulate both the RNA foci and RAN translation of the (G4C2)102 repeat, we did not detect TDP-43 proteinopathy or oxidative stress (two pathological hallmarks of ALS) under basal culture conditions. We did detect colocalisation of the SRSF1, SRSF2 and NCL with the (G4C2)n RNA foci, which is consistent with both C9ORF72-ALS/FTD CNS tissue, and other cellular models expressing the (G4C2)n repeat expansion. Therefore, we reasoned that any biochemical and transcriptomic effects we detected in this cellular model system will represent early changes in response to the presence of the (G4C2)102 repeat. Further, and in agreement with previous studies, the (G4C2)102 repeat reduced NSC34 cell viability and growth rate. However, this could have been caused by either the (G4C2)n RNA foci or the RNA translation proteins produced by the (G4C2)102 transcripts. This is consistent with the literature, showing a gain-of-toxic function associated with the (G4C2)n repeat expansion either through direct RNA toxicity or associated RAN translation protein toxicity (9,20,27,41). Interestingly, however, the proteins generated in this model are hybrids of all sense DPR proteins. There are three possible explanations for the motor neuronal toxicity observed in our model system: (1) the toxicity could be due to the (G4C2)102 RNA alone; (2) the sense DPR motifs may be toxic even when co-occurring in the same polypeptide; (3) the RAN proteins confer toxicity through the poly-GR stretches.

Once we could detect toxicity in the model, the main aim was to use the model for identifying potential therapeutic targets for C9ORF72-ALS/FTD. As a first step towards this goal, we performed transcriptomic analysis on the NSC34 (G4C2)102 cell model. Importantly, using our inducible model system, any

changes detected are likely to be the early transcriptomic changes associated with any gain-of-toxic function from the (G4C2)102 repeat. We identified the novel finding that the (G4C2)n repeat dysregulates the PI3K/Akt signalling pathway in a motor neuronal model of C9ORF72-ALS. Further, we validated this finding using transcriptomic analysis of LCM motor neurons from C9ORF72-ALS cases. This is also consistent with previous transcriptome analysis on LCM motor neurons from spinal cord of SOD1-ALS patients (54), implicating the pathway in a broader ALS context. Within the C9ORF72-ALS dataset, *PTEN* had the highest fold change of all DE genes, and previous work has shown PTEN knockdown via siRNA promotes motor neuron survival in ALS and SMA models (54,58,59). Therefore, we partially knocked down Pten in the NSC34 (G4C2)102 cells, which rescued the cells from the (G4C2)102 repeat induced toxicity. Finally, we present evidence from a GWAS study that loss of PTEN function may be protective against ALS more broadly. Taken together, the data we present suggest that PTEN may represent a potential therapeutic target in C9ORF72-ALS, but may also be relevant in other subtypes of ALS.

Materials and Methods

Generation of interrupted (G4C2)n repeat constructs

The interrupted (G4C2)102 repeat construct was generated by restriction digest and ligation. Synthesised TCGAC(G4C2)10 sense and ACGT(G2C4)10 antisense ssDNA oligonucleotides (Sigma-Aldrich) were designed so that the dsDNA produced by annealing the oligonucleotides is flanked 5′ by *Sal*I and 3′ by *Xho*I restriction sites. The ssDNA oligos were denatured at 99°C for 30 min and then annealed by reducing the temperature by 0.5 °C/min. These (G4C2)10 were ligated into *Sal*I and *Xho*I cut pcDNA6.2-GW/EmGFP-miR (*Invitrogen*), to generate pcDNA6.2-GW/EmGFP-(G4C2)10. Subsequent ligations involved *Xho*I cutting pcDNA6.2-GW/EmGFP-(G4C2)n vectors and inserting further (G4C2)10 repeats. The 5′ *Sal*I site of the inserted (G4C2)n was destroyed whilst the 3′ *Xho*I site was retained – if the insertion orientation was correct. pCMV-EmGFP-(G4C2)102 vectors containing 102 repeats were generated via this method. The exact sequence of the (G4C2)102 construct is described in Figure 1A. EmGFP was subsequently removed from these vectors by *Dra*I digestion and religation.

The (G4C2)102 construct was subcloned into pcDNA5/FRT/TO HIS (*Addgene*) using *Dra*I and *Xho*I restriction sites. The FRT/TO/(G4C2)n plasmids were then *Hind*III and *BamH*I cut, Klenow treated and subsequently PNK treated and religated. Transformations of plasmids containing the (G4C2)n repeat constructs were performed using recombination-deficient β-10 *Escherichia coli* (*NEB*) to minimise (G4C2)n repeat shrinkage. All constructs were checked by sequencing at regular intervals. Plasmids were extracted using a NucleoSpin Plasmid kit (Macherey-Nagel), according to the instructions of the manufacturer.

Generation of NSC34 (G4C2)102 cell models

Flp-In™ T-REx™ murine motor neuron-like NSC34 cells were generated in house using the Flp-In™ T-REx™ Core Kit (Invitrogen) according to the instructions of the manufacturer. Briefly, an FRT site and tetracycline repressor element were inserted independently into the NSC34 cell genome, and the best clone was selected. Then, pcDNA5/FRT/TO© and pcDNA5/FRT/TO-(G4C2)102 plasmids were stably integrated into the Flp-In™ T-REx™ NSC34 host cell line generating isogenic NSC34 sham and NSC34 (G4C2)102 cell lines, respectively. NSC34 (G4C2)102 cells had tetracycline inducible expression of the inserted (G4C2)102 construct.

Cell culture, cell line maintenance, plasmid transfection and lentiviral transduction

To maintain the cell lines, the NSC34 sham and NSC34 (G4C2)102 cells were cultured in supplemented DMEM (10% FBS, tetracycline-free (Biosera), 50 U/ml penicillin/streptomycin (Lonza Group Ltd), 100 μg/ml Hygromycin B (*Invitrogen*), 5 μg/ml Blasticidin (Invitrogen) in a 37°C/5% CO_2 incubator. Media was replaced every 2–3 days. NSC34 sham and (G4C2)102 cells that were stably transduced with the lentiviral vectors additionally had 1 μg/ml puromycin selection in the supplemented DMEM. HEK293 cells were cultured in supplemented DMEM (10% FBS (*Sigma-Aldrich*), 50 U/ml penicillin/streptomycin) in a 37°C/5% CO_2 incubator. NSC34 and HEK293 cells were seeded at an appropriate density, such that the cells were 70-80% confluent at the experimental end point.

For imaging, NSC34 cells were seeded onto gelatin coated coverslips. For each ml of media on the HEK293, 1.4 μg plasmid DNA and 5 μg PEI were diluted in 100 μL OptiMEM (*Life Technologies*), and added dropwise to the media on cells. Cells were incubated for 24 h.

For the lentiviral transduction, NSC34 sham and NSC34 (G4C2)102 cells were transduced with PTEN shRNA (mouse) lentiviral particles (Santa Cruz; sc-36326-V), cop GFP Control lentiviral particles (Santa Cruz; sc-108084), or Control shRNA lentiviral particles (Santa Cruz; sc-108080). Stable transformants were selected for using 1 μg/ml puromycin.

NSC34 (G4C2)n cell line tetracycline induction

About 0.5 μg/ml tetracycline (*Invitrogen*) was added to NSC34 (G4C2)n cell media, to induce (G4C2)n RNA expression. Tetracycline was added every 3 days (if applicable) to maintain concentration in the media.

Immunoblotting

NSC34 and HEK293 cells were grown on 6-well plates until 80% confluent. Cells were lysed in lysis buffer (150 mM NaCl, Protease Inhibitor Cocktail EDTA–free (Roche; used to according to the instructions of the manufacturer)) on ice for 20 min. Lysates were centrifuged at 17 000 × *g* at 4 °C for 5 min to clarify. Supernatant was incubated with Laemmli buffer at 95°C for 5 min. 25 μg protein from NSC34 lysate and 25 μg protein from HEK293 lysates were separated on 12% or 15% Tris-glycine SDS-polyacrylamide gels and transferred to Amersham™ Protran™ Premium 0.2 nitrocellulose membrane (GE Healthcare). Membranes were blocked in 5% milk (Marvel) in TBS-T at RT for 1 h. Membranes were then probed with mouse anti-GA (1 in 200; kindly provided by Prof Dieter Edbauer), rabbit anti-GP, rabbit anti-GR, rabbit anti-PR, rabbit anti-PA (all 1 in 2000; kindly provided by Prof Stuart Pickering-Brown), rabbit monoclonal anti-PTEN (1 in 1000, Cell Signalling Technology, #9188), mouse monoclonal anti-β actin (1 in 10 000, Abcam, AC-15), or mouse monoclonal anti-α tubulin (1 in 10 000; Sigma-Aldrich; clone DM1A), in 5% Milk/TBS-T at RT for 1 h. Membranes were washed three times in TBS-T, then probed with goat polyclonal anti-rabbit-IgG HRP (1 in 5000; DAKO) or goat polyclonal

anti-mouse-IgG HRP (1 in 5000; *DAKO*) in 5% Milk/TBT at RT for 1 h. Membranes were washed three times in TBS-T. Bands were visualised with EZ ECL chemiluminescence detection kit for HRP (Geneflow Ltd) according to the instructions of the manufacturer and imaged using a G-BOX (Syngene).

