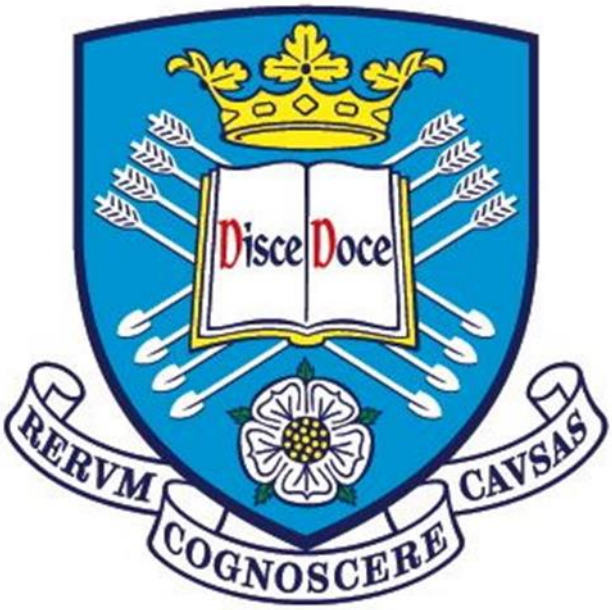
Development of a Cellular Model for C9ORF72-related Amyotrophic Lateral Sclerosis

By

Matthew John Stopford



Sheffield Institute for Translational Neuroscience

University of Sheffield

Submitted for the degree of Doctor of Philosophy (PhD)

May 2016

Acknowledgements

Firstly, I would like to acknowledge and thank my supervisors Dr. Janine Kirby, Prof. Dame Pamela Shaw, and Dr. Adrian Higginbottom for their continued patience and support throughout my PhD. I am incredibly grateful to Janine for providing both academic and personal support during my PhD. I am thankful to Pam for her guidance, but also for pushing me to develop personally and as a scientist. Also, I am very thankful to Adrian, who has been a great mentor, helping me to develop my technical skills, my ability to think critically, and has also been incredibly patient when I have found the PhD hardest. I would also like to thank the Sheffield Hospitals Charity for funding my PhD project.

There are several other individuals who have been incredibly helpful during my time in SITraN as well, who deserve acknowledgment. Dr. Guillaume Hautbergue has been optimistic and inspirational, and has taught me many techniques in the lab. Dr. Matthew Walsh has taught me several biochemical techniques, and has been incredibly supportive personally. Dr. Jonathan Cooper-Knock has been very helpful in explaining protocols, array analysis and generally supporting my work. Also, I would like to thank Dr. Paul Heath and Mrs. Catherine Gelsthorpe for their technical assistance and support with the transcriptomic work.

This PhD has genuinely been the greatest challenge of my life, and without the support of my parents and family, the Sims family, my friends and other PhD students, I would have failed. Never again.

Abstract

Background: ALS is an incurable late onset neurodegenerative disease that is characterised by progressive loss of motor neurons. A (G4C2)n repeat expansion in *C9ORF72* is the most common genetic cause of ALS, but it is unknown how the repeat causes pathogenesis, although a gain of toxic function is likely.

Aims and Objectives: 1) Generate stable, isogenic motor neuron-like NSC34 cellular models that have tetracycline-inducible (G4C2)n repeat expression. 2) Characterise the cell models for C9ORF72-ALS pathology and biochemical alterations. 3) Identify biological functions and pathways that may be transcriptionally dysregulated by (G4C2)n repeat expression early in C9ORF72-ALS pathogenesis. 4) Identify and interrogate potential therapeutic targets for C9ORF72-ALS.

Results: Stable, isogenic NSC34 cell models with tetracycline-inducible (G4C2)n expression were successfully generated. Sense RNA foci and RAN translation products were detected in the cell models. No antisense derived RNA foci or RAN translation products were detected. Expression of the (G4C2)102 caused subtle toxicity and recapitulated some aspects of C9ORF72-ALS pathology in the NSC34 (G4C2)102. The (G4C2)102 expression caused transcriptomic dysregulation in RNA metabolism, protein transport, the PI3K/Akt signalling pathway, and also caused splicing alterations. Transcriptomic dysregulation in the PI3K/Akt signalling pathway was also detected in LCM motor neurons from C9ORF72-ALS patients. Pten knock-down provided a rescue effect against the (G4C2)102 induced toxicity in the NSC34 (G4C2)102.

Conclusions: Stable, isogenic motor neuron-like cellular models that had tetracycline-inducible (G4C2)n expression were successfully generated, and allowed interrogation of the early biochemical effects associated with sense only (G4C2)n expression in a reductionist manner. Transcriptomic analysis of the NSC34 (G4C2)102 identified dysregulation in RNA splicing and the PI3K/Akt signalling pathway, which was corroborated by transcriptomic data from C9ORF72-ALS patient CNS tissue. This suggests dysregulation in these biological functions and pathways is disease relevant and an early biochemical event in C9ORF72-ALS pathogenesis. Pten is a potential therapeutic target that deserves further study.

Table of Contents

List of Figures

List of Tables

Abbreviations

ALS	Amyotrophic Lateral Sclerosis
ASO	Antisense Oligonucleotide
ATP	Adenosine Triphosphate
BAC	Bacterial Artificial Chromosome
BB	Bunina Bodies
BvFTD	Behavioural variant FTD
C9ORF72	Chromosome 9 Open Reading Frame 72
CBA	Chicken Beta Actin
CMV	Cytomegalovirus
CNS	Central Nervous System
CSF	Cerebrospinal Fluid
DAVID	Database for Annotation, Visualisation and Integrated Discovery
DE	Differentially Expressed
DPR	Dipeptide Repeat
ER	Endoplasmic Reticulum
FALS	Familial ALS
FC	Fold Change
FISH	Fluorescence *in situ* Hybridisation
FLP	Flippase
FRT	Flippase Recognition Target
FTD	Frontotemporal Dementia
FUS	Fused in Sarcoma
GEP	Gene Expression Profiling
HCI	Hyaline Conglomerate Inclusion
HEK293	Human Embryonic Kidney 293
HnRNP	Hetergeneous nuclear Ribonucleoprotein
IMPaLA	Integrated Molecular Pathway Level Analysis
iPSC	Induced Pluripotent Stem Cell
KEGG	Kyoto Encyclopedia of Genes and Genomes
LMN	Lower Motor Neurons
MN	Motor Neuron
MND	Motor Neuron Disease
MTA	Mouse Transcriptome Array
MTOR	Mammalian Target of Rapamycin
NSC34	Neuroblastoma Spinal Cord 34
PBP	Progressive Bulbar Palsy
PI3K	Phosphatidylinositol-3 Kinase
PLS	Primary Lateral Sclerosis
PMA	Progressive Muscular Atrophy
PTEN	Phosphatase and Tensin homolog
RAN	Repeat Associated Non-ATG

RIN	RNA Integrity Number
RBP	RNA Binding Protein
ROS	Reactive Oxygen Species
SALS	Sporadic ALS
ShRNA	Short Hairpin RNA
SMA	Spinal Muscular Atrophy
SOD1	Cu/Zn Super Oxide Dismutase 1
SRSF	Serine Rich Splicing Factor
TARDBP	Transactive Response DNA Binding Protein 43kDa
TDP-43	TAR DNA Binding Protein 43kDa
TetR	Tetracyline Repressor
TO	Tetracycline Operator
UBI	Ubiquitinated inclusion
UMN	Upper Motor Neurons
UPR	Unfolded Protein Response
UPS	Ubiquitin-Proteasome System
WT	Wild Type

Chapter 1. Introduction

1.1. Background

Amyotrophic lateral sclerosis (ALS) is a rapidly progressive, incurable and ultimately fatal late onset neurodegenerative disease. ALS is clinically heterogeneous, but is characterised by progressive loss of both upper and lower motor neurons. The progressive injury and death of motor neurons results in the failure of the neuromuscular system. This manifests as muscle weakness, muscle wasting, paralysis, and respiratory failure. Additionally, the pathogenesis of ALS is complex, with many genes and molecular mechanisms implicated in disease progression.

1.2. Clinical Presentation

ALS is the most common form of motor neuron disease (MND) accounting for 66% of cases, and is characterised by loss of both upper motor neurons (UMN) (including the Betz cells of the motor cortex) and lower motor neurons (LMN) (anterior horn cells of the spinal cord and brain stem nuclei) (Bäumer et al., 2014). Other less common sub-types of MND are characterised by the motor neurons initially affected. Around 25% of MND cases are classified as progressive bulbar palsy (PBP), which is confined to the bulbar musculature at presentation and affects swallowing and speech. However, a recent study showed 87% of PBP cases eventually developed into ALS, suggesting these are bulbar onset ALS cases (Karam et al., 2010). Less than 5% of MND cases are classified as primary lateral sclerosis (PLS), which affects purely the UMN (Rowland and Shneider, 2001). Finally, about 10% of MND cases are classified as progressive muscular atrophy (PMA), which affects purely the LMN at least at the time of presentation (Rowland and Shneider, 2001).

ALS has a heterogeneous clinical presentation. Motor neuron degeneration and subsequent muscle weakness progress rapidly from the initial focus in an anatomically logical manner, culminating in paralysis and respiratory failure – which is the common cause of death (Ravits and La Spada, 2009). Disease onset most commonly occurs in the

limbs and represents approximately 66% of ALS cases (Gordon, 2013). Symptoms in the limbs can be bilateral or unilateral and include muscle wasting, distal weakness, proximal weakness, fasciculation, brisk reflexes, extensor plantar reflexes and an increased tendency to trip due to foot drop (Mitchell and Borasio, 2007). Bulbar onset presents in 30% of cases and can present with dysarthria, dysphagia, wasting and fasciculation of the tongue, and facial weakness (Mitchell and Borasio, 2007). Emotional lability, characterised by episodes of involuntary laughter and crying, is also symptomatic of bulbar involvement (Mitchell and Borasio, 2007). Around 5% of cases have respiratory onset, characterised by dyspnoea and orthopnoea (Gautier et al., 2010).

In addition to motor symptoms, up to 50% of ALS patients also show evidence of cognitive impairment during the disease course (Lomen-Hoerth et al., 2003; Wheaton et al., 2007). Further, between 5 to 15% of ALS patients also have features of frontotemporal dementia (FTD) (Gordon, 2013; Lomen-Hoerth et al., 2003). Similarly, around 15% of behavioural variant FTD (bvFTD) patients also display signs of motor dysfunction, which can develop into ALS (Burrell et al., 2011; Lomen-Hoerth et al., 2002). BvFTD onset generally occurs between 50 and 60 years of age, (Achi and Rudnicki, 2012; McKhann et al., 2001; Ratnavalli et al., 2002), and is characterised by progressive degeneration of the frontal and temporal lobes of the cerebral cortex, which leads to impairment in behaviour and/or personality. These behavioural and personality changes include apathy and/or disinhibition, loss of empathy, emotional blunting, and ritualised behaviours (Lillo and Hodges, 2009). BvFTD and ALS are increasingly being considered as part of the same clinico-pathological spectrum in which different neurons are affected, due to clinical overlap, as well as similarities in neuropathology and genetic causes (Fiesel and Kahle, 2011).

As with many other neurodegenerative diseases, age is a crucial risk factor for ALS. A meta-analysis calculated the median value of the reported mean age of ALS onset from several studies, and found the average age of onset was between 63 to 65 for Europe and New Zealand, and about 59 for North America and East Asia (Marin et al., 2015). Additionally, rare juvenile onset forms of ALS exist, in which symptom onset is under 18 years of age (Shaw, 2005). Also based on a meta-analysis from studies on European and

North-American cases, the median survival time from ALS onset to death is between 25 and 40 months (Marin et al., 2015). However, whilst the disease progresses more quickly in some patients, death occurring within a year of symptom onset, between 5 and 10% of ALS patients survive over a decade after disease onset (Chiò et al., 2009).

1.3. Epidemiology

Globally, ALS incidence increases with age, and will therefore be greater in ageing populations. The life time risk of developing ALS by the age of 70 is approximately 1 in 400 (Chiò et al., 2013; Johnston et al., 2006), although this is estimated based on a study in London. ALS has a mean incidence of 2.8/100,000/year in Europe and 1.8/100,000/year in North America, and a mean prevalence of 5.4/100,000 in Europe and 3.4/100,000 in North America (Chiò et al., 2013). In addition, there are geographical differences in ALS incidence and prevalence across Europe. Both ALS incidence and prevalence are considerably lower in other countries of non-European origin, although there are only a few epidemiological studies on populations outside Europe and North America and these may under estimate ALS incidence due to sub-optimal study design (Chiò et al., 2013). However, there is an increased incidence of ALS in a region of the Western Pacific, where the incidence is 7.0/100,000/year in the Chamorro people of Guam (Waring et al., 2004) and 9.5/100,000/year in the Kii peninsula of Japan (Yoshida et al., 1998), and the increased ALS incidence is thought to be caused by neurotoxins in the diets of these populations (Cheng and Banack, 2009; Kihira et al., 2012). Lastly, multiple studies suggest an increased ALS incidence in males compared to females (M:F 1.2-1.5:1.0), although other later studies suggest the ratio may actually be more balanced (Abhinav et al., 2007; Logroscino et al., 2008; Worms, 2001).

1.4. Neuropathology

Insoluble protein aggregates found in neuronal and glial cells within the CNS are characteristic of neurodegenerative diseases. Ubiquitinated inclusions (UBIs) are the hallmark of ALS neuropathology; however other protein aggregates are often present. UBIs are observed as either filamentous skein aggregates or compact, round bodies in the degenerating motor neurons, but also in glial cells (Leigh et al., 1991). TDP-43 is the

major protein component within the UBIs in most ALS patients (excluding SOD1 and FUS related ALS cases) (Neumann et al., 2006). These UBIs stain positive for ubiquitin and phosphorylated-TDP-43 (p-TDP-43), but negative for tau and α-synuclein (Neumann et al., 2006). In ALS tissue, TDP-43 (predominantly a nuclear protein) becomes cytoplasmic. Hyperphosphorylated and C-terminal fragments of TDP-43 are also detected in the UBIs (Neumann et al., 2006). FUS-ALS pathology is characterised by FUS-positive cytoplasmic inclusions that are negative for TDP-43, tau and α-synuclein (Kwiatkowski et al., 2009; Vance et al., 2009). Mutant FUS mislocalises in the cytoplasm similarly to mutant TDP-43. In addition, SOD1-related ALS patients contain UBIs positive for SOD1 (Bruijn et al., 1998), and negative for TDP-43 (Mackenzie et al., 2007).

Bunina bodies (BB) are another type of protein aggregate observed in the degenerating LMN of ALS patients. BB are small, round, eosinophilic cytoplasmic inclusions, present in approximately 85% of ALS patients (Piao et al., 2003). BB stain positive for cystatin C (Okamoto et al., 1993) and transferrin (Mizuno et al., 2006) but stain negative for TDP-43 (Tan et al., 2007). However, the other protein components of BB remain unclear (Okamoto et al., 2008). The origin of BBs is also uncertain, but some studies suggest they could originate from the lysosomes, Golgi apparatus or endoplasmic reticulum (ER) (Okamoto et al., 2008).

Hyaline conglomerate inclusions (HCIs) are another major aggregate type observed. Hyaline (meaning glassy) refers to HCIs appearance when stained with agents such as haematoxylin and eosin. HCIs are large multifocal inclusions that contain phosphorylated and non-phosphorylated neurofilament subunits, as well as entrapped organelles and proteins (Wood et al., 2003). HCIs are associated with the SOD1 I113T (p.I114T) mutation, and are not seen in sporadic ALS cases (Ince et al., 1998). HCIs and UBIs are rarely seen in the same ALS cases, which could suggest different pathological pathways that give rise to the disease (Ince et al., 1998). However, HCIs are also seen in other neurodegenerative diseases, making them a less specific marker of ALS compared with UBIs (Sobue et al., 1990).

Recently, protein aggregates containing ubiquitin and the ubiquitin binding protein p62, but negative for TDP-43, FUS, tau, α-synuclein, α-internexin and neurofilament have been described in ALS patients (King et al., 2011). These inclusions can exist as

cytoplasmic globular or star-shaped inclusions, or smaller spherical intranuclear inclusions (Al-Sarraj et al., 2011). These inclusions can be found in both neuronal and glial cells in the granular layer, molecular layer and Purkinje cells of the cerebellum, and also in the pyramidal cells of the hippocampus. One study suggests these p62 positive, TDP-43 negative inclusions are specific to ALS and/or FTD cases that contain the (G4C2)n repeat expansion (C9ORF72-ALS/FTD) (Al-Sarraj et al., 2011). Another study confirms C9ORF72-ALS cases have much higher levels of these p62 positive, TDP-43 negative extra-motor neuronal inclusions in the hippocampus and frontal neocortex, compared to non-C9ORF72-ALS cases (Cooper-Knock et al., 2012). These inclusions could represent an alternative pathological mechanism to the TDP-43 aggregates occurring in many ALS cases.

In addition, C9ORF72-ALS/FTD patients display other exclusive neuropathology. The (G4C2)n repeat expansion is transcribed in both sense and antisense directions, which forms sense (G4C2)n and antisense (C4G2)n RNA foci respectively in the CNS of C9ORF72-ALS/FTD patients (DeJesus-Hernandez et al., 2011; Gendron et al., 2013). Sense and antisense RNA foci have been detected in neuronal cells in the frontal cortex, the hippocampal dentate fascia, cerebellar granule cells and Purkinje cells, and LMN of the spinal cord (Cooper-Knock et al., 2015b; Cooper-Knock et al., 2014b; Mizielinska et al., 2013). Additionally, sense and antisense RNA foci are present in astrocytes, microglia, and oligodendrocytes (Mizielinska et al., 2013), as well as lymphoblasts, fibroblasts and iPSC-derived neurons from C9ORF72-ALS/FTD patients (Almeida et al., 2013; Donnelly et al., 2013; Lagier-Tourenne et al., 2013). The majority of RNA foci are nuclear, but rare cytoplasmic RNA foci are also detected (Cooper-Knock et al., 2014b; Mizielinska et al., 2013). Numerous studies show the RNA foci also co-localise with various RNA binding proteins in patient CNS tissue (section 1.7.2) (Cooper-Knock et al., 2014b; Haeusler et al., 2014; Lee et al., 2013; Mori et al., 2013b; Sareen et al., 2013).

Lastly, the sense (G4C2)n and antisense (C4G2)n RNA undergoes unconventional repeat associated non-ATG (RAN) translation in all reading frames, forming aggregation-prone dipeptide repeat (DPR) proteins (section 1.5.4) (Ash et al., 2013; Mori et al., 2013c). The poly-GA, poly-GP, and poly-AP DPR protein species each form insoluble cytoplasmic aggregates, whilst the poly-GR and poly-PR DPR protein species both form insoluble

nuclear aggregates in C9ORF72-ALS/FTD patient CNS tissue (Ash et al., 2013; Mann et al., 2013; Mori et al., 2013a; Mori et al., 2013c; Zu et al., 2013). DPR pathology is highly consistent in C9ORF72-ALS/FTD patients regardless of disease phenotype, and DPR load is high in the cerebellum, all neocortical regions (frontal, motor cortex and occipital) and hippocampus, moderate in subcortical areas and minimal in LMN (Mackenzie et al., 2013). Additionally, DPR proteins are present in the p62-positive, TDP-43 negative aggregates in C9ORF72-ALS/FTD patients, and could be the cause of these seemingly pathognomonic C9ORF72-ALS/FTD aggregates (Mann et al., 2013).

1.5. Genetics

ALS has two clinically defined subtypes: familial ALS (FALS) and sporadic ALS (SALS). The majority of ALS patients (around 95%) are defined as sporadic, with no family history of ALS (Renton et al., 2014). However, twin studies estimate SALS has a heritability of 38-78% (Al-Chalabi et al., 2010). Patients with a first or second degree relative with ALS are clinically defined as familial cases, and represent an estimated 5% of total cases (Byrne et al., 2011). Incomplete evidence of inheritance can result in the incorrect assignment of familial cases as sporadic cases however, underestimating FALS as a proportion of the disease. As with other late-onset diseases, premature death from other causes in previous generations and incomplete penetrance can mask a history of family inheritance (Andersen, 2006). FALS is generally inherited in an autosomal dominant manner, but autosomal recessive and X-linked ALS causative mutations also exist (Renton et al., 2014). Currently, mutations in 24 known genes are thought to account for >68% of all FALS cases, and >11% of SALS cases (Table 1.1) (Renton et al., 2014).

Table 1.1 Genes associated with Familial ALS and Familial ALS/FTD. The genetic locus, gene name, implicated pathogenic mechanism, mode of inheritance, age of onset, disease phenotype, and reference to initial discovery are shown.

Locus	Chromosome	Gene	Implicated Pathogenic Mechanism	Mode	Onset	Phenotype	Reference
ALS 1	21q22.11	*SOD1*	Oxidative stress; UPS; autophagy	AD (AR)	Adult	ALS	(Rosen, 1993)
ALS 2	2q33.1	*ALSIN*	Endosomal trafficking	AR	Juvenile	ALS, HSP	(Hadano et al., 2001; Yang et al., 2001)
ALS 3	18q21	*Unknown*	Unknown	AD	Adult	ALS	(Hand et al., 2002)
ALS 4	9q34.13	*SETX*	RNA metabolism	AD	Juvenile	ALS, AOA2	(Chen et al., 2004)
ALS 5	15q21.1	*SPG11*	DNA damage repair; axonal growth	AR	Juvenile	ALS, HSP	(Orlacchio et al., 2010)
ALS 6	16p11.2	*FUS*	RNA metabolism; stress granules	AD (AR)	Adult	ALS, FTD	(Kwiatkowski et al., 2009; Vance et al., 2009)
ALS 7	20p13	*Unknown*	Unknown	AD	Adult	ALS	(Sapp et al., 2003)
ALS 8	20q13.32	*VAPB*	Vesicle trafficking; UPR; ER stress	AD	Adult	ALS, SMA	(Nishimura et al., 2004)
ALS 9	14q11.2	*ANG*	RNA metabolism, Angiogenesis	AD	Adult	ALS, FTD, PD	(Greenway et al., 2006)
ALS 10	1p36.22	*TARDBP*	RNA metabolism	AD	Adult	ALS, FTD, PD	(Sreedharan et al., 2008)
ALS 11	6q21	*FIG4*	Endosomal trafficking	AD	Adult	ALS, CMT	(Chow et al., 2009)
ALS 12	10p13	*OPTN*	Autophagy	AD (AR)	Adult	ALS, POAG	(Maruyama et al., 2010)
ALS 13	12q24.12	*ATXN2*	Endocytosis, RNA translation	AD	Adult	ALS, FTD, SCA2	(Elden et al., 2010)
ALS 14	9p13.3	*VCP*	Autophagy	AD	Adult	IBMPFD/ALS, FTD, CMT, HSP	(Johnson et al., 2010)
ALS 15	Xp11.21	*UBQLN2*	UPS; autophagy	XD	Adult	ALS/FTD	(Deng et al., 2011)
ALS 16	9p13.3	*SIGMAR1*	UPR; ER stress; proteasome	AD	Juvenile	ALS/FTD	(Al-Saif et al., 2011; Luty et al., 2010)
ALS 17	3p11.2	*CHMP2B*	Endosomal trafficking; autophagy	AD	Adult	ALS, FTD	(Parkinson et al., 2006)
ALS 18	17p13.2	*PFN1*	Cytoskeleton; axonal growth	AD	Adult	ALS, FTD	(Wu et al., 2012)
ALS 19	2q34	*ERBB4*	Neuronal development	AD	Adult	ALS	(Takahashi et al., 2013)
ALS 20	12q13.13	*HNRNPA1*	RNA metabolism	AD	Adult	IBMPFD/ALS	(Kim et al., 2013)
ALS 21	5q31.2	*MATR3*	RNA metabolism	AD	Adult	ALS, VCPDM	(Johnson et al., 2014)
ALS 22	2q35	*TUBA4A*	Cytoskeleton	AD	Adult	ALS	(Smith et al., 2014)
ALS-FTD1	9p21.2	*C9ORF72*	RNA metabolism; endosomal trafficking; autophagy	AD	Adult	ALS/FTD	(DeJesus-Hernandez et al., 2011; Renton et al., 2011)
ALS-FTD2	22q11.23	*CHCHD10*	Mitochondrial maintenance	AD	Adult	ALS/FTD	(Bannwarth et al., 2014)
ALS-FTD3	5q35.3	*SQSTM1*	Autophagy, protein degradation	AD	Adult	ALS/FTD, PDB	(Fecto et al., 2011)
ALS-FTD4	12q14.2	*TBK1*	Autophagy, neuroinflammation	AD	Adult	ALS/FTD	(Cirulli et al., 2015; Freischmidt et al., 2015)

1.5.1. SOD1

The first mutations associated with ALS were found in Cu/Zn superoxide dismutase 1 (*SOD1*) (Rosen, 1993). There are over 170 known missense mutations that are found in each of the five exons of *SOD1* (Andersen, 2006; Renton et al., 2014; Su et al., 2014). However, reliable genetic evidence linking the mutation to ALS pathogenicity still remains for a portion of these. *SOD1* mutations account for 12-20% of FALS cases and 2-7% of SALS cases (www.alsod.iop.kcl.ac.uk), however there is considerable phenotypic variability caused by different *SOD1* mutations. For example, the A4V (p.A5V) mutation is autosomal dominant and causes a particularly aggressive form of ALS, typically leading to death within 1 year of disease onset (Cudkowicz et al., 1997). Whilst, the D90A (p.D91A) mutation is recessive and causes a slowly progressive form of ALS where patients survive over 10 years from disease onset on average (Andersen et al., 1996).

SOD1 mutations were first proposed to cause ALS via a loss of function. SOD1 is ubiquitously expressed and converts superoxide radicals into molecular oxygen and hydrogen peroxide, and therefore, SOD1 loss of function was hypothesised as a cause of oxidative stress (Rosen, 1993). However, whilst most *SOD1* mutations do reduce the SOD1 dismutase activity, there is no correlation between the disease severity and the SOD1 dismutase activity (Al-Chalabi et al., 2012). Further, the autosomal dominant inheritance of most *SOD1* mutations indicates a toxic gain of function in mutant SOD1. WT SOD1 has a high propensity to misfold, which can be induced by oxidation and demetallation (Rakhit and Chakrabartty, 2006; Rakhit et al., 2004). In addition, ALS causative mutations in SOD1 also increase the protein's propensity to misfold and aggregate (Stathopulos et al., 2006; Turner and Talbot, 2008; Vassall et al., 2006). Aggregated SOD1 is suggested to affect a range of cellular functions that could lead to neurotoxicity and ALS, and include oxidative stress, mitochondrial dysfunction, glutamate-mediated excitotoxicity, and axonal transport defects (Joyce et al., 2011). Further, misfolded and/or aggregated mutant SOD1 can be taken up by neuronal cells, and nucleate misfolding and aggregation of soluble endogenous mutant and wildtype SOD1 (Münch et al., 2011; Sundaramoorthy et al., 2013). This seeding of SOD1 aggregation can then propagate from cell to cell in a prion-like manner, and could explain the anatomical progression of ALS.

1.5.2. TARDBP

Mutations in the transactive response DNA binding protein 43kDa (*TARDBP*) gene, which encodes transactive response DNA binding protein 43kDa (TDP-43), also cause ALS and/or bvFTD (Gitcho et al., 2008; Kabashi et al., 2008; Sreedharan et al., 2008; Yokoseki et al., 2008). Mutations in *TARDBP* account for 4-5% of FALS and about 1% of SALS (Millecamps et al., 2010). There are currently 40 known mutations in the *TARDBP* gene that are associated with ALS, and these primarily occur in the C-terminus of TDP-43 which is important for ribonucleoprotein binding and splicing (Millecamps et al., 2010; Renton et al., 2014). *TARDBP* mutations are also autosomal dominant and associated with a classical ALS phenotype.

TDP-43 is a multifunctional heterogeneous nuclear ribonucleoprotein involved in a range of different RNA metabolism processes including transcription, mRNA splicing, RNA transport, mRNA stability and stress granule dynamics, and miRNA biogenesis (Buratti and Baralle, 2008, 2012). TDP-43 is expressed ubiquitously and is predominantly localised in the nucleus, where it performs the majority of its functions. However, TDP-43 also shuttles between the nucleus and cytoplasm (Ayala et al., 2008; Winton et al., 2008). In the majority of ALS patients, TDP-43 becomes mislocalised from the nucleus to the cytoplasm, where it then becomes ubiquitinated, phosphorylated, truncated and aggregated (section 1.4) (Neumann et al., 2006). Taken together with the fact that TARDBP mutations cause ALS, TDP-43 is strongly implicated in the pathogenesis of most ALS cases.

However, despite TDP-43 aggregates being the pathological hallmark of ALS, it is unclear whether TDP-43 aggregation causes toxicity through gain and/or loss of function in ALS pathogenesis. Several studies have shown that loss of nuclear TDP-43 causes splicing defects in cellular and animal models, as well as in motor neurons from TARDBP-ALS patients (De Conti et al., 2015; Highley et al., 2014; Ling et al., 2015). Importantly, TDP-43 represses splicing of nonconserved cryptic exons and maintains intron integrity (Ling et al., 2015). A loss of this feature of TDP-43 could provide an important link to C9ORF72-ALS, in which the intron containing a (G4C2)n repeat expansion is aberrantly licenced for export to the cytoplasm (section 1.7.3).

In addition to a loss of TDP-43 function, the cytoplasmic TDP-43 aggregates could also gain a toxic function. TDP-43 has a propensity to misfold and aggregate, and recent studies show misfolded TDP-43 can seed further TDP-43 misfolding and aggregation, which can also spread along axons and from cell to cell in a prion-like manner (Feiler et al., 2015; Zhu et al., 2014). Similarly to misfolded SOD1, this gain of prion-like function could explain the anatomically logical progression of motor neuron degeneration in the majority of ALS cases. However, a dominant negative mechanism proposes simultaneous gain and loss of TDP-43 function, in which the misfolded TDP-43 drives further TDP-43 aggregation and subsequently nuclear loss of TDP-43 and TDP-43 function.

1.5.3. FUS

Fused in Sarcoma (FUS) shares functional homology with TDP-43. Mutations in *FUS* account for 4-5% FALS and slightly less than 1% SALS, and similarly to *TARDBP*, these mutations cluster in the RNA binding domain at the C terminus of the protein (Deng et al., 2014; Kwiatkowski et al., 2009; Renton et al., 2014; Vance et al., 2009). There are currently over 50 *FUS* mutations associated with ALS (Deng et al., 2014). Most of these mutations have an autosomal dominant mode of inheritance, although *FUS* mutations with a recessive pattern of inheritance were found in a family of Cape Verdean origin (Kwiatkowski et al., 2009). In addition, *FUS* mutations are associated with a classical ALS phenotype, with an earlier disease onset (Deng et al., 2014).

FUS is involved in a range of RNA and DNA metabolic processes including transcription, mRNA splicing, mRNA transport, stress granule formation, miRNA biogenesis, and genome integrity (Deng et al., 2014). FUS is predominantly localised in the nucleus, but also shuttles between the nucleus and cytoplasm (Zinszner et al., 1997). FUS related ALS patients contain FUS aggregates (section 1.4), and similarly to TDP-43, it is unknown precisely how these aggregates confer neurotoxicity in the MN. However, a dominant negative mechanism of FUS mislocalisation, misfolding and aggregation, leading to depleted nuclear function is likely (Deng et al., 2014). Lastly, misfolded FUS is also hypothesised to spread in a prion-like manner similar to TDP-43 (King et al., 2012; Polymenidou and Cleveland, 2011), but there are currently no functional experiments to prove this hypothesis.

1.5.4. C9ORF72

In 2011, two groups found a (G4C2)n hexanucleotide repeat expansion within a non-coding region of Chromosome 9 open reading frame 72 (*C9ORF72*) that causes both ALS and FTD (Figure 1.1) (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Expansions of >30 (G4C2)n repeats are considered pathogenic (Beer et al., 2015; Byrne et al., 2014), but expansions of 200-5000 repeats are commonly detected in ALS patients (Cooper-Knock et al., 2014a). The (G4C2)n repeat expansion is the most common mutation associated with ALS, and accounts for about 8% of SALS and 39% of FALS cases (Majounie et al., 2012). In addition, the (G4C2)n repeat expansion accounts for about 7% of sporadic FTD and 25% of familial FTD cases (Majounie et al., 2012). Further, there is a greater incidence of dementia in C9ORF72-ALS patients compared to non-C9ORF72-ALS patients (Cooper-Knock et al., 2012; Cooper-Knock et al., 2014a; Gijselinck et al., 2012; Stewart et al., 2012). It is currently unclear how the (G4C2)n repeat expansion causes ALS and/or FTD pathogenesis, although three mutually inclusive hypotheses have been suggested; C9ORF72 haploinsufficiency, RNA toxicity, and dipeptide repeat (DPR) protein toxicity (section 1.7).

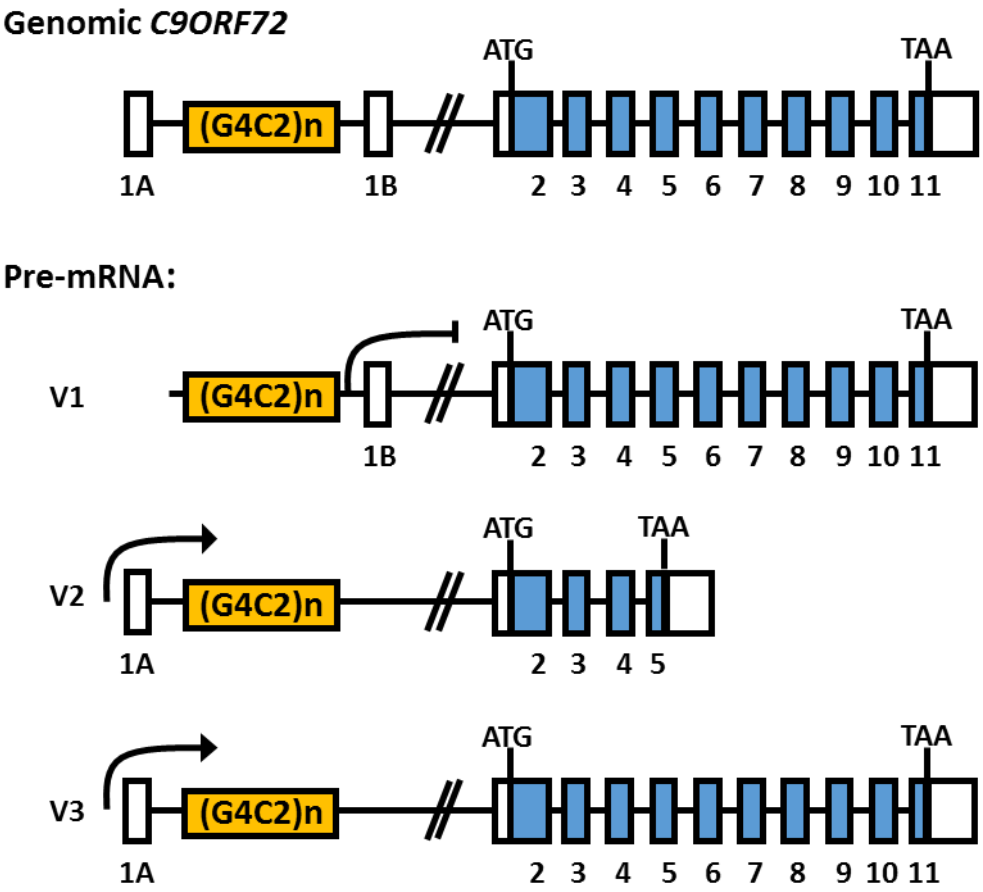


Figure 1.1 Genomic C9ORF72 and transcript variants. The line represents introns and boxes represent exons - blue are translated and white are untranslated. The (G4C2)n repeat expansion is shown in an orange box, and is located between exons 1A and 1B. Transcript variants 1 and 3 encode C9ORF72 long form, whilst variant 2 encodes C9ORF72 short form. The transcription start site (arrow) is upstream of exon 1A and the (G4C2)n repeat in variants 2 and 3. Whilst in variant 1, the (G4C2)n repeat is upstream of the transcription start site and is thought to inhibit transcription (flat head arrow).