RNA fish

A 5' TYE-563-labelled LNA (16-mer fluorescent)-incorporated DNA probe was used against the sense and antisense RNA hexanucleotide repeat (Exiqon Inc., batch numbers 607323 and 515905, respectively). Slides with tissue were fixed in 4% paraformaldehyde for 10 min. Before use, formalin fixed paraffin-embedded tissue sections were deparaffinized. Slides with NSC34 and HEK293 cells were fixed and permeabilised in 4% PFA/0.2% Triton X-100 for 20 min. For the RNAse-treated control, slides were incubated with 10 μg/ml RNAse A at 37°C for 30 min. Slides were blocked with hybridization solution [50% formamide, 2 × saline sodium citrate (SSC), 100 mg/ml dextran sulphate, 50 mM sodium phosphate pH 7.0] at 66°C for 1 h and then incubated with 400 ng/ml of denatured probe in hybridization solution at 66 °C overnight. After hybridization, slides were washed once in 2 × SSC/0.1% Tween-20 at RT for 5 min and three times in 0.1 × SSC at 65°C for 10 min. NSC34 cells that were subsequently dual stained by immunocytochemistry (ICC) were first irradiated on ice with 0.3 J/cm^2 UV, washed three times with PBS, and then ICC staining was performed. Slides were mounted with mounting medium containing DAPI (Vector Labs Inc.). All solutions were made with DEPC-treated water.

Immunocytochemistry and immunohistochemistry

Slides with NSC34 cells were fixed and permeabilised in 4% PFA/0.2% Triton X-100 at RT for 20 min. Slides were incubated with rabbit polyclonal anti-TDP-43 (1 in 200, Proteintech, 10782-2-AP), rabbit polyclonal anti-SRSF1 (1 in 200, Abcam ab38017), mouse monoclonal anti-SRSF2 (1 in 200, Abcam ab11826), or rabbit polyclonal anti-NCL (1:200; Proteintech; 10556-1-AP) in 2% BSA in PBS at RT for 1 h. Slides were washed three times in PBS, then incubated with goat anti-rabbit IgG H&L (1 in 1000; Alexa Fluor® 488; *Abcam*) or goat anti-mouse IgG H&L (1 in 1000; Alexa Fluor® 488; Abcam) in 2% BSA in PBS at RT for 1 h. Formalin fixed paraffin-embedded tissue sections were deparaffinized and mouse anti-NCL (1 in 150 in 5% BSA; Abcam; ab136649) stained slides were first antigen retrieved by heat in trisodium citrate pH 6.5 for 20 min, then completed as for ICC. Slides of NSC34 and tissue sections were mounted with mounting medium containing DAPI (Vector Labs Inc.).

Microscopy imaging

RNA foci and RBP: RNA foci were visualised using a Leica SP5 confocal microscope system with a X63/1.4 oil immersion objective lens. The presence of foci was assessed within a high-resolution (1433 mm^2 per image, 511 × 511 pixels) z-stack made up of images at 0.13 μm intervals through the entire nuclear volume of the cell under consideration. RNA foci were quantified manually. The same imaging was used for RNA foci and RBP (SRSF1, SRSF2 and NCL) co-staining, and for each RBP foci co-stain 50 NSC34 (G4C2)102 cells were analysed for co-localisation.

NCL: NCL staining in the NSC34 cells and tissue was visualised using a Leica SP5 confocal microscope system with an X63/1.4 oil immersion objective lens. (3775 mm^2 per image, 511 × 511 pixels) z-stack made up of images at 0.5-mm intervals through the entire nuclear volume of the cell under consideration. To quantify the NCL area relative to the nuclear area, we used the analysis previously described by Haeusler *et al.* 2014 (28). Briefly, a threshold of 50–100 was set in FIJI to measure the nucleolar NCL area, relative to the nuclear area (defined by DAPI staining). Data are means ± SD $n = 3$ for NSC34 cells.

TDP-43: TDP-43 staining in NSC34 cells were visualised using a Nikon DS Ri1 Eclipse.

MTT assay

NSC34 cells were grown for 7 days on a 96-well plate. 0.5 μg/ml tetracycline was added to media, and re-added every 3 days. After 7 days incubation, 0.5 μg/ml MTT reagent (Sigma-Aldrich) was added to media. Plates were incubated in a 37°C/5% CO_2 incubator for 90 min. One volume SDS/DMF (20% sodium dodecyl sulphate in 50% dimethyl formamide, pH 4.7) was added to lyse cells. An absorbance of 595 nm of wells was assessed using a PHERAstar FS plate reader (BMG Labtech Ltd).

DCF assay

NSC34 cells were grown on 96-well plates for 5 days. 0.5 μg/ml tetracycline was added to media and re-added every 3 days. After 5 days incubation, media was replaced with supplemented phenol-red free DMEM (Lonza) + 10% FBS, tetracycline free (*Biosera*). Cytosolic reactive oxygen species levels were measured using 6-carboxy-2′,7′-dichlorodihydrofluorescein diacetate, di(acetoxymethyl ester) (DCF; Molecular Probes™, Life Technologies, C-2938) fluorescence. DCF was added to NSC34 cells at 10 μM, and cells were incubated at 37 °C for 90 min. The fluorescence of oxidised DCF was read at Ex485 nm/Em520 nm using a PHERAstar FS plate reader (BMG labtech Ltd). Cells were then freeze-thawed, and cell number was measured by adding 1.5 μM Ethidium homodimer-1 (EthD1, Molecular Probes™, Life Technologies) to the medium. Fluorescence was measured at Ex570 nm/Em610 nm. Raw DCF data were then normalised to EthD1 reading of cell number.

RNA preparation for microarray

NSC34: Three replicate sets of NSC34 cells were cultured for 5 days with or without tetracycline. Media and tetracycline (where necessary) were replenished every 3 days. RNA was extracted using TRI-Reagent® (Zymo Research) and a Direct-zol™ RNA MiniPrep kit (Zymo Research) according to the instructions of the manufacturer. RNA quantity and quality was assessed on the Nanodrop spectrophotometer (Thermofisher) and a Bioanalyser (Agilent), respectively, to ensure all samples were of comparable and sufficient quality to proceed.

Laser capture microdissected (LCM) Motor Neuron Microarray: Refer to Cooper-Knock *et al.* 2015 (30).

Microarray hybridisation

NSC34: Total RNA (500 ng) was linearly amplified using the GeneChip® WT PLUS Reagent kit (Affymetrix) according to the instructions of the manufacturer, to produce ss-cDNA. ss-cDNA (5.5 μg) was fragmented and labelled with biotin, then applied to the GeneChip® Mouse Whole Transcript 1.0 ST Array (Affymetrix), according to the instructions of the manufacturer. Three chips were used for each NSC34 cell line and tetracycline

induction condition. Array washing and staining was performed in the Fluidics Station 400 according to the Affymetrix protocol. Arrays were scanned using the Affymetrix GeneChip® 3000 7G scanner. The Affymetrix Gene Chip Command Console (Affymetrix) was used to monitor scanning and to convert the raw image file into a cell intensity file ('.CEL').

LCM Motor Neurons: Refer to Cooper-Knock *et al.* 2015 (30). RNA (20–25 ng) was linearly amplified using the Affymetrix Two Cycle cDNA synthesis protocol to produce biotin-labelled copy RNA. Copy RNA (15 μg) was fragmented for 15 min and hybridized to the Human Genome U133 Plus 2.0 GeneChips, according to Affymetrix protocols. Array washing and staining was performed as above.

Gene expression data analysis

NSC34 :For the mouse microarrays, RMA normalisation was performed. 61 transcripts within the murine PI3K/Akt signalling pathway (KEGG) were measured on the Mouse Whole Transcript microarray platform in the NSC34 cells (Supplementary Material, Table S2). PCA analysis without statistical manipulation revealed a distinct separation between NSC34 (G4C2)102 and NSC34 sham cells (Supplementary Material, Fig. S3). In order to interrogate this signal further, expression of all transcripts was ranked according to *t*-statistic in a comparison between NSC34 (G4C2)102 and NSC34 sham cell lines. The combined rank-product (60) of transcripts from the PI3K/Akt signalling pathway was compared to 100 000 random permutations of the same number of transcripts. In short, the product of (rank $_{\text{transcript X}}$/total number of transcripts) was calculated where transcript X was each member of the PI3K/Akt pathway in turn; then the probability that this rank-product occurred by 'chance' was determined by comparison with random selections of the same number of transcripts where each rank is equally probable. Differential expression of individual transcripts (Supplementary Material, Table S2) was examined by t-test and fold change.

LCM Motor Neurons: Data from the laser captured microdissection motor neurons were normalised using the *PUMA* package (61,62). Around 283 transcripts within the human PI3K/Akt signalling pathway (KEGG) were measured on the Human U133-Plus 2.0 microarray platform in the laser-captured motor neurons (Supplementary Material, Table S1). Expression of all transcripts was ranked according to *t*-statistic in a comparison between motor neurons with and without expansion of C9ORF72. The combined rank-product (60) of transcripts from the PI3K/Akt signalling pathway was compared with 100 000 random permutations of the same number of transcripts. Differential expression of individual transcripts (Supplementary Material, Table S1) was examined by *t*-test and fold change.