The C9ORF72-ALS and classic ALS phenotypes are indistinguisable (Cooper-Knock et al., 2012; Cooper-Knock et al., 2014a; Gijselinck et al., 2012; Murray et al., 2011). It is not clear whether the age of C9ORF72-ALS disease onset differs from non-C9ORF72-ALS, however disease duration is shorter in C9ORF72-ALS (Cooper-Knock et al., 2014a; Millecamps et al., 2012). The (G4C2)n repeat expansion is thought to be non-penetrant in those under 35 years of age, 50% penetrant by 58 years of age, and approaching full penetrance at 80 years of age (Benussi et al., 2014; Galimberti et al., 2014; Majounie et al., 2012). Additionally, genetic analysis shows there is a greater than expected incidence of the ALS associated mutations (including the (G4C2)n repeat expansion) co-occurring with a second ALS-linked mutation (Bury et al., 2015; Chiò et al., 2012b; van Blitterswijk

et al., 2012a). This suggests that a second genetic 'hit' (oligogenic inheritance) may be necessary for ALS pathogenesis (and more explicity, for C9ORF72-ALS pathogenesis) and could also influence age of onset, symptom presentation and/or disease progression.

Additionally, the (G4C2)n repeat expansion is unstable. Within individuals, the (G4C2)n repeat length varies between different cell types, and even between different regions of the brain, which suggests somatic instability (Beck et al., 2013; Nordin et al., 2015; van Blitterswijk et al., 2013b). Further, families which carry the (G4C2)n repeat expansion have shown evidence of genetic anticipation, with the average age of ALS and/or FTD onset lower in successive generations (Benussi et al., 2014; Chiò et al., 2012a). In addition, the (G4C2)n the repeat DNA is liable to both expansions and contractions in a length-dependent manner in both *E.coli* and HEK293 cells (Thys and Wang, 2015). However, there is no direct evidence, from southern blotting for example, to confirm whether the potential genetic anticipation in C9ORF72-ALS/FTD patients is due to expansion of the (G4C2)n repeat between generations.

The *C9ORF72* gene expresses three transcript variants (V1, V2, and V3), which in turn code for two different C9ORF72 protein isoforms (long and short). V2 and V3 include exon 1a, and contain the (G4C2)n repeat within intron 1, whilst V1 includes exon 1b and is transcribed downstream of the (G4C2)n repeat, which is therefore excluded from the mRNA transcript (DeJesus-Hernandez et al., 2011) (Figure 1.1). V1 and V3 encode the 481 amino acid C9ORF72 long form, whilst V2 encodes the 222 amino acid C9ORF72 short form (DeJesus-Hernandez et al., 2011). The murine *C9ORF72* homolog is highly expressed at transcript level in neuronal cells throughout the CNS, including the cortex and spinal motor neurons, however expression appears much lower in glial cells (Atkinson et al., 2015; Suzuki et al., 2013). C9ORF72 protein shares homology with differentially expressed in normal and neoplasia (DENN) domain, which function as Rab-GTPase GDP/GTP exchange factors (GEF), and are involved in membrane trafficking (Levine et al., 2013; Zhang et al., 2012). *In vitro* studies further implicate C9ORF72 protein in membrane trafficking, and suggest the protein regulates endocytosis and autophagy (Farg et al., 2014). Additionally, C9ORF72 is suggested to localise in the nucleus, cytoplasm and synapses in human CNS tissue, although the specificity of currently available C9ORF72-antibodies is questionable and casts doubt over the validity

of these findings (Atkinson et al., 2015; Cooper-Knock et al., 2012; Snowden et al., 2012; Stewart et al., 2012; Xiao et al., 2015).

1.5.5. Genes Involved in RNA Metabolism

In addition to *TARDBP* and *FUS*, there are rarer mutations in several other genes involved in RNA metabolism that are associated with ALS. These genes include senataxin (*SETX*), angiogenin (*ANG*), ataxin 2 (*ATXN2*), heterogeneous nuclear ribonucleoprotein A1 (*HNRNPA1*), and matrin 3 (*MATR3*). This implicates dysregulated RNA metabolism as a key mechanism in ALS pathogenesis and corroborates the RNA toxicity hypothesis suggested for C9ORF72-ALS (section 1.7.2).

Mutations in *SETX* are associated with a juvenile onset form of ALS that has a much longer and slower disease progression (Chen et al., 2004; Hirano et al., 2011). SETX is a DNA/RNA helicase that plays a role in transcription, R-loop resolution, and DNA damage repair (Skourti-Stathaki et al., 2011). Currently it is unclear how *SETX* mutations cause ALS however.

There are at least 20 known mutations in *ANG* that are associated with ALS, and have been found in both FALS and SALS patients (Greenway et al., 2006; Kirby et al., 2013; Pan et al., 2015). ANG is a member of the pancreatic ribonuclease A superfamily, and is neuroprotective to motor neurons. However, disease associated mutant ANG loses this neuroprotective function (Subramanian et al., 2008). Importantly, ANG binds RNA species with G-quadruplex secondary structure and induces stress granule assembly, which is critical for the neuroprotective effect (Ivanov et al., 2014). Interestingly, the sense RNA expressed from the C9ORF72 (G4C2)n repeat expansion also forms a G-quadruplex secondary structure (Fratta et al., 2012; Haeusler et al., 2014; Reddy et al., 2013b), which may impair the neuroprotective function of ANG.

ATXN2 contains a (CAG)n repeat that encodes a polyglutamine (polyQ) tract in the ATXN protein. An intermediate polyQ tract (31-33 repeats) is associated with ALS, whilst longer polyQ expansions (>36 repeats) cause spinocerebellar ataxia 2 (Elden et al., 2010; Neuenschwander et al., 2014). ATXN2 is an RNA binding protein implicated in RNA metabolism (Satterfield and Pallanck, 2006; Shibata et al., 2000), as well as membrane trafficking (Nonis et al., 2008; Ralser et al., 2005) and stress granule function (Anderson

and Kedersha, 2006). ATXN2 interacts with TDP-43 in an RNA-dependent manner, and ATXN2 actually exacerbates TDP-43 induced toxicity in both yeast and *Drosophila* models (Elden et al., 2010). In addition, ATXN2 interacts with FUS, and ATXN2 containing intermediate length polyQ tracts also exacerbate mutant FUS induced toxicity in cellular models (Farg et al., 2013).

Mutations in *HNRNPA1* are also associated with ALS (Kim et al., 2013). Similarly to the other ALS associated genes involved in RNA metabolism, HnRNPA1 binds TDP-43 (Buratti et al., 2005; Ling et al., 2010). Additionally, HnRNPA1 is involved in RNA processing, splicing, and transport.

Lastly, there are seven known mutations in *MATR3* that associate with ALS; two of those were found in FALS patients, and five were found in apparent SALS patients (Fifita et al., 2015; Johnson et al., 2014; Leblond et al., 2016; Lin et al., 2015; Millecamps et al., 2014). MATR3 is a nuclear matrix protein that can bind both DNA and RNA, but also directly interacts with TDP-43 (Ling et al., 2010). MATR3 is also suggested to have a range of DNA and RNA metabolic processes including chromatin organisation, DNA replication and repair, as well as transcription, RNA stability, RNA processing and RNA transport (Salton et al., 2011).

1.5.6. Genes Involved in Protein Transport and Degradation

In addition to RNA metabolism, there are also rare mutations in several genes involved in protein transport and/or degradation that are associated with ALS. Taken together with the fact that protein aggregation is a hallmark of ALS (and many other neurodegenerative diseases), protein degradation is particularly likely to be critical in ALS pathogenesis. Genes that are associated with ALS and are involved in endosomal transport include alsin (*ALS2*), vesicle-associated membrane protein-associated protein B (*VAPB*), charged multivesicular body protein B (*CHMP2B*), and phosphoinositide 5-phosphatase (*FIG4*), whilst those involved in the ubiquitin-proteasome system (UPS) include ubiquilin 2 (*UBQLN2*), sequestosome 1 (*SQSTM1*), and sigma non-opoid intracellular receptor 1 (*SIGMAR1*), and finally, those involved in autophagy include optineurin (*OPTN*), valosin containing protein (*VCP*), and TANK-binding kinase (*TBK1*).

Additionally, this supports the proposed hypotheses of C9ORF72 haploinsufficiency (section 1.7.1) and/or DPR aggregate toxicity in C9ORF72-ALS (section 1.7.3).

Mutations in *ALS2* are associated with an autosomal recessive, juvenile onset form of ALS (Hadano et al., 2001; Yang et al., 2001). The majority of these mutations are thought to cause a loss of function. *ALS2* encodes alsin, which is a Rab5 guanine nucleotide exchange factor (GEF) involved in the regulation of endocytosis and thereby implicates dysregulation of endocytic trafficking in ALS pathogenesis (Lai et al., 2009).

Several mutations in *VAPB* are also associated with a slowly progressive form of ALS, however the pathogenicity of some of these mutations requires validation (Chen et al., 2010; Ingre et al., 2013b; Kabashi et al., 2013; Nishimura et al., 2004; van Blitterswijk et al., 2012b). VAPB is an integral ER protein and is involved in intracellular trafficking, the unfolded protein response (UPR), and regulates ER-mitochondria interactions (Lev et al., 2008; Stoica et al., 2014). Additionally, the P56S VAPB mutation protein associated with ALS disrupts the UPR, anterograde axonal transport of mitochondria and calcium homeostasis (De Vos et al., 2012; Kanekura et al., 2006; Mórotz et al., 2012).

Mutations in *CHMP2B* are associated with ALS that has a predominantly lower motor neuron phenotype (Cox et al., 2010; Parkinson et al., 2006). CHMP2B is highly expressed throughout the CNS, and is part of the ESCRT-III complex (endosomal sorting complex required for transport III) that sorts endosomal cargo for recycling or degradation in the lysosome (Cox et al., 2010; Skibinski et al., 2005). Additionally, mutant CHMP2B disrupts autophagy in transiently transfected cellular models (Cox et al., 2010).

FIG4 mutations have been found in both FALS and SALS cases, and are thought to cause a loss of function (Chow et al., 2009). FIG4 regulates PI(3,5)P2 levels, which mediates retrograde trafficking of endosomal vesicles to the trans-golgi network (Chow et al., 2009). Interestingly, *FIG4* knock-down causes neurodegeneration in mice and motor defects in *Drosophila* models (Chow et al., 2007; Kyotani et al., 2016).

There are at least ten mutations in *UBQLN2* that are associated with dominantly inherited, X-linked ALS that can have either juvenile or adult onset (Deng et al., 2011; Gellera et al., 2013; Williams et al., 2012). UBQLN2 is a ubiquitin-like protein that functions in ubiquitin proteasome system (UPS) that is important for degrading and

recycling misfolded or damaged proteins. Importantly, several studies show mutant UBQLN2 impairs this proteolytic pathway (Chang and Monteiro, 2015; Deng et al., 2011; Osaka et al., 2015), but may also impair RNA metabolism as well (Gilpin et al., 2015).

SQSTM1/p62 plays an important role in protein degradation via the UPS and the autophagy pathway (Bjørkøy et al., 2006; Seibenhener et al., 2004), and is found in the characteristic ubiquitinated aggregates in ALS patients including the p62-positive, TDP-43 negative aggregates specific to C9ORF72-ALS pathology (section 1.4). Multiple mutations in *SQSTM1* were found in both FALS and SALS patients, further implicating SQSTM1/p62 and impaired protein degradation in ALS pathogenesis (Fecto et al., 2011; Kwok et al., 2014; Teyssou et al., 2013). SQSTM1/p62 knock-down in a zebrafish model impaired autophagy and caused motor defects, which could be rescued with the autophagy inducer rapamycin (Lattante et al., 2015).

Missense mutations in *SIGMAR1* cause an autosomal recessive, juvenile onset form of ALS (Al-Saif et al., 2011). Addition mutations in the 3'UTR of *SIGMAR1* are suggested to affect mRNA stability and are also associated with ALS (Luty et al., 2010), but their role in ALS pathogenesis remains unclear. Sigma-1 receptor (Sig-1R) is an ER receptor chaperone that is involved in calcium transport between the ER and mitochondria via the IP3 receptor, counteracts ER stress, and is neuroprotective (Fukunaga et al., 2015; Hayashi and Su, 2007). In addition, the p.E102Q Sig-R1 mutant protein reduces mitochondrial ATP production, inhibits proteasome activity and causes mitochondrial injury, and also aggravates ER-stress induced death in neuronal cell lines (Al-Saif et al., 2011; Fukunaga et al., 2015).

Several different types of mutation in *OPTN* are associated with either autosomal dominant or recessive ALS (Beeldman et al., 2015; Goldstein et al., 2016; Maruyama et al., 2010; van Blitterswijk et al., 2012c). OPTN regulates a range of cellular functions including protein degradation via the UPS and the autophagy pathway, and interacts with several other proteins associated with ALS including UBQLN2 and TBK1 (Gilpin et al., 2015; Morton et al., 2008).

VCP mutations are associated with both FALS and SALS (Abramzon et al., 2012; Johnson et al., 2010; Koppers et al., 2012). VCP is an AAA+-ATPase and facilitates both UPS and

autophagy (Meyer and Weihl, 2014). Additionally, mutant VCP is suggested to cause mitochondrial uncoupling and a reduction in ATP production (Bartolome et al., 2013).

Loss of function mutations in *TBK1* are thought to cause ALS via haploinsufficiency (Cirulli et al., 2015; Freischmidt et al., 2015). TBK1 regulates both immune signalling and autophagy (Weidberg and Elazar, 2011). TBK1 also binds and phosphorylates the ALS-related proteins OPTN and SQSTM1 (Pilli et al., 2012; Wild et al., 2011), strongly implicating protein degradation and autophagy in ALS pathogenesis.

1.5.7. Genes Involved in Axonal Transport and Cytoskeleton

There also several rare mutations in genes that are involved in axonal transport and cytoskeleton that are associated with ALS associated: implicating these similar biological functions in ALS pathogenesis. Motor neurons have long axons, and efficient transport of protein, lipid, RNA and organelle cargo along the cytoskeleton is necessary for correct cellular function. ALS associated genes that are involved in axonal transport and/or the cytoskeleton include spatacsin (*SPG11*), profilin 1 (*PFN1*), tubulin alpha 4A (*TUBA4A*).

Several loss of function mutations in *SPG11* are associated with an autosomal recessive, juvenile onset form of ALS with slow disease progression (Daoud et al., 2012; Orlacchio et al., 2010). The exact function of SPG11 is unknown, but SPG11 colocalises with the cytoskeleton and synaptic vesicles in neuronal axons and dendrites. Additionally, iPSC-derived neuronal cells from *SPG11* patients displayed axonal instability by downregulation of acetylated tubulin and impaired cargo transport (Pérez-Brangulí et al., 2014).

Mutations in *PFN1* are a rare cause a FALS, and the p.E117G variant is now recognised only as a moderate ALS risk factor, rather than fully penetrant (Fratta et al., 2014; Ingre et al., 2013a; Smith et al., 2015; Tiloca et al., 2013; van Blitterswijk et al., 2013a; Wu et al., 2012). PFN1 regulates actin filament dynamics, but is also implicated in stress granule regulation (Figley et al., 2014). ALS-linked mutant PFN1 variants are less stable, have a propensity to aggregate, impair stress granule dynamics, and reduce axon outgrowth and growth cone size (Boopathy et al., 2015; Figley et al., 2014; Wu et al., 2012). Interestingly, other ALS-linked mutations in *TARDBP*, *FUS*, *HNRNPA1*, and *VCP*

also display altered stress granule dynamics suggesting another important mechanism in ALS pathogenesis (Figley et al., 2014).

Lastly, at least five loss of function mutations in *TUBA4A* are associated with a classical ALS phenotype (Smith et al., 2014). TUBA4A is an important component of cytoskeletal microtubules, and ALS-linked mutant TUBA4A has an increased propensity to aggregate and impairs microtubule dynamics and stability through a dominant-negative mechanism (Smith et al., 2014).

1.6. General ALS Pathomechanisms

Extensive evidence shows multiple pathogenic mechanisms exist in ALS. However, in the majority of ALS cases, separating the initial toxic insult and exacerbating secondary pathways that drive ALS progression is difficult. Pathomechanisms suggested include oxidative stress, excitotoxicity, mitochondrial dysfunction, impaired axonal transport, dysregulated RNA metabolism, impaired protein homeostasis, and glial involvement.

1.6.1. Oxidative Stress

Oxidative stress arises from an imbalance between reactive oxygen species (ROS) production and removal, and/or an impaired ability of the biological system to repair oxidative damage. ROS cause oxidative damage to DNA, RNA, proteins and lipids, and also disrupt redox sensitive cellular signalling. Cumulative oxidative damage in post-mitotic motor neurons may eventually overwhelm the homeostatic cellular mechanisms that cope with other toxic insults such as ALS-linked mutations, resulting in neuronal death later in older age (Turner et al., 2013). ALS-causative *SOD1* mutations are thought to cause aberrant free radical handling in a gain of function mechanism, implicating oxidative stress in ALS pathogenesis (Barber and Shaw, 2010; Rosen, 1993). In addition, CNS tissue and biosamples from ALS patients show increased levels of oxidative damage (Chang et al., 2008; Ferrante et al., 1997; Fitzmaurice et al., 1996; Shaw et al., 1995; Shibata et al., 2001).

1.6.2. Glutamate Excitotoxicity

Excitotoxicity is a process in which excessive stimulation of glutamate receptors causes a large influx of calcium ions into the post-synaptic neurons, and results in toxicity.

Glutamate levels are elevated in the CSF of ALS patients, and correlates with disease severity (Spreux-Varoquaux et al., 2002). Motor neurons are also inherently vulnerable to glutamate excitotoxicity due to low glutamate receptor 2 (GluR2) expression (which limits the calcium permeability of AMPA receptors) and limited calcium-buffering capacity (Van Damme et al., 2007; Vanselow and Keller, 2000). Additionally, astrocytes express excitatory amino acid transporter 2 (EAAT2), which removes glutamate from the synaptic cleft. EAAT2 levels are reduced in both ALS patients and animal models (Bruijn et al., 1997; Rothstein et al., 1992), which is suggested to cause prolonged glutamatergic stimulation of the motor neurons, implicating both excitotoxicity and glial cell involvement in ALS pathogenesis. Lastly, the antiglutamatergic drug riluzole is the only effective treatment available for ALS, and improves patient survival by three months (Andreadou et al., 2008; Bensimon et al., 1994).

1.6.3. Mitochondrial Dysfunction

Mitochondria are critical for cell survival and generate ATP, buffer intracellular calcium and act as gatekeepers of the intrinsic apoptotic pathway. Impaired mitochondrial morphology and function (including ATP production, calcium buffering, mitochondria-dependent apoptosis, respiratory complexes and redox balance) are described in ALS spinal motor neurons, as well as various cellular and animal models of ALS (Bowling et al., 1993; Cozzolino and Carrì, 2012; Dal Canto and Gurney, 1994; Ferri et al., 2006; Fujita et al., 1996; Grosskreutz et al., 2010; Menzies et al., 2002a; Sasaki et al., 2007; Shi et al., 2010). Impaired ATP production would damage motor neuron function due to their high metabolic demand (Menzies et al., 2002b), whilst impaired calcium buffering could lead to excitotoxicity (section 1.6.2). Additionally, inefficient mitochondrial function increases ROS production, which in turn causes oxidative damage to mitochondrial components and leads to greater mitochondrial dysfunction – a vicious cycle that could be central to motor neuron degeneration in ALS (Kaal et al., 2000; Robberecht, 2000).

1.6.4. Impaired Axonal Transport

Motor neurons have very long axons, and depend on efficient transport of protein, lipids, RNA and organelle cargoes along the axons via the cytoskeleton and associated motor proteins. Several ALS associated genes are involved in axonal transport and/or

the cytoskeleton (section 1.5.7), suggesting impaired axonal transport plays a role in ALS pathogenesis. Also, impaired axonal transport is measured in several different genetic models of ALS (Alami et al., 2014; De Vos et al., 2007; Mórotz et al., 2012), and is suggested to be an early pathogenic event (De Vos et al., 2008). Impaired mitochondrial transport also supports a 'dying-back' axonopathy model of ALS disease progression (Shi et al., 2010).

1.6.5. Dysregulated RNA Metabolism

Aberrant RNA metabolism is strongly implicated in ALS. Firstly, due to the number of ALS associated genes that are involved in RNA metabolism (section 1.5.5). Secondly, because the RNA processing protein TDP-43 forms the characteristic neuronal inclusions in most ALS cases (section 1.4), and a loss of TDP-43 function is strongly suggested to play a role in ALS pathogenesis (section 1.5.2). Additionally, the (G4C2)n repeat expansion in C9ORF72, which is the most common cause of ALS, is thought to exert direct RNA toxicity by disrupting RNA metabolism (section 1.7.2).

1.6.6. Impaired Protein Homeostasis

Insoluble protein aggregates are the pathological hallmark of ALS (section 1.4), suggesting impaired protein homeostasis. Cytoplasmic protein aggregates have recently been suggested to cause nucleocytoplasmic defects (Woerner et al., 2016), although the exact mechanism is still unknown. Other suggested toxic mechanisms associated with aggregated SOD1, TDP-43, FUS, and DPR proteins are described elsewhere (sections 1.5.1, 1.5.2, 1.5.3, and 1.7.3). Compellingly, SOD1, TDP-43 and DPR inclusions are all ubiquitinated, which suggests that the UPS and autophagy protein degradation systems are impaired in motor neurons in the majority of ALS cases. However, FUS-positive inclusions are not ubiquitinated, and suggests impairment to protein homeostasis may have less relevance in FUS-ALS. In addition, there is extensive evidence of ER stress in ALS patient CNS tissue, as well as cellular and animal models of ALS, suggesting an increased burden and inability to cope with misfolded proteins (Matus et al., 2013). Further, multiple other ALS associated genes are involved in protein trafficking or protein degradation (via the UPS and/or the autophagy), strongly implicating impaired protein homeostasis in ALS pathogenesis (section 1.5.6).

1.6.7. Non-Cell Autonomous Toxicity and Neuroinflammation

Although ALS is characterised by motor neuron cell degeneration, various studies implicate glial cell involvement in disease progression. Astrocytes derived from either SOD1-ALS, C9ORF72-ALS or SALS fibroblasts are toxic toward wildtype neurons in co-culture (Meyer et al., 2014), and murine SOD1-ALS models further suggest astrocytes and microglia are involved in non-cell autonomous toxicity (Beers et al., 2006; Boillée et al., 2006a; Boillée et al., 2006b; Clement et al., 2003; Nagai et al., 2007; Yamanaka et al., 2008b). Additionally, astrocytes and microglia are in an activated state, with an increase in pro-inflammatory cytokines, in both CNS tissue from ALS patients and SOD1 mouse models, suggesting neuroinflammation may contribute to ALS pathogenesis (Alexianu et al., 2001; Engelhardt and Appel, 1990; Engelhardt et al., 1993; Hall et al., 1998; Nagy et al., 1994; Philips and Robberecht, 2011).

1.7. C9ORF72 (G4C2)n Repeat Expansion Specific Pathomechanisms

It is currently unknown how the (G4C2)n repeat expansion within *C9ORF72* initiates ALS, and/or interacts with other pathomechanisms (section 1.6), but there are three mutually inclusive hypotheses: C9ORF72 haploinsufficiency, RNA toxicity, and dipeptide repeat (DPR) protein toxicity (Figure 1.2).

1.7.1. C9ORF72 Haploinsufficiency

Various reports demonstrate *C9ORF72* mRNA is reduced in CNS tissue, lymphoblast cells, and iPSC-derived neurons of patients containing the (G4C2)n repeat expansion, suggesting C9ORF72 haploinsufficiency as a pathogenic mechanism (Figure 1.2A) (Almeida et al., 2013; Belzil et al., 2013; Ciura et al., 2013; DeJesus-Hernandez et al., 2011; Donnelly et al., 2013; Mori et al., 2013c; Waite et al., 2014). Disrupted *C9ORF72* transcription could be caused by hypermethylation of the C9ORF72 locus and/or disruption of the core promoter region (Gijselinck et al., 2012; Gijselinck et al., 2015; Xi et al., 2013). In addition, one study shows reduced C9ORF72 protein in the frontal cortex of patients with the repeat expansion (Waite et al., 2014).

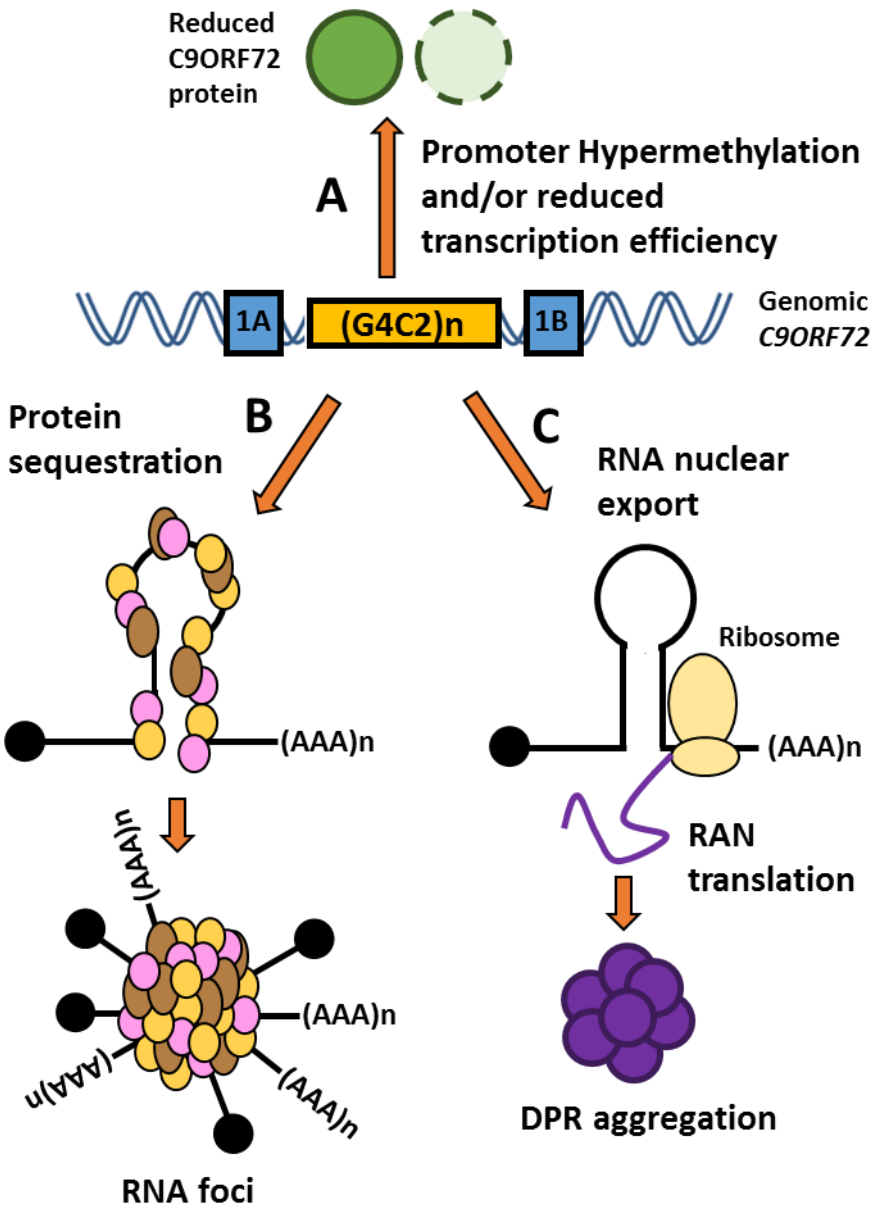


Figure 1.2 Pathomechanisms associated with C9ORF72 (G4C2)n repeat expansion. A) Transcriptional silencing of *C9ORF72* via hypermethylation or transcription abortion results in less C9ORF72 mRNA and protein, which could result in C9ORF72 haploinsufficiency. **B)** (G4C2)n and/or (C4G2)n repeat RNA binds and sequesters RNA binding proteins, and could result in dysregulated RNA metabolism and gene expression. **C)** The (G4C2)n and (C4G2)n repeat RNA is exported from the nucleus, and undergoes repeat associated non-ATG (RAN) translation producing dipeptide repeat (DPR) proteins. The DPR proteins aggregate and could cause toxicity.

The function of C9ORF72 protein is currently unknown, although it may play a role in endocytosis and autophagy (section 1.5.4) (Farg et al., 2014; Levine et al., 2013). Knock out of C9ORF72 homologues in zebrafish and *C. elegans* caused defects in motor function (Ciura et al., 2013; Therrien et al., 2013). However, when C9ORF72 knock-down

(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 (Koppers et al., 2015; Lagier-Tourenne et al., 2013). Furthermore, in human patients, the phenotype is no more severe in individuals with (G4C2)n repeat expansions in both *C9ORF72* alleles (Cooper-Knock et al., 2013; Fratta et al., 2013), and there are no known coding mutations in *C9ORF72* that result in disease (Harms et al., 2013). Taken together, these findings argue against haploinsufficiency as the only disease mechanism in *C9ORF72*-ALS, however, a partial loss of C9ORF72 protein function may disrupt autophagy, which could then exacerbate RNA toxicity and/or DPR toxicity.

1.7.2. RNA Toxicity

The second hypothesis suggests the RNA transcribed from the repeat may confer trans-dominant toxicity. The (G4C2)n repeat expansion is transcribed in both sense and antisense directions, and forms sense (G4C2)n and antisense (C4G2)n RNA foci respectively. These RNA foci are predominantly nuclear, but rare cytoplasmic RNA foci are also detected in post mitotic cells which suggests the intronic (G4C2)n and (C4G2)n containing transcripts are abberantly licensed for nuclear export (Cooper-Knock et al., 2015b; DeJesus-Hernandez et al., 2011; Gendron et al., 2013; Mizielinska et al., 2013; Zu et al., 2013). The RNA foci are present in C9ORF72-ALS/FTD brain and spinal cord (DeJesus-Hernandez et al., 2011; Gendron et al., 2013; Mizielinska et al., 2013; Zu et al., 2013), and also in neuronal cells differentiated from C9ORF72-ALS/FTD patient derived iPSC (Almeida et al., 2013; Donnelly et al., 2013; Sareen et al., 2013). Interestingly, the relative abundancy of sense and antisense RNA foci varies in different neuronal populations in the CNS (Cooper-Knock et al., 2015b). Further, antisense RNA foci, but not sense RNA foci, are associated with TDP-43 nuclear loss in the motor neurons suggesting the antisense RNA foci and/or derived DPR (section 1.7.3) are the pathogenic species in C9ORF72-ALS (Cooper-Knock et al., 2015b).

RNA foci potentially bind and sequester essential RNA binding proteins (RBP) conferring a trans-dominant toxicity by disrupting RNA processing, as seen in other repeat expansion disorders (Figure 1.2B) (Renoux and Todd, 2012). Therefore, several groups have performed *in vitro* pull downs using (G4C2)n and/or (C4G2)n repeat RNA to identify candidate RBP (Cooper-Knock et al., 2014b; Donnelly et al., 2013; Haeusler et al., 2014;

Lee et al., 2013; Mori et al., 2013b; Rossi et al., 2015; Xu et al., 2013). Several RBP were found to bind the repeat RNA including ADARB2, ALYREF, HnRNP A3, HnRNP H, NCL, PURA, SRSF1, and SRSF2. Subsequently, many of these RBP were shown to co-localise with the (G4C2)n and/or (C4G2)n RNA foci in either cell models or C9ORF72-ALS post mortem tissue, suggesting *in vivo* sequestration (Cooper-Knock et al., 2014b; Donnelly et al., 2013; Haeusler et al., 2014; Lee et al., 2013; Mori et al., 2013b; Rossi et al., 2015; Xu et al., 2013). Further, transcriptomic analysis of C9ORF72-ALS/FTD patient CNS tissue showed dysregulation in RNA splicing and processing, which may be caused by sequestration of RBP (Cooper-Knock et al., 2015a; Prudencio et al., 2015). In addition, several cellular and animal models have shown a gain of function toxicity from the (G4C2)n repeat, but these will be reviewed in section 1.8. Lastly, dysregulated RNA metabolism is strongly implicated in ALS pathogenesis (section 1.6.5), indicating RNA toxicity is likely involved in C9ORF72-ALS/FTD pathogenesis.

1.7.3. Dipeptide Repeat (DPR) Protein Toxicity

The last hypothesis suggests translation products from the repeat expansion may confer direct toxicity. The (G4C2)n and (C4G2)n repeat RNA also undergoes unconventional repeat associated non-ATG (RAN) translation in both the sense and antisense directions, forming aggregation-prone DPR proteins (Figure 1.2C) (Ash et al., 2013; Mori et al., 2013a; Mori et al., 2013c; Zu et al., 2013). Interestingly, the (G4C2)n repeat expansion causes intron 1 retention in *C9ORF72* mRNA, and could explain how the intronic (G4C2)n repeat becomes aberrantly licensed for nuclear export to the cytoplasm, where it then undergoes RAN translation (Niblock et al., 2016). A similar mechanism could also exist in the antisense (C4G2)n repeat containing transcripts. RAN translation occurs in all reading frames yielding five species of DPR: (GA)n and (GR)n are translated from the sense transcript, (AP)n and (PR)n from the antisense transcript, and (GP)n is translated from both sense and antisense transcripts (Figure 1.3).

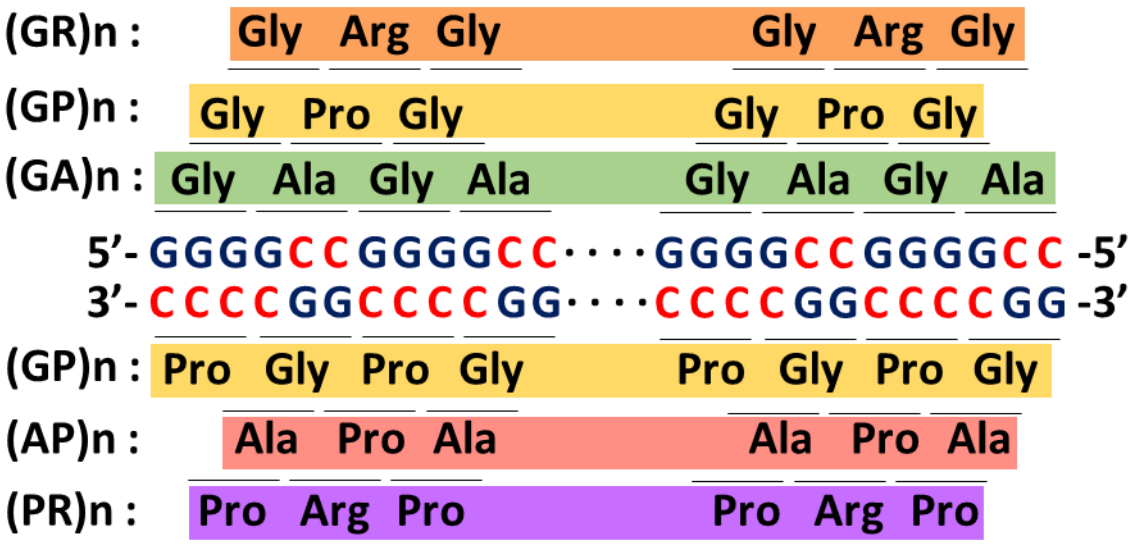


Figure 1.3 Schematic representation of DPR proteins generated by RAN translation from sense (G4C2)n and antisense (C4G2)n repeats in all reading frames.