GWAS analysis

GWAS data were examined using PLINK (http://pngu.mgh.harvard.edu/purcell/plink/ (63)). Genomes were screened from 4890 ALS patients and 5191 controls including 5649 males and 4432 females. All genomes were founders i.e. no two individuals were related to each other. Association analysis was performed and adjusted for multiple testing using the Benjamini and Benjamini & Hochberg (1995) step-up false discovery rate (FDR) control (64). At a significance level of FDR < 0.05, rs202004587 was significantly associated with risk of ALS. Repeating this analysis for a random selection of 500 genes, with approximately the same number of SNPs as PTEN, further suggested that this finding was significant ($p < 0.05$).

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We are grateful to Prof. Dieter Edbauer and Prof. Stuart Pickering Brown for kindly supplying antibodies to the dipeptide repeat proteins.

Conflict of Interest statement. None declared.

Funding

This work was supported in part by the European Community's Seventh Framework Programme [FP7/2007-2013] under the EuroMOTOR project [Grant agreement no. 259867 to J.K., P.J.S. and A.C.]; P.J.S. is also supported as a National Institute for Health Research Senior Investigator and by the Medical Research Council; a Motor Neuron Disease Association/Medical Research Council Lady Edith Wolfson Fellowship award [MR/K003771/1 to J.C.K.]; and the Sheffield Hospitals Charitable Trust (Grant no. 131425 to J.K. and M.S.). The Italian Ministry of Health (Ricerca Sanitaria Finalizzata 2010, Grant RF-2010-2309849), the Agenzia Italiana per la Ricerca sulla SLA (ARISLA) (Sardinials project) (to A.C.). This work was supported in part by the Intramural Research Programs of the NIH, National Institute on Aging (Z01-AG000949-02 to B.J.T.). The work was also supported by the Agency of Toxic Substances and Disease Registry, Center for Disease Control (to B.J.T.). Funding for Open Access was provided by the University of Sheffield.

References

1. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.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.
2. Renton, A.E., Majounie, E., Waite, A., Simon-Sanchez, 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.
3. Beer, A.M., Cooper-Knock, J., Higginbottom, A., Highley, J.R., Wharton, S.B., Ince, P.G., Milano, A., Jones, A.A., Al-Chalabi, A., Kirby, J., *et al.* (2015) Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology. *Amyotroph. Lateral Scler. Frontotemporal Degener*, **16**, 249–251.
4. Byrne, S., Heverin, M., Elamin, M., Walsh, C. and Hardiman, O. (2014) Intermediate repeat expansion length in C9orf72 may be pathological in amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Frontotemporal Degener*, **15**, 148–150.
5. Cooper-Knock, J., Shaw, P.J. and Kirby, J. (2014) The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathol.*, **127**, 333–345.
6. 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.
7. Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A. and Kabashi, E. (2013) Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. *Ann. Neurol.*, **74**, 180–187.
8. Belzil, V.V., Bauer, P.O., Prudencio, M., Gendron, T.F., Stetler, C.T., Yan, I.K., Pregent, L., Daughrity, L., Baker, M.C., Rademakers, R., *et al.* (2013) Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol.*, **126**, 895–905.
9. Donnelly, C.J., Zhang, P.W., Pham, J.T., Haeusler, A.R., Heusler, A.R., Mistry, N.A., Vidensky, S., Daley, E.L., Poth, E.M., Hoover, B., *et al.* (2013) RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron*, **80**, 415–428.
10. 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.e1775-1779. e1713.
11. 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.
12. Therrien, M., Rouleau, G.A., Dion, P.A. and Parker, J.A. (2013) Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in *C. elegans*. *PLoS One*, **8**, e83450.
13. Sullivan, P.M., Zhou, X., Robins, A.M., Paushter, D.H., Kim, D., Smolka, M.B. and Hu, F. (2016) The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol. Commun.*, **4**, 51.
14. Yang, M., Liang, C., Swaminathan, K., Herrlinger, S., Lai, F., Shiekhattar, R. and Chen, J.F. (2016) A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci. Adv.*, **2**, e1601167.
15. Webster, C.P., Smith, E.F., Bauer, C.S., Moller, A., Hautbergue, G.M., Ferraiuolo, L., Myszczynska, M.A., Higginbottom, A., Walsh, M.J., Whitworth, A.J., *et al.* (2016) The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J.*, **35**, 1656–1676.
16. Burberry, A., Suzuki, N., Wang, J.Y., Moccia, R., Mordes, D.A., Stewart, M.H., Suzuki-Uematsu, S., Ghosh, S., Singh, A., Merkle, F.T., *et al.* (2016) Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci. Transl. Med.*, **8**, 347ra393.
17. Atanasio, A., Decman, V., White, D., Ramos, M., Ikiz, B., Lee, H.C., Siao, C.J., Brydges, S., LaRosa, E., Bai, Y., *et al.* (2016) C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci. Rep.*, **6**, 23204.
18. O'Rourke, J.G., Bogdanik, L., Yáñez, A., Lall, D., Wolf, A.J., Muhammad, A.K., Ho, R., Carmona, S., Vit, J.P., Zarrow, J., *et al.* (2016) C9orf72 is required for proper macrophage and microglial function in mice. *Science*, **351**, 1324–1329.
19. Sudria-Lopez, E., Koppers, M., de Wit, M., van der Meer, C., Westeneng, H.J., Zundel, C.A., Youssef, S.A., Harkema, L., de Bruin, A., Veldink, J.H., *et al.* (2016) Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta Neuropathol.*, **132**, 145–147.
20. 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.
21. Koppers, M., Blokhuis, A.M., Westeneng, H.J., Terpstra, M.L., Zundel, C.A., Vieira de Sá, R., Schellevis, R.D., Waite, A.J., Blake, D.J., Veldink, J.H., *et al.* (2015) C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann. Neurol.*, **78**, 426–438.
22. Gendron, T.F., Bieniek, K.F., Zhang, Y.J., Jansen-West, K., Ash, P.E., 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.
23. Renoux, A.J. and Todd, P.K. (2012) Neurodegeneration the RNA way. *Prog. Neurobiol.*, **97**, 173–189.
24. Cooper-Knock, J., Walsh, M.J., Higginbottom, A., Robin Highley, J., Dickman, M.J., Edbauer, D., Ince, P.G., Wharton, S.B., Wilson, S.A., Kirby, J., *et al.* (2014) Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain*, **137**, 2040–2051.
25. Lee, Y.B., Chen, H.J., Peres, J.N., Gomez-Deza, J., Attig, J., Stalekar, 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.
26. Mori, K., Lammich, S., Mackenzie, I.R.A., Forne, I., Zilow, S., Kretzschmar, H., Edbauer, D., Janssens, J., Kleinberger, G., Cruts, M., *et al.* (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**, 413–423.
27. 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.
28. Haeusler, A.R., Donnelly, C.J., Periz, G., Simko, E.A., 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.
29. 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.
30. Cooper-Knock, J., Bury, J.J., Heath, P.R., Wyles, M., Higginbottom, A., Gelsthorpe, C., Highley, J.R., Hautbergue, G., Rattray, M., Kirby, J., *et al.* (2015) C9ORF72 GGGGCC expanded repeats produce splicing dysregulation which correlates with disease severity in amyotrophic lateral sclerosis. *PLoS One*, **10**, e0127376.
31. Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P., *et al.* (2015) Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.*, **18**, 1175–1182.
32. 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.* (2013) The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science*, **339**, 1335–1338.
33. 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.
34. 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.* (2013) Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol.*, **126**, 881–893.
35. Zu, T., Liu, Y., Bañez-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.
36. Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P., Wang, T., Kim, J., Yun, J., Xie, Y. and McKnight, S.L. (2014) Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science*, **345**, 1139–1145.
37. Tao, Z., Wang, H., Xia, Q., Li, K., Jiang, X., Xu, G., Wang, G. and Ying, Z. (2015) Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Hum. Mol. Genet.*, **24**, 2426–2441.
38. Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S.S., Shi, Y., Lin, S., Shneider, N.A., Monaghan, J., Pandey, U.B., *et al.* (2014) Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron*, **84**, 1213–1225.
39. Yang, D., Abdallah, A., Li, Z., Lu, Y., Almeida, S. and Gao, F.B. (2015) FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. *Acta Neuropathol.*, **130**, 525–535.
40. May, S., Hornburg, D., Schludi, M.H., Arzberger, T., Rentzsch, K., Schwenk, B.M., Grässer, F.A., Mori, K., Kremmer, E., Banzhaf-Strathmann, J., *et al.* (2014) C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathol.*, **128**, 485–503.
41. Mizielinska, S., Grönke, S., Niccoli, T., Ridler, C.E., Clayton, E.L., Devoy, A., Moens, T., Norona, F.E., Woollacott, I.O., Pietrzyk, J., *et al.* (2014) C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science*, **345**, 1192–1194.
42. Zhang, Y.J., Jansen-West, K., Xu, Y.F., Gendron, T.F., Bieniek, K.F., Lin, W.L., Sasaguri, H., Caulfield, T., Hubbard, J., Daughrity, L., *et al.* (2014) Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol.*, **128**, 505–524.
43. Tran, H., Almeida, S., Moore, J., Gendron, T.F., Chalasani, U., Lu, Y., Du, X., Nickerson, J.A., Petrucelli, L., Weng, Z., *et al.* (2015) Differential toxicity of nuclear RNA foci versus dipeptide repeat proteins in a drosophila model of C9ORF72 FTD/ALS. *Neuron*, **87**, 1207–1214.
44. Cooper-Knock, J., Higginbottom, A., Stopford, M.J., Highley, J.R., Ince, P.G., Wharton, S.B., Pickering-Brown, S., Kirby, J., Hautbergue, G.M. and Shaw, P.J. (2015) Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol.*, **130**, 63–75.
45. Davidson, Y., Robinson, A.C., Liu, X., Wu, D., Troakes, C., Rollinson, S., Masuda-Suzukake, M., Suzuki, G., Nonaka, T., Shi, J., *et al.* (2015) Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9orf72 is linked to TDP-43 pathology and not associated with aggregated forms of dipeptide repeat proteins. *Neuropathol. Appl. Neurobiol.*, **3**, 242–254.
46. Gomez-Deza, J., Lee, Y.B., Troakes, C., Nolan, M., Al-Sarraj, S., Gallo, J.M. and Shaw, C.E. (2015) Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. *Acta Neuropathol. Commun.*, **3**, 38.
47. Mackenzie, I.R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., Weng, S.M., Haass, C., Kretzschmar, H.A., Edbauer, D., *et al.* (2013) Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathol.*, **126**, 859–879.
48. Mackenzie, I.R., Frick, P., Grässer, F.A., Gendron, T.F., Petrucelli, L., Cashman, N.R., Edbauer, D., Kremmer, E., Prudlo, J., Troost, D., *et al.* (2015) Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta Neuropathol*, **130**, 845–861.
49. Chew, J., Gendron, T.F., Prudencio, M., Sasaguri, H., Zhang, Y.J., Castanedes-Casey, M., Lee, C.W., Jansen-West, K., Kurti, A., Murray, M.E., *et al.* (2015) Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science*, **348**, 1151–1154.
50. O'Rourke, J.G., Bogdanik, L., Muhammad, A.K., Gendron, T.F., Kim, K.J., Austin, A., Cady, J., Liu, E.Y., Zarrow, J., Grant, S., *et al.* (2015) C9orf72 BAC transgenic mice display typical pathologic features of ALS/FTD. *Neuron*, **88**, 892–901.
51. Peters, O.M., Cabrera, G.T., Tran, H., Gendron, T.F., McKeon, J.E., Metterville, J., Weiss, A., Wightman, N., Salameh, J., Kim, J., *et al.* (2015) Human C9ORF72 hexanucleotide expansion reproduces RNA foci and dipeptide repeat proteins but not neurodegeneration in BAC transgenic mice. *Neuron*, **88**, 902–909.
52. Jiang, J., Zhu, Q., Gendron, T.F., Saberi, S., McAlonis-Downes, M., Seelman, A., Stauffer, J.E., Jafar-Nejad, P., Drenner, K., Schulte, D., *et al.* (2016) Gain of toxicity from ALS/FTD-linked repeat expansions in C9ORF72 is alleviated by antisense oligonucleotides targeting GGGGCC-containing RNAs. *Neuron*, **90**, 535–550.
53. Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O., Borchelt, D.R., Yachnis, A.T. and Ranum, L.P. (2016) C9orf72 BAC mouse model with motor deficits and neurodegenerative features of ALS/FTD. *Neuron*, **90**, 521–534.
54. Kirby, J., Ning, K., Ferraiuolo, L., Heath, P.R., Ismail, A., Kuo, S.W., Valori, C.F., Cox, L., Sharrack, B., Wharton, S.B., *et al.* (2011) Phosphatase and tensin homologue/protein kinase B pathway linked to motor neuron survival in human superoxide dismutase 1-related amyotrophic lateral sclerosis. *Brain*, **134**, 506–517.
55. Nalls, M.A., Bras, J., Hernandez, D.G., Keller, M.F., Majounie, E., Renton, A.E., Saad, M., Jansen, I., Guerreiro, R., Lubbe, S., *et al.* (2015) NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. *Neurobiol. Aging*, **36**, 1605.e1607–e1612.

56. Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G. and Kasprzyk, A. (2009) BioMart – biological queries made easy. *BMC Genomics*, **10**, 22.
57. Milella, M., Falcone, I., Conciatori, F., Cesta Incani, U., Del Curatolo, A., Inzerilli, N., Nuzzo, C.M., Vaccaro, V., Vari, S., Cognetti, F., *et al.* (2015) PTEN: multiple functions in human malignant tumors. *Front. Oncol.*, **5**, 24.
58. Ning, K., Drepper, C., Valori, C.F., Ahsan, M., Wyles, M., Higginbottom, A., Herrmann, T., Shaw, P., Azzouz, M. and Sendtner, M. (2010) PTEN depletion rescues axonal growth defect and improves survival in SMN-deficient motor neurons. *Hum. Mol. Genet.*, **19**, 3159–3168.
59. Little, D., Valori, C.F., Mutsaers, C.A., Bennett, E.J., Wyles, M., Sharrack, B., Shaw, P.J., Gillingwater, T.H., Azzouz, M. and Ning, K. (2015) PTEN depletion decreases disease severity and modestly prolongs survival in a mouse model of spinal muscular atrophy. *Mol. Ther.*, **23**, 270–277.
60. Breitling, R., Armengaud, P., Amtmann, A. and Herzyk, P. (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett.*, **573**, 83–92.
61. Pearson, R.D., Liu, X., Sanguinetti, G., Milo, M., Lawrence, N.D. and Rattray, M. (2009) puma: a Bioconductor package for propagating uncertainty in microarray analysis. *BMC Bioinformatics*, **10**, 211.
62. Rattray, M., Liu, X., Sanguinetti, G., Milo, M. and Lawrence, N.D. (2006) Propagating uncertainty in microarray data analysis. *Brief Bioinform.*, **7**, 37–47.
63. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
64. Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N. and Golani, I. (2001) Controlling the false discovery rate in behavior genetics research. *Behav. Brain Res.*, **125**, 279–284.

Supplementary Materials

Members of the NeuroX consortium are:

Robert Bowser (Division of Neurology, Barrow Neurological Institute), Maura Brunetti (Molecular Genetics Unit, Department of Clinical Pathology, Azienda Sanitaria Ospedaliera Ospedale Infantile Regina Margherita-Santa Anna, Turin), Luigi Ferrucci (Longitudinal Studies Section, Clinical Research Branch, National Institute on Aging, National Institutes of Health, Baltimore) Pietro Fratta (Department of Neurodegenerative Disease, University College London, London), John Hardy (Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, University College London, London), Hannu Laaksovirta (Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki), Francesco Landi (Department of Gerontology, Geriatrics, and Rehabilitative Medicine, Catholic University of Sacred Heart, Rome), Nicholas Maragakis (Brain Science Institute, Department of Neurology, Johns Hopkins University, Baltimore), Michael Nalls (Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda), Richard Orrell (Department of Clinical Neuroscience, Institute of Neurology, University College London, London), Lyle Ostrow (Brain Science Institute, Department of Neurology, Johns Hopkins University, Baltimore), Gabriella Restagno (Molecular Genetics Unit, Department of Clinical Pathology, Azienda Sanitaria Ospedaliera Ospedale Infantile Regina Margherita-Santa Anna, Turin), Ekaterina Rogaeva (Tanz Centre for Research of Neurodegenerative Diseases, Division of Neurology, Department of Medicine, University of Toronto, Toronto), Jeffrey Rothstein (Brain Science Institute, Department of Neurology, Johns Hopkins University, Baltimore), Andrew Singleton (Molecular Genetics Section, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda) Pentti Tienari (Department of Neurology, Helsinki University Central Hospital and Molecular Neurology Programme, Biomedicum, University of Helsinki, Helsinki), Juan Troncoso (Department of Neuropathology, Johns Hopkins University Baltimore), Lorne Zinman (Division of Neurology, Department of Internal Medicine, Sunnybrook Health Sciences Centre, University of Toronto, Toronto), Rita Guerreiro (Department of Molecular Neuroscience, UCL Institute of Neurology, London), Jose Bras (Department of Molecular Neuroscience, UCL Institute of Neurology, London), John Powell (King's College London, Institute of Psychiatry, Department of Neuroscience, London), Michelle K. Lupton (King's College London, Institute of Psychiatry, Department of Neuroscience, London and QIMR Berghofer Medical Research Institute, Genetic Epidemiology, Brisbane), Katie Lunnon (University of Exeter Medical School, RILD, Exeter), Olaf Ansorge (Department of Neuropathology, Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital), Laura Parkkinen (Oxford Parkinson's Disease Centre, Nuffield Department of Clinical Neurosciences, University of Oxford).