DPR form insoluble inclusions in C9ORF72-ALS/FTD patient CNS tissue (section 1.4) (Ash et al., 2013; Mori et al., 2013a; Mori et al., 2013c; Zu et al., 2013). DPR proteins are toxic in cultured cells, and cause neurodegeneration in *Drosophila* models (section 1.8) (Kwon et al., 2014; May et al., 2014; Mizielinska et al., 2014; Tao et al., 2015; Wen et al., 2014; Yang et al., 2015; Zhang et al., 2014; Zu et al., 2013). 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 (Kwon et al., 2014; Mizielinska et al., 2014; Tao et al., 2015; Wen et al., 2014; Yang et al., 2015). 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 (Mizielinska et al., 2014; Tran et al., 2015). However in C9ORF72-ALS patients, DPR load is much lower in spinal motor neurons compared to other unaffected regions of CNS, and TDP-43 inclusions rarely co-localise with DPR suggesting they may not be the primary toxic insult in motor neuron degeneration (Davidson et al., 2015; Gomez-Deza et al., 2015; Mackenzie et al., 2013; Mackenzie et al., 2015). In fact, the expression level of DPR is likely to be very low in C9ORF72-ALS/FTD patient CNS, and the toxicity observed in model systems may be artificial due to high expression levels (Davidson et al., 2015).

1.8. C9ORF72-ALS Cellular and Animal Models

Over the last five years, many cellular and animal models of C9ORF72-ALS have been generated, and as a result, we have a much better understanding of how the (G4C2)n repeat expansion may cause C9ORF72-ALS pathogenesis. The models described are either loss or gain of function, and these are summarised with their major findings in Table 1.2. Loss of function models were used to investigate the function of C9ORF72 protein, and the potential effect of C9ORF72 haploinsufficiency. Currently, far fewer studies have interrogated the C9ORF72 loss of function hypothesis, and the most sophisticated loss of function models suggest C9ORF72 haploinsufficiency is not the major toxic insult in C9ORF72-ALS. The gain of function models however, have been used to study the biology and toxicity of the (G4C2)n repeat by studying the RNA, DPR or both. Initial studies looking at the (G4C2)n repeat showed clear gain of function toxicity, whilst more recent studies have used more sophisticated models to try and separate out the relative contributions of each potentially toxic species.

Table 1.2 Cellular and animal models of C9ORF72-ALS. The models are divided into loss of C9ORF72 function, and gain of (G4C2)n function, which are then separated into models that interrogate the (G4C2)n repeat, the RNA only or the DPR only. A brief summary of the model is included, the main findings discovered using the model, and the reference.

Model	Main Findings	Reference
Loss of function models		
Cellular		
C9ORF72 siRNA transfected into Neuro2a and SHSY5Y cells	Impaired endosomal trafficking and autophagy.	(Farg et al., 2014)
C9ORF72 shRNA lentiviral transduced into rat cortical and motor neurons	No toxicity.	(Wen et al., 2014)
Animal		
C9ORF72 knock out *C.elegans*	Age dependent motor defects and degeneration.	(Therrien et al., 2013)
C9ORF72 knock-down Zebrafish	Motor neuron axononopathy and motor function defects.	(Ciura et al., 2013)
Mouse with *C9ORF72* sense ASO delivered to CNS	No C9ORF72-ALS/FTD associated pathology, no motor defects, and minimal transcriptomic changes.	(Lagier-Tourenne et al., 2013)
Mouse with glial and motor neuron *C9ORF72* knock out	No motor neuron degeneration or motor defects.	(Koppers et al., 2015)
C9ORF72 knock out mouse	Robust immune phenotype. Mild motor impairment.	(Atanasio et al., 2016)
C9ORF72 knock out mouse	Impaired macrophage and microglial function.	(O'Rourke et al., 2016)
(G4C2)n Gain of function models		
Cellular		
(G4C2)n repeats transfected into Neuro2a cells	Toxicity, which was rescued by PURA overexpression.	(Xu et al., 2013)
(G4C2)n repeats transfected into HEK293, HeLa, SYSHY5Y and mouse cortical neurons	Neurotoxicity. Sense RNA foci colocalised with HnRNP H, SRSF1, and SRSF2.	(Lee et al., 2013)
(G4C2)n constructs tranfected into HEK293 cells	Length dependent expression of DPR via RAN translation.	(Mori et al., 2013c)
(C4G2)n repeats transfected into HEK293 cells	Antisense RNA foci and DPR produced.	(Gendron et al., 2013)
(C4G2)n repeats transfected into HEK293 cells	Antisense repeats cause toxicity.	(Zu et al., 2013)

(G4C2)n repeats transfected into HeLa and NSC34 cells	RNA foci colocalised with eIF2a, FUS, HnRNP H, and ILF3. Altered PURA distribution, stress granule dynamics, translation inhibition and mRNA nuclear retention.	(Rossi et al., 2015)
(G4C2)n constructs transfected into HEK293 and HeLa cells	Nuclear retention of mRNA.	(Freibaum et al., 2015)
(G4C2)n constructs transfected into S2 cells	Disrupted nucleocytoplasmic transport	(Zhang et al., 2015)
(G4C2)n constructs transfected into HEK293 cells	Length dependent repeat instability. Whole (G4C2) units gained or lost during expansion or contractions. Repeat caused replication fork stalling.	(Thys and Wang, 2015)
Methylated and unmethylated (G4C2)66 transfected into HEK293 cells	Methylation of the repeat reduced RNA foci and DPR production.	(Bauer, 2016)
Animal		
Drosophila expressing (G4C2)30 in the nervous system	Embryonic lethality	(Xu et al., 2013)
Drosophila expressing (G4C2)30 in the motor neurons	Motor defects.	(Xu et al., 2013)
Drosophila expressing (G4C2)30 in the eyes	Eye degeneration, which was rescued by PURA. Genetic manipulation of nucleocytoplasmic transport modifies toxicity.	(Xu et al., 2013; Zhang et al., 2015)
Drosophila expressing (G4C2)n in the eyes	Neurodegeneration associated with DPR (RNA only models showed no neurodegeneration).	(Mizielinska et al., 2014)
Drosophila expressing a *C9ORF72* minigene with (G4C2)n in intronic context	Sense RNA foci, but no DPR in glia and motor neurons. No neurodegeneration or splicing defects.	(Tran et al., 2015)
Drosophila expressing (G4C2)36	Developmental lethality.	(Tran et al., 2015)
Drosophila expressing (G4C2)36 in the eyes	Cytoplasmic RNA foci and DPR in glia and neurons. Neurodegeneration - taken together with intronic model, toxicity is associated with DPR.	(Tran et al., 2015)
Drosophila expressing (G4C2)n in the eyes	Length dependent neurodegeneration	(Freibaum et al., 2015)
Drosophila expressing (G4C2)n in the motor neurons or nerves system	Developmental lethality, locomotor defects, NMJ abnormalities.	(Freibaum et al., 2015)
Drosophila expressing (G4C2)n in the salivary glands	Nucleocytoplasmic defects and mRNA nuclear retention.	(Freibaum et al., 2015)
Mouse with tetracycline inducible, HnRNP promoter driven (G4C2)80	Ubiquitin positive inclusions, but no TDP-43 or DPR pathology.	(Hukema et al., 2014)
Mouse expressing *C9ORF72* gene containing approx. (G4C2)500	Sense and antisense RNA foci, and DPR throughout CNS. No alteration to survival, motor function or cognitive function.	(Peters et al., 2015)

Mouse expressing *C9ORF72* gene containing approx. (G4C2)800	Sense and antisense RNA foci, and DPR throughout CNS. DPR inclusions accumulate with age. No neurodegeneration or motor defects.	(O'Rourke et al., 2015)
Mouse expressing (G4C2)66 in CNS via AAV delivery	Sense RNA foci, sense derived DPR inclusions, phosphorylated TDP-43 inclusions in cortex and hippocampus. Neuronal cell loss and astrogliosis. Behavioural phenotype similar to C9ORF72-ALS/FTD.	(Chew et al., 2015)
(G4C2)n RNA only models		
Cellular		
(G4C2)n intronic constructs transduced via LV into rat cortical, motor and hippocampal neurons	(G4C2)42 repeats in intronic context caused toxicity to cortical and motor neurons independently of RAN translation.	(Wen et al., 2014)
Animal		
Drosophila expressing RNA only (G4C2)n constructs in the eyes	No neurodegeneration.	(Mizielinska et al., 2014)
DPR only models		
Cellular		
Synthetic (PR)20 and (GR)20 applied to U2OS cells and astrocytes	(PR)20 and (GR)20 cause toxicity, localise to nucleolus and impair rRNA biogenesis	(Kwon et al., 2014)
DPR constructs tranfected into HEK293 cells, or transduced via LV into rat hippocampal and cortical neurons	(GA)n formed p62 positive inclusions, and caused neurotoxicity.	(May et al., 2014)
DPR constructs transfected into NSC34 cells, rat cortical and motor neurons, and human iPSC derived neurons	(GR)n and (PR)n caused neurotoxicity.	(Wen et al., 2014)
DPR constructs transfected into HEK293 cells and mouse cortical neurons	(GA)n formed ubiquitinated p62 positive inclusions, caused toxicity, ER stress, and impaired the proteasome.	(Zhang et al., 2014)
DPR constructs transfected into HEK293 and NSC34 cells	Only arginine rich DPR were toxic. (PR)n and (GR)n caused nucleolar stress, and inhibit stress granule formation.	(Tao et al., 2015)
DPR constructs transduced via LV into rat hippocampal and cortical neurons	(PR)n and (GR)n form nucleolar inclusions, whilst GA forms p62 positive inclusions.	(Schludi et al., 2015)
(PR)n and (GR)n transformed into yeast	Length dependent toxicity of (PR)n and (GR)n. Genetic screening suggested nucleocytoplasmic transport modifies (PR)n induced toxicity.	(Jovičić et al., 2015)
(PR)50 transduced via LV into mouse cortical neurons	(PR)50 caused toxicity which could be partly rescued by *KPNA3* cotransduction.	(Jovičić et al., 2015)
Synthetic (GA)15 applied to, and (GA)80 constructs transfected into BE2C and Neuro2a cells	(GA)15 was neurotoxic, and could be transmitted between cells.	(Chang et al., 2016)

Animal		
Drosophila expressing DPR in the eyes	Only arginine rich DPR cause emrbyonic lethality and neurodegeneration.	(Mizielinska et al., 2014)
Drosophila expressing (PR)50, (GA)50 or (PA)50 in the eyes	(PR)50 caused neurodegeneration.	(Wen et al., 2014)
Drosophila expressing (PR)50, (GA)50 or (PA)50 in the motor neurons	(PR)50 caused developmental lethality.	(Wen et al., 2014)
Drosophila expressing (PR)80, (GA)80 or (PA)80 in the eyes	(PR)80 and (GR)80 cause lethality and neurodegeneration. (GR)80 suppresses Notch signalling.	(Yang et al., 2015)
Drosophila expressing DPR in the eyes	(GR)n caused developmental lethality, and neurodegeneration. Genetic screening suggested nucleocytoplasmic transport modifies (GR)n induced toxicity.	(Freibaum et al., 2015)
C9ORF72-ALS patient derived models		
Fibroblasts	Transcriptomic changes. *C9ORF72* sense ASO ablated sense RNA foci, but did not restore transcriptomic changes.	(Lagier-Tourenne et al., 2013)
iPSC derived motor neurons	RNA foci colocalised with HnRNP A1 and PURA. Transcriptomic changes, which could be ameliorated with *C9ORF72* sense ASO. Altered electrophysiology.	(Sareen et al., 2013)
iPSC derived neurons	RNA foci colocalised with ADARB2. Transcriptomic changes and susceptibility to glutamte mediated toxicity, both of which were restored by *C9ORF72* sense ASO.	(Donnelly et al., 2013)
iPSC derived neurons	Repeat instability. No TDP-43 mislocalisation. Cells were more susceptible to stress	(Almeida et al., 2013)
Fibroblasts, lymphoblasts, iPSC derived motor neurons	Morphological and functional signs of nucleolar stress	(Haeusler et al., 2014)
iPSC derived motor neurons	Altered electrophysiology.	(Devlin et al., 2015)
iPSC derived neurons	*C9ORF72* mRNA variant 2 was reduced from repeat expansion allele. Repeat expansion did not affect *C9ORF72* mRNA splicing.	(Tran et al., 2015)
iPSC derived neurons	Disrupted nuclear import, which was rescued with *C9ORF72* sense ASO	(Zhang et al., 2015)
Fibroblast derived iNeurons	Disrupted nucleocytoplasmic transport.	(Jovičić et al., 2015)
iPSC derived cortical neurons	Nuclear retention of mRNA.	(Freibaum et al., 2015)
iPSC derived neurons	Notch signalling was dysregulated.	(Yang et al., 2015)
iPSC derived neurons and motor neurons	Repeat instability. Repeat was progressively methylated during differentiation.	(Esanov et al., 2016)

1.8.1. Loss of Function Models

Several studies show a reduction in *C9ORF72* expression in C9ORF72-ALS patients (section 1.7.1), and therefore a loss of C9ORF72 function may cause or contribute to C9ORF72-ALS pathogenesis. Initial zebrafish and *C. elegans* loss of function models showed knock out of *C9ORF72* homologues caused motor function defects (Ciura et al., 2013; Therrien et al., 2013). Additionally, *C9ORF72* knock-down via siRNA transfection in Neuro2a and SHSY5Y neuronal cell lines caused defects in endosomal trafficking and autophagy (Farg et al., 2014). However, in a more sophisticated loss of function model, in which *C9ORF72* knock-down was confined to the glia and motor neurons of adult mice, neither motor neuron degeneration or motor defects were observed (Koppers et al., 2015). Also, in different *C9ORF72* knock out mouse models, either no or mild motor defects were described (Atanasio et al., 2016; O'Rourke et al., 2016). Further, *C9ORF72* knock-down by antisense oligonucleotide (ASO) treatment does not cause toxicity in cell and animal models (Donnelly et al., 2013; Lagier-Tourenne et al., 2013; O'Rourke et al., 2015; Sareen et al., 2013; Zhang et al., 2015) and taken together suggests C9ORF72 loss of function is not the primary toxic insult derived from the (G4C2)n repeat expansion in *C9ORF72*. However, we cannot rule out a contributory effect of C9ORF72 loss of function in ALS pathogenesis, since no studies have looked at a potential additive effect of C9ORF72 loss of function combined with a (G4C2)n gain of function toxicity.

1.8.2. Gain of Function Models

1.8.2.1. Toxicity associated with the (G4C2)n

Multiple studies have used cellular and animal models to interrogate the gain of function toxicity associated with the (G4C2)n. The earliest and most basic models were cell lines transfected with (G4C2)n repeat constructs. Interrupted and pure (G4C2)n repeat constructs cause toxicity when transfected into neuronal cell lines, but not in HEK293 cells (Lee et al., 2013; Xu et al., 2013). In (G4C2)n transfected neuronal cells, various RNA binding proteins were shown to colocalise with RNA foci, many of which have been validated in C9ORF72-ALS CNS tissue (Lee et al., 2013; Rossi et al., 2015; Xu et al., 2013). Additionally, a construct expressing 31 G4C2 repeats (abbreviated as (G4C2)31) was transfected in NSC34 and HeLa cells, and affected PURA and FMRP distribution, and

inhibited stress granule formation, translation and mRNA export (Rossi et al., 2015). Transfection of antisense (C4G2)66 constructs into HEK293 cells also caused toxicity (Zu et al., 2013). However, the (G4C2)n expression level is much higher in these models than in C9ORF72-ALS patients, and only short term effects can be studied, making transfection models fairly poor at replicating the actual physiological effects of the (G4C2)n repeat.

Several groups have also generated *Drosophila* and mouse models that express the (G4C2)n. *Drosophila* models that expressed an interrupted (G4C2)30 repeat specifically in the eyes or motor neurons displayed neurodegeneration and motor defects respectively (Xu et al., 2013). Additionally, the eye neurodegeneration could be rescued by PURA overexpression in the *Drosophila*. In a separate study, *Drosophila* models with either eye or pan-neuronal expression of either (G4C2)8 or (G4C2)58 repeats displayed length dependent eye degeneration or motor defects with neuromuscular junction defects respectively (Freibaum et al., 2015). Additionally, in both the (G4C2)30 and (G4C2)58 *Drosophila* models with eye specific expression, genetic screening identified nucleocytoplasmic transport as a potent modifier of (G4C2)n induced toxicity (Freibaum et al., 2015; Zhang et al., 2015). However, the (G4C2)n repeat is overexpressed in these *Drosophila* models, and the physiological relevance of fly eye degeneration to human motor neuron degeneration is questionable. Although, the defects in nucleocytoplasmic transport identified in the *Drosophila* models have also been validated in patient derived neuronal cells (section 1.8.2.2).

In addition to the *Drosophila* (G4C2)n gain of function models described above, four mouse models have also been generated to study (G4C2)n gain of function toxicity. One mouse model expressed a (G4C2)80 repeat from a tetracycline inducible HnRNP promoter, and displayed ubiquitin positive inclusions in the CNS, but no TDP-43 inclusions, DPR pathology, neurodegeneration or behavioural defects (Hukema et al., 2014). Additionally, two other mouse models were generated that carry a patient derived *C9ORF72* gene containing either 500 or 800 (G4C2)n repeats (O'Rourke et al., 2015; Peters et al., 2015). Sense and antisense RNA foci as well as DPR were detected throughout the CNS, but there was no neurodegeneration, or alteration in survival, motor function or cognitive function in either of these mouse models (O'Rourke et al.,

2015; Peters et al., 2015). Although the aforementioned mouse models do not support a (G4C2)n gain of function toxicity, another mouse model replicates several aspects of C9ORF72-ALS/FTD pathology (Chew et al., 2015). A (G4C2)66 repeat expressed from a strong promoter was delivered to the mouse CNS via adeno-associated virus, and brains of 6-month old mice subsequently displayed sense RNA foci, and sense derived DPR inclusions throughout the CNS. Additionally, the brains contained TDP-43 inclusions, cortical and cerebellar degeneration, and astrogliosis. Further, these mice also developed behavioural abnormalities similar to the clinical symptoms of C9ORF72-ALS/FTD patients, including hyperactivity, anxiety, antisocial behaviour, and motor defects (Chew et al., 2015). Therefore, the mouse models suggest high level expression of the (G4C2)n repeat expansion is capable of causing neurodegeneration and C9ORF72-ALS/FTD pathology. Equally, since ALS develops in older age, the (G4C2)n repeat may only cause toxicity in the aged human CNS environment, which is not replicated in the aforementioned mouse models.

1.8.2.2. Patient Derived Cellular Models

More advanced cellular models of C9ORF72-ALS are derived from patient cells. Some of the more sophisticated models convert fibroblasts to iPSC, which can be subsequently differentiated to neuronal or motor neuronal cultures. Whilst the (G4C2)n expression is certainly much closer to physiologically relevant levels, these models are not reductionist and are affected by (potential) C9ORF72 haploinsufficiency, (G4C2)n/(C4G2)n RNA, and all DPR species simultaneously. Similarly to transfected cells, C9ORF72-ALS patient derived cell lines show colocalisation between (G4C2)n RNA foci and various RNA binding proteins, including ADAR2B, HnRNPA1, and PURA (Donnelly et al., 2013; Sareen et al., 2013). Additionally, many functional defects are described in C9ORF72-ALS patient derived cells including nucleolar stress, susceptibility to glutamate mediated excitotoxicity and other cellular stresses, transcriptomic alterations, altered electrophysiology, and impaired nucleocytoplasmic transport (Almeida et al., 2013; Devlin et al., 2015; Donnelly et al., 2013; Freibaum et al., 2015; Haeusler et al., 2014; Lagier-Tourenne et al., 2013; Sareen et al., 2013; Zhang et al., 2015). Further, in many of these studies, ASO targeted to the *C9ORF72* sense strand were used, which ablated sense derived RNA foci and DPR, and also partly or completely ameliorated the

functional defects. This strongly argues for a gain of function toxicity, and against a loss of function toxicity.

1.8.2.3. DPR Toxicity Models

Upon discovery of the (G4C2)n RAN translation products in C9ORF72-ALS, several groups have examined the toxicity of the DPR proteins. DPR constructs that are independent of the (G4C2)n repeat sequence have been transfected into various cell lines including HEK293 and NSC34 cells, or transduced via lentiviral vectors into primary neuronal cultures from rodents. The findings suggest that all the DPR species possess neurotoxicity, although the arginine containing (GR)n and (PR)n species appear to be the most toxic (Jovičić et al., 2015; Kwon et al., 2014; May et al., 2014; Schludi et al., 2015; Tao et al., 2015; Wen et al., 2014; Zhang et al., 2014). Further, the arginine rich DPR proteins are thought to mimic the SR domain found in splicing factors, which localises proteins to the nucleolus (Kwon et al., 2014). However, the (GR)n and (PR)n lack serine residues, so cannot be phosphorylated, and are therefore retained in the nucleolus. As a result, the arginine rich DPR proteins are suggested to cause nucleolar stress and RNA splicing dysregulation (Kwon et al., 2014).

Similarly in *Drosophila* models, expression of the arginine rich DPR proteins in the eye caused neurodegeneration, and is suggested to disrupt Notch signalling (Freibaum et al., 2015; Yang et al., 2015). In addition, two elegant studies using *Drosophila* models showed that the toxicity of the (G4C2)n repeats was dependent on the production of DPR, and not (G4C2)n RNA foci (Mizielinska et al., 2014; Tran et al., 2015). In the first study, *Drosophila* expressing either pure (G4C2)36 or 103 repeats in the eye displayed neurodegeneration, whilst flies expressing 'RNA only' repeats (which contained stop codons interspersed throughout the construct) did not develop neurodegeneration (Mizielinska et al., 2014). Importantly, the pure repeats produced DPR proteins whilst the 'RNA only' repeats did not (Mizielinska et al., 2014). Additional DPR only *Drosophila* models showed the arginine rich DPR species caused the most severe eye neurodegeneration (Mizielinska et al., 2014). *Drosophila* expressing (G4C2)160 in an intronic context displayed nuclear sense RNA foci in the glutamatergic neurons and glia cells, but displayed no signs of neurodegeneration or toxicity (Tran et al., 2015). However, a polyadenylated (G4C2)36 repeat mRNA expressed in a different *Drosophila*

model displayed far greater toxicity, and this was associated with (G4C2)n RNA export and much greater DPR expression in the affected cells (Tran et al., 2015).

1.8.3. Other Genetic Phenomena Associated with the (G4C2)n Repeat

Lastly, other pure biological aspects of the (G4C2)n repeat have been studied in various cellular and animal models. Several C9ORF72-ALS patient derived models show somatic instability in the (G4C2)n, specifically during differentiation of fibroblasts to iPSC, and from iPSC to neuronal and motor neuronal cultures (Almeida et al., 2013; Esanov et al., 2016; Sareen et al., 2013). Further, HEK293 cells transfected with (G4C2)n repeat constructs showed length dependent instability, that whole (G4C2) units were lost or gained via contraction and expansions respectively, and also replication efficiency of the (G4C2)n decreased in a length dependent manner (Thys and Wang, 2015). C9ORF72-ALS patient derived cell models also show reduced expression of C9ORF72 mRNA variant 2 from the allele containing the repeat expansion (Almeida et al., 2013; Donnelly et al., 2013; Tran et al., 2015). The reduced C9ORF72 transcription is suggested to be caused by (G4C2)n expansion methylation, and in HEK293 cells that were transfected with methylated or unmethylated (G4C2)66 constructs, methylation of the (G4C2)n repeat reduced both RNA foci and DPR proteins (Bauer, 2016).

1.9. Overall Aims and Objectives

ALS is currently incurable, and therefore therapeutic agents which may slow or even prevent disease progression are vital. A (G4C2)n repeat expansion in C9ORF72 is the most common genetic cause of ALS, but it is still not fully understood how the (G4C2)n repeat expansion leads to motor neuron injury and ALS pathogenesis. Therefore, by using cellular models of C9ORF72-ALS, we may discover important therapeutic targets that could be directly translated into therapies for C9ORF72-ALS patients, and even more broadly for ALS patients as a whole.

The overall aims of this project were to:

1) Generate motor neuron like cellular models with tetracycline inducible expression of (G4C2)n repeats.

2) Characterise the cellular models for features of ALS and C9ORF72-ALS pathology, and interrogate biochemical effects of (G4C2)n repeat expression.

3) Perform transcriptomic analysis to identify biological pathways and functions that may be dysregulated by (G4C2)n expression.

4) Interrogate any potential therapeutic targets in rescue assays.

Chapter 2. Materials and Methods

2.1. Materials

Unless stated otherwise, all general materials and reagents were purchased from Thermo Fisher Scientific® Inc. Solutions were autoclaved in an MP25 autoclave (Rodwell) at 121°C, 15psi for 15mins where stated.

2.1.1. General Materials

Diethylpyrocarbonate (DEPC) was purchased from Applichem. Dextran Sulphate, Sodium salt, molecular biology grade was purchased from Calbiochem. Bromophenol blue powder; Chloroform; Glycerol; Methanol; and Sodium Chloride (NaCl), analytical grade; as well as Sodium Dodecyl Sulphate (SDS), general purpose grade were purchased from Fisher Scientific. Amersham™ Protran™ supported 0.2µm nitrocellulose membrane was purchased from GE Healthcare Life Sciences. Glycine; HEPES free acid; Tris (molecular biology grade); and Tris-HCl (molecular biology grade) were purchased from Melford Laboratories Ltd. Immobilon®-P Transfer Membrane, PVDF, 0.45µm was purchased from Merk Millipore. Acetic acid; Ammonium Persulphate (APS); β-Mercaptoethanol; Dimethylformamide (DMF); Dithiothreitol (DTT); Formamide; Sodium Citrate Dehydrate; Sodium Phosphate Monobasic (NaH_2PO_4); Sodium Phosphate Dibasic (Na_2HPO_4); TEMED; Triton™ X-100; Trizma® base; and Tween® 20 were purchased from Sigma-Aldrich. EDTA, Disodium salt, Dihydrate was purchased from Thermo Fisher Scientific® Inc. Oxoid™ Phosphate Buffered Saline Tablets (Dulbecco A) were purchased from Thermo Scientific.

2.1.2. General Buffers and Solutions

2.1.2.1. Phosphate Buffered Saline (PBS)

The PBS consisted of 137mM NaCl, 3mM KCl, 8mM Na_2HPO_4, and 1.5mM KH_2PO_4, at pH 7.3: 10 Oxoid™ PBS tablets were dissolved in 1L dH_2O, and the solution was autoclaved.

2.1.2.2. 1X Tris Acetate EDTA (TAE) Buffer

The 1X TAE buffer consisted of 40mM Tris, 40mM Acetate, and 1mM EDTA, at pH 8.0.

2.1.2.3. 20X Saline Sodium Citrate (SSC) Buffer

The 20X SSC buffer consisted of 3M NaCl, and 0.3M Sodium Citrate, at pH 7.0, and was DEPC treated.

2.1.2.4. 1M Sodium Phosphate Buffer pH 7.0

The 1M Sodium Phosphate buffer was comprised of 39% (v/v) 1M NaH_2PO_4, and 61% (v/v) 1M Na_2HPO_4, at pH 7.0, and was DEPC treated.

2.1.2.5. Diethylpyrocarbonate (DEPC) Treatment of Solutions

Buffers and solutions used for RNA work were pre-treated with DEPC to inhibit RNase activity. 0.001 volumes of DEPC was added to the solution and stirred for at least 1h. The solution was then autoclaved to degrade the DEPC.

2.1.3. Molecular Biology Materials

HyperLadder™ I, HyperLadder™ II, HyperLadder™ III, HyperLadder™ IV, and HyperLadder™ V molecular weight markers and agarose (molecular grade) were purchased from Bioline. T4 Polynucleotide Kinase (PNK); Kinase 10X Reaction Buffer; alkaline phosphatase, shrimp (SAP); and 10X SAP Buffer were purchased from Boehringer Ingelheim. LB Broth, Miller (molecular genetics granular); and LB Agar, Miller (molecular genetics powder) were purchased from Fisher Scientific® Inc. DNA Polymerase I, large fragment (Klenow), and 10X Klenow Buffer were purchased from Invitrogen™. NucleoSpin Plasmid Kit (No Lid) was purchased from Macherey-Nagel. 5-α Competent *E.coli* (High Efficiency); β-10 Competent *E.coli* (High Efficiency); and SOC Outgrowth Media were purchased from New England BioLabs®. T4 DNA Ligase and 10X Ligase Buffer were purchased from Promega. QIAquick Gel Extraction Kit was purchased from QIAGEN. dATP; ethidium bromide; carbenicillin; spectinomycin; and synthesised TCGAC(G4C2)10 sense and ACGT(G2C4)10 antisense ssDNA oligonucleotides were purchased from Sigma-Aldrich. *Bam*HI, *Dra*I, *Eco*RV, *Hind*III, *Nco*I, *Sal*I, and *Xho*I FastDigest restriction enzymes; 10X FastDigest Green Buffer; 2X ReddyMix PCR Master Mix were purchased from Thermo Fisher Scientific® Inc. The Translate tool and Compute pI/Mw tool were freely available at http://web.expasy.org/translate/ and http://web.expasy.org/compute_pi/ respectively.

Table 2.1 Primers used in generation of the cell models. All primers used for molecular biology methods were purchased from Sigma-Aldrich.

Primer	Sequence
CMV-Forward	5'-CGCAAATGGGCGGTAGGCGTG
BGH-Reverse	5'-TAGAAGGCACAGTCGAGG

Table 2.2 Plasmid origins and manipulations. The individuals who provided the plasmid or performed the manipulation are referred to by their initials: Prof Stuart Wilson (SW); Dr Adrian Higginbottom (AH); Matthew Stopford (MS); Dr Padraig Mulchay (PM); Dr Kurt De Vos (KDV); and Dr Adrian Isaacs (AI).

Plasmid	Manipulations	Company/ Origin
(G4C2)n construction vectors		
pcDNA6.2-GW/EmGFP-miR	N/A	Gifted by SW
pcDNA6.2-GW/EmGFP-miR-(G4C2)10	(G4C2)10 construct inserted into *Sal*I and *Xho*I sites.	Generated by AH
pcDNA6.2-GW/EmGFP-miR-(G4C2)51	(G4C2)51 construct inserted into *Sal*I and *Xho*I sites.	Generated by AH
pcDNA6.2-GW/EmGFP-miR-(G4C2)102	(G4C2)102 construct inserted into *Sal*I and *Xho*I sites.	Generated by AH
pcDNA6.2-(G4C2)10	EmGFP removed from pcDNA6.2-GW/EmGFP-miR-(G4C2)10 via *Dra*I digest and relegation.	Generated by MS
pcDNA6.2-(G4C2)51	EmGFP removed from pcDNA6.2-GW/EmGFP-miR-(G4C2)10 via *Dra*I digest and relegation.	Generated by MS
pcDNA6.2-(G4C2)102	EmGFP removed from pcDNA6.2-GW/EmGFP-miR-(G4C2)10 via *Dra*I digest and relegation.	Generated by MS
pcDNA6.2-(C4G2)102	(G4C2)102 construct was inserted in the antisense orientation.	Generated by PM
FRT vectors		
pcDNA5/FRT/TO-GFP	pcDNA5/FRT/TO-GFP was cut with BamHI and XhoI, end filled and ligated.	Addgene
pcDNA5/FRT/TO	N/A	Invitrogen™
pcDNA5/FRT/TO-HIS	6 x HIS Tag inserted into *Hind*III and *Bam*HI restriction sites.	Addgene
pcDNA5/FRT/TO-HIS-(G4C2)10	(G4C2)10 cut from pcDNA6.2-(G4C2)10 using *Dra*I and *Xho*I restriction sites, and inserted into EcoRV and XhoI sites in pcDNA5/FRT/TO-HIS.	Generated by MS

pcDNA5/FRT/TO-HIS-(G4C2)51	(G4C2)51 cut from pcDNA6.2-(G4C2)51 using *Dra*I and *Xho*I restriction sites, and inserted into EcoRV and XhoI sites in pcDNA5/FRT/TO-HIS.	Generated by MS
pcDNA5/FRT/TO-HIS-(G4C2)102	(G4C2)102 cut from pcDNA6.2-(G4C2)102 using *Dra*I and *Xho*I restriction sites, and inserted into EcoRV and XhoI sites in pcDNA5/FRT/TO-HIS.	Generated by MS
pcDNA5/FRT/TO-(G4C2)10	HIS tag removed from pcDNA5/FRT/TO-HIS-(G4C2)10 by *Hind*III and *Bam*HI digest and relegation.	Generated by MS
pcDNA5/FRT/TO-(G4C2)51	HIS tag removed from pcDNA5/FRT/TO-HIS-(G4C2)51 by *Hind*III and *Bam*HI digest and relegation.	Generated by MS
pcDNA5/FRT/TO-(G4C2)102	HIS tag removed from pcDNA5/FRT/TO-HIS-(G4C2)102 by *Hind*III and *Bam*HI digest and relegation.	Generated by MS
Vectors for Flp-In™ T-REx™ cell line generation		
pPGKFLPobpA	N/A	Addgene
pFRT/*lac*Zeo	N/A	Invitrogen™
pcDNA6/TR	N/A	Invitrogen™
DPR expression vectors		
p(GA)68	N/A	Generated by KDV
p(AP)100, p(GR)100, p(PR)100	N/A	Gifted by AI

2.1.4. Cell Culture Materials

Gelatin powder was purchased from BDH biochemical. Tetracycline-free foetal bovine serum (FBS), South American origin, sterile filtered was purchased from Biosera. Puromycin dihydrocholride; and BpV (Phen) were purchased from Calbiochem. NSC34 cells were kindly gifted by Prof. Neil Cashman. Blasticidin S HCl powder; Flp-In™ T-REx™ Core Kit; Flp-In™ T-REx™ HEK293 cells; hygromycin B; Lipofectamine® 2000; Tetracycline hydrochloride powder; and Zeocin™ were purchased from Invitrogen™. Phenol red-free DMEM; Penicillin/Streptomycin; and 10X Trypsin, with versene were purchased from Lonza. 6-carboxy-2',7'-dichlorodihydrofluorescein diacetate, di(acetoxymethyl ester) (DCFDA); and ethidium homodimer-1 (EthD1) were purchased from Molecular Probes™, Life Technologies. Dulbecco's modified eagle medium (DMEM); FBS; menadione sodium

bisuphate; Polyethylenimine (PEI); and 5, 10, 15, 20-tetrakis(1-methyl-4-pyridino)porphyrin tetra(p-toluenesulfonate) (TMPyP) were purchased from Sigma-Aldrich. Opti-MEM® Reduced Serum Medium, no phenol red was purchased from Thermo Fisher Scientific® Inc.

2.1.5. Biochemical Assay Materials

2X Brilliant III qPCR Master Mix; and 2X Brilliant III SYBR Green qPCR Master Mix were purchased from Agilent Technologies Ltd. Protein Assay Dye Reagent Concentrate was purchased from Bio-Rad®. BioScript™ Reverse Transcriptase kit (which contains BioScript™ Reverse Transcriptase; and 5X Reaction Buffer); and RiboSafe RNase Inhibitor were purchased from Bioline. EZ-RNA kit (which contains denaturing solution, and extraction and phase separation solution) was purchased from Biological Industries. 5' TYE-563-labelled LNA sense (5'-CCCCGGCCCCGGCCCC) and (5'-GGGGCCGGGGCCGGGG) DNA probes (batch numbers 607323 and 515905 respectively) were purchased from Exiqon Inc. Bovine serum albumin (BSA) powder, Fraction V, microbiological grade were purchased from Fisher Scientific. Streptavidin Sepharose High Performance beads were purchased from GE Healthcare. Ultra-Pure ProtoGel® 30% (w/v) Acrylamide, 0.8% (w/v) Bis-Acrylamide Stock Solution (37.5:1), protein and sequencing electrophoresis grade; Prestained Blue Protein Ladder; and EZ-ECL Kit were purchased from Geneflow Ltd. Original dried skimmed milk powder was purchased from Marvel. DNase I recombinant, RNase-free; 10X DNase I Buffer; and Phosphatase inhibitor tablets (PhosSTOP™) (used at concentration stated by manufacturer) were purchased from Roche. Cyclohexamide, Paraformaldehyde (PFA), crystalline, reagent grade; SIGMAFAST™ Protease Inhibitor Cocktail (PIC) tablets, EDTA free (used at concentration stated by manufacturer); RNase A; Sodium Azide; and Thiozolyl Blue Tetrazolium Bromide (MTT) powder; were purchased from Sigma-Aldrich. TRIzol® LS Reagent was purchased from Thermo Fisher Scientific® Inc. Hard Set mounting medium with DAPI was purchased from Vector Laboratories Inc.

2.1.5.1. qRT-PCR Primers

Table 2.3 qRT-PCR primer details. Target gene and primer sequences are shown.

Primer	Sequence
Actb Forward	5’-CGGTTCCGATGCCCTGAGGCTCTT
Actb Reverse	5’-CGTCACACTTCATGATGGAATTGA
Gapdh Forward	5’-ATGGTGAAGGTCGGTGTGAA
Gapdh Reverse	5’-TGGCAACAATCTCCACTTTGC
Mouse 18S rRNA Forward	5’-GATGGTAGTCGCCGTGCC
Mouse 18S rRNA Reverse	5’-GCCTGCTGCCTTCCTTGG
Mouse 28S rRNA Forward	5’-AGAGGTAAACGGGTGGGGTC
Mouse 28S rRNA Reverse	5’-GGGGTCGGGAGGAACGG
Mouse 45S rRNA Forward	5’-CGTAGGGAAGTCGGTCGTTC
Mouse 45S rRNA Reverse	5’-GAGGGGGCTCCAGACATCC
Mouse 5.8S rRNA Forward	5’-ACTCGGCTCGTGCGTC
Mouse 5.8S rRNA Reverse	5’-CCGACGCTCAGACAGG

Table 2.4 PrimeTime® qPCR assay (Integrated DNA Technologies®) product details. Target gene, assay ID, RefSeq Accession number(s), gene region, primers sequences and probe sequences are shown.

Gene	PrimeTime® Assay ID	Ref Seq	Region	Primer Sequence	Probe Sequence
Actb	Mm.PT.58.33540333	NM_007393	Exon 1-2	5'-ATGCCGGAGCCGTTGTC-3' 5'-GCGAGCACAGCTTCTTTG-3'	5'-/56-FAM/CCGCCACCA/ZEN/GTTCGCCATG/3IABkFQ/-3'
Akt1	Mm.PT.58.8333433	NM_009652 NM_001165894	Exon 2-3	5'-GCCGTTCCTTGTAGCCAAT-3' 5'-GACGTAGCCATTGTGAAGGAG-3'	5'-/56-FAM/TATTCCCCT/ZEN/CGTTTGTGCAGCCA/3IABkFQ/-3'
Mtor	Mm.PT.58.28403918	NM_020009	Exon 46-47	5'-TGCATCACTCGTTCATCCTG-3' 5'-AAGTCATCACATCCAAGCAGA-3'	5'-/56-FAM/CCCATTGCT/ZEN/GCCCATCAGAGTCA/3IABkFQ/-3'
Pten	Mm.PT.56a.13345002	NM_008960	Exon 8-9	5'-TCATTGTCACTAACATCTGGAGTC-3' 5'-GCCAACCGATACTTCTCTCC-3'	5'-/56-FAM/TTGAACTGC/ZEN/TAGCCTCTGGATTTGATGG/3IABkFQ/-3'
Srsf1	Mm.PT.58.32558636.g	NM_173374 NM_001078167	Exon 2-3	5'-CATACATCACCTGCCTCACG-3' 5'-GGTCCGAGAACAGAGTGGT-3'	5'-/56-FAM/TGATCCTTT/ZEN/AAGTCCTGCCAGCTTCC/3IABkFQ/-3'
Srsf2	Mm.PT.58.7770373	NM_011358	Exon 2-3	5'-CCCAATGTCCTCTGTTAAGCC-3' 5'-CGAAGCGAGAGTCCAAGTCTA-3'	5'-/56-FAM/CAAGAGCCC/ZEN/ACCCAAGTCTCCAG/3IABkFQ/-3'
Trp53	Mm.PT.58.42581447	NM_011640 NM_001127233	Exon 8-10	5'-CTCCCGGAACATCTCGAAG-3' 5'-CAGGGAGCGCAAAGAGAG-3'	5'-/56-FAM/CGCCTCTCC/ZEN/CCCGCAAAAGAA/3IABkFQ/-3'

Table 2.5 Primary antibodies used for Immunoblotting. Antibody specificity, species, type, dilution, blocking agent (5% (w/v) Milk/TBST (M) or 5% (w/v) BSA/TBST (B)) and source are shown. The individuals who provided the antibody are referred to by their initials: Prof Stuart Wilson (SW); Prof Stuart Pickering-Brown (SPB); and Prof Dieter Edbauer (DE).

Antibody Specificity	Species/type	Dilution/blocking agent	Source/product number/clone number
Anti-β-actin	Mouse monoclonal	1/10000 (M)	Abcam, ab6276, clone AC-15
Anti-Akt	Rabbit polyclonal	1/1000 (M)	Cell Signalling, #9272
Anti-NCL	Rabbit polyclonal	1/2000 (M)	Proteintech, 10556-1-AP
Anti-p53	Mouse monoclonal	1/250 (M)	BD Pharmingen™, 554147
Anti-Phospho-Akt (Ser473)	Rabbit monoclonal	1/1000 (B)	Cell Signalling, #4060
Anti-Phospho-PTEN (Ser380)	Rabbit polyclonal	1/2000 (B)	Cell Signalling, #9551
Anti-poly(AP)	Rabbit	1/5000 (M)	SPB
Anti-poly(GA)	Mouse	1/500 (M)	DE
Anti-poly(GP)	Rabbit	1/5000 (M)	SPB
Anti-poly(GR)	Rabbit	1/5000 (M)	SPB
Anti-poly(PR)	Rabbit	1/5000 (M)	SPB
Anti-PTEN	Rabbit monoclonal	1/1000 (M)	Cell Signalling, #9188
Anti-RPL19	Rabbit polyclonal	1/2000 (M)	SW
Anti-RPL29	Rabbit polyclonal	1/500 (M)	Sigma-Aldrich, R0655
Anti-α-tubulin	Mouse monoclonal	1/10000 (M)	Sigma-Aldrich, ab7291, clone DM1A

Table 2.6 Secondary antibodies used for Immunoblotting. Antibody specificity, species, type, dilution, blocking agent (5% (w/v) Milk/TBST (M) or 5% (w/v) BSA/TBST (B)) and source are shown.

Antibody	Dilution	Source/product number/ clone number
Polyclonal Goat Anti-Mouse-IgG HRP	1/10000 (M)	Dako, P 0447
Polyclonal Goat Anti-Rabbit-IgG HRP	1/10000 (M)	Dako, P 0448

Table 2.7 Primary antibodies used for Immunocytochemistry. Antibody specificity, species, type, dilution (in 2% (w/v) BSA/PBS) and source are shown.

Antibody	Species/type	Dilution	Source/product number/clone number
HnRNP A1	Mouse monoclonal	1/500	Abcam, ab5832
HnRNP F/H	Mouse monoclonal	1/1000	Abcam, ab10689
SRSF1	Rabbit polyclonal	1/200	Abcam, ab38017
SRSF2	Mouse monoclonal	1/200	Abcam, ab11826
NCL	Rabbit polyclonal	1/200	Protientech, 10556-1-AP
Pur-α	Rabbit polyclonal	5µg/mL	Abcam, ab79936
ALYREF	Mouse monoclonal	1/500	Sigma-Aldrich, A9979
TDP-43	Rabbit polyclonal	1/200	Proteintech, 10782-2-AP

Table 2.8 Secondary antibodies used for Immunocytochemistry. Antibody specificity, species, type, dilution (in 2% (w/v) BSA/PBS) and source are shown.

Antibody	Dilution	Source/product number/clone number
Goat Anti-Rabbit IgG H&L (AlexaFluor® 488) preadsorbed	1/1000	Abcam, ab150081
Goat Anti-Rabbit IgG H&L (AlexaFluor® 594) preadsorbed	1/1000	Abcam, ab150084
Goat Anti-Mouse IgG H&L (AlexaFluor® 488) preadsorbed	1/1000	Abcam, ab150117
Goat Anti-Mouse IgG H&L (AlexaFluor® 594) preadsorbed	1/1000	Abcam, ab150120

2.1.6. Gene Expression Profiling Materials

The GeneChip® WT PLUS Reagent Kit (containing First-Strand Enzyme, First-Strand Buffer, Second-Strand Enzyme, Second-Strand Buffer, IVT Enzyme, IVT Buffer, 1mg/mL HeLa total RNA (Control RNA), 2nd-Cycle Primers, 2nd-Cycle ss-cDNA Enzyme, 2nd-Cycle ss-cDNA Buffer, RNase H, Nuclease-free water, magnetic Purification Beads, Poly-A Control Stock, Poly-A Control Dilution Buffer, 10X cDNA Fragmentation Buffer, 10units/µL UDG, 1,000 units/µL apurinic/apyrimidinic endonuclease (APE) 1, 5X TdT

Buffer, 30units/µL TdT, 5mM DNA Labelling Reagent, RNase-free water, 20X Eukaryotic Hybridisation Controls (*bioB*, *bioC*, *bioD*, *cre*), 3nM Control Oligo B2); the GeneChip® Hybridisation, Wash and Stain Kit (containing DMSO, 2X Hybridisation Mix, Stain Cocktail 1, Stain Cocktail 2, Array Holding Buffer, Wash Buffer A, and Wash Buffer B); GeneChip® Mouse Transcriptome Array 1.0 cartridges; GeneChip® Scanner 3000; GeneChip® Fluidics Station 450; GeneChip® Command Console Software; and Affymetrix® Expression Console; Affymetrix® Transcriptomics Analysis Console (TAC) Software, were all purchased from Affymetrix® Ltd. 5X Novex® Hi-Density TBE Sample Buffer; and 4-20% TBE Gel, 1.0mm were purchased from Life Technologies. 10mg/mL NeutrAvidin was purchased from Thermo Scientific. The Direct-zol™ RNA Kit (which contains TRI-Reagent®, Zymo-Spin™ II Columns, collection tubes, Wash Buffer, 1units/µL DNAse I, 10X DNAse I Reaction Buffer, DNAse/RNAse-free water, RNA Wash Buffer, Direct-zol RNA PreWash) was purchased from Zymo Research. The Database for Annotation and Visualisation and Integrated Discovery (DAVID; bioinformatics resources 6.7) functional annotation clustering tool for gene ontology enrichment analysis was freely available at https://david.ncifcrf.gov/. The Integrated Molecular Pathway Level Analysis (IMPaLA; version 9; build January 2015) over-representation tool was freely available online at http://impala.molgen.mpg.de/.

2.2. Methods

2.2.1. Molecular Biology Methods

2.2.1.1. Restriction Digests

Restriction sites were utilised during plasmid manipulation, construction and screening. ≤ 1 µg plasmid DNA was digested with FastDigest restriction enzyme according to manufacturer's instructions.

2.2.1.2. Blunt Ending DNA Fragments

During plasmid cloning procedures that required blunt end ligation, Klenow was used to blunt 5' and 3' overhangs on DNA fragments and linearised plasmids. Klenow has 5' to 3' polymerase activity and 3' to 5' exonuclease activity, and can therefore blunt both 5'

and 3' overhangs. ≤ 1µg of restriction digested plasmid was incubated with 20µL Klenow reaction solution (0.25units/µL Klenow, 1X Klenow Buffer in dH_2O) for 1h at 37°C.

2.2.1.3. 5' End Phosphorylation of DNA Fragments

Blunt ended insert fragments were 5' phosphorylated with T4 polynucleotide kinase (PNK) during ligation of blunt ended DNA fragments into linearised, blunt ended plasmids. 5' phosphorylation increases the efficiency of blunt ended ligation. ≤ 1µg of DNA fragment was incubated with 20µL PNK reaction solution (0.35units/µL PNK, 500µM ATP, 1X Kinase Reaction Buffer in dH_2O) for 30mins at 37°C. The PNK was inactivated by incubating for 15mins at 65°C.

2.2.1.4. 5' End Dephosphorylation of DNA Fragments

5' phosphate groups are necessary for ligation. Therefore, to minimise re-ligation of the linearised plasmid during ligation procedures, shrimp alkaline phosphatase (SAP) was used to remove the 5' phosphate groups from the linearised plasmid. ≤ 1µg of linearised plasmid was incubated with 20µL SAP reaction solution (0.05units/µL SAP, 1 X SAP Buffer in dH_2O) for 30mins at 37°C. The SAP was inactivated by incubating for 15mins at 65°C.

2.2.1.5. Ligation of DNA Fragments

T4 DNA Ligase was used to ligate DNA fragments and linearised plasmids. DNA fragments containing (G4C2)n constructs were incubated for 10mins at 70°C then snap cooled on ice for 5mins prior to ligation. ≤ 1µg of the linearised plasmid and DNA fragment insert (if applicable) were incubated with T4 DNA Ligase reaction solution (0.15units/µL T4 DNA Ligase, 1X Ligase Buffer in dH_2O) overnight at room temp. The linearised plasmid and the DNA fragment to be inserted were mixed in a 1:3 ratio of absolute number of DNA molecules. In addition, re-ligation controls were set up that contained the same reagents described above, excluding the DNA fragment insert, and were used to measure the background number of *E.coli* colonies produced after transformation.

2.2.1.6. Agarose Gel Electrophoresis

Agarose gel electrophoresis was used to separate linear DNA fragments produced by restriction digest or PCR. 1.5% agarose gels were prepared by adding 1.5g agarose powder to 100mL 1X TAE buffer. The solution was then heated in a microwave until the agarose powder had dissolved. Ethidium bromide was added at a final concentration of 100ng/mL to the agarose solution, and the solution was then poured into a cassette. A gel comb was also placed into the agarose solution, and the solution was left for least 30mins to cool and set. Agarose gels were then placed into an electrophoresis tank containing 1X TAE buffer. 2µL of a HyperLadder™ molecular weight marker was loaded into the first well, and was used to size the DNA fragments on the gel. Restriction digest and PCR samples were loaded into consecutive wells. The gel was run at 100-120V for 45-60mins. Agarose gels were imaged using the GENi UV light imaging system (Syngene).

2.2.1.7. DNA Extraction from Agarose Gel

After agarose gel electrophoresis, the DNA bands were visualised using a UV transilluminator, and the desired DNA bands were manually cut out using a scalpel. The DNA was then extracted and purified using the QIAquick Gel Extraction Kit according to manufacturer's instructions.

2.2.1.8. LB Agar Plates

Sterile LB agar (10g/L Tryptone, 5g/L Yeast extract, 10g/L NaCl, 15g/L Agar) was prepared by resuspending 40g LB agar powder in 1L dH_2O, and autoclaving. The LB agar was then heated in a microwave until the LB agar had completely melted. The LB agar was left to cool for 5 mins, and then in sterile conditions either 50µg/mL carbenicillin or 50µg/mL spectinomycin selection was added to the melted LB agar. Still in sterile conditions, the LB agar solution was mixed and poured onto 10cm petri dishes.

2.2.1.9. Plasmid Transformation into Competent E.coli

Plasmid DNA or ligation mixtures were transformed into competent *E.coli* to clone the plasmid. ≤ 50ng of circular plasmid or the ligation mixture (section 2.2.1.5) were pre-chilled on ice. 20µL competent *E.coli* were added to the DNA using a large-bore pipette

tip, and incubated for 10mins on ice. *E.coli* were heat shocked for 30s at 42°C and then incubated for 5mins on ice. In sterile conditions, 100µL SOC outgrowth media was added to the *E.coli* and they were incubated for 30mins at 37°C on a shaker. Also in sterile conditions, the SOC outgrowth media containing the *E.coli* was then spread onto an LB agar plate containing 50µg/mL carbenicillin or 50µg/mL spectinomycin selection (dependent on the plasmid's antibiotic resistance) and incubated over-night at 37°C. Plasmids that contained the (G4C2)n constructs were transformed into β-10 competent *E.coli* (K12 strain, DH10B™ derivative) because they have reduced recombinase activity, which minimised the rearrangement of the repetitive (G4C2)n DNA sequences. All other plasmids were transformed into 5-α competent *E.coli* (K12 strain, DH5α™ derivative).

2.2.1.10. Miniprep of Plasmid DNA

Sterile LB Broth (10g/L Tryptone, 5g/L Yeast extract, 10g/L NaCl) was made by resuspending 25g LB Broth powder in 1L dH_2O, and autoclaving. *E.coli* colonies containing the plasmid of interest were picked and grown in LB Broth with either 50µg/mL carbenicillin or 50µg/mL spectinomycin selection (dependent on the plasmid's antibiotic resistance) overnight at 37°C on a shaker. The plasmids were then purified using a NucleoSpin Plasmid Kit (No Lid) according to manufacturer's instructions. DNA concentration was determined with a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific® Inc).

2.2.1.11. Colony PCR Screen

PCR was used to screen for successful pcDNA5/FRT/TO-HIS-(G4C2)n and pcDNA5/FRT/TO-(G4C2)n plasmid ligations in the transformed *E.coli* colonies. Each *E.coli* colony was picked and then incubated with 6µL PCR reaction solution (0.42µM CMV-Forward primer, 0.42µM BGH-Reverse primer, 1X ReddyMix PCR Master Mix in dH_2O). Samples were incubated for 3mins at 95°C, followed by 35 amplification cycles of 30s at 92°C, 30s at 55°C, and 60s at 72°C, followed by a final extension step of 5mins at 72°C.

2.2.2. Cell Culture Methods

2.2.2.1. HEK293 Cell Line Maintenance

HEK293 cells were cultured in 10mL supplemented DMEM (10% (v/v) FBS, 50units/mL penicillin/streptomycin) in 10cm plates in a 37°C / 5% CO_2 incubator. HEK293 cells were split every 3-4 days. To split the HEK293 cells, the media was removed, and cells were washed in PBS. 1mL 1X trypsin was added to the cells, and cells were returned to incubator for 4mins. The plate was tapped to dislodge trypsinised cells from the plate, and 9mL supplemented DMEM was added. Cells were resuspended, and 1mL cell suspension was added to a new 10cm plate. Supplemented DMEM was added to a final volume of 10mL.

2.2.2.2. HEK293 Sham and HEK293 (G4C2)n Cell Line Generation

pcDNA5/FRT/TO, pcDNA5/FRT/TO-(G4C2)10, pcDNA5/FRT/TO-(G4C2)51, and pcDNA5/FRT/TO-(G4C2)102 were each co-transfected with pPGKFLPobpA into Flp-In™ T-REx™ HEK293 cells. 24h post-transfection, the HEK293 cells were split onto 4 new plates. Supplemented tetracycline-free DMEM was added containing 15µg/mL blasticidin and 150µg/mL hygromycin to select for transformed cells. 5 days post-transfection, media was replenished with 50% HEK293 conditioned / 50% fresh media with selection. 15 days post-transfection, visible colonies of HEK293 cells were picked and transferred to a 48 well plate. The HEK293 colonies were grown on the 48 well plate until 60% confluent. The clones were then screened for Zeocin™ sensitivity, and blasticidin/hygromycin resistance. Clones that were blasticidin/hygromycin-resistant but Zeocin™-sensitive were expanded and cryopreserved.

2.2.2.3. HEK293 Sham and HEK293 (G4C2)n Cell Line Maintenance

HEK293 sham and HEK293 (G4C2)n cells were cultured in 10mL supplemented DMEM (10% (v/v) tetracycline-free FBS, 50units/mL penicillin/streptomycin) with 15µg/mL blasticidin, and 150µg/mL hygromycin, in 10cm plates in a 37°C / 5% CO_2 incubator. HEK293 cells were split every 3-4 days, as described in section 2.2.2.1.

2.2.2.4. NSC34 Cell Line Maintenance

NSC34 cells were cultured in 10mL supplemented DMEM (10% (v/v) FBS, 50units/mL penicillin/streptomycin), in 10cm plates in a 37°C / 5% CO_2 incubator Media was removed and replenished every 2-3 days. NSC34 cells were split every 3-4 days. To split the NSC34 cells, the media was removed and 10mL supplemented DMEM was added to the plate. The NSC34 cells were dislodged from the plate into suspension by pipetting the media up and down. 1mL of the cell suspension was added to a new 10cm plate. Supplemented DMEM with the relevant selection agents was added to a final volume of 10mL.

2.2.2.5. Flp-In™ T-REx™ NSC34 Cell Line Generation

A Flp-In™ T-REx™ NSC34 host cell line was generated using the Flp-In™ T-REx™ core kit according to the manufacturer's instructions. Briefly, pFRT/*lac*Zeo was transfected into WT NSC34 cells, and Zeocin™ selection was added to the NSC34 media to select for stable transformants. Zeocin™-resistant clones were picked, expanded, and cryopreserved. Genomic DNA was then extracted from these Flp-In™ NSC34 clones, and screened using Southern Blotting. Flp-In™ NSC34 clones that contained only one genomic FRT site were further analysed for the transcriptional activity of that FRT site.

Flp-In™ NSC34 clones were transfected with pcDNA5/FRT/GFP, and hygromycin selection was added to the Flp-In™ NSC34 media to select for stable transformants. Flp-In™ NSC34 clones were also negatively screened with Zeocin™ to check the insertions were not random. These Flp-In™ NSC34 GFP cells were then screened using a fluorescent microscope for GFP expression levels. Clones that expressed the highest GFP levels were transfected with pcDNA6/TR, and blasticidin/hygromycin selection was added to Flp-In™ NSC34 GFP media to select for stable transformants. These mixed populations of Flp-In™ T-REx™ NSC34 GFP cells were screened for GFP repression at basal levels, and increased GFP expression with tetracycline induction.

Once a Flp-In™ clone with repressible GFP expression was identified, the non-GFP T4E2A Flp-In™ T-REx™ NSC34 clonal line was transfected with pcDNA6/TR, and blasticidin/Zeocin™ selection was added to the Flp-In™ T-REx™ NSC34 media to select

for stable transformants. Blasticidin/Zeocin™-resistant clones were picked, expanded, and cyropreserved. Flp-In™ T-REx™ NSC34 clones were transfected with pcDNA5/FRT/GFP, and blasticidin/hygromycin selection was added to the Flp-In™ T-REx™ NSC34 GFP media to select for stable transformants. Blasticidin/hygromycin-resistant Flp-In™ T-REx™ NSC34 GFP clones were also screened for Zeocin™ sensitivity. Blasticidin/hygromycin-resistant, Zeocin™-sensitive Flp-In™ T-REx™ NSC34 GFP clones were then screened for GFP expression ± tetracycline induction. Flp-In™ T-REx™ NSC34 clonal line B10-2 had modest GFP expression at basal levels, and high GFP expression when induced with tetracycline. Therefore B10-2 was selected as the parental Flp-In™ T-REx™ NSC34 host cell line for generating the isogenic, tetracycline-inducible Flp-In™ T-REx™ NSC34 sham and (G4C2)n cell lines.

2.2.2.6. Flp-In™ T-REx™ NSC34 Cell Line Maintenance

All NSC34 cells were grown in 10cm plates in a 37°C / 5% CO_2 incubator. Flp-In™ T-REx™ NSC34 cells were cultured in 10mL supplemented tetracycline-free DMEM with 20µg/mL Zeocin™, and 5µg/mL blasticidin, in 10cm plates in a 37°C / 5% CO_2 incubator. Media was removed and replenished every 2-3 days. The media that was removed was centrifuged at 400 x g for 4mins, and the media was transferred to a new flask and stored as conditioned media. NSC34 cells were split every 3-4 days, as described in section 2.2.2.4.

2.2.2.7. NSC34 Sham and NSC34 (G4C2)n Cell Line Generation

pcDNA5/FRT/TO, pcDNA5/FRT/TO-(G4C2)10, pcDNA5/FRT/TO-(G4C2)51, and pcDNA5/FRT/TO-(G4C2)102 were each co-transfected with pPGKFLPobpA into Flp-In™ T-REx™ NSC34 cells. 48h post-transfection, the media was replenished, and the NSC34 cells were split onto 3 new 10cm plates. 5µg/mL blasticidin and 100µg/mL hygromycin were added to the media to select for transformed clones. The media containing selection was replenished every 3 days. 10 days post-transfection, media was replenished with 50% conditioned / 50% fresh media with selection. Visible colonies of NSC34 cells were picked and transferred to a 48 well plate. The NSC34 colonies were grown on the 48 well plate until 60% confluent. The clones were then screened for Zeocin™ sensitivity, and blasticidin/hygromycin resistance. Clones that were

blasticidin/hygromycin-resistant but Zeocin™-sensitive were expanded and cryopreserved.

2.2.2.8. NSC34 Sham and NSC34 (G4C2)n Cell Line Maintenance

NSC34 sham and NSC34 (G4C2)n cells were cultured in 10mL supplemented tetracycline-free DMEM with 5µg/mL blasticidin, and 100µg/mL hygromycin B, in 10cm plates in a 37°C / 5% CO_2 incubator. Media was removed and replenished every 2-3 days. NSC34 cells were split every 3-4 days, as described in section 2.2.2.4. NSC34 sham and NSC34 (G4C2)n cells were seeded onto gelatin-coated coverslips for RNA FISH and ICC methods.

2.2.2.9. Cryopreservation of Cell Lines

Media was removed from cells, and cells were resuspended in fresh media (either via trypsinising for HEK293 cellls (section 2.2.2.1) or dislodging the cells by pipetting for NSC34 cells (section 2.2.2.4)). Cells were centrifuged at 400 x g for mins, and then supernatant was removed. Approximately 5-10 x 10^6 Cells were resuspended in 800 µL 10% (v/v) DMSO/ 90% (v/v) FBS, and transferred to a cryogenic vials. Cryogenic vials were placed in a CoolCell® SV2 (Biocision), which was then incubated for 2h at -80°C. The cryogenic vials were then transferred to liquid nitrogen for long term storage.

2.2.2.10. Tetracycline Induction of Cell Lines

10µg/mL tetracycline was added to HEK293 sham and HEK293 (G4C2)n cell media to induce expression of the construct. Unless stated otherwise, 0.5µg/mL tetracycline was added to NSC34 sham and NSC34 (G4C2)n cell media to induce expression of the construct. Tetracycline was added every 3 days (if applicable) to maintain concentration in the media.

2.2.2.11. Plasmid Transfection of Cell Lines

pcDNA5/FRT/TO and pPGKFLPobpA transfections – For the Flp-In™ T-REx™ HEK293 cells, 10µL Lipofectamine® 2000, 6µg pPGKFLPobpA, and 4 µg pcDNA5/FRT/TO vector were used per 10cm plate. For the Flp-In™ T-REx™ NSC34 cells, 15µL Lipofectamine® 2000, 9µg pPGKFLPobpA, and 6µg pcDNA5/FRT/TO vector were used per 10cm plate. Lipofectamine® 2000 was mixed with 1.5mL Opti-MEM®, and separately, the

pPGKFLPobpA and pcDNA5/FRT/TO plasmids were mixed with 1.5mL Opti-MEM®. Both mixtures were incubated for 10mins at room temperature. The two mixtures were then mixed together, and incubated for 10mins at room temperature. Media was removed from cells, and 3mL DNA/Lipofectamine® mixture was added to the cells. The cells were incubated for 6h, and then 10mL supplemented tetracycline-free DMEM containing blasticidin was added to cells. The media was removed and replenished 24h post transfection for the Flp-In™ T-REx™ HEK293 cells, and 48h post transfection for the Flp-In™ T-REx™ NSC34 cells.

(C4G2)102, (GA)68, (GR)100, (AP)100, (PR)100 – 14µg plasmid DNA and 50µg PEI were used per 1mL Opti-MEM®. Firstly, the plasmid DNA and PEI were mixed with Opti-MEM®, and then incubated for 15mins at room temperature. The DNA/PEI mix was then added dropwise to HEK293 cells. 100µL of DNA/PEI mix was added per 1mL media on cells.

2.2.2.12. Stable Lentiviral Transduction of Cell Lines

NSC34 sham and (G4C2)102 cells were transduced with an estimated 0.1-0.2 Multiplicity of Infection (MOI) PTEN shRNA, GFP, or control shRNA lentiviral particles. 24h post-transduction, media was removed and supplemented tetracycline-free DMEM with 5µg/mL blasticidin, 100µg/mL hygromycin, and 1µg/mL puromycin selection was added. Blasticidin/hygromycin/puromycin-resistant cells were expanded and cryopreserved.

2.2.2.13. MTT Cell Viability Assay

The number of viable cells was measured using MTT reagent, which is metabolised to an insoluble purple formazan salt. Cells were grown in triplicate wells on 96 well plates for up to 7 days. 0.5µg/mL MTT reagent was added to media. Plates were incubated for 90mins in a 37 °C / 5% CO_2 incubator. 1 volume SDS/DMF lysis buffer (20% (w/v) SDS, 50% DMF (v/v), pH 4.7) was added to lyse cells. 595nm absorbance of wells was measured using a PHERAstar FS plate reader (BMG labtech Ltd.).

2.2.2.14. EthD1 Cell Death Assay

Dead cells were measured using EthD1, which is a cell-impermeant, high affinity nucleic acid stain that emits red fluorescence when bound to DNA. Cells were grown in triplicate wells on 96 well plates for up to 5 days. 1µM EthD1 was then added, and cells were incubated for 30mins at 37°C. The fluorescence of DNA-bound EthD1 was measured at Ex570nm/Em610nm using a PHERAstar FS plate reader (BMG labtech Ltd.). Cells were then freeze-thawed, and cell number was measured by adding an extra 1µM fresh EthD1 to the medium. Fluorescence was measured at Ex570nm/Em610nm. EthD1 measurements of dead cells were then calculated as a percentage of the EthD1 measurements of total cell number to give values for % cells dead in each condition.

2.2.2.15. DCF Reactive Oxygen Species Assay

Cytosolic Reactive Oxygen Species (ROS) levels were measured using DCF fluorescence. Cells were grown in triplicate wells on 96 well plates for up to 5 days. Media was then replaced with supplemented phenol-red free DMEM (10% (v/v) tetracycline-free FBS). 10µM DCFDA was then added, and cells were incubated for 90mins at 37°C. The fluorescence of oxidised DCF was measured at Ex485nm/Em520nm using a PHERAstar FS plate reader (BMG labtech Ltd.). Cells were then freeze-thawed, and cell number was measured by adding 1.5µM EthD1 to the medium. Fluorescence was measured at Ex570nm/Em610nm. Raw DCF data were then normalised to EthD1 measurement of cell number.

2.2.2.16. Growth Curve

1.5 x 10^6 cells were seeded onto a 10cm plate, and incubated for 4 days. Media was then removed, and cells were washed off in PBS. Cells were centrifuged at 400 x g for 4mins. PBS was removed and cell pellet was vortexed briefly by flicking. The cells were resuspended in 10mL media, and then counted using a haemocytometer. 1.5 x 10^6 cells were then re-seeded onto a new 10cm plate. This counting procedure was repeated and recorded every 4 days up to 16 days cell growth.

2.2.3. Biochemical Methods

2.2.3.1. SDS-Polyacrylamide Gel Preparation

Resolving gels of the desired acrylamide % (w/v) were prepared by mixing reagents described in Table 2.9. Spacer plates with 1.0 mm integrated spacers (Bio-Rad) and short plates (Bio-Rad) were assembled on the Mini-PROTEAN® Tetra Cell Casting Stand and clamps (Bio-Rad) with gaskets (Bio-Rad). Gels were poured into glass plates, and a layer of isopropanol was layered on top of the gels. The gels were left for at least 15 mins to set, and then the isopropanol was removed. 5 % Stacking gels were prepared by mixing reagents described in Table 2.9, and then poured onto the set resolving gels in the glass plates. 1.0 mm 15-well Mini-PROTEAN® Combs (Bio-Rad) were then inserted into the stacking gels, and left for at least 15 mins to set.

Table 2.9 Composition of 5% Stacking gels, 12% Resolving gels and 15% Resolving gels.

	5% Stacking gel	**12% Resolving gel**	**15% Resolving gel**
dH_2O	5.8 mL	3.5 mL	2.5 mL
30 % (w/v) Acrylamide	1.7 mL	4.0 mL	5.0 mL
Resolving buffer (1.5 M Trizma®, 13.9 mM SDS, pH 8.8, filtered)	-	2.5 mL	2.5 mL
Stacking buffer (0.5 M Trizma®, 13.9 mM SDS, pH 6.8, filtered)	2.5 mL	-	-
10 % (w/v) APS	50 µL	50 µL	50 µL
TEMED	20 µL	20 µL	10 µL

2.2.3.2. Sucrose Gradient Preparation

Sucrose was dissolved in Sucrose Gradient Buffer (SGB) (20mM Tris, 140mM KCl, 5mM MgCl, pH 8.0, DEPC treated) to produce 7%, 17%, 27%, 37%, and 47% (w/v) sucrose in SGB. The sucrose/SGB solutions were filtered using a Minisart filter with 0.2µm pores. 2.1mL of each sucrose/SGB solution was layered carefully on top of one another in centrifuge tubes (lowest to highest sucrose % solutions from top to bottom respectively) to produce a 7-47% (w/v) sucrose gradient. Sucrose gradients were stored overnight at 4°C.

2.2.3.3. Cell Lysis for Immunoblotting

Media was removed, and cells were then washed with ice cold PBS. 150µL ice cold IP lysis buffer (150mM NaCl, 50mM HEPES, 1mM EDTA, 1mM DTT, 0.5% (v/v) Triton™ X-100, PIC, pH 8.0) was added per well of the 6-well plate. PhosSTOP™ was additionally added to the IP lysis buffer if specific phosphorylated protein isoforms were probed for in immunoblotting. Cells were scraped into IP lysis buffer using a cell scraper, and incubated for 15mins on ice. To clarify, the lysate was then centrifuged at 17,000 x g for 5mins at 4°C. The lysate supernatant was reserved whilst the pellet of debris was discarded.

2.2.3.4. Bradford Assay

Bradford assays were used to quantify the protein concentration of cell lysates. Firstly, the protein assay dye reagent concentrate was diluted in dH_2O in a 1:4 ratio to make a working concentration of Bradford reagent. 1µL cell lysate was then added to 1mL Bradford reagent, and mixed. The Bradford reagent and protein sample was then loaded into a polystyrene cuvette with 1cm path length, and the optical density shift at A_{595nm} (OD_{595nm}) of the sample relative to a blank control was measured using a WPA S1200 Diode Array Spectrophotometer (Biochrom®). The concentration of the protein lysate was then calculated and converted to µg/mL using the Beer-Lambert law ($OD_{595nm} = \varepsilon cl$; where $\varepsilon = 1/15$, and $l = 1cm$).