Members of the ITALSGEN consortium are:

Francesco O. Logullo (Ancona); Isabella Simone (Bari); Giancarlo Logroscino (Bari and Tricase, LE); Fabrizio Salvi, Ilaria Bartolomei (Bologna); Giuseppe Borghero, Maria Rita Murru, Emanuela Costantino, Carla Pani, Roberta Puddu, Carla Caredda, Valeria Piras, Stefania Tranquilli, Stefania Cuccu, Daniela Corongiu, Maurizio Melis, Antonio Milia, Francesco Marrosu, Maria Giovanna Marrosu, Gianluca Floris, Antonino Cannas, Stefania

Cuccu, Stefania Tranquilli (Cagliari); Margherita Capasso (Chieti); Claudia Caponnetto, Gianluigi Mancardi, Paola Origone, Paola Mandich (Genova); Francesca L. Conforti (Mangone, CS); Gabriele Mora, Kalliopi Marinou, Riccardo Sideri (Milan, IRCCS Maugeri Foundation); Silvana Penco, Lorena Mosca (Milan, Niguarda Ca' Granda Hospital); Christian Lunetta (NeuroMuscular OmniCenter, NEMO, Milan); Giuseppe Lauria Pinter (Milan, Besta Neurological Institute), Massimo Corbo (Milan, Casa di Cura del Policlinico); Nilo Riva, Paola Carrera (Milan, IRCCS San Raffaele Scientific Institute); Paolo Volanti (Mistretta, ME); Jessica Mandrioli, Nicola Fini, Antonio Fasano (Modena); Lucio Tremolizzo, Alessandro Arosio, Carlo Ferrarese (Monza); Francesca Trojsi, Gioacchino Tedeschi, Maria Rosaria Monsurrò, Giovanni Piccirillo, Cinzia Femiano (Napoli); Anna Ticca (Nuoro); Enzo Ortu (Ozieri); Vincenzo La Bella, Rossella Spataro, Tiziana Colletti (Palermo); Mario Sabatelli, Marcella Zollino, Amelia Conte, Marco Luigetti, Serena Lattante, Giuseppe Marangi (Rome, Catholic University of Sacred Heart); Marialuisa Santarelli (Rome, San Filippo Neri Hospital); Antonio Petrucci (Rome, San Camillo Forlanini Hospital); Maura Pugliatti, Angelo Pirisi, Leslie D. Parish, Patrizia Occhineri (Sassari), Fabio Giannini, **Stefania Battistini**, Claudia Ricci, Michele Benigni (Siena); Tea B. Cau, Daniela Loi (Tempio-Olbia); Andrea Calvo, Cristina Moglia, Maura Brunetti, Marco Barberis, Gabriella Restagno, Federico Casale, Giuseppe Marrali, Giuseppe Fuda, Irene Ossola, Stefania Cammarosano, Antonio Canosa, Antonio Ilardi, Umberto Manera, Davide Bertuzzo (Torino), Raffaella Tanel (Trento); Fabrizio Pisano (Veruno, NO).

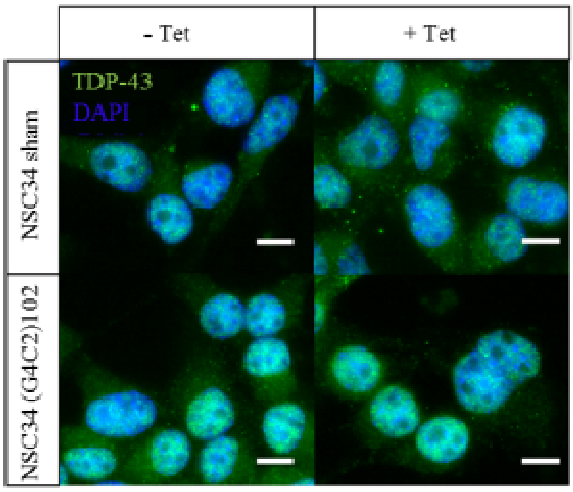

Fig. S1. NSC34 (G4C2)102 cells do not show TDP-43 mislocalisation. NSC34 sham and (G4C2)102 cells were cultured with or without 0.5 µg/mL tetracycline for 9 days. Cells were stained for TDP-43 (Green) and Dapi (Blue). Scale bar = 10 µm.

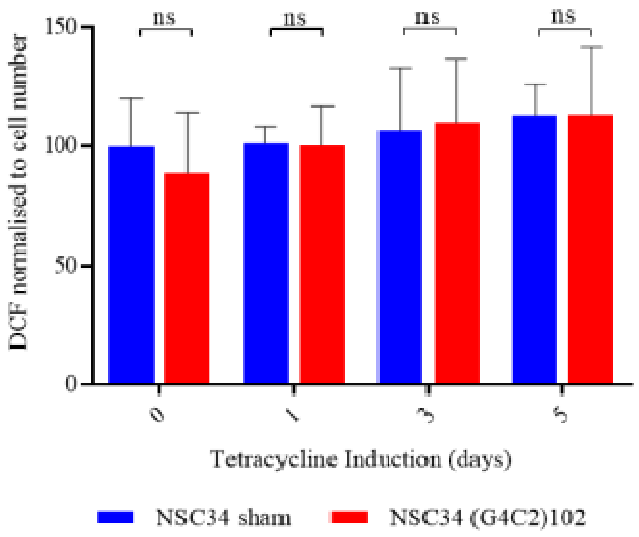

Fig. S2. NSC34 (G4C2)102 cells do not show signs of oxidative stress. NSC34 sham and (G4C2)102 cells were cultured for 5 days, and were induced for various lengths of time with 0.5 µg/mL tetracycline. The level of reactive oxygen species (ROS) in the NSC34 cells was measured using the DCF assay, and normalised to cell number (measured after cell lysis using EthD1 fluorescence assay). (Two-way ANOVA with Tukey's post hoc test; data shown are mean and SD; n = 3).

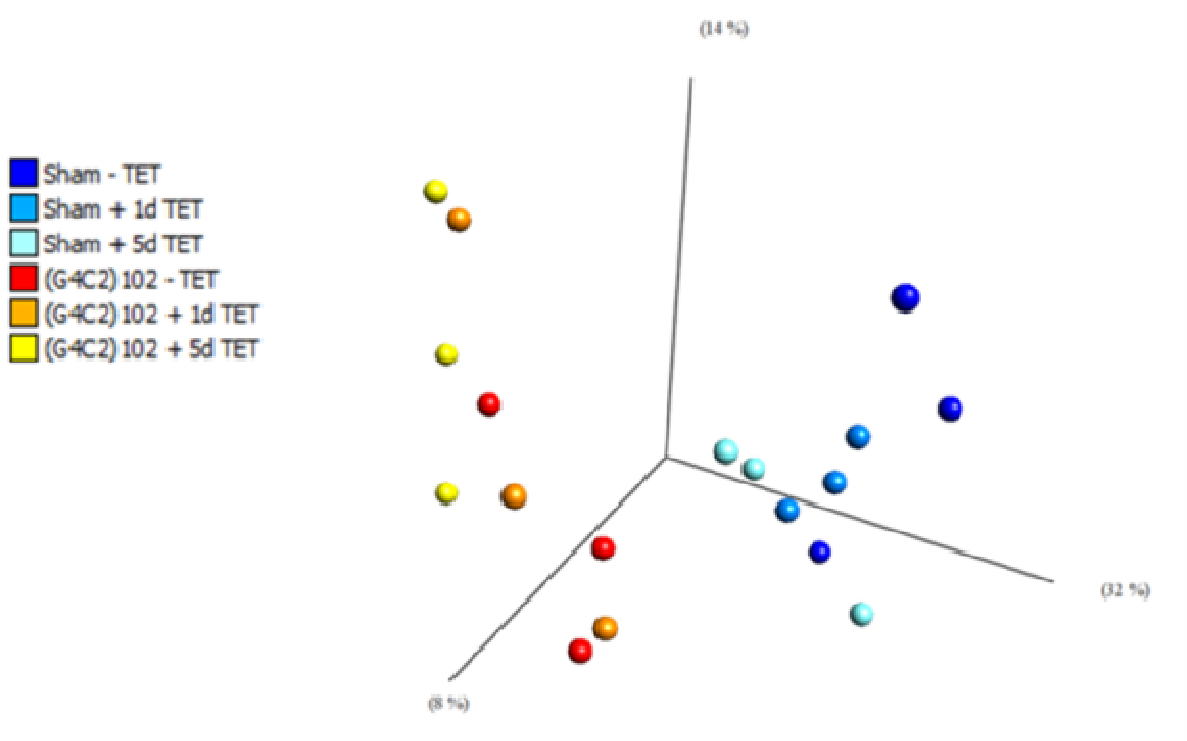

Fig. S3. PCA analysis of NSC34 sham and NSC34 (G4C2)102 cells without statistical manipulation.

Transcript ID	P-value	Fold Change	Log2 Fold Change	Gene Symbol
228006_at	1.35E-05	11.25456	3.49243775	PTEN
201125_s_at	0.001193	5.584296	2.481375415	ITGB5
212777_at	0.019792	3.356961	1.747155776	SOS1
209341_s_at	0.116312	3.168838	1.663953907	IKBKB
202686_s_at	0.012691	3.109265	1.636573582	AXL
201124_at	0.023576	2.868891	1.520493156	ITGB5
213093_at	0.00332	2.802556	1.486743201	PRKCA
226731_at	0.0494	2.734129	1.451081313	ITGA1
203685_at	0.064852	2.704587	1.435408306	BCL2
202723_s_at	0.027603	2.606658	1.38220131	FOXO1
203809_s_at	0.078663	2.553882	1.352691868	AKT2
210482_x_at	0.143572	2.434591	1.283679427	MAP2K5
1555612_s_at	0.06008	2.428381	1.27999479	G6PC
1552610_a_at	0.189486	2.334724	1.223252011	JAK1
217399_s_at	0.164479	2.210499	1.144372082	FOXO3
202887_s_at	0.000448	2.168607	1.116768628	DDIT4
223196_s_at	0.090703	2.102367	1.072014536	SESN2
208536_s_at	0.336397	2.082498	1.05831511	BCL2L11
203379_at	0.143807	2.082037	1.057995707	RPS6KA1
226068_at	0.077805	2.037049	1.026480684	SYK
202284_s_at	0.164219	2.018647	1.013388649	CDKN1A
225691_at	0.079002	2.010563	1.007599543	CDK12
236664_at	0.277998	2.005314	1.003828157	AKT2
201834_at	0.326416	2.005067	1.003650446	PRKAB1
202830_s_at	0.148225	2.001074	1.000774519	SLC37A4
212590_at	0.152239	1.874144	0.906231807	RRAS2
232068_s_at	0.204519	1.867146	0.900834742	TLR4
222999_s_at	0.045031	1.829493	0.871443896	CCNL2
201739_at	0.016888	1.828987	0.871044821	SGK1
202340_x_at	0.348162	1.82347	0.866686464	NR4A1
202530_at	0.429409	1.769648	0.823462423	MAPK14
209666_s_at	0.283035	1.739357	0.798554074	CHUK
226441_at	0.095141	1.715889	0.778956229	MAP3K2
204054_at	0.392798	1.693505	0.760012246	PTEN
202431_s_at	0.401722	1.67664	0.745572954	MYC
202847_at	0.372926	1.671302	0.740972448	PCK2
221060_s_at	0.50849	1.666396	0.736731281	TLR4
236459_at	0.245049	1.629361	0.704306282	PRKCE
224889_at	0.512673	1.627228	0.702416409	FOXO3
209342_s_at	0.381914	1.612545	0.689339421	IKBKB
239201_at	0.227475	1.597199	0.675544074	CDK15
1558143_a_at	0.05218	1.576673	0.656883478	BCL2L11
204297_at	0.245594	1.573533	0.654007435	PIK3C3
202853_s_at	0.006202	1.559587	0.641164034	RYK

210449_x_at	0.501364	1.558745	0.640384932	MAPK14
212589_at	0.333773	1.554202	0.636174023	RRAS2
212628_at	0.240387	1.535592	0.618794949	PKN2
208712_at	0.329863	1.512912	0.597328074	CCND1
209364_at	0.437234	1.493794	0.578981209	BAD
206398_s_at	0.490597	1.488985	0.57432922	CD19
202449_s_at	0.366843	1.442837	0.528908325	RXRA
202426_s_at	0.556319	1.440839	0.526909137	RXRA
241453_at	0.417711	1.431025	0.517048876	PTK2
207540_s_at	0.466116	1.425089	0.511052022	SYK
201234_at	0.506568	1.42441	0.510364469	ILK
241387_at	0.440439	1.417247	0.503091215	PTK2
225066_at	0.269049	1.401803	0.487283617	PPP2R2D
241722_x_at	0.557553	1.38925	0.47430624	MCL1
212332_at	0.25134	1.352265	0.435377901	RBL2
208824_x_at	0.057664	1.346003	0.428681625	CDK16
208823_s_at	0.337467	1.336739	0.418717805	CDK16
200797_s_at	0.166945	1.333536	0.415256772	MCL1
227073_at	0.507719	1.317354	0.397643079	MAP3K2
225690_at	0.513019	1.306153	0.385323901	CDK12
206854_s_at	0.567931	1.285591	0.362431734	MAP3K7
226310_at	0.425082	1.283579	0.360172092	RICTOR
211333_s_at	0.610405	1.280122	0.356281311	FASLG
206952_at	0.527055	1.279746	0.355857497	G6PC
206853_s_at	0.338928	1.268168	0.342745879	MAP3K7
208820_at	0.335075	1.255543	0.328311439	PTK2
202161_at	0.679044	1.23813	0.308162801	PKN1
209184_s_at	0.246192	1.236105	0.305801297	IRS2
202210_x_at	0.596252	1.235289	0.304848605	GSK3A
219226_at	0.714163	1.230401	0.299128581	CDK12
211087_x_at	0.610452	1.226702	0.294784821	MAPK14
211561_x_at	0.609939	1.21891	0.285591606	MAPK14
202724_s_at	0.764604	1.218777	0.28543418	FOXO1
212312_at	0.705952	1.208077	0.272712412	BCL2L1
215195_at	0.710539	1.20301	0.266648635	PRKCA
201984_s_at	0.428889	1.201119	0.264379092	EGFR
225606_at	0.705461	1.199045	0.261885804	BCL2L11
204906_at	0.331858	1.197615	0.260164196	RPS6KA2
229705_at	0.032404	1.189092	0.249860341	PIK3C3
205271_s_at	0.806096	1.187002	0.247322366	CDK20
207163_s_at	0.426152	1.186063	0.246180643	AKT1
1569272_at	0.674759	1.179942	0.238715946	PIK3C3
235254_at	0.753921	1.169931	0.226423445	MAP3K2
223195_s_at	0.710065	1.155934	0.209059027	SESN2
214265_at	0.557851	1.13662	0.184750006	ITGA8
204924_at	0.772744	1.132702	0.179768356	TLR2

228177_at	0.828615	1.12985	0.176131252	CREBBP
204131_s_at	0.513469	1.12735	0.172935488	FOXO3
1556655_s_at	0.868312	1.124945	0.169854468	CDK12
231017_at	0.600747	1.122861	0.167179346	STK11
228248_at	0.756222	1.117075	0.159726051	RICTOR
202160_at	0.393015	1.115457	0.1576349	CREBBP
238733_at	0.789257	1.10799	0.147944861	MDM2
226979_at	0.575744	1.102505	0.140785199	MAP3K2
203984_s_at	0.783146	1.094592	0.130393218	CASP9
205386_s_at	0.873715	1.092642	0.127820785	MDM2
208641_s_at	0.758982	1.087062	0.120434226	RAC1
240964_at	0.878754	1.086979	0.120324068	PTEN
226156_at	0.867426	1.085308	0.118104524	AKT2
225471_s_at	0.86043	1.083219	0.11532495	AKT2
203836_s_at	0.849016	1.078407	0.108901766	MAP3K5
237451_x_at	0.718962	1.076788	0.106734237	CASP9
220587_s_at	0.822453	1.07642	0.106241102	MLST8
202685_s_at	0.908978	1.066797	0.093285673	AXL
211372_s_at	0.883678	1.062037	0.086834029	IL1R2
227627_at	0.876863	1.04972	0.070004558	SGK3
225160_x_at	0.802144	1.048663	0.068551126	MDM2
244616_x_at	0.909744	1.04365	0.061637969	MDM2
227426_at	0.870694	1.040067	0.056676468	SOS1
211832_s_at	0.912271	1.039894	0.056436477	MDM2
203837_at	0.792837	1.036344	0.051502965	MAP3K5
217492_s_at	0.924599	1.035607	0.050476622	PTEN
1557970_s_at	0.9432	1.033772	0.047918032	RPS6KA2
212629_s_at	0.935661	1.032482	0.04611663	PKN2
212719_at	0.886709	1.020152	0.028784126	PHLPP1
209185_s_at	0.973334	1.00973	0.01396957	IRS2
226299_at	0.985456	1.008872	0.012743145	PKN3
202670_at	0.988163	1.007332	0.01053925	MAP2K1
229711_s_at	0.968365	1.006463	0.009294136	MDM2
225694_at	0.994284	1.004493	0.00646751	CDK12
632_at	0.9922	1.003997	0.005754958	GSK3A
224891_at	0.994465	1.001816	0.002617558	FOXO3
208640_at	0.996	1.001424	0.002052936	RAC1
207239_s_at	0.999889	1.000062	8.94443E-05	CDK16
213012_at	0.959725	0.981467	-0.026988334	NEDD4
223049_at	0.907224	0.980561	-0.028320713	GRB2
204756_at	0.936182	0.974529	-0.037222977	MAP2K5
215394_at	0.955286	0.970561	-0.043109205	PIK3C3
224999_at	0.937243	0.968813	-0.045709871	EGFR
205798_at	0.938054	0.964411	-0.052279989	IL7R
212331_at	0.890071	0.96321	-0.054077725	RBL2
232876_at	0.931585	0.959169	-0.060143063	MAPK14