2.2.3.5. Polysome Profiling and Ribosome Fractionation

1 X 10^6 NSC34 cells were seeded onto 10cm plates, and were grown for 3 days. The media was then removed, and cells were split onto 2 new 10 cm plates in fresh media. The cells were incubated for a further 3 days. The media was then removed, and polysomes were stalled by washing cells with ice cold 0.1mg/mL cyclohexamide in PBS. Cells from the 2 X 10 cm plates were then lysed and scraped in 500µL Sucrose Gradient (SG) lysis buffer (0.5mM DTT, 1% (v/v) Triton X-100, PIC, 80units/µL RiboSafe RNase Inhibitor, 100µg/mL cyclohexamide, in SGB) for 10mins on ice. The lysate was centrifuged at 2,400 x g for 5mins at 4°C. The supernatant was transferred to a new Eppendorf tube, and centrifuged again at 9,600 x g for 5mins at 4°C. The supernatant was transferred to a fresh tube, and the protein concentration was assessed by Bradford assay. 3mg stalled polysome protein lysate was loaded onto the 7-47% (w/v) sucrose gradients and centrifuged at 35,000rpm for 3h at 4°C in an SW41 Beckman rotor. The sucrose gradients were separated from top to bottom into 500µL fractions using an in house collection system coupled with an ÄKTA pure FPLC machine (GE Healthcare Life Sciences). UV-absorbance of the sucrose gradient was also measured to profile the ribosomal subunits, ribosomes and polysomes.

2.2.3.6. SDS-Polyacrylamide Gel Electrophoresis

Cell lysates were mixed with 4X Laemmli buffer (228mM Tris-HCl, 38% (v/v) glycerol, 277mM SDS, 0.038% (w/v) bromophenol blue, 5% (v/v) β-mercaptoethanol, pH 6.8) and boiled for 5mins at 95°C to denature the proteins. 12% or 15% SDS-Polyacrylamide gels were loaded into a Mini-PROTEAN® Tetra Vertical Electrophoresis Cell (Bio-Rad), and the apparatus was filled with running buffer (25mM Tris, 3.5mM SDS, 20mM glycine). For general immunoblotting methods, 25µg of denatured protein was loaded per well onto the SDS-polyacrylamide gels. For the polysome profiling method, 40µL of denatured sucrose fractions were loaded per well onto the SDS-polyacrylamide gels. 2µL prestained protein ladder was loaded as a molecular weight marker in one well per gel. Gel electrophoresis was performed at 50V for 30mins, then 150V for approximately 1.5h until the dye front had reached the bottom of the gel. Gels were then removed from the electrophoresis cell, and assembled with transfer buffer (47.9mM Tris, 38.6mM glycine,

1.38mM SDS, 20% (v/v) methanol)-saturated Whatman paper, and transfer buffer-saturated membrane in a semi-dry transfer apparatus. Nitrocellulose membranes were used routinely for immunoblotting, however, methanol pre-soaked PVDF membranes were used specifically when membranes were going to be probed with anti-AP, anti-GA, anti-GR, anti-GP, and/or anti-PR antibodies. Electrophoretic transfer of the proteins from the gels to the membranes was performed at 0.15A / gel transferred for 1h. Membranes were then stained with Ponceau stain (0.1% (w/v) Ponceau S, 5% (v/v) acetic acid), and trimmed.

2.2.3.7. Immunoblotting

For general immunoblotting, membranes were blocked in 5% (w/v) milk/Tris Buffered Saline, with Tween® 20 (TBST) (20mM Tris, 137mM NaCl, 0.2% (v/v) Tween® 20, pH 7.6) for 1h at room temperature on a roller. However, membranes were blocked in 5% (w/v) BSA/TBST for 1h at room temperature on a roller, if specific phosphorylated protein isoforms were being probed for. The membranes were incubated with primary antibody (see Error! Reference source not found. for dilutions, blocking agent and source) in the d esignated blocking agent for 1h at room temperature or over-night at 4°C on a roller. The membranes were then washed 3 times in TBST for 15mins at room temperature. The membranes were then incubated with secondary antibody conjugated to horseradish peroxidase (HRP) (see Table 2.6 for dilutions and source) in 5% milk/TBST for 1h at room temperature on a roller. The membranes were then washed 3 times in TBST for 15mins at room temperature. The membranes were then incubated with ECL for 1mins and imaged using a G:BOX (Syngene).

2.2.3.8. RNA Fluorescence in situ Hybridisation (FISH)

Coverslips with NSC34 and HEK293 cells were fixed and permeabilised in 4% (w/v) PFA / 0.2% (v/v) Triton™ X-100 in PBS for 20mins at room temperature. For the RNAse treated control, slides were incubated with 10µg/mL RNAse A in PBS for 30mins at 37°C. Coverslips were blocked with hybridisation solution (50% (v/v) formamide, 2X saline sodium citrate (SSC), 100mg/ml dextran sulphate, 50mM sodium phosphate pH 7.0) for 1h at 66°C. The LNA probes were incubated at 80°C for 75s, then snap cooled on ice for 5mins to denature the DNA secondary structure. Coverslips were then incubated with

400ng/ml denatured probe in hybridisation solution overnight at 66°C. After hybridization, slides were washed once in 2X SSC / 0.1% Tween® 20 for 5mins at room temperature and three times in 0.1X SSC for 30mins at 65°C. NSC34 cells that were subsequently dual stained by immunocytochemistry (ICC), were first irradiated on ice with 0.3J/cm^2 UV, washed 3 times with PBS, and then ICC staining was performed. Coverslips were mounted with mounting medium containing DAPI. All solutions were made with DEPC-treated water.

2.2.3.9. Immunocytochemistry (ICC)

Coverslips with NSC34 cells were fixed and permeabilised in 4% (w/v) PFA / 0.2% (v/v) Triton X-100 at room temperature for 20mins. Slides were incubated with primary antibody in 2% (w/v) BSA/PBS at room temperature for 1h (see Table 2.7 for antibody dilution and source). Coverslips were washed 3 times in PBS for 15mins at room temperature. Coverslips were then incubated with secondary antibody in 2% BSA/PBS at room temperature for 1h (see Table 2.8 for antibody dilution and source). Coverslips were mounted with mounting medium containing DAPI.

2.2.3.10. Microscopy Imaging and Image Analysis

RNA foci and RNA Binding Protein (RBP) co-stain – Both sense (G4C2)n and antisense (C4G2)n RNA foci were visualised using an SP5 confocal microscope system (Leica) with a X63/1.4 oil immersion objective lens. The presence of foci was assessed within a high resolution (1433mm^2 per image, 511 X 511 pixels) z-stack made up of images at 0.13µm intervals through the entire nuclear volume of the cell under consideration. The same imaging was used for sense (G4C2)n RNA foci and ALYREF, NCL, PURA, SRSF1, and SRSF2 co-staining. For sense (G4C2)n RNA foci counts, 20 cells were imaged per condition in 3 independent experiments, and RNA foci were quantified manually. For the sense (G4C2)n RNA foci and RBP co-stain, 50 NSC34 (G4C2)102 cells were analysed for co-localisation, and co-localisation was quantified manually.

NCL – NCL staining in the NSC34 cells was visualised using an SP5 confocal microscope system (Leica) with a X63/1.4 oil immersion objective lens. The NCL staining was assessed within a high resolution (3775 mm^2 per image, 511 X 511 pixels) z-stack made

up of images at 0.5µm intervals through the entire nuclear volume of the cell under consideration. To quantify the area of the nucleolus relative to the nuclear area, the analysis previously described by Haeusler et al. 2014 was employed. 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). 25 cells were imaged per condition in 3 independent experiments.

TDP-43 – TDP-43 staining in NSC34 cells were visualised using a LV100ND microscope (Nikon) fitted with a DS Ri1 Eclipse camera (Nikon).

2.2.3.11. RNA Extraction for qRT-PCR

Sucrose gradient fractions – 750µL TRIzol® LS Reagent was mixed with 250µL sucrose gradient fraction (section 2.2.3.5). 150µL chloroform was added, and samples were shaken vigorously for 15s. Samples were incubated for 5mins at room temperature. Samples were centrifuged at 12,000 x g for 10mins at room temperature. 500µL aqueous phase was transferred to a new Eppendorf, and 5µg glycogen, 50µL 3M sodium acetate pH 5.0, and 100µL isopropanol was added. The samples were incubated overnight at -20°C. Samples were then centrifuged at 12,000 x g for 20mins at room temperature. The supernatant was removed and the RNA pellet was washed in 70% (v/v) ethanol. Samples were centrifuged at 12,000 x g for 5mins at room temperature, and the wash was removed. The RNA pellet was air dried for 5mins at 37°C. The RNA pellet was resuspended in 50µL DNase I solution (1X DNase I buffer, and 0.2units/µL DNase I recombinant, RNase-free in dH_2O) and incubated for 30mins at 37°C, then for 5mins at 70°C. RNA yield and quality were assessed using the NanoDrop™ 1000 (Thermo Fisher Scientific® Inc.).

NSC34 cell pellets for qRT-PCR – NSC34 sham and (G4C2)102 cells were grown identically to those used for microarray transcriptomic analysis (section 2.2.4.1), but stored as cell pellets in RNA Later. RNA was extracted from these cell pellets, and used for the qRT-PCR validation experiments. 100µL denaturing solution was added to the cell pellet, and the solution was pipetted up and down. Samples were incubated for 10mins at room temperature. 150µL extraction and phase separation solution was added, and samples were shaken vigorously for 15s. Samples were then incubated for 10mins at room

temperature and then centrifuged at 12,000 x g for 10mins at room temperature. 100µL aqueous phase was transferred to a new eppendorf, and 100µL isopropanol was added. The samples were incubated over-night at -20°C. Samples were then centrifuged at 17,000 x g for 20mins at room temperature. The supernatant was removed and the RNA pellet was washed in 70% (v/v) ethanol. Samples were centrifuged at 17,000 x g for 5mins at room temperature, and the wash was removed. The RNA pellet was air dried for 5mins at 37°C. The RNA pellet was resuspended in 25µL DNase I solution (1X DNase I buffer, and 0.2units/µL DNase I recombinant, RNase-free in dH_2O) and incubated for 30mins at 37°C, then for 5mins at 70°C. RNA yield and quality were assessed using the NanoDrop™ 1000 (Thermo Fisher Scientific® Inc.).

2.2.3.12. cDNA Synthesis for qRT-PCR

Total RNA isolated from NSC34 cells was reverse transcribed into single-stranded complementary DNA (ss-cDNA). cDNA synthesis was performed using the BioScript™ Reverse Transcriptase kit according to manufacturer's instructions. 2µg total RNA was mixed with 1µL 40µM random hexamer primers, 1µL 10mM dNTP mix, and a total reaction volume (14µL) made up in DEPC-treated dH_2O. The samples were incubated for 5mins at 70°C for mins, then incubated for at least 1mins on ice. 1X RT buffer, 2units/µL RiboSafe RNase Inhibitor, 10units/µL BioScript Reverse Transcriptase, to a total volume reaction volume (20µL) made up in DEPC-treated dH_2O. A no-RT control was also prepared as described, without the addition of BioScript Reverse Transcriptase to check for the presence of potentially contaminating genomic DNA. Samples were mixed by gentle pipetting, and then incubated for 10mins at 25°C, then 60mins at 42°C, then 5mins at 85°C. In each reaction there is an approximate 1:1 conversion of RNA to cDNA. cDNA samples were diluted to approximately 25ng/µL cDNA in DEPC-treated dH_2O. Samples were stored at -20°C prior to qRT-PCR.

2.2.3.13. Quantitative Real Time PCR (qRT-PCR)

qRT-PCR was performed in triplicate 10µL reaction volumes with a no-RT control and a no template control (NTC), on 96-well qRT-PCR plates. Each well contained 25ng of cDNA, forward and reverse primers (Table 2.3) at optimised concentrations, 2X Brilliant III SYBR Green qPCR Master Mix and dH_2O. Plates were briefly centrifuged to collect

samples at the bottom of the wells. Samples were run on a 3 step profile on a Stratagene Mx3000P™ Real Time Thermal Cycler (Agilent Technologies Ltd.). Samples were incubated for 10mins at 95°C, then 40 cycles of 30s at 95°C, 60s at 60°C, and 60s at 72°C. A final cycle was used to determine the primer dissociation curve.

2.2.3.14. PrimeTime® qPCR Assays

PrimeTime® qPCT assays were performed in triplicate 10µL reaction volumes with a no-RT control and a NTC, on 96-well qRT-PCR plates. Each well contained 25ng of cDNA, 500nM forward and reverse primer, and 250nM probe (Table 2.4) resuspended in TE buffer (10mM Tris-HCl, 1mM EDTA, pH 7.5), 2X Brilliant III qPCR Master Mix and dH_2O. Plates were briefly centrifuged to collect samples at the bottom of the wells. Samples were run on a 3 step profile on a Stratagene Mx3000P™ Real Time Thermal Cycler (Agilent Technologies Ltd.). Samples were incubated for 10mins at 95°C, then 40 cycles of 30s at 95°C, 60s at 60°C, and 60s at 72°C. A final cycle used to determine the primer dissociation curve was also run.

2.2.4. Gene Expression Profiling Methods

2.2.4.1. RNA Extraction for Gene Expression Profiling Analysis

0.5 x 10^6 NSC34 cells were seeded per well on a 6 well plate and were grown for 5 days until 80% confluent. Tetracycline was added to cells for either the final 5 days, final 3 days, final day or not at all. Media (and tetracycline if applicable) was replenished after 3 days. After 5 days growth, RNA was extracted from the cells using the Direct-zol™ RNA Kit. Briefly, media was removed and 1mL TRI-Reagent® was added to the cells. The solution was pipetted up and down to mix well. The samples were incubated for 10mins at room temperature. The samples were then centrifuged at 16,000 x g for 1mins at room temperature. The supernatant was romoved and reserved in a new eppendorf, whilst the particulate matter was discarded. 1mL ethanol was added to each of the samples, and then they were briefly vortexed. 700µL of each of the samples were loaded into Zymo-Spin™ II Columns, in collection tubes, and samples were centrifuged at 16,000 x g for 1mins at room temperature. The flow-throughs were discarded, and the remaining samples were loaded and centrifuged on the respective columns as previously

described until all of the sample solutions had been centrifuged through the columns. 1 mL Wash Buffer was then loaded onto each of the columns, and they were centrifuged at 16,000 x g for 1mins at room temperature. 80µL DNAse I Reaction Mix (5µL of 1units/µL DNAse I, 8µL of 10X DNAse I Reaction Buffer, 3µL DNAse/RNAse-free water, 64µL RNA Wash Buffer with ethanol added) was added directly to each of the columns, and incubated for 15mins at room temperature. The columns were centrifuged at 12,000 x g for 1mins at room temperature. 400µL Direct-zol RNA PreWash was added to each of the columns, and the columns were then centrifuged at 16,000 x g for 1mins at room temperature. The flow-throughs were discarded and then the Direct-zol RNA PreWash step was repeated. 700µL RNA Wash Buffer was added to each of the columns, and the columns were then centrifuged at 16,000 x g for 1mins at room temperature. The flow-throughs were discarded and then the columns were centrifuged again at 16,000 x g for 2mins at room temperature to completely remove the Wash Buffer. 30µL DNase/RNase-free water was added to the matrix and incubated for 5mins at room temperature to elute the RNA. The columns were centrifuged at 16,000 x g for 1mins at room temperature to collect the eluted RNA samples. The RNA samples were snap frozen using liquid nitrogen, and stored at -80°C.

2.2.4.2. RNA Yield and Quality Assessment

RNA concentration and purity were assessed using the NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific® Inc.). The ratios of absorbance at 260nm and 280nm, as well as 260nm and 230nm are used to assess RNA yield and purity. All RNA samples had $A_{260\ nm/280\ nm}$ values of 2.0 ± 0.2, and $A_{260\ nm/230\ nm}$ values of 2.2 ± 0.2, indicating high purity. In addition, the RNA integrity was measured using a Nanochip and an Agilent 2100 Bioanalyser (Agilent Technologies Ltd.). Briefly, an electropherogram was produced to assess the RNA for the 18S and 28S rRNA peaks, the amount of RNA degradation, and provide a RNA Integrity Number (RIN) based on these variables.

2.2.4.3. WT PLUS Amplification and Labelling Process

The GeneChip® WT PLUS Reagent Kit was used to generate amplified and biotinylated sense-strand DNA targets from total RNA. The RNA was amplified and labelled in 3

separate batches; A, B and C (which are used in the array names). All manipulations took place on ice, unless stated otherwise.

2.2.4.3.1. Poly-A RNA Control Preparation

Poly-A RNA controls were prepared to include with each RNA sample, and provided exogenous positive controls to monitor the entire target preparation. Poly-A Control Stock was diluted 1 in 20, then 1 in 50, then 1 in 50, then 1 in 2 in Poly-A Control Dil Buffer to a final dilution of 1 in 100,000. 2 µL of the final Poly-A Control mRNA dilution was then added to 500ng RNA sample in dH_2O to a final volume of 5µL.

2.2.4.3.2. First-Strand cDNA Synthesis

Total RNA was reverse transcribed using primers that contained a T7 promoter sequence, producing single-stranded complementary DNA (ss-cDNA) with a T7 promoter sequence at the 5' end. Firstly, the First-Strand Master Mix was prepared by mixing 42µL First-Strand Buffer and 10.5µL First-Strand Enzyme. 5µL First-Strand Master Mix was mixed with each of the 5µL total RNA samples. The samples were incubated for 1h at 25°C, then for 1h at 42°C, then for 5mins at 4°C. The samples were centrifuged briefly to collect first-strand cDNA/RNA at bottom of the tube.

2.2.4.3.3. Second-Strand cDNA Synthesis

The ss-cDNA was then converted to double-stranded cDNA (ds-cDNA), which acted as a template for *in vitro* transcription. RNase H and DNA polymerase simultaneously degraded the RNA and synthesised second-strand cDNA. The Second-Strand Master Mix was prepared by mixing 189µL Second-Strand Buffer and 21µL Second-Strand Enzyme. 20µL Second-Strand Master Mix was mixed with each of the (10µL) First-Strand cDNA/RNA samples for a total volume of 30µL. The samples were incubated for 1h at 16°C, then for 10mins at 65°C, then for 5mins at 4°C. The samples were centrifuged briefly to collect ds-cDNA at bottom of the tube.

2.2.4.3.4. Complementary RNA Synthesis by in vitro Transcription

Complementary RNA (cRNA) was synthesised and amplified by *in vitro* transcription (IVT) of the ds-cDNA template using the T7 RNA Polymerase. The IVT Master Mix was

prepared at room temperature by mixing 252µL IVT Buffer and 63µL IVT Enzyme. 30µL IVT Master Mix was mixed with each of the (30µL) ds-cDNA samples at room temperature for a total volume of 60µL. The samples were incubated for 16h at 40°C, then at 4°C. The samples were centrifuged briefly to collect the cRNA at bottom of the tube, and stored at -20°C.

2.2.4.3.5. cRNA Purification

The cRNA samples were purified to remove enzymes, salts, inorganic phosphates, and unincorporated nucleotides to prepare the cRNA for Second-cycle ss-cDNA synthesis. The magnetic purification beads were briefly vortexed to resuspend, and then 100µL beads were added to each (60µL) cRNA sample, mixed by pipetting, and placed on a U-bottom plate at room temperature. The samples were incubated for 10mins at room temperature to allow the cRNA to bind to the purification beads. The plate with samples was then placed onto a magnetic stand to capture the beads for 5mins at room temperature. The supernatant was removed, and the beads were washed 3 times with 80% (v/v) ethanol for 1.5mins at room temperature. The supernatant was removed and the beads were air dried for 5mins at room temperature. The plate was then removed from the magnetic stand. Nuclease-free water was heated to 65°C, and then 27µL was added to each well to elute the cRNA. The beads and water were mixed and then incubated for 1mins at room temperature. The plate with samples was placed onto the magnetic stand to capture the beads, and incubated for 5 min at room temperature. The eluted cRNA was transferred to new nuclease-free tubes, and placed on ice. The cRNA yield and size distribution were assessed and then cRNA was stored at -20°C.

2.2.4.3.6. cRNA Yield and Size Distribution Assessment

cRNA concentration was assessed using the NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific® Inc.). The cRNA size distribution was assessed using a Nanochip and an Agilent 2100 Bioanalyser (Agilent Technologies Ltd.).

2.2.4.3.7. Second-Cycle Single-Stranded cDNA Synthesis

cRNA was reverse transcribed using 2nd-Cycle Primers producing sense-strand cDNA, which contained dUTP at a fixed ratio relative to dTTP. 15µg cRNA was diluted in

nuclease-free water to a final volume of 24µL. 4µL of 2nd-Cycle Primers were added. The samples were incubated for 5mins at 70°C, then for 5mins at 25°C, then for 2mins at 4°C. The 2nd-Cycle ss-cDNA Master Mix was prepared by mixing 84µL 2nd-Cycle ss-cDNA Buffer and 42µL 2nd-Cycle ss-cDNA Enzyme. 12µL 2nd-Cycle ss-cDNA Master Mix was mixed with each of the (28µL) cRNA/2nd-Cycle Primers samples for a total volume of 40µL. The samples were incubated for 10mins at 25°C to allow the primers to anneal, then for 90mins at 42°C for cDNA synthesis, then 10mins at 70°C to inactivate the enzyme, and then samples were stored at 4°C.

2.2.4.3.8. RNA Hydrolysis Using RNase H

The cRNA template was hydrolysed by RNAse H, leaving ss-cDNA. 4µL RNase H was added to each of the (40µL) 2nd-Cycle ss-cDNA samples for a total volume of 44µL. The samples were incubated for 45mins at 37°C, followed by 5mins at 95°C, and then 2mins at 4°C. 11µL nuclease-free water was added to each (44µL) hydrolysed 2nd-Cycle ss-cDNA samples for a total volume of 55µL. Samples were stored at -20°C.

2.2.4.3.9. Second-Cycle ss-cDNA Purification

The 2nd-cycle ss-cDNA was purified to remove enzymes, salts, and unincorporated dNTPs to prepare the ss-cDNA for fragmentation and labelling. The magnetic purification beads were briefly vortexed to resuspend, and then 100µL beads were added to each (55µL) ss-cDNA sample, mixed by pipetting, and placed on a U-bottom plate at room temperature. 150µL 100% ethanol was added to each (155µL) ss-cDNA/beads sample. Samples were mixed well by pipetting up and down. Samples were then incubated for 20mins at room temperature, during this time the ss-cDNA bound to the magnetic purification beads. The plate with samples was then placed onto the magnetic stand to capture the beads and bound ss-cDNA, and incubated for 5mins at room temperature. The supernatant was removed, and the beads were washed 3 times with 80% (v/v) ethanol for 1.5mins at room temperature. The supernatant was removed and the beads were air dried for 5mins at room temperature. The plate was then removed from the magnetic stand. Nuclease-free water was heated to 65°C, and then 30µL was added to each well to elute the ss-cDNA. The beads and water were mixed and then incubated for 1mins at room temperature. The plate with samples was placed onto the magnetic

stand to capture the beads, and incubated for 5mins at room temperature. The purified ss-cDNA was transferred to new nuclease-free tube, and placed on ice. The ss-cDNA yield and size distribution was assessed and then the ss-cDNA was stored at -20°C.

2.2.4.3.10. ss-cDNA Yield and Size Distribution Assessment

ss-cDNA concentration was assessed using the NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific® Inc.). The ss-cDNA size distribution was assessed using a Nanochip and an Agilent 2100 Bioanalyser (Agilent Technologies Ltd.).

2.2.4.3.11. ss-cDNA Framentation and Labelling

The ss-cDNA was fragmented by uracil-DNA glycosylase (UDG) and apurinic/apyrimidinic endonuclease 1 (APE 1) at the unnatural dUTP residues, breaking the DNA strand. The fragmented cDNA was then labelled by terminal deoxynucleotidyl transferase (TdT) using the Affymetrix propriety DNA Labelling Reagent that is covalently linked to biotin. 5.5µg ss-cDNA was prepared in 31.2µL nuclease-free water. The Fragmentation Master Mix was prepared by mixing 105µL nuclease-free water, 50.4µL 10X cDNA Fragmentation Buffer, 10.5µL 10units/µL UDG, and 10.5µL 1,000 units/µL APE 1. 16.8µL Fragmentation Master Mix was mixed with each of the (31.2µL) ss-cDNA samples for a total volume of 48µL. The samples were incubated for 1h at 37°C, then for 2mins at 93°C, and then at 4°C. The Labelling Master Mix was prepared by mixing 126µL 5X TdT Buffer, 10.5µL 5mM DNA Labelling Reagent, and 21µL 30units/µL TdT. 15µL Labelling Master Mix was mixed with each of the (45µL) fragmented ss-cDNA samples for a total volume of 60µL. The samples were incubated for 1h at 37°C, then for 10mins at 70°C, and then for 5mins at 4°C. 2µL of each of the fragmented and labelled ss-cDNA samples were removed for Gel-shift analysis, and the remaining samples were stored at -20°C.

2.2.4.3.12. Gel-Shift Assay

The efficiency of the biotin-labelling procedure was assessed by adding NeutrAvidin (or PBS as a negative control) to the ss-cDNA, and then performing gel electrophoresis. NeutrAvidin binds to biotin and therefore retards the biotin-labelled ss-cDNA movement during gel electrophoresis, causing a shift in molecular weight between the NeutrAvidin +ve and –ve treated ss-cDNA. 10mg/mL NeutrAvidin was diluted to 2mg/mL in PBS. 1µL

of the fragmented and labelled ss-cDNA samples were aliquoted to nuclease-free tubes, and then incubated for 2mins at 70°C. For the +ve gel-shift samples, 5µL 2mg/mL NeutrAvidin was added to the fragmented and labelled ss-cDNA samples, whilst for the -ve gel-shift samples, 5µL PBS was added to the fragmented and labelled ss-cDNA samples. The samples were then mixed and incubated for 5mins at room temperature. 2.5µL 5X Novex® Hi-Density TBE Sample Buffer was mixed with each sample. Samples were loaded onto a 4-20% TBE Gel, 1.0 mm. Gel electrophoresis was performed at 100 V for 2h. The gel was stained in 0.001% (w/v) ethidium bromide in 1X TAE buffer for 30 mins at room temperature on a shaker. The gels were imaged using the GENi UV light imaging system (Syngene).

2.2.4.4. WT Array Hybridisation

The fragmented and labelled ss-cDNA was hybridised onto GeneChip® Mouse Transcriptome Array 1.0 using the GeneChip® Hybridisation, Wash and Stain Kit. Three separate batches of arrays were hybridised, washed, and stained on separate days and the arrays contained in each batch are listed in **Appendix 4**. All manipulations took place at room temperature, unless stated otherwise. The 20X Eukaryotic Hybridisation Controls (*bioB*, *bioC*, *bioD*, *cre*) were incubated for 5mins at 65°C. The Hybridisation Master Mix was prepared by mixing 40.7µL 3nM Control Oligo B2, 121µL 20X Eukaryotic Hybridisation Controls (*bioB*, *bioC*, *bioD*, *cre*), 1210µL 2X Hybridisation Mix, 169.4µL DMSO, and 240.9µL Nuclease-free water. Hybridisation cocktails were prepared by mixing 162µL Hybridisation Master Mix with 58µL (5.2µg) of each of the fragmented and labelled ss-cDNA samples, for total volumes of 220µL. The hybridisation cocktail was then incubated for 5mins at 95°C, then for 5 min at 45°C. 200µL of the hybridisation cocktail was loaded onto GeneChip® Mouse Transcriptome Array 1.0 cartridges. The arrays were then incubated in a GeneChip® Hybridisation Oven 640 (Affymetrix® Ltd) with 60rpm rotation for 16h at 45°C.

2.2.4.5. Array Washing, Staining and Scanning

Array washing and staining was performed on the GeneChip® Fluidics Station 450 (Affymetrix® Ltd) according using the Whole Transcript (WT) Sense Target Labelling Manual for fluidics protocol FS450_0001 (Affymetrix® Ltd). After washing and staining,

the arrays were scanned on the GeneChip® Scanner 3000 (Affymetrix® Ltd) according to the Affymetrix® GeneChip® Command Console® User Manual, to generate CEL files containing the image of raw probe signal intensities.

2.2.4.6. Bioinformatics Analysis

2.2.4.6.1. Affymetrix® Expression Console Software

CEL files containing the raw probe signal intensities were exported from the GeneChip® Command Console. Gene-level RMA-sketch normalisation and signal summarisation were performed in the Affymetrix® Expression Console (EC) software, converting the CEL files to RMA-GENE-FULL CHP files. In addition, Exon-level Alt Splice analysis normalisation and signal summarisation were performed in EC, converting the CEL files to RMA-ALT-SPLICE CHP files. Quality Control was performed in EC. Hybridisation controls, poly-A labelling controls, AUC, and signal distribution assessed in EC and were used for Quality Control assessment of the arrays.

2.2.4.6.2. Qlucore Omics Explorer

Gene-level expression analysis was performed in Qlucore Omics Explorer (version 3.0). All Affymetrix® RMA-GENE-FULL CHP files were opened in Qlucore Omics explorer, and normalised using the RMA-sketch method. Differentially expressed (DE) transcripts were defined when P value < 0.01 (using Student's T-test) and fold change (FC) ≥ 1.2. PCA plots and Heat Maps were generated in Qluocore Omics explorer. Lists of DE transcripts were exported to Microsoft Excel 2010.

2.2.4.6.3. Gene Ontology Enrichment Analysis

Gene Ontology enrichment analysis was performed using the Database for Annotation and Visualisation and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/). The Entrez Gene IDs of the DE transcripts were imported to DAVID, and functional annotation clustering was subsequently performed using the GOTERM_BP_FAT gene ontology and KEGG_PATHWAY pathway terms, applying a *Mus musculus* background, and filtering using medium stringency. Functional clusters with DAVID enrichment scores > 1.30 (equivalent to a P value < 0.05) were considered statistically significant.

2.2.4.6.4. Pathway Enrichment Analysis

Integrated pathway-level enrichment analysis was performed by importing the gene symbols of DE transcripts to the Integrated Molecular Pathway Level Analysis (IMPaLA) over-representation (enrichment) tool (http://impala.molgen.mpg.de/).

2.2.4.6.5. Affymetrix® Transcriptomics Analysis Console Software

Differential splicing analysis was performed in the Affymetrix® Transcriptomics Analysis Console (TAC) Software. All Affymetrix® RMA-ALT-SPLICE CHP files were opened in TAC. Genes were defined as differentially spliced when at least one differential splicing event had a False Discovery Rate (FDR) corrected P value < 0.05 (using One-Way ANOVA) and Splicing Index (SI) ≤ -2.0 or ≥ 2.0.

Chapter 3. Generation of a Motor Neuron-Like Cell Model of C9ORF72-Amyotrophic Lateral Sclerosis with Tetracycline-Inducible (G4C2)n Repeat Expression

3.1. Introduction

The initial aim of the project was to generate and characterise stable, isogenic cell models with tetracycline-inducible expression of the (G4C2)n repeat constructs. Plasmids containing different lengths of interrupted (G4C2)n repeat constructs and an FRT site were engineered. The commercially available Flp-In™ T-REx™ HEK293 cells, and Flp-In™ T-REx™ NSC34 cells that were generated in house by Dr. Adrian Higginbottom, were used. These cell lines utilise the Flp-In™ system (Figure 3.1) which allowed genomic site specific insertion of the (G4C2)n repeats into both HEK293 cells and motor neuron-like NSC34 cells. These cell lines also utilise the T-REx™ system (Figure 3.1), which allow for tetracycline-inducible expression of the (G4C2)n repeats. RNA Fluorescence *in situ* Hybridisation (FISH) and immunoblotting were used to confirm tetracycline-inducible expression of the (G4C2)n repeats at RNA and protein levels respectively. Finally, MTT cell viability assays, EthD1 cell death assays, and growth curves were used to assess the (G4C2)n expression for toxicity in the NSC34 cells.

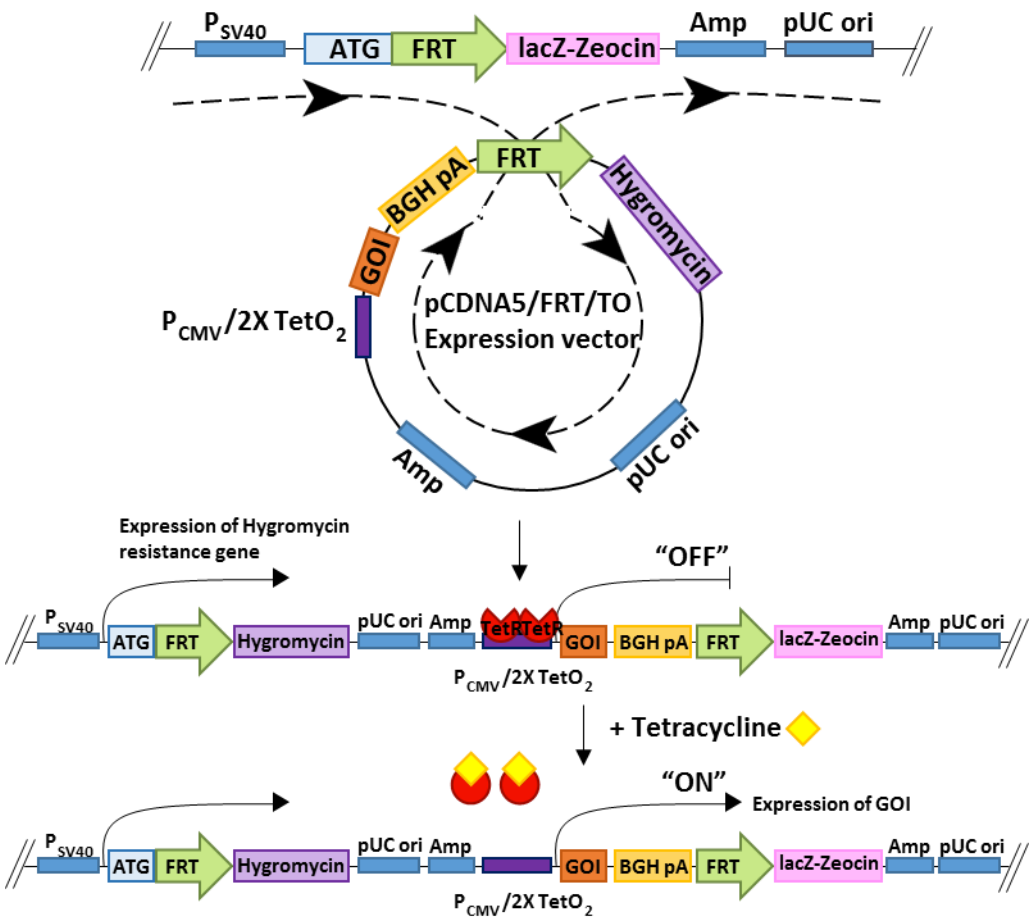


Figure 3.1 Flp-In™ and T-REx™ systems used to generate isogenic cell lines with tetracycline-inducible expression of the gene of interest. The pcDNA5/FRT/TO expression vector containing the gene of interest (GOI) is co-transfected with a FLP recombinase into the Flp-In™ T-REx™ mammalian cell line. FLP recombinase catalyses homologous recombination between the genomic FRT site and the FRT site within the pcDNA5/FRT/TO expression vector. Successful integration of the pcDN5/FRT/TO plasmid confers hygromycin resistance and Zeocin™ sensitivity to the cell line. The gene of interest is repressed by the Tet repressor element (*TetR*). Addition of tetracycline to the cells induces expression of the GOI by de-repressing the CMV/$TetO_2$ promoter. Figure is adapted from Invitrogen™.