201983_s_at	0.924495	0.95877	-0.060743327	EGFR
1555804_a_at	0.894588	0.950584	-0.073113976	MAP3K19
205498_at	0.898394	0.949704	-0.074450165	GHR
200980_s_at	0.803407	0.946482	-0.079353026	PDHA1
226218_at	0.908871	0.945496	-0.08085674	IL7R
211808_s_at	0.891156	0.940032	-0.089218226	CREBBP
202288_at	0.894346	0.931247	-0.102764222	MTOR
200798_x_at	0.772837	0.928812	-0.106541483	MCL1
209112_at	0.594059	0.921765	-0.117529106	CDKN1B
214328_s_at	0.75804	0.915552	-0.127286267	HSP90AA1
208743_s_at	0.855563	0.911983	-0.132921163	YWHAB
201389_at	0.856486	0.910729	-0.134906271	ITGA5
208456_s_at	0.918453	0.908647	-0.138208164	RRAS2
212912_at	0.705757	0.906253	-0.142014229	RPS6KA2
204247_s_at	0.783234	0.902163	-0.148539976	CDK5
217718_s_at	0.693721	0.89881	-0.153911919	YWHAB
210655_s_at	0.789241	0.897205	-0.156490434	FOXO3
204369_at	0.845732	0.895233	-0.159664877	PIK3CA
201452_at	0.760296	0.89119	-0.166195051	RHEB
226046_at	0.835819	0.883273	-0.179068683	MAPK8
239300_at	0.363481	0.883176	-0.179227127	PIK3C3
210775_x_at	0.847061	0.881818	-0.181447169	CASP9
225363_at	0.559081	0.877415	-0.188668724	PTEN
218852_at	0.815605	0.876718	-0.189815226	PPP2R3C
44654_at	0.501269	0.876607	-0.189997896	G6PC3
210969_at	0.692941	0.87551	-0.191804438	PKN2
209390_at	0.55561	0.873355	-0.195359897	TSC1
210211_s_at	0.74048	0.869027	-0.202527094	HSP90AA1
1567458_s_at	0.433362	0.866275	-0.207103012	RAC1
202647_s_at	0.659401	0.863871	-0.211112201	NRAS
1560074_at	0.59213	0.859693	-0.218106536	PRKCA
243492_at	0.763459	0.852679	-0.229925369	THEM4
214172_x_at	0.313143	0.850473	-0.23366266	RYK
201437_s_at	0.43703	0.846859	-0.239806311	EIF4E
1552611_a_at	0.443595	0.846177	-0.240968623	JAK1
224985_at	0.509339	0.846049	-0.241186874	NRAS
229664_at	0.386319	0.839849	-0.251798132	MAPK8
215037_s_at	0.609737	0.836868	-0.256928012	BCL2L1
217542_at	0.617322	0.831476	-0.266253473	MDM2
204798_at	0.727843	0.826783	-0.27441937	MYB
211814_s_at	0.774846	0.826533	-0.274855673	CCNE2
211550_at	0.625894	0.821579	-0.283528789	EGFR
1555864_s_at	0.597819	0.820961	-0.284614407	PDHA1
1552559_a_at	0.646832	0.81826	-0.289368766	CDK15
1553118_at	0.413062	0.816151	-0.293091998	THEM4
226312_at	0.452194	0.813619	-0.297574725	RICTOR

211938_at	0.290751	0.803602	-0.31544694	EIF4B
211937_at	0.356438	0.80297	-0.316582007	EIF4B
204053_x_at	0.560409	0.80144	-0.319333578	PTEN
201835_s_at	0.355711	0.796135	-0.328915007	PRKAB1
217557_s_at	0.631736	0.794321	-0.33220595	MDM2
200979_at	0.285155	0.792685	-0.335180418	PDHA1
235980_at	0.266705	0.792475	-0.335562671	PIK3CA
221759_at	0.518122	0.791827	-0.336742833	G6PC3
214660_at	0.677173	0.775384	-0.367017129	ITGA1
211536_x_at	0.63566	0.768882	-0.37916589	MAP3K7
212688_at	0.216451	0.764708	-0.387019128	PIK3CB
209269_s_at	0.673538	0.762087	-0.391972389	SYK
215152_at	0.580627	0.755046	-0.405363554	MYB
217717_s_at	0.373991	0.753917	-0.407522391	YWHAB
217373_x_at	0.477403	0.752907	-0.409456422	MDM2
210865_at	0.377927	0.74328	-0.428022306	FASLG
1861_at	0.551345	0.73509	-0.444007199	BAD
201453_x_at	0.337992	0.734954	-0.444274139	RHEB
232274_at	0.329023	0.73347	-0.447190136	CCNL2
201648_at	0.248252	0.728557	-0.456886246	JAK1
211027_s_at	0.518768	0.726536	-0.460893809	IKBKB
217289_s_at	0.484965	0.726016	-0.461926752	SLC37A4
200796_s_at	0.597017	0.725817	-0.462322247	MCL1
211711_s_at	0.077039	0.721442	-0.47104468	PTEN
207005_s_at	0.486609	0.720751	-0.472427162	BCL2
221772_s_at	0.375017	0.716974	-0.480007292	PPP2R2D
205403_at	0.491429	0.716384	-0.481194979	IL1R2
220357_s_at	0.604314	0.708024	-0.49812983	SGK2
204531_s_at	0.434037	0.702258	-0.509926941	BRCA1
1552734_at	0.421191	0.702117	-0.510216635	RICTOR
216976_s_at	0.534127	0.701251	-0.511997172	RYK
230573_at	0.346966	0.696043	-0.52275166	SGK2
209953_s_at	0.289378	0.691039	-0.533160961	CDC37
206248_at	0.262083	0.68902	-0.537382235	PRKCE
237891_at	0.466151	0.677031	-0.562706201	MDM2
221695_s_at	0.600041	0.673253	-0.570779342	MAP3K2
212780_at	0.025322	0.669516	-0.578809562	SOS1
1565483_at	0.422473	0.662606	-0.593776828	EGFR
227633_at	0.40547	0.662048	-0.594992275	RHEB
242674_at	0.173354	0.656777	-0.606524489	EIF4E
231228_at	0.339066	0.649862	-0.621794705	BCL2L1
206665_s_at	0.435532	0.648406	-0.625030654	BCL2L1
213404_s_at	0.144982	0.646752	-0.628715485	RHEB
217620_s_at	0.336637	0.646357	-0.629596871	PIK3CB
231854_at	0.461979	0.64623	-0.629880368	PIK3CA
226101_at	0.163161	0.646057	-0.630266639	PRKCE

214621_at	0.372588	0.641953	-0.639460419	GYS2
215075_s_at	0.141904	0.634678	-0.65590326	GRB2
201436_at	0.135331	0.634276	-0.656817341	EIF4E
208711_s_at	0.455611	0.633879	-0.657720622	CCND1
211370_s_at	0.484713	0.632944	-0.659850233	MAP2K5
41657_at	0.343057	0.630361	-0.665749816	STK11
225715_at	0.252504	0.628748	-0.669446189	RPTOR
222343_at	0.362422	0.628394	-0.670258689	BCL2L11
232086_at	0.386168	0.626462	-0.674701094	PIK3C3
221427_s_at	0.04162	0.617323	-0.695902551	CCNL2
224341_x_at	0.361249	0.616781	-0.697169771	TLR4
220038_at	0.391962	0.613958	-0.703788129	SGK3
211851_x_at	0.461197	0.611764	-0.708952883	BRCA1
1560359_at	0.491519	0.607575	-0.718865587	ITGA1
1553096_s_at	0.054393	0.605402	-0.724034654	BCL2L11
221180_at	0.288631	0.59205	-0.756209075	MAP3K19
1553088_a_at	0.108437	0.591752	-0.756935418	BCL2L11
211968_s_at	0.252139	0.59163	-0.757232885	HSP90AA1
239188_at	0.346612	0.590814	-0.759224082	PPP2R3C
229253_at	0.073477	0.587559	-0.767194367	THEM4
203684_s_at	0.405007	0.584138	-0.775618855	BCL2
242071_x_at	0.254398	0.578616	-0.789321877	ITGA8
206341_at	0.334766	0.578428	-0.789790704	IL2RA
211537_x_at	0.392671	0.567562	-0.817150095	MAP3K7
215735_s_at	0.118292	0.564778	-0.824244203	TSC2
211969_at	0.289239	0.556204	-0.846313975	HSP90AA1
226048_at	0.094545	0.544801	-0.876198743	MAPK8
225697_at	0.061816	0.53981	-0.889476392	CDK12
1565484_x_at	0.360314	0.53701	-0.896979141	EGFR
1556654_at	0.163266	0.534599	-0.903470956	CDK12
207821_s_at	0.339514	0.532389	-0.909447332	PTK2
210226_at	0.258313	0.531457	-0.911975126	NR4A1
201020_at	0.081573	0.525793	-0.92743316	YWHAH
209239_at	0.119177	0.500991	-0.997143408	NFKB1
235011_at	0.263437	0.491702	-1.024143872	MAP3K2
211607_x_at	0.118234	0.48165	-1.053942929	EGFR
237718_at	0.166519	0.480893	-1.056212169	EIF4E
206923_at	0.080503	0.480587	-1.057130471	PRKCA
1552798_a_at	0.177667	0.473998	-1.077047123	TLR4
239092_at	0.159601	0.470743	-1.086988453	ITGA8
240437_at	0.10184	0.457427	-1.12838657	CASP9
1560689_s_at	0.162751	0.451964	-1.145720232	AKT2
210984_x_at	0.196761	0.439456	-1.18620937	EGFR
205034_at	0.015597	0.422482	-1.243038219	CCNE2
1554826_at	0.232229	0.416916	-1.262171355	CDK15
211453_s_at	0.032224	0.414763	-1.269640894	AKT2

201435_s_at	0.089872	0.410499	-1.284549387	EIF4E
211269_s_at	0.173879	0.403133	-1.31067221	IL2RA
1555780_a_at	0.236837	0.39096	-1.354907085	RHEB
211297_s_at	0.031508	0.383244	-1.383664889	CDK7
204132_s_at	0.02473	0.360665	-1.471268668	FOXO3
210671_x_at	0.043001	0.357881	-1.482448142	MAPK8
235666_at	0.063203	0.308998	-1.694330595	ITGA8
210477_x_at	0.029171	0.288981	-1.790953454	MAPK8
204292_x_at	0.007038	0.222178	-2.170212126	STK11

Table S1. Differentially expressed transcripts within the human PI3K/Akt signalling pathway (KEGG) measured on the Human Genome microarray platform in the Laser Captured Microdissected motor neurons. Transcript ID, Fold change and p-value are included for a comparison between C9ORF72-ALS patients and control LCM MNs.

Transcript ID	P-value	Fold Change	Log2 Fold Change	Gene Symbol
ENSMUST00000061673	2.18E-08	1.41441	0.50020038	ITGA1
ENSMUST00000146023	5.21E-08	1.33624	0.418179151	BCL2L11
ENSMUST00000013807	4.82E-03	1.19164	0.252948457	PTEN
ENSMUST00000135810	9.37E-05	1.18084	0.239813497	PHLPP1
ENSMUST00000061512	2.32E-03	1.13641	0.184483432	MAP3K19
ENSMUST00000017290	3.04E-02	1.13287	0.179982317	BRCA1
ENSMUST00000160662	4.83E-04	1.13155	0.178300335	CDK15
ENSMUST00000027243	2.31E-03	1.12736	0.172948285	IL1R2
ENSMUST00000033500	1.25E-02	1.12504	0.169976296	ERAS
ENSMUST00000019469	1.37E-02	1.11416	0.155956427	G6PC
ENSMUST00000112751	2.62E-04	1.10624	0.145664414	BCL2
ENSMUST00000018012	1.10E-03	1.1019	0.139993302	SGK2
ENSMUST00000019615	4.00E-02	1.09648	0.132879497	CDC37
ENSMUST00000120135	5.31E-03	1.08072	0.111992789	SYK
ENSMUST00000048096	9.17E-02	1.07976	0.110710678	TLR4
ENSMUST00000125346	1.38E-02	1.0793	0.110095929	PKN3
ENSMUST00000020308	4.67E-01	1.06268	0.08770723	DDIT4
ENSMUST00000129514	1.13E-01	1.04996	0.070334367	RXRA
ENSMUST00000144883	3.93E-01	1.02521	0.035919456	STK11
ENSMUST00000053764	3.96E-01	1.02437	0.034736907	FOXO1
ENSMUST00000166384	4.79E-01	1.01442	0.020655095	SGK3
ENSMUST00000106499	9.58E-01	1.00188	0.00270972	GRB2
ENSMUST00000094361	9.16E-01	0.99356	-0.009321002	HSP90AA1
ENSMUST00000023779	4.55E-01	0.98085	-0.027895571	NR4A1
ENSMUST00000156857	3.03E-01	0.97905	-0.030545555	TSC1
ENSMUST00000049822	9.89E-02	0.93663	-0.094448847	THEM4
ENSMUST00000023165	2.95E-01	0.93459	-0.097594494	CREBBP
ENSMUST00000001780	3.63E-02	0.91375	-0.130128594	AKT1
ENSMUST00000125583	5.31E-04	0.88349	-0.17871429	RPTOR
ENSMUST00000107539	2.63E-02	0.88127	-0.182344001	CDK12
ENSMUST00000108344	3.26E-05	0.87959	-0.185096893	AKT2
ENSMUST00000128660	3.52E-06	0.85127	-0.232311306	CASP9
ENSMUST00000115028	4.62E-06	0.82873	-0.271025946	ITGB5
ENSMUST00000114758	1.80E-04	0.81702	-0.2915567	MAPK14
ENSMUST00000031486	2.93E-09	0.80861	-0.30648405	PRKAB1
ENSMUST00000128157	9.07E-06	0.80835	-0.306948007	PPP2R2D
ENSMUST00000043812	1.66E-04	0.80476	-0.313369496	PKN2
ENSMUST00000115812	4.61E-07	0.80411	-0.314535223	PIK3C3
ENSMUST00000107739	6.30E-08	0.79307	-0.334479884	ITGA3
ENSMUST00000019109	3.23E-06	0.78673	-0.346059497	YWHAH
ENSMUST00000029803	6.06E-04	0.78327	-0.352418392	EIF4E
ENSMUST00000110036	2.59E-04	0.78045	-0.357621887	PTK2
ENSMUST00000030724	7.98E-07	0.78036	-0.357788265	SESN2
ENSMUST00000095806	1.54E-09	0.77172	-0.3738506	MAP3K5

ENSMUST00000005066	2.95E-09	0.75827	-0.399216449	MAP2K1
ENSMUST00000093962	1.15E-07	0.74789	-0.419102001	CCND1
ENSMUST00000108243	1.70E-05	0.74781	-0.419256331	PIK3CA
ENSMUST00000070334	8.98E-07	0.74624	-0.422288401	G6PC3
ENSMUST00000097275	2.89E-10	0.74093	-0.432590846	PRKCE
ENSMUST00000037607	7.29E-06	0.73346	-0.447209806	MAP3K7
ENSMUST00000018470	3.42E-05	0.73286	-0.448390472	YWHAB
ENSMUST00000071739	3.92E-06	0.72087	-0.472188984	GSK3A
ENSMUST00000037947	1.34E-06	0.71222	-0.489605146	MCL1
ENSMUST00000175883	9.14E-08	0.70524	-0.503813791	RYK
ENSMUST00000115364	1.06E-09	0.67775	-0.561174887	CDK16
ENSMUST00000111945	1.34E-08	0.66852	-0.580957373	MAPK8
ENSMUST00000034740	7.86E-06	0.64571	-0.631041725	NEDD4
ENSMUST00000145430	9.93E-07	0.63284	-0.660087304	CCNE2
ENSMUST00000028106	8.25E-07	0.61691	-0.696868063	ITGA8
ENSMUST00000119901	1.53E-07	0.54694	-0.870545519	CDKN1A
ENSMUST00000033662	6.12E-10	0.50029	-0.999163479	PDHA1

Table S2. Differentially expressed transcripts within the murine PI3K/Akt signalling pathway (KEGG) measured on the Mouse Transcriptome microarray platform in the NSC34 cells. Probe ID, Fold change and p-value are included for a comparison between NSC34 (G4C2)102 and NSC34 sham cells.

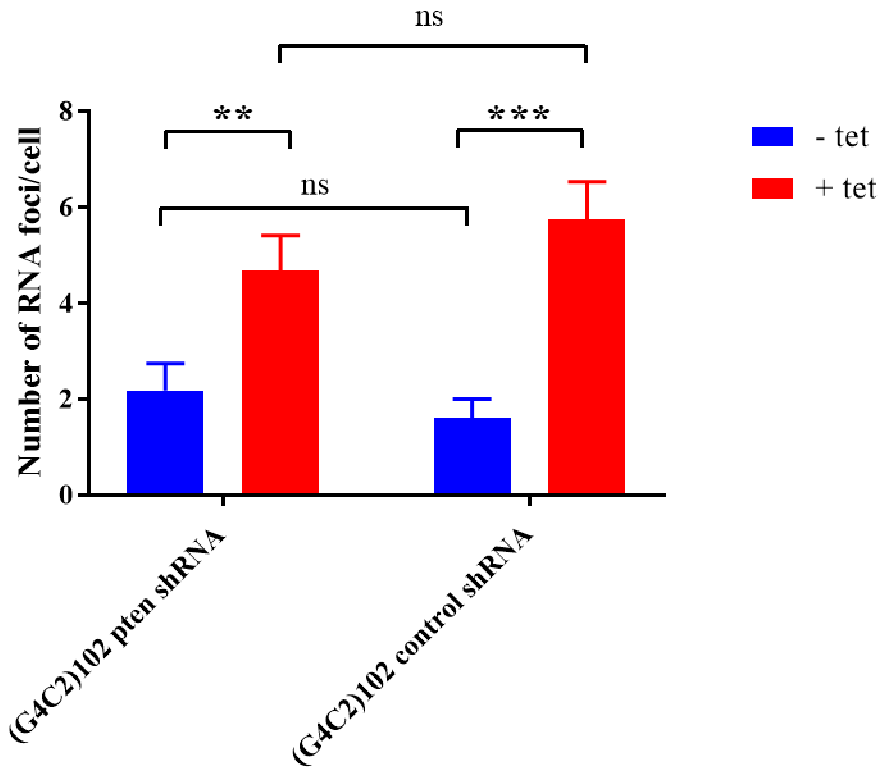

Fig. S4. Pten knockdown does not affect the number of (G4C2)n RNA foci. NSC34 (G4C2)102 control shRNA and NSC34 (G4C2)102 Pten shRNA cells were cultured for 3 days, with (or without) 0.5 µg/mL tetracycline. Cells were stained with a Locked Nucleic Acid (C4G2)3 sense probe and Dapi, and the average number of RNA foci per cell was counted (**P<0.01; ***P < 0.001; Two-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; n = 3).