Interrupted (G4C2)n repeat constructs were generated to model the pathogenic (G4C2)n repeat expansion, because the GC pure repeat is technically very challenging to manipulate. PCR can only process through a small number of repeats before failing. Also, the (G4C2)n repeats form secondary structure in the DNA, making it challenging to clone via restriction digestion and ligation. In addition, the minimum number of (G4C2)n repeats required to cause ALS and/or FTD is unknown, the longest number of (G4C2)n repeats were generated as possible. The only technically feasible method of generating

larger (G4C2)n repeat constructs, was via restriction digest and ligation of smaller (G4C2)n repeats. Synthesised TCGAC(G4C2)10 sense and ACGT(G2C4)10 antisense ssDNA oligonucleotides were designed such that the dsDNA produced by annealing the oligonucleotides was flanked 5' by *Sal*I and 3' by *Xho*I cut restriction sites. It is important to note that the *Sal*I and *Xho*I restriction enzymes produce compatible cohesive ends. These (G4C2)10 were ligated into *Xho*I cut plasmid. Subsequent ligations involved digesting these vectors with *Xho*I and inserting further (G4C2)10 repeats. The 5' *Sal*I site of the inserted (G4C2)10 was destroyed whilst the 3' *Xho*I site was retained – if the insertion orientation was correct. Vectors containing (G4C2)10, interrupted (G4C2)51 and interrupted (G4C2)102 constructs were generated via this method (Figure 3.2).

$(G4C2)10 = (GGGGCC)_{10}$

$(G4C2)51 = (GGGGCC)_{12}TCGAC(GGGGCC)_{17}TCGAC(GGGGCC)_{10}$
$TCGAC(GGGGCC)_{12}$

$(G4C2)102 = (GGGGCC)_{12}TCGAC(GGGGCC)_{17}TCGAC(GGGGCC)_{10}TCGAC$
$(GGGGCC)_{24}TCGAC(GGGGCC)_{17}TCGAC(GGGGCC)_{10}TCGAC(GGGGCC)_{12}$

Figure 3.2 Sequences of the (G4C2)10, (G4C2)51 and (G4C2)102 constructs.

The Flp-In™ and T-REx™ systems were used in both HEK293 and NSC34 cells to generate stable and isogenic cell models with tetracycline-inducible expression of the interrupted (G4C2)n repeats (Figure 3.1). A Flp recombination target (FRT) site, and a Tn*10*-encoded tetracycline resistance repressor element (*TetR*) were inserted into the mammalian cell line's genome at specific locations independently of one another. The FRT site is a target for Flp recombinase-targeted integration. The *TetR* expresses a tet repressor (TetR) protein that binds the Tet Operator (TO) and inhibits expression from the CMV/TO hybrid promoter. Tetracycline binds the TetR protein and derepresses gene expression from the CMV/TO promoter.

Plasmids containing an FRT site, in addition to the (G4C2)n repeat constructs under control of the CMV/TO promoter were also generated. These (G4C2)n plasmids were then integrated via Flp recombinase-mediated recombination into the genomic FRT site

in the mammalian cell lines. In addition, sham control cell lines were generated by integrating empty plasmids. Each of the resultant HEK293 cell lines were isogenic with one another, and each of the resultant NSC34 cell lines were isogenic with one another, and all cell lines had tetracycline-inducible expression of the (G4C2)n repeats.

3.2. Aims and Objectives for Cell Model Generation and Characterisation

1) Clone the (G4C2)10, (G4C2)51, and (G4C2)102 constructs into the pcDNA5/FRT/TO vector.
2) Integrate the pcDNA5/FRT/TO-(G4C2)n plasmids into commercially available Flp-In™ T-REx™ HEK293 cell line to validate (G4C2)n expression.
3) Integrate the pcDNA5/FRT/TO-(G4C2)n plasmids into Flp-In™ T-REx™ NSC34 cell line.
4) Characterise the NSC34 (G4C2)n cells for tetracycline-inducible expression of (G4C2)n RNA.
5) Characterise the NSC34 (G4C2)n cells for antisense (C4G2)n RNA expression.
6) Characterise the NSC34 (G4C2)n cells for RAN translation products.
7) Investigate whether (G4C2)n expression causes NSC34 cell death or affects NSC34 cell viability.

3.3. Results

3.3.1. pcDNA5/FRT/TO-(G4C2)n Construction

To stably integrate the (G4C2)n constructs into the genome of the host Flp-In™ T-REx™ cell line, vectors containing the (G4C2)n constructs (under control of a CMV/TO hybrid promoter), an FRT site, and a gene for hygromycin resistance were first generated. The pcDNA6.2-(G4C2)n plasmids were cut using *Dra*I and *Xho*I restriction enzymes, and then the (G4C2)n constructs were purified via agarose gel electrophoresis and subsequent gel extraction. pcDNA5/FRT/TO-HIS was cut using *Eco*RV and *Xho*I restriction enzymes, and the linearised vector backbone was purified via agarose gel electrophoresis and gel

extraction. The linearised pcDNA5/FRT/TO-HIS vector backbone was then treated with Shrimp Alkaline Phosphatase (SAP) to reduce the chance of the vector religating. The DNA fragments containing the (G4C2)n constructs were then ligated into the pcDNA5/FRT/TO-HIS vector backbone in a forced orientation, due to the *Xho*I-cut and blunt ends of the DNA fragments. During initial cloning attempts, the ligation reactions containing the (G4C2)n repeats actually produced fewer *E.coli* colonies after transformation than the re-ligation control reactions. It was proposed that one end of the (G4C2)n repeat fragment was binding the cut pcDNA5/FRT/TO-HIS vector, but the other end was buried or unavailable to bind due to secondary structure formed by the (G4C2)n repeat, thus inhibiting the ligation reaction. Importantly, it was discovered that heating the fragments containing the (G4C2)n constructs to 70 °C for 10 mins, followed by snap cooling on ice, dramatically improved the ligation efficiency. This heating and snap cooling step is suggested to denature and thus remove secondary structure inherent in the (G4C2)n repeat, allowing efficient ligation. The ligation mixtures were then transformed into β-10 competent *E.coli*, and subsequent colonies were screened for successful (G4C2)n construct insertion using PCR (Figure 3.3). The re-ligation control produced a PCR product of approximately 300 bp (Figure 3.3). Unexpectedly however, successful insertion of the (G4C2)10, (G4C2)51, or (G4C2)102 repeats into pcDNA5/FRT/TO-HIS all produced PCR products of a similar size around 400 bp (Figure 3.3). This increase in PCR product size is not as large as expected because the PCR fails to completely process the full (G4C2)n repeat, and demonstrates why the (G4C2)n repeat could not be cloned via PCR.

Colonies that tested positive for the (G4C2)n construct insertion into the pcDNA5/FRT/TO-HIS backbone vector were cultured and miniprepped. To size the (G4C2)n constructs, the pcDNA5/FRT/TO-HIS-(G4C2)n plasmids were digested using *BamH*I and *Xho*I restriction enzymes and run on an agarose gel (Figure 3.4). The correct size fragments for successful insertion of the (G4C2)10, (G4C2)51, and (G4C2)102 constructs were 110 bp, 400 bp, and 750 bp respectively. pcDNA5/FRT/TO-HIS-(G4C2)n plasmids containing the correct size (G4C2)n insert (Figure 3.4) were sequenced using the CMV-Forward primer. This confirmed the (G4C2)n repeat lengths were correct in the

pcDNA5/FRT/TO-HIS-(G4C2)10, pcDNA5/FRT/TO-HIS-(G4C2)51, and pcDNA5/FRT/TO-HIS-(G4C2)102 plasmids.

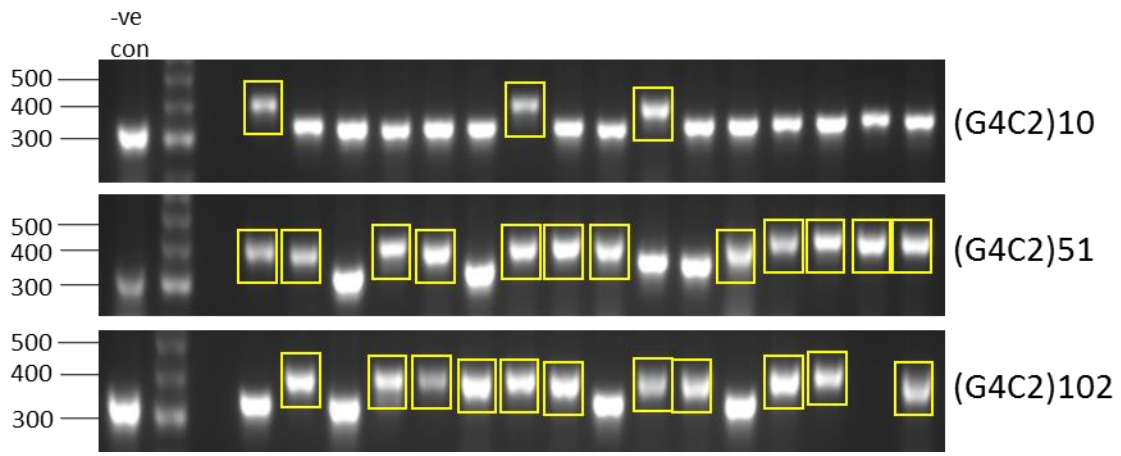


Figure 3.3 Colony PCR screen for (G4C2)n ligation into pcDNA5/FRT/TO-HIS backbone vector. DNA fragments containing the (G4C2)10, (G4C2)51, and (G4C2)102 constructs were each ligated into pcDNA5/FRT/TO-HIS, and subsequently transformed into β10 *E.coli.* Colonies were then picked and screened using PCR and gel electrophoresis. Colonies that produced a band shift (boxed in yellow) contain the (G4C2)n construct. A religated pcDNA5/FRT/TO-HIS was used as a negative control. Molecular weight markers are indicated (bp).

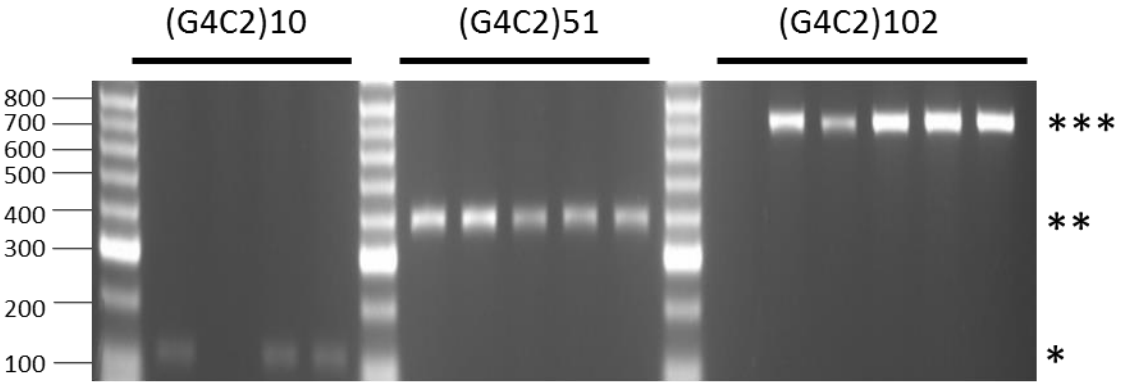


Figure 3.4 Restriction digest of pcDNA5/FRT/TO-HIS-(G4C2)n plasmids to size the (G4C2)n constructs. The (G4C2)n constructs were cut from the pcDNA5/FRT/TO-HIS-(G4C2)n plasmids using *BamH*I and *Xho*I restriction enzymes. Agarose gel electrophoresis was performed to size the (G4C2)n constructs. The correct size for bands containing (G4C2)10, (G4C2)51, and (G4C2)102 are shown with *, **, and *** respectively. Molecular weight markers are indicated (bp).

The pcDNA5/FRT/TO-HIS-(G4C2)n vectors contain a HIS tag 5' of, and adjacent to, the (G4C2)n construct. The HIS tag contains an ATG start codon and could therefore lead to

conventional translation of the (G4C2)n repeat. To prevent conventional translation of the (G4C2)n constructs in the mammalian cell models, the HIS tag and the associated ATG start codon were removed. The pcDNA5/FRT/TO-HIS-(G4C2)n vectors were cut using *Hind*III and *BamH*I restriction enzymes, blunt ended using DNA polymerase I (Klenow), treated with PNK, religated and then transformed into β-10 competent *E.coli*. Colonies containing potential pcDNA5/FRT/TO-(G4C2)n plasmids were cultured and miniprepped, then the plasmids were screened using *Nco*I restriction digestion and agarose gel electrophoresis (Figure 3.5). An *Nco*I site resides in the HIS tag, therefore plasmids containing the HIS tag produced 3 bands from the *Nco*I digest, whilst plasmids with the HIS tag removed produced 2 bands. A pcDNA5/FRT/TO-HIS plasmid was digested using *Nco*I, and served as a negative control for HIS tag removal (Figure 3.5). pcDNA5/FRT/TO-(G4C2)n plasmids that had the HIS tag removed were sequenced using the CMV-Forward primer to confirm the removal of the HIS tag and ATG start codon, as well as to sequence and confirm the exact length of the (G4C2)n construct. pcDNA5/FRT/TO-(G4C2)10, pcDNA5/FRT/TO-(G4C2)51, and pcDNA5/FRT/TO-(G4C2)102 were all successfully generated, and the sequencing chromatograms for the (G4C2)n constructs are included in Appendices 1-3.

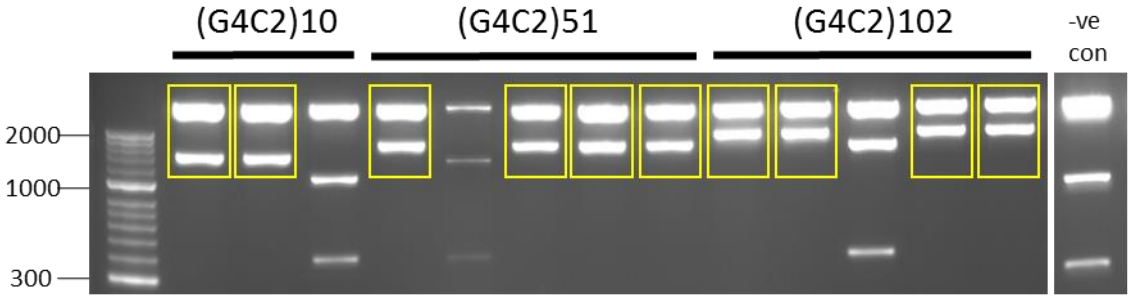


Figure 3.5 Restriction digest of pcDNA5/FRT/TO-(G4C2)n plasmids to screen for HIS tag removal. pcDNA5/FRT/TO-(G4C2)n plasmids were cut using *Nco*I, and agarose gel electrophoresis was performed to check for the removal of the HIS tag and the associated ATG start codon. Two bands indicate removal of HIS tag (boxed in yellow), whilst three bands indicate the HIS tag is still present. pcDNA5/FRT/TO-HIS was used as a negative control. Molecular weight markers are indicated (bp).

3.3.2. Flp-In™ T-REx™ HEK293 (G4C2)n Cell Generation and Characterisation

The pcDNA5/FRT/TO-(G4C2)n plasmids were first stably transformed into the commercially available Flp-In™ T-REx™ HEK293 cells, to test these plasmids in the Flp-In™ and T-REx™ systems. pPGKFLPobpA (which expresses a codon optimised FLP DNA-recombinase) was co-transfected with each of the pcDNA5/FRT/TO, pcDNA5/FRT/TO-(G4C2)10, pcDNA5/FRT/TO-(G4C2)51, and pcDNA5/FRT/TO-(G4C2)102 plasmids separately. The FLP DNA recombinase catalysed recombination between the FRT site on the pcDNA5/FRT/TO-(G4C2)n plasmids and the genomic FRT site in the Flp-In™ T-REx™ HEK293 cells, stably inserting the empty sham vector or tetracycline-inducible (G4C2)n repeat constructs. The media was removed and replenished 24 h post transfection. Blasticidin and hygromycin selection was added to the media to select for cells which had pcDNA5/FRT/TO-(G4C2)n stably integrated into the genomic FRT site. A lot of the transfected Flp-In™ T-REx™ HEK293 cells were dead after 5 days of blasticidin/hygromycin selection. Therefore, the media was removed and replaced with conditioned media (containing blasticidin/hygromycin selection) to help support the successfully transformed Flp-In™ T-REx™ HEK293 cells. Colonies of hygromycin-resistant Flp-In™ T-REx™ HEK293 (G4C2)n cells were visible after 15 days of selection. These transformed colonies were then picked and expanded in media containing blasticidin/hygromycin selection on 48 well plates. When cells were at a sufficient confluency, they were screened for Zeocin™-sensitivity and blasticidin/hygromycin resistance. Zeocin™-resistance would suggest the pcDNA5/FRT/TO-(G4C2)n plasmid had integrated randomly into the genome, and not at the genomic FRT site in that particular clonal line. Each clone of the Flp-In™ T-REx™ HEK293 (G4C2)n cells were split into two separate wells, and each population was grown in either blasticidin/Zeocin™ or blasticidin/hygromycin selection. Clones that were sensitive to Zeocin™, but grew well in the blasticidin and hygromycin were expanded and frozen down.

3.3.2.1. Flp-In™ T-REx™ HEK293 (G4C2)n Cells Express (G4C2)n RNA Foci

RNA FISH was used to detect (G4C2)n RNA expression in the HEK293 (G4C2)n cells. The (G4C2)n repeat is transcribed and forms characteristic RNA foci in the CNS of C9ORF72-ALS/FTD patients (DeJesus-Hernandez et al., 2011; Renton et al., 2011). RNA FISH was

performed using a sense-specific fluorescently labelled (C4G2)3 Locked Nucleic Acid probe, to check whether the Flp-In™ T-REx™ HEK293 (G4C2)n cells expressed the (G4C2)n repeat RNA. A confocal microscope was use to image RNA foci (section 2.2.3.10).

(G4C2)n RNA foci were not detected in the HEK293 sham cells ± tetracycline induction (Figure 3.6). (G4C2)n RNA foci were detected in the HEK293 (G4C2)102 cells ± tetracycline induction, however, there were more (G4C2)n RNA foci in tetracycline induced HEK293 (G4C2)102 cells (Figure 3.6). In addition, the RNA foci were visibly larger (Figure 3.6). The (G4C2)n RNA foci were predominantly nuclear, although rare cytoplasmic (G4C2)n RNA foci were also detected (Figure 3.6). Additionally, when HEK293 (G4C2)102 cells were treated with RNAse A prior to the RNA FISH stain, the (G4C2)n RNA foci were ablated (Figure 3.6). These observations were consistent in three experimental repeats.

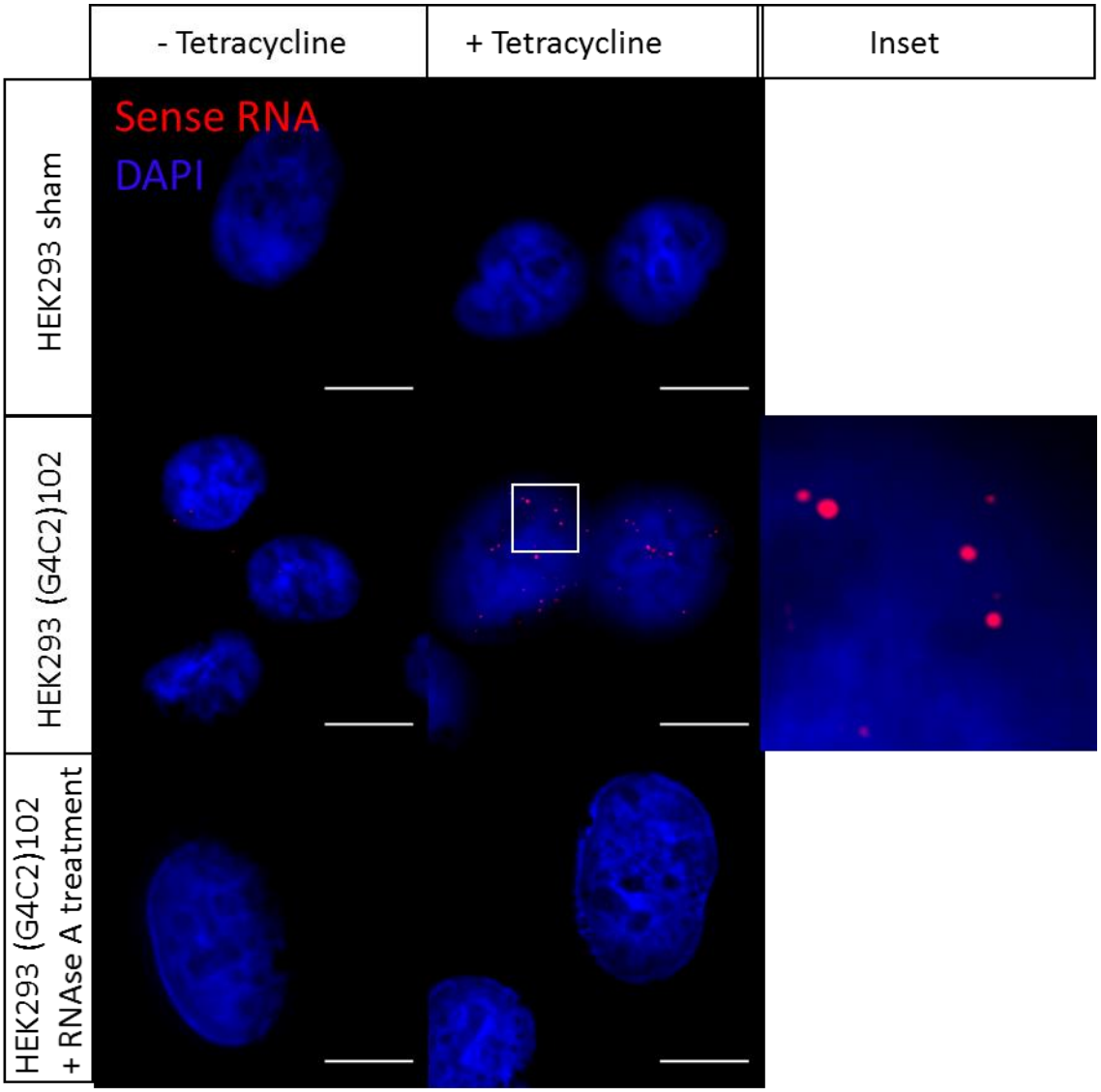


Figure 3.6 HEK293 (G4C2)n cells have tetracycline-inducible (G4C2)n RNA expression, which forms RNA foci. HEK293 sham and HEK293 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. RNAse treated control cells were additionally treated with RNAse A after PFA fixation and prior to RNA FISH staining. Cells were stained with a fluorescently labelled Locked Nucleic Acid (C4G2)3 sense probe (Red) and Dapi (Blue). RNA foci were imaged using a confocal microscope within a high resolution z-stack through the entire nuclear volume. The images shown are projections of the z-stack to show all RNA foci imaged in the cells. Foci magnified 5X inset. Scale bar = 10µm.

3.3.3. Flp-In™ T-REx™ NSC34 (G4C2)n Cell Generation

To generate isogenic motor neuron-like NSC34 cell lines with tetracycline-inducible (G4C2)n repeat expression, a Flp-In™ T-REx™ NSC34 host cell line first had to be generated using the Flp-In™ T-REx™ core kit. pFRT/*lac*Zeo and pcDNA6/TR plasmids were independently and stably transfected into the WT NSC34 cell line. pFRT/*lac*Zeo

introduced a single FRT site into the NSC34 genome, which acted as an isogenic target site for pcDNA5/FRT/TO-(G4C2)n vector integration. pcDNA6/TR introduced a *TetR* element into the NSC34 cell line, which constitutively and stably expresses the TetR protein. The TetR protein binds the CMV/TO hybrid promoter and represses transcription of the (G4C2)n constructs.

3.3.3.1. Flp-In™ T-REx™ NSC34 Host Cell Generation

The work in section 3.3.3.1 was exclusively carried out by Dr Adrian Higginbottom. First, it was determined that 2.5µg/mL of blasticidin, 75µg/mL of hygromycin and 20µg/mL of Zeocin™ were the minimum respective concentrations that wild type (WT) NSC34 cells were sensitive to. pFRT/*lac*Zeo was transfected into WT NSC34 cells to introduce a target FRT site within the NSC34 genome. Zeocin™ selection was added to the NSC34 media to select for Zeocin™-resistant Flp-In™ NSC34 cells stably transfected with pFRT/*lac*Zeo. Zeocin™-resistant clones were picked, expanded, and banked. Genomic DNA was then extracted from these Zeocin™-resistant Flp-In™ NSC34 clones, and screened using Southern Blotting to test for the number of integrated FRT sites (Figure 3.7). The fibroblast negative control did not contain a genomic FRT site, whilst the Flp-In™ T-REx™ HEK293 positive control did contain one genomic FRT site (Figure 3.7). Flp-In™ NSC34 clones that contained only one genomic FRT site were further analysed for the transcriptional activity of that FRT site.

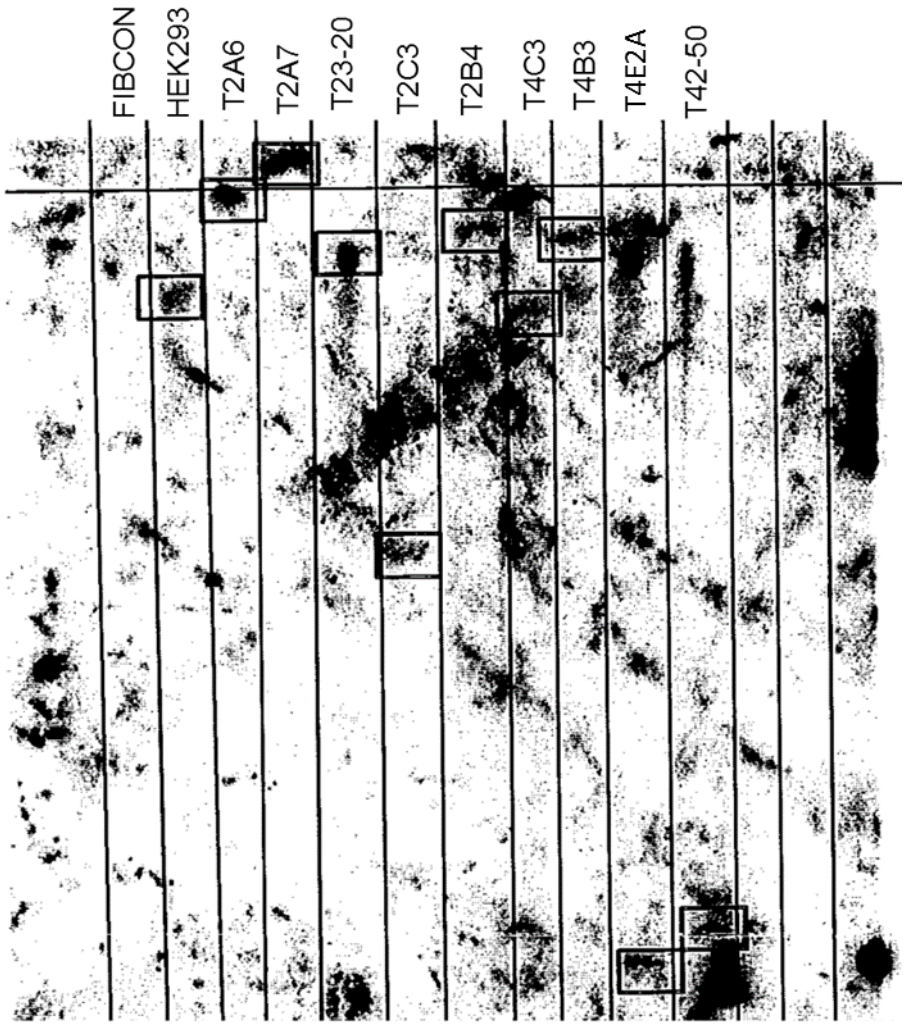


Figure 3.7 Zeocin™-resistant Flp-In™ NSC34 clones were screened for single FRT integrations using Southern Blot. DNA bands containing a positive FRT fragment from the genomic DNA are boxed. Fibroblast (FIBCON) was used as a negative control, and Flp-In™ T-REx™ HEK293 (HEK293) was used as a positive control. Image courtesy of Dr. Adrian Higginbottom.

Flp-In™ NSC34 clones that contained only one genomic FRT site were transfected with pcDNA5/FRT/TO-GFP to test the transcriptional activity of the FRT site. The level of GFP expression was used to determine what effect the chromosomal position of the FRT site had on its transcriptional activity. Hygromycin selection was added to Flp-In™ NSC34 media to select for hygromycin-resistant Flp-In™ NSC34 GFP cells stably transfected with pcDNA5/FRT/TO-GFP. These Flp-In™ NSC34 GFP cells were then screened using a fluorescent microscope for GFP expression levels. Clones TC3, T42, T23 and T4E2A expressed the highest GFP levels and were transfected again with pcDNA6/TR. This

would determine which Flp-In™ NSC34 clone had a transcriptionally active FRT site also capable of being repressed by the *TetR* element. Blasticidin and hygromycin selection were added to Flp-In™ NSC34 GFP media to select for blasticidin/hygromycin-resistant Flp-In™ T-REx™ NSC34 GFP cells stably transfected with pcDNA6/TR. These clonal mixtures of stable Flp-In™ T-REx™ NSC34 GFP cells were screened for GFP expression ± tetracycline induction. Flp-In™ T-REx™ NSC34 clonal line T4E2A had the greatest increase in GFP expression upon tetracycline induction (Figure 3.8), indicating the FRT site was in a transcriptionally active genomic location, which was also capable of repression by the *TetR* element.

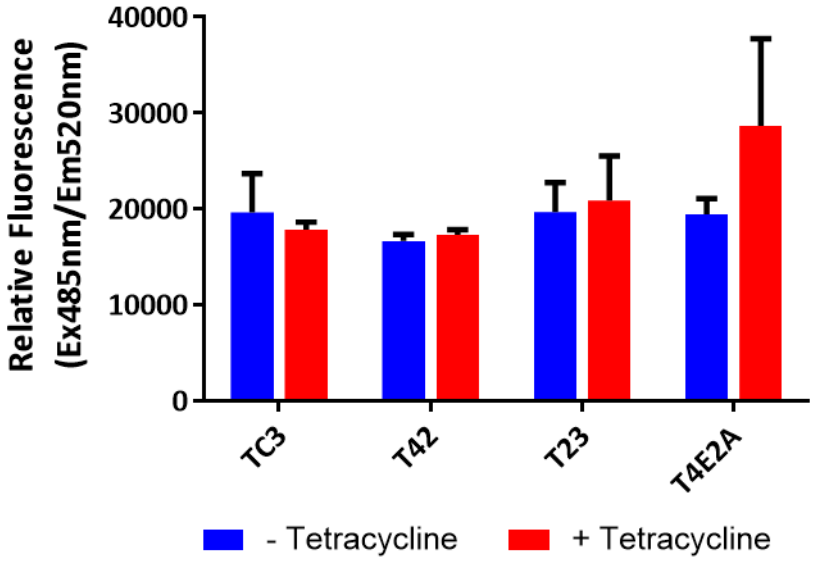


Figure 3.8 Screen for Flp-In™ NSC34 cells with transcriptionally active but repressible FRT sites. Flp-In™ NSC34 cells were transfected with and selected for stable integration of pcDNA5/FRT/TO-GFP and subsequently pcDNA6/TR. Cells were then cultured for 2 days ±5µg/mL tetracycline, and then Ex485/Em520nm fluorescence was measured to assess GFP expression levels. Data shown are mean and SD of technical replicates; n=1.

pcDNA6/TR was transfected into the T4E2A Flp-In™ NSC34 clonal line to stably introduce the *TetR* element. Blasticidin and Zeocin™ selection was added to the Flp-In™ T-REx™ NSC34 media to select for blasticidin/Zeocin™-resistant Flp-In™ T-REx™ NSC34 cells stably transfected with pcDNA6/TR. Blasticidin/Zeocin™-resistant clones were picked, expanded, and banked. Flp-In™ T-REx™ NSC34 clones were transfected with pcDNA5/FRT/TO-GFP to identify a clone which had low basal GFP expression and high

GFP expression upon tetracycline induction. Blasticidin/hygromycin selection was added to the Flp-In™ T-REx™ NSC34 GFP media to select for blasticidin/hygromycin-resistant Flp-In™ T-REx™ NSC34 GFP cells stably transfected with pcDNA5/FRT/TO-GFP. Blasticidin/hygromycin-resistant Flp-In™ T-REx™ NSC34 GFP clones were also screened for Zeocin™ sensitivity, which indicates the pcDNA5/FRT/TO-GFP inserted at the genomic FRT site. Blasticidin/hygromycin-resistant, Zeocin™-sensitive Flp-In™ T-REx™ NSC34 GFP clones were then screened for GFP expression ± tetracycline induction. Flp-In™ T-REx™ NSC34 clonal line E3 expressed very little GFP at basal levels, but there was also very little increase in GFP expression upon tetracycline induction (Figure 3.9). Clonal line C2 had high GFP expression at basal levels, and very high GFP expression when induced with tetracycline (Figure 3.9). It was decided that the E3 clone was too weak an inducer, and the C2 clone expressed the GFP too strongly at basal levels. Therefore neither E3 nor C2 were used as the parental Flp-In™ T-REx™ NSC34 line. However, the B10-2 clone had modest GFP expression at basal levels because the fluorescence intensity was similar to the background fluorescence seen in WT NSC34 cells (Figure 3.9). In addition, the B10-2 clone had high GFP expression when induced with tetracycline (Figure 3.9). Therefore clonal line B10-2 was selected as the parental Flp-In™ T-REx™ NSC34 host cell line for generating the isogenic, tetracycline-inducible Flp-In™ T-REx™ NSC34 (G4C2)n cell lines, as well as other lines with different genes of interest.

3.3.3.2. Stable Integration of pcDNA5/FRT/TO-(G4C2)n Plasmids into Flp-In™ T-REx™ NSC34 Cells

The same method described for the Flp-In™ T-REx™ HEK293 cells (section 3.3.3) was used to integrate the (G4C2)n repeat expansion vectors into the motor neuron-like NSC34 cell lines. However, the sensitivity of the Flp-In™ T-REx™ NSC34 cells to the hygromycin selection agent had to be determined first, because they were built in house. The Flp-In™ T-REx™ NSC34 cells were seeded onto a 24 well plate and left for 24 h. After 24 h, hygromycin was added to the cells at various concentrations. The cells were left for 7 days in the selection media, and then observed. The NSC34 cells were completely devastated in media containing ≥100µg/mL hygromycin, partially devastated

at 75µg/mL and relatively unaffected at ≤50µg/mL. Therefore, 100µg/mL hygromycin was sufficient to detect stably transfected cells with resistance to hygromycin.

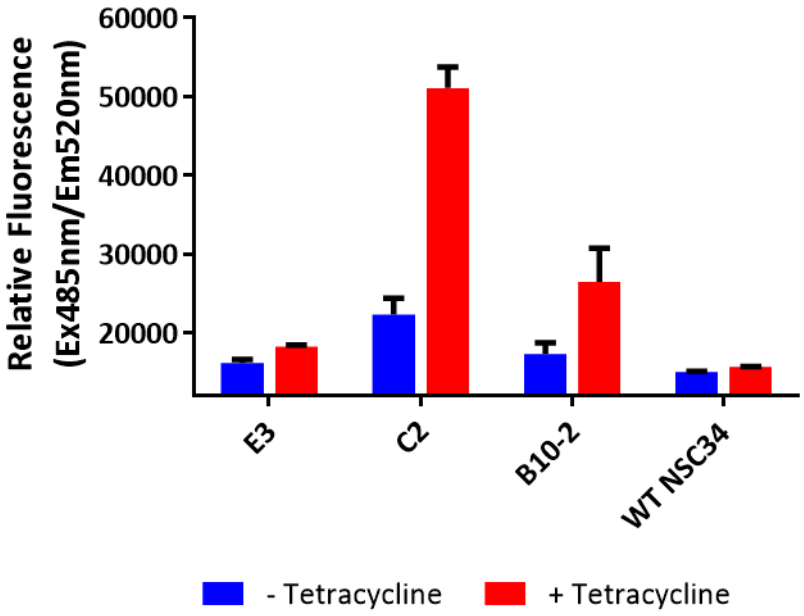


Figure 3.9 Screen for Flp-In™ T-REx™ NSC34 cells with low basal GFP expression and high tetracycline-inducible GFP expression. Flp-In™ T-REx™ NSC34 cells were transfected with and selected for stable integration of pcDNA5/FRT/TO-GFP. Cells were then cultured for 2 days ±10µg/mL tetracycline, and then Ex4850/Em520nm fluorescence was measured to assess GFP expression levels. Data shown are mean and SD of technical replicates; n=1.

The pcDNA5/FRT/TO-(G4C2)n plasmids were stably transformed into the Flp-In™ T-REx™ NSC34 cells. pPGKFLPobpA was co-transfected with each of the pcDNA5/FRT/TO, pcDNA5/FRT/TO-(G4C2)10, pcDNA5/FRT/TO-(G4C2)51, and pcDNA5/FRT/TO-(G4C2)102 plasmids separately. 48 h post-transfection, the media was replenished, the NSC34 cells were split onto 3 new plates and 5µg/mL blasticidin and 100µg/mL hygromycin were added to the media to select for transformed clones. The media and dead floating cells were removed and replaced with fresh selection media every 3 days. 50% conditioned/50% fresh media with selection was added to the plates once massive cell death had occurred and only very few cells remained alive on the plates. This was to ensure survival of any transformed colonies remaining on the plates. Once distinct colonies were visible on the plate, they were picked and transferred to a 48 well plate containing blasticidin and hygromycin selection. The NSC34 colonies were grown in the

48 well (each colony in a separate well) until 60% confluent. They were then split 50/50 into two 48 wells: one containing Zeocin™, the other containing blasticidin/hygromycin. Clones that were blasticidin/hygromycin-resistant but Zeocin™-sensitive were expanded and cryopreserved.

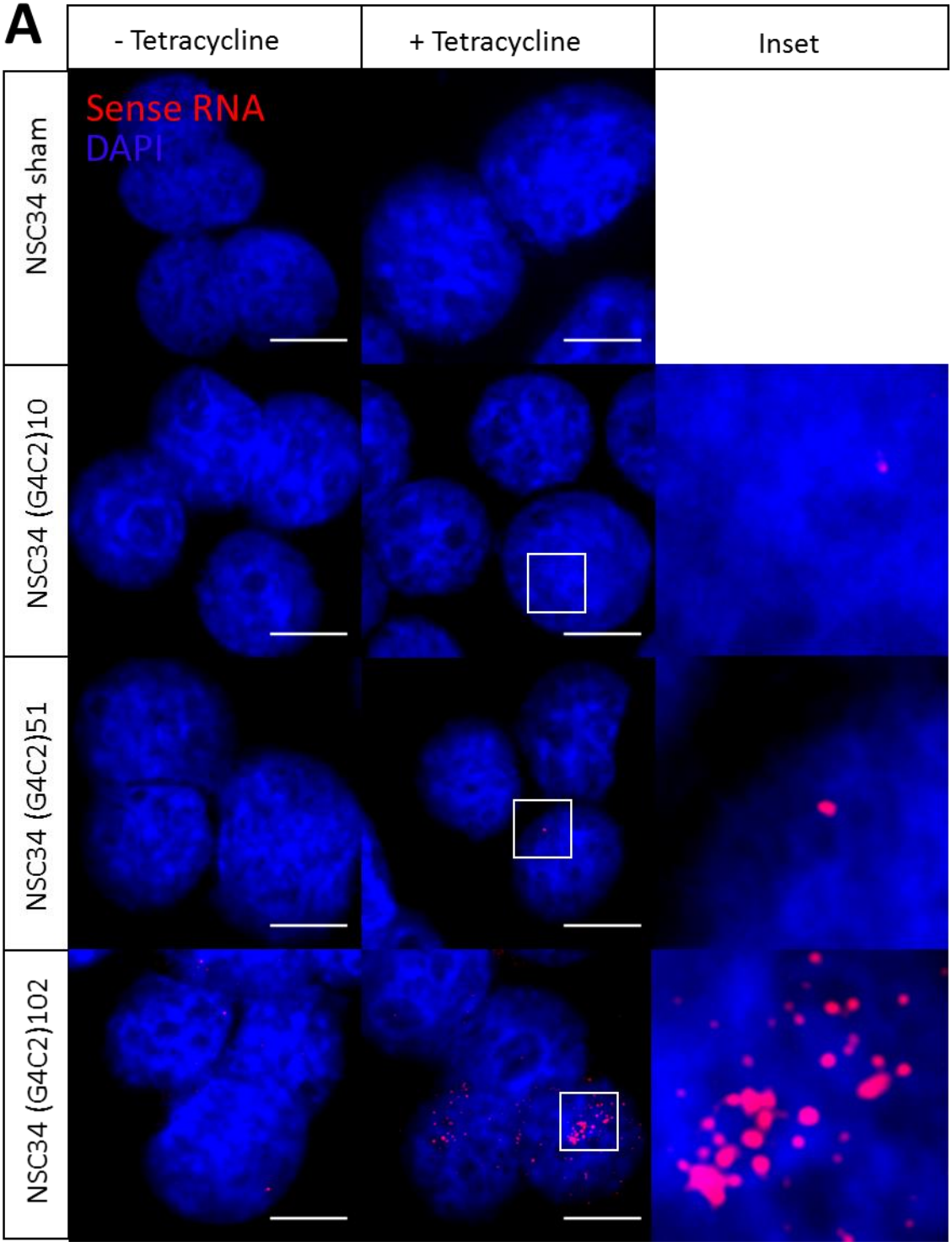
3.3.4. Characterisation of Flp-In™ T-REx™ NSC34 (G4C2)n Cells

The Flp-In™ T-REx™ NSC34 (G4C2)n cells were then characterised for expression of the (G4C2)n repeat constructs. The (G4C2)n repeat expansion is transcribed bidirectionally and forms both sense (G4C2)n and antisense (G4C2)n RNA foci in C9ORF72-ALS/FTD patient CNS (DeJesus-Hernandez et al., 2011; Gendron et al., 2013; Renton et al., 2011). In addition, the (G4C2)n and (C4G2)n repeat transcripts are translated via RAN translation to form DPR proteins (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013a; Mori et al., 2013c). Therefore, RNA FISH was performed (using both sense and antisense probes), and immunoblotting (using anti-DPR antibodies), to detect expression of the (G4C2)n repeats in the NSC34 cells. In addition, the concentration of tetracycline was optimised to minimise toxicity, but retain maximal induction of (G4C2)n expression. From now on, the Flp-In™ T-REx™ NSC34 (G4C2)n cells will be referred to as NSC34 (G4C2)n.

3.3.4.1. NSC34 (G4C2)n Cells Express (G4C2)n RNA Foci

RNA FISH was performed using the sense-specific fluorescent probe to characterise the NSC34 (G4C2)n cells for tetracycline-inducible (G4C2)n RNA expression. The number of sense (G4C2)n RNA foci was dependent on the (G4C2)n repeat size (Figure 3.10). NSC34 sham cells contained 0.10±0.10 and 0.10±0.10 (G4C2)n RNA foci per cell, without and with tetracycline respectively. NSC34 (G4C2)10 cells contained 0.04±0.04 and 0.89±1.13 RNA foci per cell, without and with tetracycline respectively. NSC34 (G4C2)51 cells contained 0.63±0.24 (G4C2)n and 1.91±1.31 (G4C2)n RNA foci per cell, without and with tetracycline respectively. NSC34 (G4C2)102 cells contained 7.12±3.71 (G4C2)n and 17.69±5.37 (G4C2)n RNA foci per cell, without and with tetracycline respectively. However, whilst the number of (G4C2)n RNA foci increased in the NSC34 (G4C2)10, NSC34 (G4C2)51 and NSC34 (G4C2)102 cells when induced with 10μg/mL tetracycline (Figure 3.10), this increase was only significant (using a Two-Way ANOVA with Tukey's

multiple comparisons post-hoc test) in the NSC34 (G4C2)102 cells (P<0.01). Additionally, the size of the RNA foci visibly correlates with the size of the (G4C2)n repeat within the NSC34 cells (Figure 3.10). Also, whilst the vast majority of RNA foci were nuclear, very rare cytoplasmic RNA foci were also observed in NSC34 (G4C2)102 cells (Figure 3.11). Finally, RNAse A treatment ablated foci in the NSC34 (G4C2)102 cells (Figure 3.11).



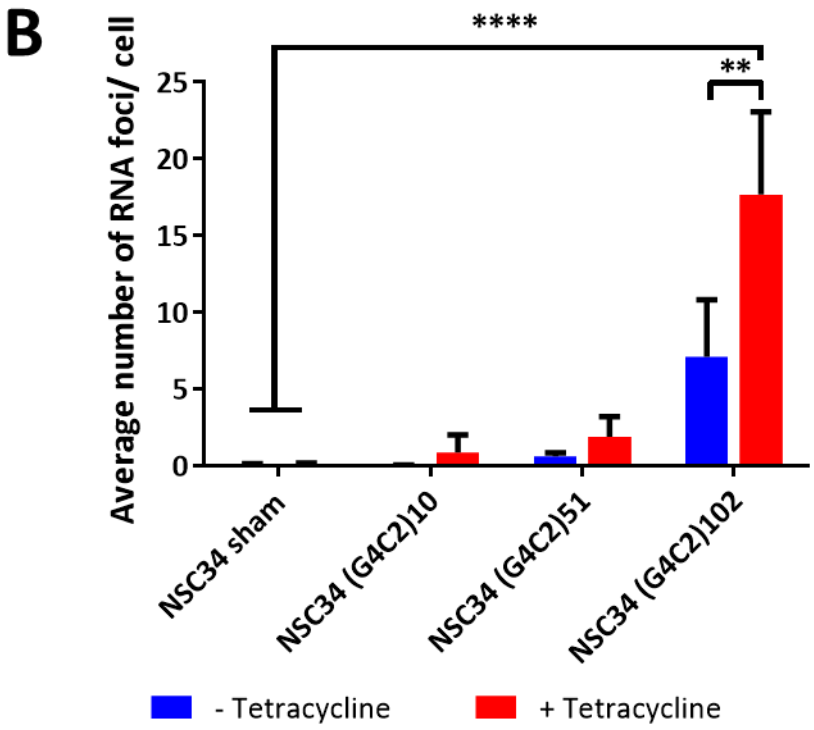


Figure 3.10 NSC34 (G4C2)n cells have tetracycline-inducible (G4C2)n RNA expression, which forms RNA foci. A) NSC34 sham, NSC34 (G4C2)10, NSC34 (G4C2)51, and NSC34 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. Cells were stained with a fluorescently labelled Locked Nucleic Acid (C4G2)3 sense probe (Red) and Dapi (Blue). RNA foci were imaged using a confocal microscope within a high resolution z-stack through the entire nuclear volume. The images shown are projections of the z-stack to show all RNA foci imaged in the cells. Foci magnified 5 X inset. Scale bar = 10µm. **B)** Average number of RNA foci per cell (**P<0.01; ****P<0.0001; Two-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; n=3).

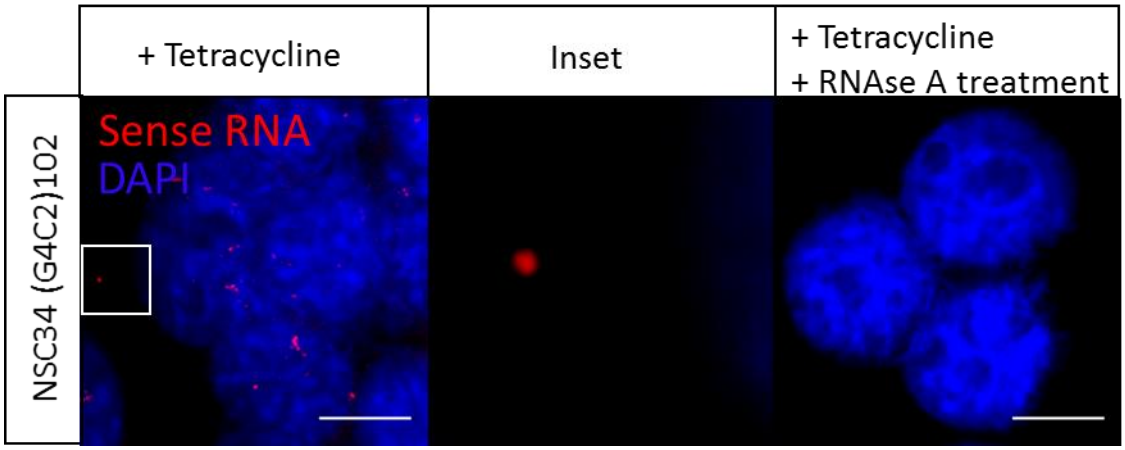


Figure 3.11 Foci are ablated by RNAse A treatment. NSC34 (G4C2)102 cells were cultured for 3 days with 10µg/mL tetracycline. Cells were additionally treated with RNAse A prior to RNA FISH staining. Cells were stained with a fluorescently labelled Locked Nucleic Acid (C4G2)3 sense probe (Red) and Dapi (Blue). RNA foci were imaged using a confocal microscope within a high resolution z-stack through the entire nuclear volume. The images shown are projections of the z-stack to show all RNA foci imaged in the cells. Foci magnified 5X inset. Scale bar = 10µm.

3.3.4.2. NSC34 (G4C2)n Cells Do Not Express Antisense (C4G2)n RNA Foci

The (G4C2)n repeat expansion is also transcribed from a cryptic promoter in the antisense direction in C9ORF72-ALS/FTD patients, and forms characteristic antisense (C4G2)n RNA foci in the CNS of C9ORF72-ALS/FTD patients (Gendron et al., 2013). The (G4C2)n repeat is under control of a sense orientated CMV/TO promoter in the NSC34 (G4C2)n cells, and should not be transcribed in the antisense direction. However, the exact insertion site of the FRT during the cloning process is unknown, and a promoter running in the antisense orientation is possible. Therefore, RNA FISH was performed using an antisense-specific fluorescently labelled (G4C2)3 Locked Nucleic Acid probe, to confirm the absence of antisense (C4G2)n RNA foci in the NSC34 (G4C2)102 cells. No antisense (C4G2)n RNA foci were detected in either the NSC34 sham or the NSC34 (G4C2)102 cells (Figure 3.12). As a positive control for the antisense-specific FISH probe, (C4G2)102 constructs were transiently transfected into HEK293 cells (Figure 3.12).

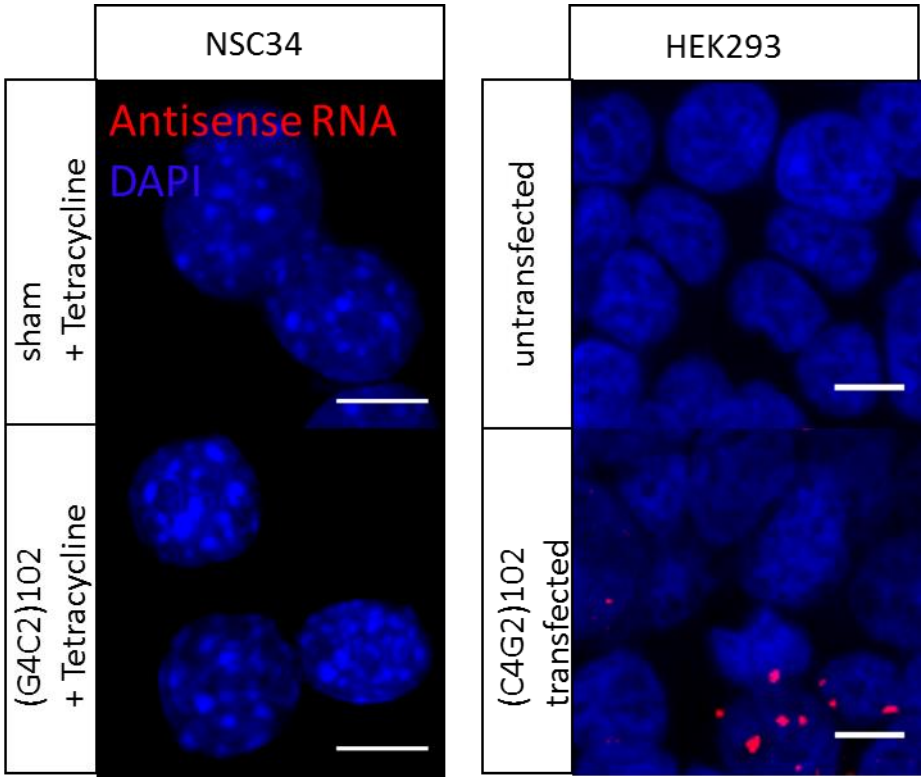


Figure 3.12 NSC34 (G4C2)n cells do not transcribe (G4C2)n in the antisense direction. NSC34 sham and NSC34 (G4C2)102 cells were cultured for 3 days with 10µg/mL tetracycline. HEK293 cells transfected with a (C4G2)102 plasmid contain (C4G2)102 foci, and serve as a positive control for the antisense Locked Nucleic Acid probe. Cells were stained with a fluorescently labelled Locked Nucleic Acid (G4C2)3 antisense probe (Red) and Dapi (Blue). RNA foci were imaged using a confocal microscope within a high resolution z-stack through the entire nuclear volume. The images shown are projections of the z-stack to show all RNA foci imaged in the cells. Scale bar = 10µm.

3.3.4.3. The (G4C2)102 Repeat RNA Is Translated in Both the NSC34 and HEK293 cells

The (G4C2)n repeat expansion undergoes Repeat Associated Non-ATG (RAN) translation in cells throughout the CNS of C9ORF72-ALS/FTD patients (Ash et al. 2013; Gendron et al. 2013; Mori et al. 2013a; Mori et al. 2013b). The sense (G4C2)n and antisense (C4G2)n RNA are both translated in all three reading frames, generating five DPR proteins. (GA)n and (GR)n are produced specifically from the sense (G4C2)n RNA, (PA)n and (PR)n are produced specifically from the antisense (C4G2)n RNA, and (GP)n is produced from both sense and antisense RNA (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013a; Mori et al., 2013c). Therefore, bioinformatics and immunoblotting were used to investigate whether the (G4C2)102 repeats also undergo RAN translation in the NSC34 and HEK293 cells.

The pcDNA5/FRT/TO-(G4C2)102 plasmid was integrated into the Flp-In™ T-REx™ NSC34 and Flp-In™ T-REx™ HEK293 genomes, and therefore, the plasmid sequence was used to predict (G4C2)102 translation products in both cell lines. The pcDNA5/FRT/TO-(G4C2)102 plasmid sequence was imported to the Translate tool in ExPASy (http://web.expasy.org/translate/). There were no ATG start codons in any reading frame preceding the (G4C2)102 repeat (Figure 3.13). RAN translation does not require an ATG start codon to initiate translation however. Predicted (G4C2)102 RAN translation products would contain all three sense DPR motifs ((GA)n, (GR)n, and (GP)n), due to the 5 bp TCGAC interruptions between the (G4C2)n repeats (Figure 3.13). Some of the sequence flanking the (G4C2)n repeats could also be translated and included in these predicted translation products as well (Figure 3.13). Therefore, the maximum molecular weight of the RAN translation proteins were calculated by exporting the full peptide sequences, from each frame (Figure 3.13), to the Compute pI/Mw tool in ExPASy (http://web.expasy.org/compute_pi/). The predicted maximum molecular weights for the three reading frames were 20.6, 20.3, and 20.4kDa respectively.

Cells were immunoblotted using anti-DPR antibodies to check whether the (G4C2)n repeats do undergo RAN translation. NSC34 sham and (G4C2)102 were cultured for 7 days ±0.5µg/mL tetracycline (section 3.3.4.4), and HEK293 sham and (G4C2)102 were cultured for 3 days ±10µg/mL tetracycline, prior to immunoblotting. WT HEK293 cells were also transiently transfected with (GA)68, (GR)100, (AP)100, and (PR)100 expression plasmids, to act as positive controls for the respective antibodies. Unfortunately, no (GP) positive control was available.

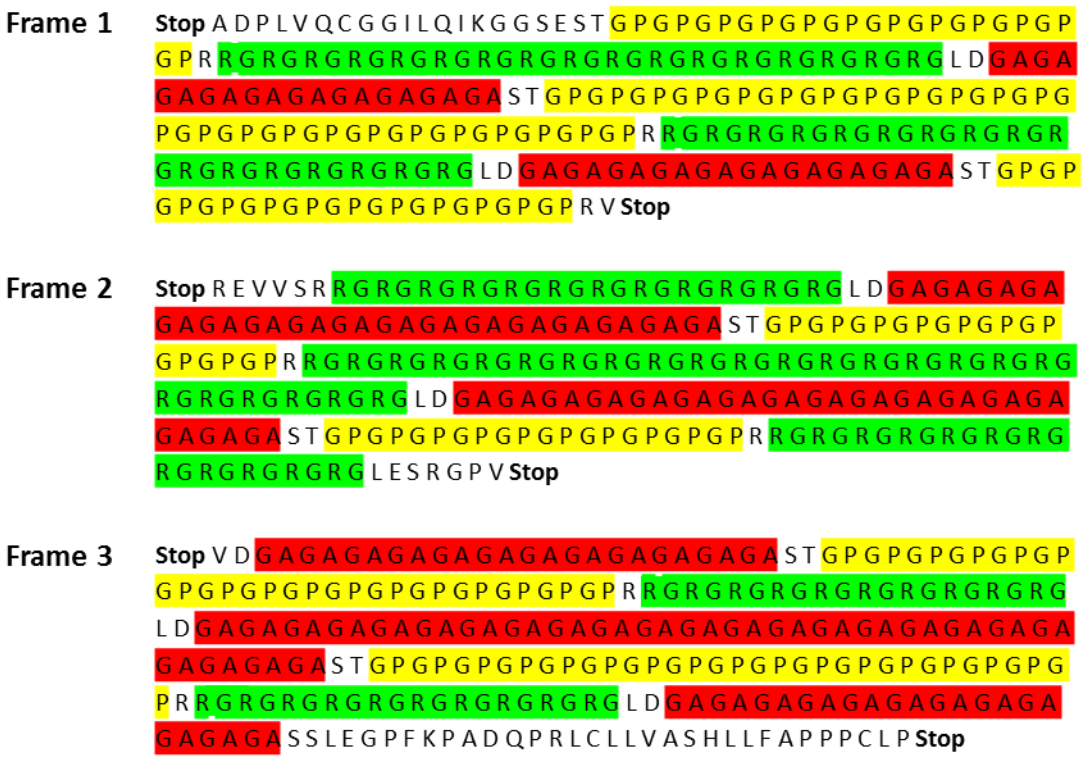


Figure 3.13 Schematic of the predicted (G4C2)102 RAN translation products in NSC34 (G4C2)102 and HEK293 (G4C2)102 cells. The translation products of the (G4C2)n in the NSC34 (G4C2)102 and HEK293 (G4C2)102 cells were predicted by importing the pcDNA5/FRT/TO-(G4C2)102 plasmid sequence into the Translate tool in ExPASy. Translation products (including Stop codons) from (G4C2)102 RNA construct in all three reading frames with GA, GP, and GR repeats highlighted in red, yellow, and green respectively.

3.3.4.3.1. Anti-GA Detects Tetracycline-Inducible Protein(s) in NSC34 (G4C2)102 and HEK293 (G4C2)102 Cells

The anti-GA antibody detected a protein band at approximately 40kDa in the NSC34 sham and NSC34 (G4C2)102 cells ± tetracycline, and also in the WT HEK293 ± (GA)68 transfection (Figure 3.14A). The anti-GA antibody detected a protein at approximately 15kDa specifically in the positive control WT HEK293 transfected with (GA)68 (Figure 3.14A). This band was not detected in NSC34 sham or NSC34 (G4C2)102 ± tetracycline, or the untransfected WT HEK293 cells (Figure 3.14A). However, the anti-GA antibody detected a 27kDa band in the NSC34 (G4C2)102 ± tetracycline, and a further 24kDa band in the NSC34 (G4C2)102 + tetracycline (Figure 3.14A). Neither of these bands at 24 or 27kDa were detected in the NSC34 sham ± tetracycline (Figure 3.14A). Further, the

27kDa band was more abundant in the NSC34 (G4C2)102 induced with tetracycline compared to the NSC34 (G4C2)102 without tetracycline (Figure 3.14A).

The anti-GA antibody detected a protein band at approximately 63kDa in the HEK293 sham and HEK293 (G4C2)102 ± tetracycline (Figure 3.14B). However, the anti-GA antibody detected a 24 and 27kDa band in the HEK293 (G4C2)102 – tetracycline, with the 27kDa band being more abundant (Figure 3.14B). Further, the anti-GA antibody detected as many as 10 different bands ranging from 17 to 35kDa in the HEK293 (G4C2)102 + tetracycline (Figure 3.14B). Also, the bands at 24 and 27kDa were much more abundant (Figure 3.14B). None of these bands were detected in the HEK293 sham ± tetracycline (Figure 3.14B).

3.3.4.3.2. Anti-GR Detects Tetracycline-Inducible Protein(s) in NSC34 (G4C2)102 and HEK293 (G4C2)102 Cells

The anti-GR antibody detected protein bands at approximately 18.5 and 50kDa in the NSC34 sham and NSC34 (G4C2)102 cells ± tetracycline, and also in the WT HEK293 ± (GR)100 transfection (Figure 3.15A). Also, there was a band in the WT HEK293 ± (GR)100 transfection at approximately 80kDa (Figure 3.15A). However, the anti-GR antibody detected protein bands at approximately 30, 35, and >100kDa specifically in the WT HEK293 transfected with (GR)68, which was not detected in NSC34 sham or NSC34 (G4C2)102 ± tetracycline, or the untransfected WT HEK293 cells (Figure 3.15A). Also, the anti-GR antibody detected 24 and 27kDa bands in the NSC34 (G4C2)102 + tetracycline (Figure 3.15A). Neither of these bands at 24 or 27kDa was detected in the NSC34 sham ± tetracycline, or NSC34 (G4C2)102 - tetracycline (Figure 3.15A).

The anti-GR antibody detected protein bands at approximately 19, 30, 55, and 65kDa in the HEK293 sham and HEK293 (G4C2)102 ± tetracycline (Figure 3.15B). However, a 24 and 27kDa band were detected only in the HEK293 (G4C2)102 + tetracycline (Figure 3.15B).

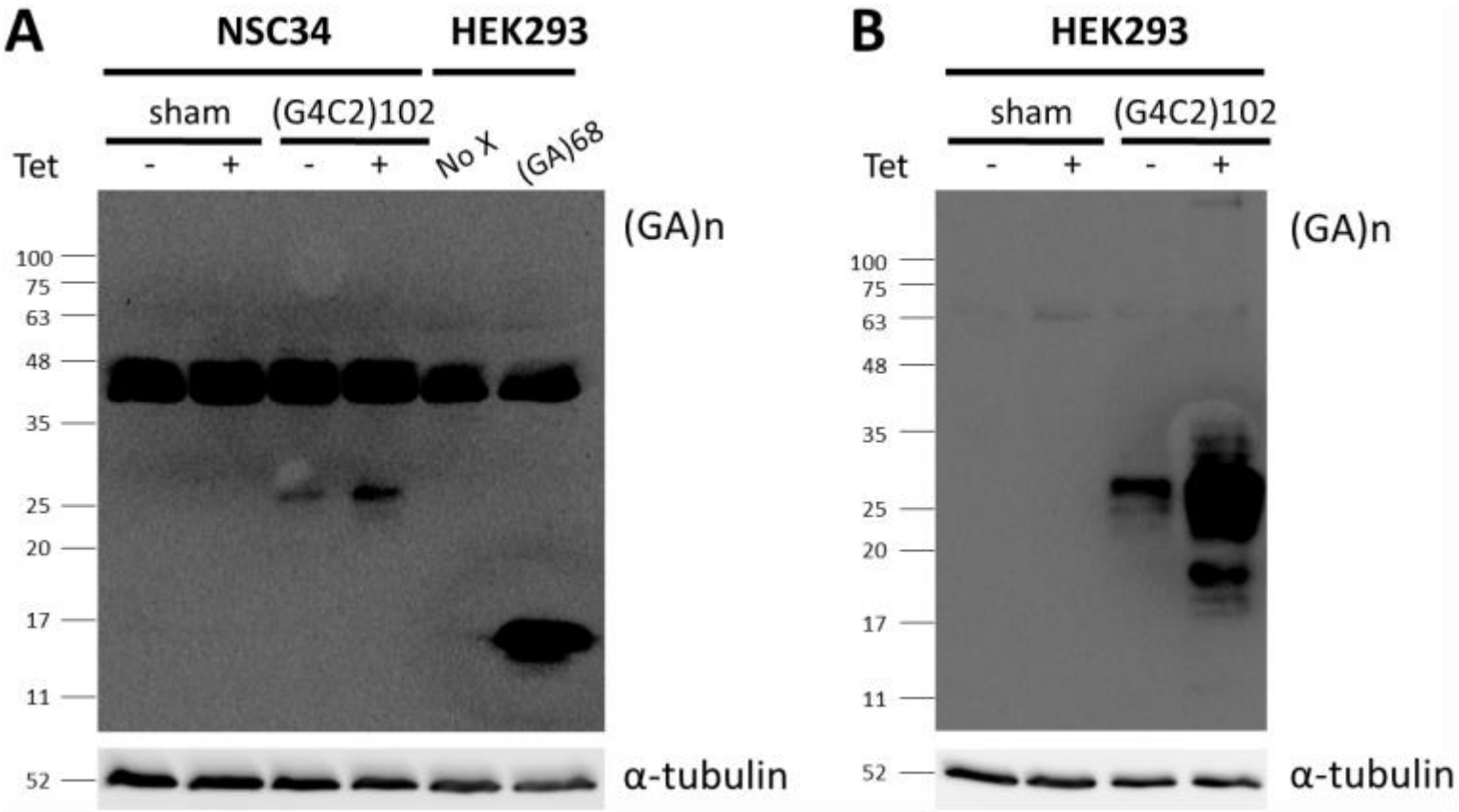


Figure 3.14 The anti-GA antibody detects RAN translation proteins from the (G4C2)102 RNA in the NSC34 (G4C2)102 and HEK293 (G4C2)102 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days ±0.5µg/mL tetracycline. HEK293 cells were transfected with (GA)68 constructs to serve as a positive control for the anti-GA antibody. **B)** HEK293 sham and HEK293 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. Cells were lysed and immunoblotted with anti-GA and anti-tubulin. Molecular weight markers are indicated (kDa).

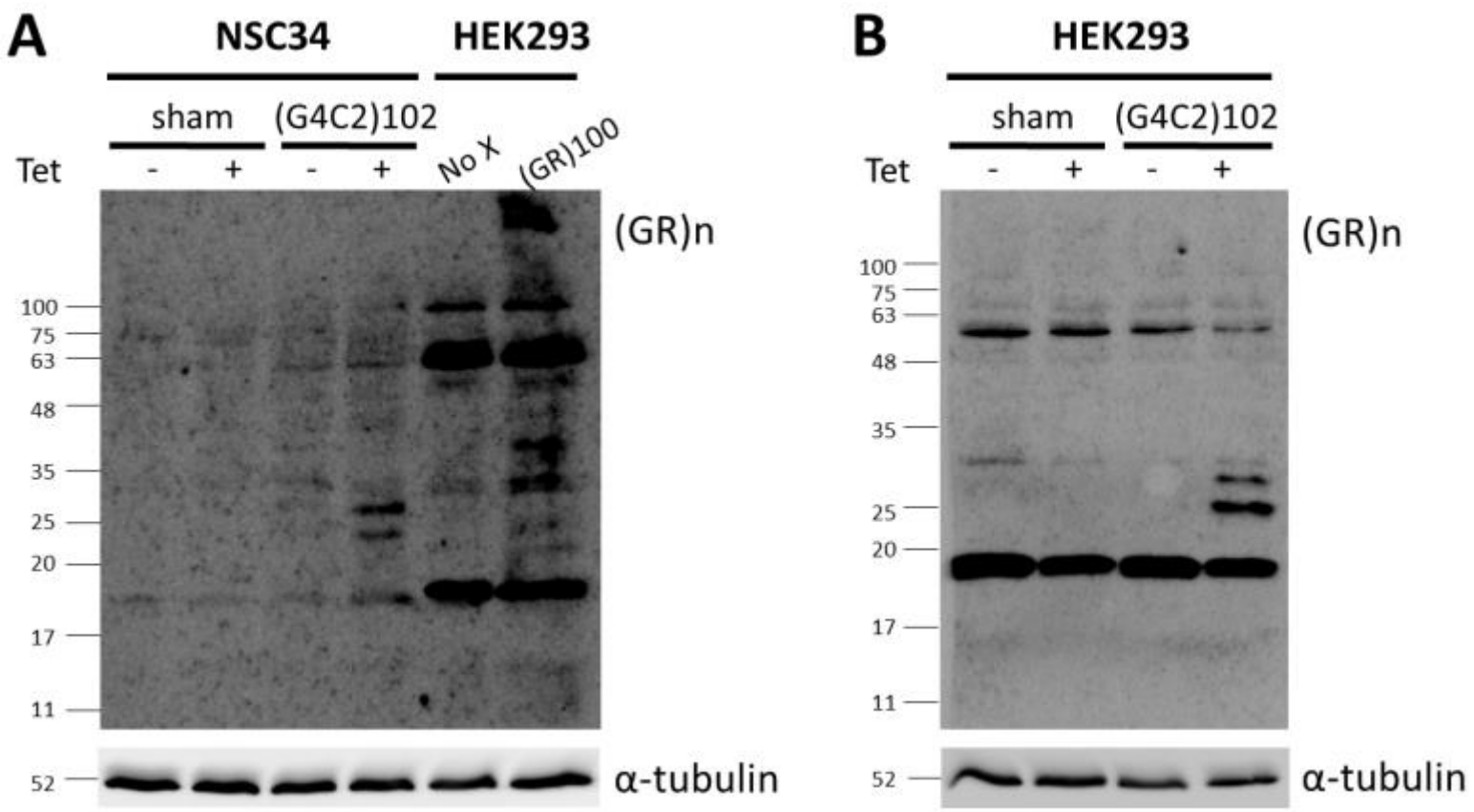


Figure 3.15 The anti-GR antibody detects RAN translation proteins from the (G4C2)102 RNA in the NSC34 (G4C2)102 and HEK293 (G4C2)102 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days ±0.5µg/mL tetracycline. HEK293 cells were transfected with (GR)100 constructs to serve as a positive control for the anti-GR antibody. **B)** HEK293 sham and HEK293 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. Cells were lysed and immunoblotted with anti-GR and anti-tubulin. Molecular weight markers are indicated (kDa).

3.3.4.3.3. Anti-GP Detects Tetracycline-Inducible Protein(s) in NSC34 (G4C2)102 and HEK293 (G4C2)102 Cells

The anti-GP antibody detected a range of bands >65kDa in the NSC34 sham and NSC34 (G4C2)102 cells ± tetracycline (Figure 3.16A). However, a 27kDa band was detected specifically in the NSC34 (G4C2)102 ± tetracycline, which was more abundant in the NSC34 (G4C2)102 with tetracycline (Figure 3.16A). In addition, a 24kDa band was specifically detected in the NSC34 (G4C2)102 + tetracycline (Figure 3.16A). Neither of these bands were detected in the NSC34 sham ± tetracycline (Figure 3.16A).

The anti-GP antibody detected a range of protein bands >75kDa in the HEK293 sham and HEK293 (G4C2)102 ± tetracycline (Figure 3.16B). However, the anti-GP antibody detected a 27kDa band in the HEK293 (G4C2)102 ± tetracycline, with the 27kDa band being more abundant in the HEK293 induced with tetracycline compared to those without (Figure 3.16B). Further, the anti-GP antibody specifically detected a band at 25kDa, and 4 different bands ranging from 15 to 19kDa in the HEK293 (G4C2)102 + tetracycline (Figure 3.16B). None of these bands were detected in the HEK293 sham ± tetracycline (Figure 3.16B).

3.3.4.3.4. Anti-AP Does Not Detect Proteins Specifically in NSC34 (G4C2)102, but Does Detect a Large Protein in HEK293 (G4C2)102 Cells

The anti-AP antibody detected an abundant protein band >100kDa, and 2 smaller protein bands at approximately 75 and 100kDa in the WT HEK293 transfected with (AP)100 (Figure 3.17A). The anti-AP antibody did not detect any other proteins in either the NSC34 sham or NSC34 (G4C2)102 ± tetracycline, or the untransfected WT HEK293 cells (Figure 3.17A).

The anti-AP antibody did detect a protein band >100kDa in the HEK293 (G4C2)102 ± tetracycline (Figure 3.17B). The abundance of the protein also seems increased by tetracycline induction (Figure 3.17B).

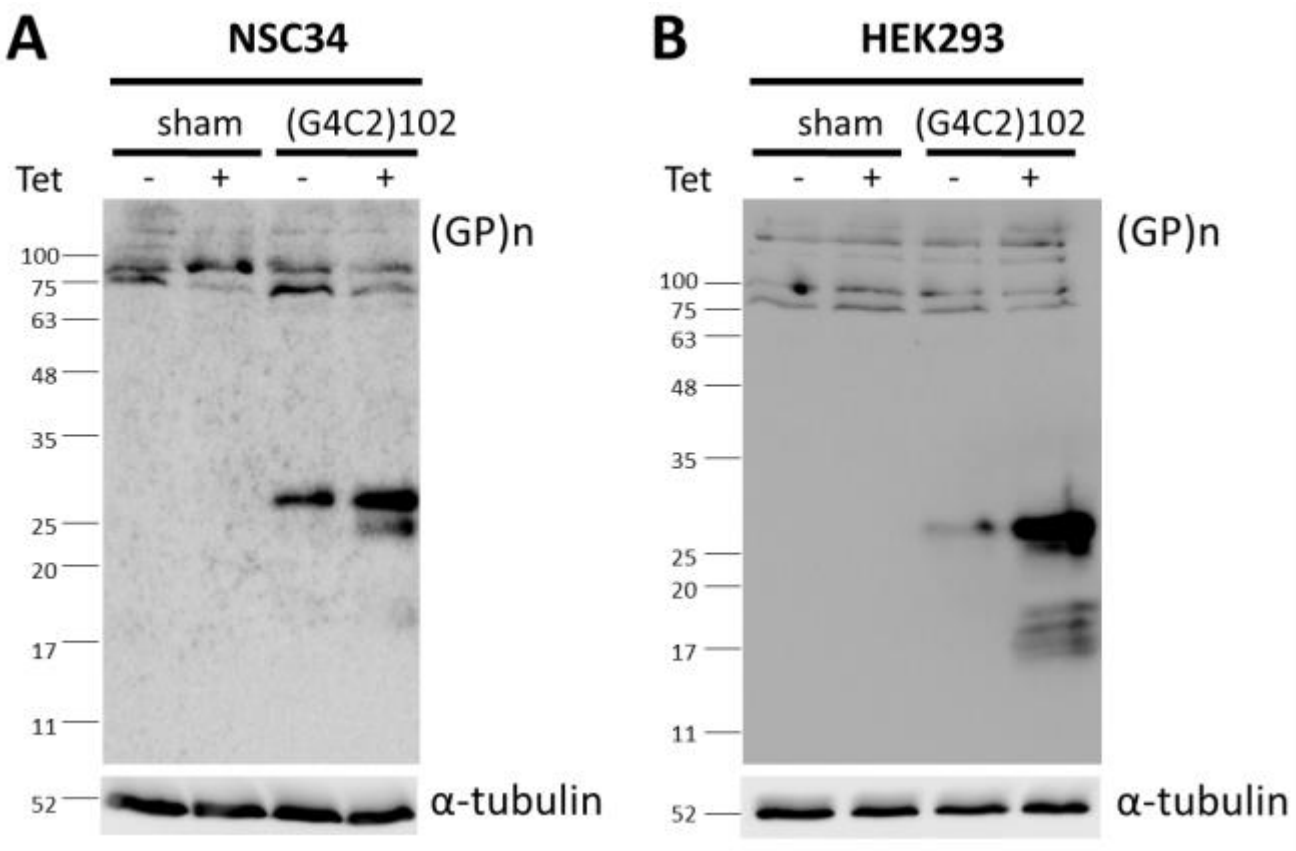


Figure 3.16 The anti-GP antibody detects RAN translation proteins from the (G4C2)102 RNA in the NSC34 (G4C2)102 and HEK293 (G4C2)102 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days ±0.5µg/mL tetracycline. **B)** HEK293 sham and HEK293 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. Cells were lysed and immunoblotted with anti-GP and anti-tubulin. Molecular weight markers are indicated (kDa).

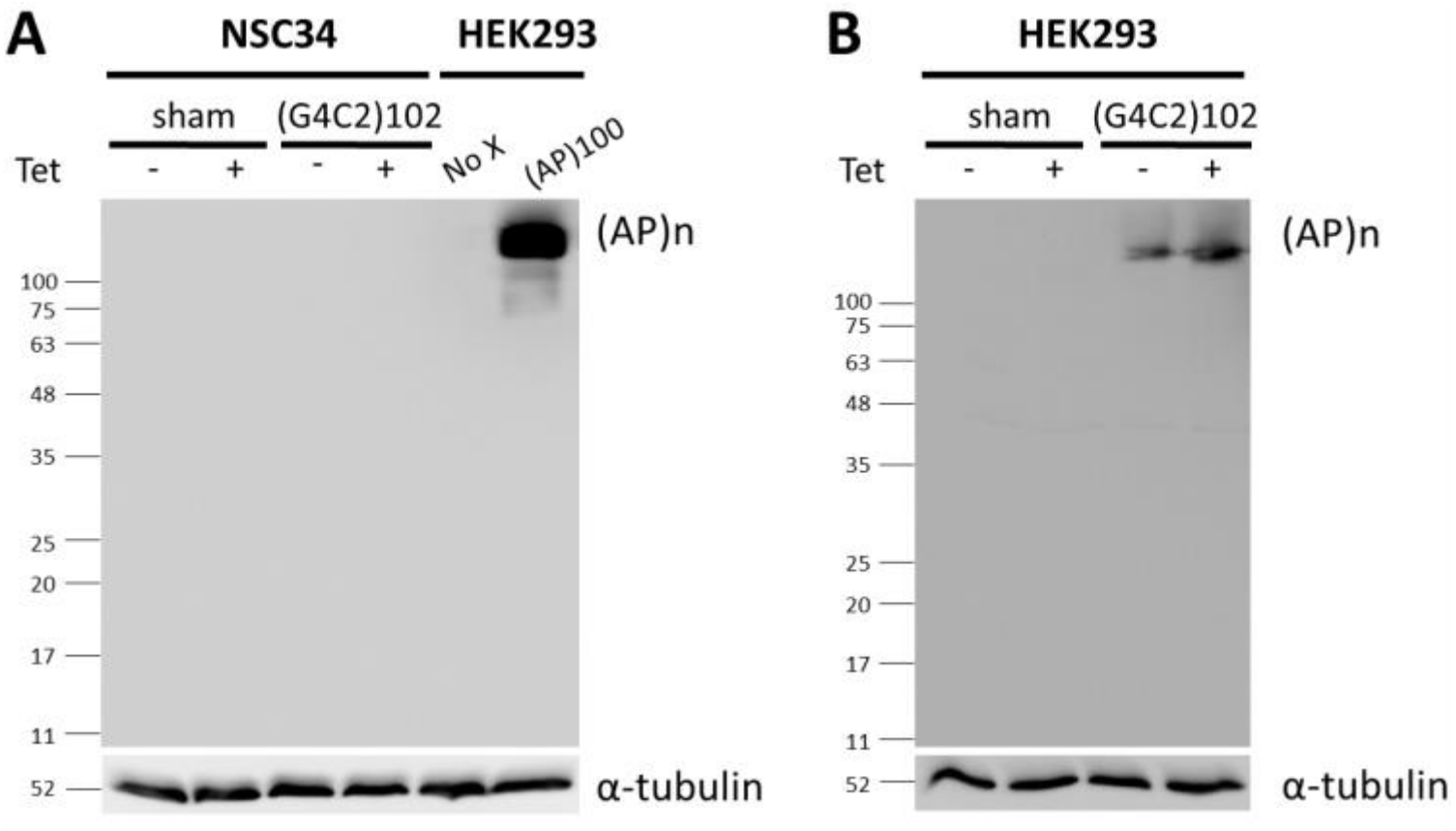


Figure 3.17 The anti-AP antibody does not detects RAN translation proteins the NSC34 (G4C2)102but does detect protein in the HEK293 (G4C2)102 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days ±0.5µg/mL tetracycline. HEK293 cells were transfected with (AP)100 constructs to serve as a positive control for the anti-AP antibody. **B)** HEK293 sham and HEK293 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. Cells were lysed and immunoblotted with anti-AP and anti-tubulin. Molecular weight markers are indicated (kDa).

3.3.4.3.5. Anti-PR Does Not Detect Proteins Specifically in NSC34 (G4C2)102 or HEK293 (G4C2)102 Cells

The anti-PR antibody detected a range of different sized protein bands consistently and at low abundance in both the NSC34 sham and NSC34 (G4C2)102 ± tetracycline, and the untransfected WT HEK293 (Figure 3.18A). The anti-PR antibody also detected a range of protein bands, that were much more abundant and represented by a 'smear' on the membrane, in the WT HEK293 transfected with (PR)100 (Figure 3.18A). The anti-PR antibody did not detect any proteins specifically in the NSC34 (G4C2)102 cells.

The anti-PR antibody detected a protein band >100kDa in the HEK293 sham and HEK293 (G4C2)102 ± tetracycline (Figure 3.18B). In addition, the abundance of the protein did not seem affected by tetracycline induction (Figure 3.18B) The anti-PR antibody did not detect any proteins specifically in the HEK293 (G4C2)102 cells.

3.3.4.3.6. Summary of the Protein Species Detected Using Anti-DPR Antibodies in the NSC34 (G4C2)102 and HEK293 (G4C2)102 Cells

Several protein bands were detected specifically in the NSC34 (G4C2)102 cells (with or without tetracycline induction) that were not detected in the NSC34 sham cells (regardless of tetracycline induction), suggesting that these proteins were derived from RAN translation of the (G4C2)102 repeat RNA. The same was found when comparing the HEK293 (G4C2)102 and HEK293 sham cells, although additional protein bands were detected in the HEK293 (G4C2)102 cells compared to the NSC34 (G4C2)102 cells. The RAN translation protein species, and their molecular weights, that were specifically detected in either the NSC34 (G4C2)102 and/or HEK293 (G4C2)102 cells are summarised in Table 3.1.

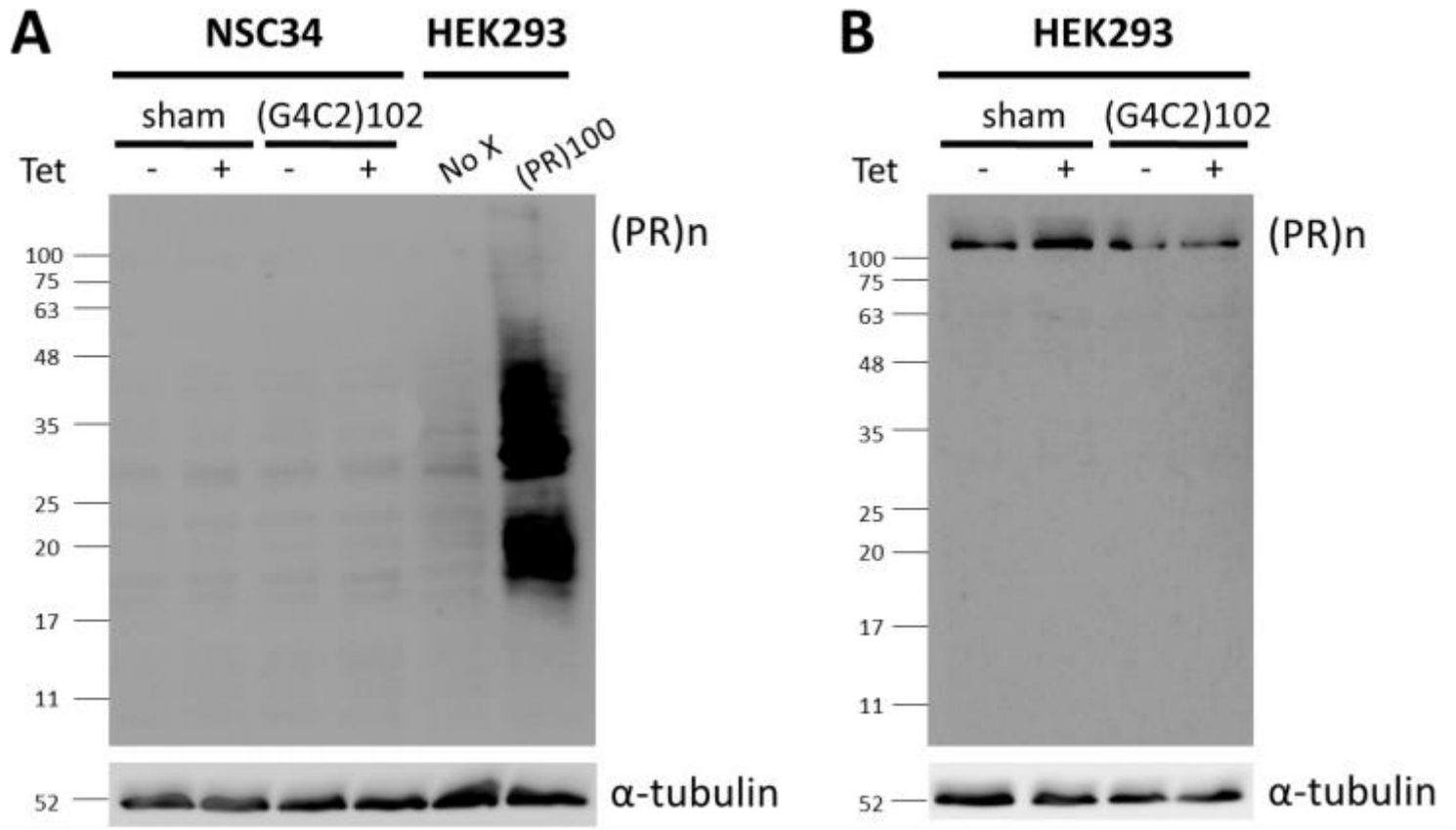


Figure 3.18 The anti-PR antibody does not detect RAN translation proteins in the NSC34 (G4C2)102 or HEK293 (G4C2)102 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days ±0.5µg/mL tetracycline. HEK293 cells were transfected with (PR)100 constructs to serve as a positive control for the anti-PR antibody. **B)** HEK293 sham and HEK293 (G4C2)102 cells were cultured for 3 days ±10µg/mL tetracycline. Cells were lysed and immunoblotted with anti-PR and anti-tubulin. Molecular weight markers are indicated (kDa).

Table 3.1 Summary of the protein species that are specifically detected in cells containing the (G4C2)102 repeat construct. The antibody used, the molecular weight (kDa), and the cell line that the protein bands were detected in are shown. * denotes protein bands that were only detected when the cell line was induced with tetracycline.

Antibody (Sense (S) and/or Antisense (AS) derived)	**Molecular weight (kDa) of RAN translated protein species specifically detected in the following cells**	
	NSC34 (G4C2)102	**HEK293 (G4C2)102**
Anti-GA (S)	24*, 27	Multiple bands ranging from 17-35*, including 24, and 27
Anti-GR (S)	24*, 27*	24*, 27*
Anti-GP (S and AS)	24*, 27	15*, 16*, 17*, 19*, 25*, 27
Anti-AP (AS)	None	>100
Anti-PR (AS)	None	None

3.3.4.4. 10 µg/mL Tetracycline Reduces NSC34 Cell Viability

Initially, an MTT cell viability assay was performed on the NSC34 sham and NSC34 (G4C2)102 cells to assess whether the (G4C2)102 expression reduces NSC34 cell viability. For the MTT assay, the NSC34 cells were grown for 7 days, and induced with 10 µg/mL tetracycline for increasing lengths of time. The viability of the NSC34 sham cells was significantly reduced by 32.38±4.71% ($P<0.0001$), 43.61±7.79% ($P<0.0001$), and 51.83±6.16% ($P<0.0001$) after 5, 6 and 7 days of 10µg/mL tetracycline induction respectively, compared to non-induced NSC34 sham cells (Figure 3.19A). Similarly the viability of the NSC34 (G4C2)102 cells was significantly reduced by 35.78±6.97% ($P<0.0001$), 52.56±3.14% ($P<0.0001$), and 60.31±8.07% ($P<0.0001$) after 5, 6 and 7 days of 10µg/mL tetracycline induction respectively, compared to non-induced NSC34 (G4C2)102 cells (Figure 3.19A). Therefore, it appears 10µg/mL tetracycline is toxic to the NSC34 cells after prolonged exposure, although there is no detectable toxicity prior to 5 days induction.

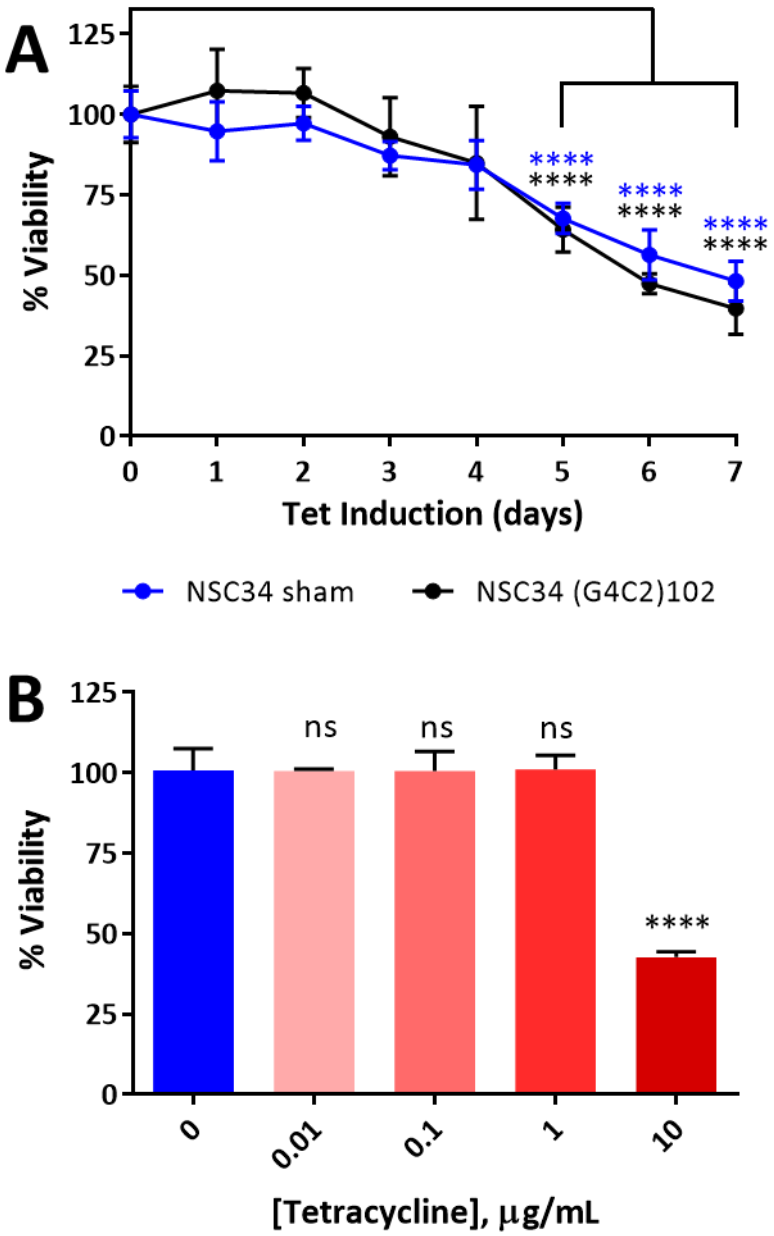


Figure 3.19 10 µg/mL tetracycline reduces NSC34 cell viability. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days, and were induced for various lengths of time with 10µg/mL tetracycline. Cell viability was measured using an MTT assay. (****P<0.0001; Two-way ANOVA with Tukey's multiple comparisons post hoc test; data shown are mean and SD; n=3). **B)** NSC34 sham cells were cultured for 7 days with various concentrations of tetracycline. (****P<0.0001; One-way ANOVA with Dunnett's multiple comparisons post hoc test; data shown are mean and SD; n=3).

The toxicity of a range of lower concentrations of tetracycline was tested to find a non-toxic concentration that would be taken forward for use in future experiments. NSC34 sham cells were grown for 7 days, and incubated with a range of tetracycline concentrations. As before, 10µg/mL tetracycline reduced NSC34 sham cell viability by

57.37±1.67% (P<0.0001) compared to NSC34 sham cells treated with no tetracycline (Figure 3.19B). However, ≤1µg/mL tetracycline caused no significant reduction in NSC34 sham cell viability compared to NSC34 sham cells treated with no tetracycline (Figure 3.19B).

To confirm that the (G4C2)n RNA expression can still be induced using lower tetracycline concentrations, NSC34 (G4C2)102 cells were induced with a range of tetracycline concentrations for 3 days, and then RNA FISH was performed. RNA foci were counted in 25 cells for each condition (Figure 3.20). NSC34 (G4C2)102 cells without tetracycline treatment contained an average of 12.32 sense (G4C2)n RNA foci per cell, whilst NSC34 (G4C2)102 cells treated with 0.1, 0.5, 1, and 10µg/mL tetracycline contained an average of 35.12, 26.96, 28.6, and 23.88 RNA foci per cell respectively. The average number of RNA foci was at least doubled for all tetracycline concentrations, and therefore, it was concluded that ≥0.1µg/mL tetracycline is sufficient to induce (G4C2)n RNA transcription in the NSC34 (G4C2)n cells. From here on, 0.5µg/mL tetracycline was used to induce (G4C2)n expression in the NSC34 cells, as this concentration of tetracycline does not affect NSC34 sham cell viability after 7 days in the MTT assay.

3.3.4.5. (G4C2)102 Expression Reduces NSC34 Cell Viability

Cell viability assays were repeated using a suitable concentration of tetracycline that did not affect NSC34 sham cell viability, but still achieved maximum induction of (G4C2)n RNA expression. For the MTT assay, the NSC34 cells were grown for 7 days, and induced with 0.5µg/mL tetracycline for increasing lengths of time. The viability of the NSC34 (G4C2)102 cells was reduced by 29.9±8.6% (P<0.01) after 7 days tetracycline induction compared to NSC34 sham cells (Figure 3.21). However, there was no significant reduction in NSC34 (G4C2)10 or NSC34 (G4C2)51 cell viability after 7 days tetracycline induction. In addition, tetracycline did not reduce NSC34 sham cell viability (Figure 3.21).

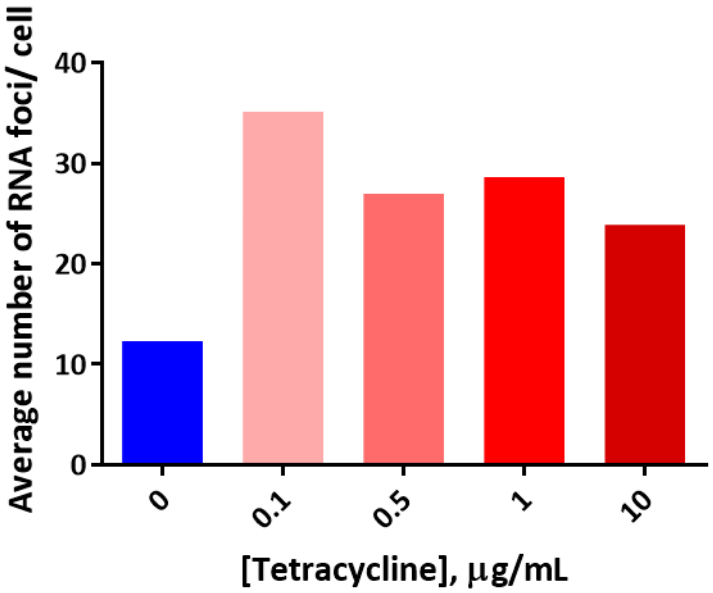


Figure 3.20 ≥0.1µg/mL tetracycline induces increased (G4C2)102 transcription in NSC34 (G4C2)102 cells. NSC34 (G4C2)102 cells were cultured for 3 days with various concentrations of tetracycline. Cells were stained with a fluorescently labelled Locked Nucleic Acid (C4G2)3 sense probe (Red) and Dapi (Blue), and the number of RNA foci was counted in 25 cells per condition. (Data shown are mean; n=1).

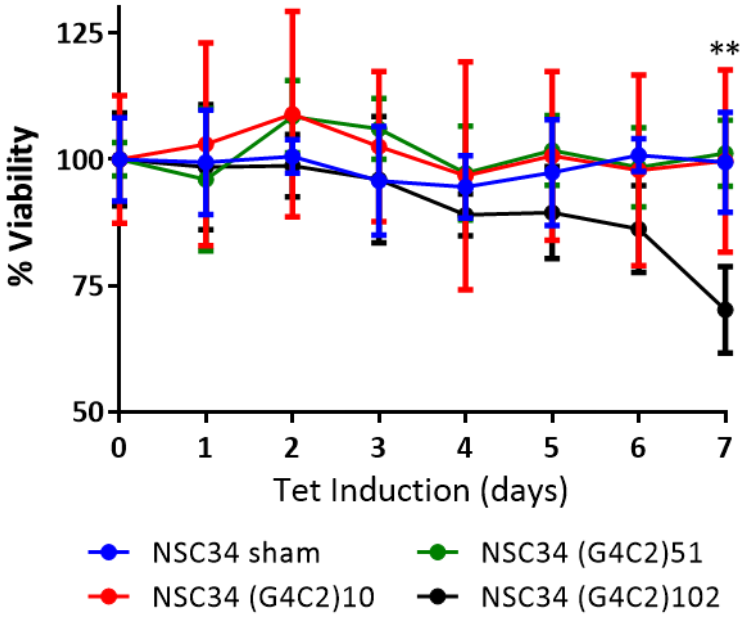


Figure 3.21 (G4C2)102 expression reduces NSC34 cell viability. NSC34 sham, NSC34 (G4C2)10, NSC34 (G4C2)51, and NSC34 (G4C2)102 cells were cultured for 7 days, and were induced for various lengths of time with 0.5µg/mL tetracycline. Cell viability was measured using an MTT assay (**P<0.01; Two-way ANOVA with Tukey's post hoc test; data shown are mean and SD; n=3).

3.3.4.6. (G4C2)102 Expression Does Not Cause NSC34 Cell Death

A cell death assay was then used to confirm whether the reduced NSC34 cell viability was caused by increased cell death. Ethidium homodimer (EthD1) fluorescent dye is a cell-impermeant, high affinity nucleic acid stain that emits red fluorescence when bound to DNA. EthD1 fluorescence is therefore proportional to the number of dead cells. For the EthD1 assays, the NSC34 sham and NSC34 (G4C2)102 cells were grown for 7 days ±0.5µg/mL tetracycline. The amount of dead cells was then assessed using the EthD1 fluorescence. The cells were then freeze-thawed to lyse the cells, and a second EthD1 fluorescence assay was performed. The amount of dead cells was normalised to the total number of cells. There was no significant difference in % dead cells in the NSC34 sham – tet and NSC34 sham + tet cells, with 7.01±6.42% and 10.11±6.27% dead cells respectively (Figure 3.22). Neither was there any significant difference in % dead cells in the NSC34 (G4C2)102 – tet and NSC34 (G4C2)102 + tet cells, with 27.43±11.73% and 28.78±7.65% dead cells respectively (Figure 3.22). Finally, the % dead cells was not significantly increased in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells ± tet.

3.3.4.7. (G4C2)102 Expression Reduces NSC34 Cell Growth Rate

If induction of the (G4C2)102 expression did not increase NSC34 cell death, the other explanation for reduced cell viability is reduced growth rate. Therefore, a growth curve was also performed. The NSC34 sham and NSC34 (G4C2)102 cells were cultured for 16 days in total, ±0.5µg/mL tetracycline. Viable cells were counted every 4 days, and $1.5x10^6$ cells were re-seeded and cultured. There was no significant difference in the number of viable cells between all conditions at days 4 or 8 (Figure 3.23). However, at days 12 and 16 there were significantly fewer viable NSC34 (G4C2)102 cells that were induced with tetracycline, compared to NSC34 sham ± tetracycline and NSC34 (G4C2)102 cells without tetracycline induction. There were only 66.8±26.5% ($P<0.001$) and 52.4±11.6% ($P<0.0001$) NSC34 (G4C2)102 + tetracycline compared to NSC34 sham + tetracycline at days 12 and 16 respectively (Figure 3.23). There were no significant differences in the number of NSC34 sham ± tetracycline and NSC34 (G4C2)102 without tetracycline induction.

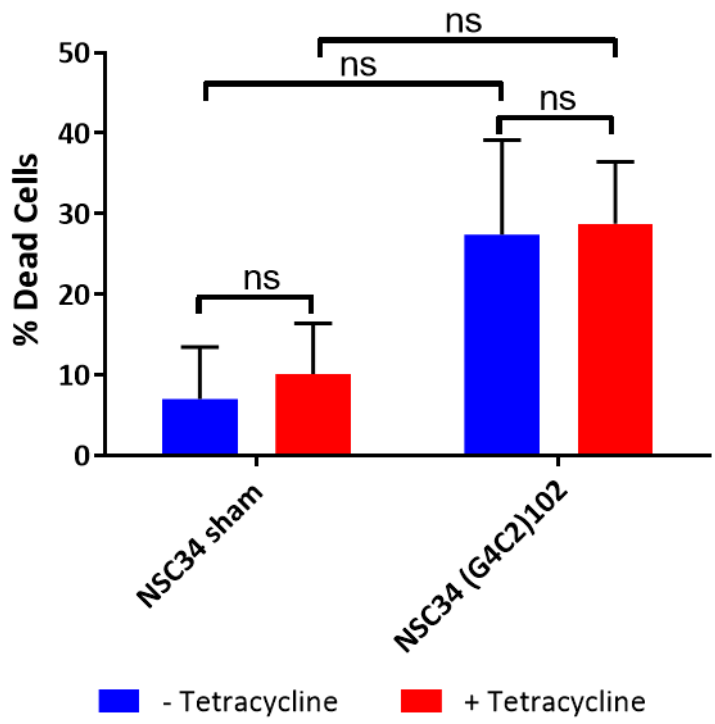


Figure 3.22 (G4C2)102 expression does not increase NSC34 cell death. NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days ±0.5µg/mL tetracycline. The number dead cells was measured using an EthD1 fluorescence assay, and the number of total cells was measured by lysing the cells and repeating the EthD1 fluorescence assay. The number of dead cells was normalised to the number of total cells to calculate % dead cells (Two-way ANOVA with Tukey's post hoc test; data shown are mean and SD; n=3).

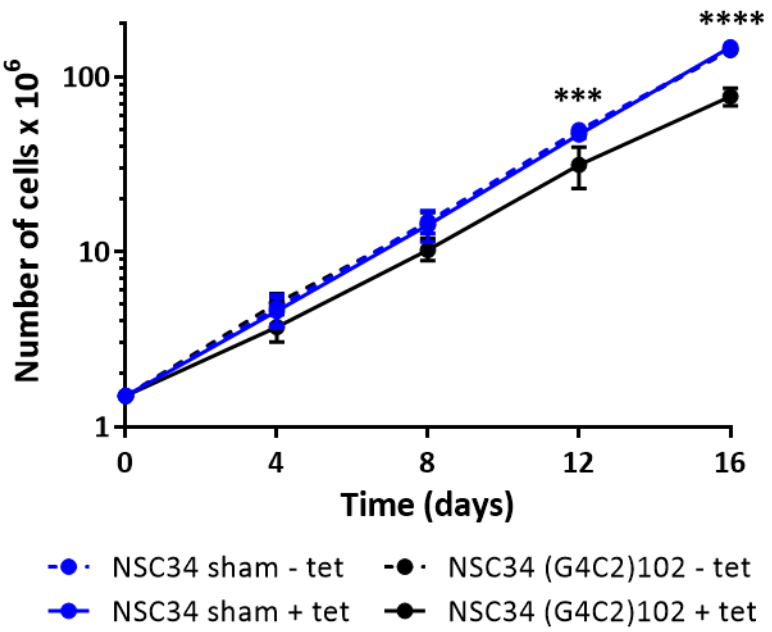


Figure 3.23 (G4C2)102 expression reduces NSC34 cell growth rate. NSC34 sham and NSC34 (G4C2)102 cells were cultured for 16 days ±0.5µg/mL tetracycline. The cells were counted every 4 days, and then 1.5x10^6 cells were reseeded. (***P<0.001; ****P<0.0001; Two-way ANOVA with Tukey's post hoc test; data shown are mean and SD; n=4).

3.4. Discussion

The first aim was to generate stable cell lines with tetracycline-inducible (G4C2)n repeat expression. To do this, the Flp-In™ and T-REx™ systems were used. Importantly, Flp-In™ T-REx™ HEK293 cells were used as a proof of principle, and showed the (G4C2)n constructs were compatible with the Flp-In™ T-REx™ systems. However, the NSC34 (G4C2)n cells were the priority model of C9ORF72-ALS in this project, and were therefore characterised in greater detail. A Flp-In™ T-REx™ NSC34 cell line was generated in house to use as a motor neuron-like cell model, as this was not commercially available. In addition, the (G4C2)n repeat expansion had to be cloned *in vitro*, and subsequently integrated into the Flp-In™ T-REx™ NSC34 cell host line. This required generating the (G4C2)n repeat constructs and then inserting them into a plasmid containing an FRT site. After stably transfecting the (G4C2)n plasmids into the Flp-In™ T-REx™ NSC34 cells, the cells were characterised for (G4C2)n expression, at RNA and protein levels. Finally, the tetracycline-inducible (G4C2)n repeat expression was

confirmed, and the tetracycline concentration had to be optimised to reduce toxicity, but maintain induction.

3.4.1. Cloning the (G4C2)n Repeat

The (G4C2)n hexanucleotide repeat DNA can form abnormal non B-form stable secondary structures. The sense (G4C2)n and antisense (C4G2)n DNA strands are both able to form G-quadruplex secondary structures (Haeusler et al., 2014; Zamiri et al., 2015), whilst the (C4G2)n DNA strand can also form i-motifs and hairpin secondary structures (Kovanda et al., 2015). Additionally, RNA transcribed from the (G4C2)n repeat in either direction can form stable secondary structures with the template (G4C2)n DNA, forming RNA:DNA hybrid R-loops (Haeusler et al., 2014; Reddy et al., 2014). These non B-form secondary structures can interfere with normal cellular processes such as transcription and replication (Reddy et al., 2014; Thys and Wang, 2015). This makes the (G4C2)n repeat DNA unstable in both *E.coli* and mammalian cells, and the repeats are liable to both expansions and contractions in a length-dependent manner (Thys and Wang, 2015). Additionally, the G-quadruplexes cause polymerase slippage during DNA replication, and the (G4C2)n repeats reduce replication efficiency in a length-dependent manner (Thys and Wang, 2015).

The interference with biochemical processes makes the (G4C2)n repeat DNA technically challenging to clone and manipulate. Currently available PCR methods cannot amplify the (G4C2)n repeats when they are above a certain length. In our hands, PCR of the (G4C2)n repeat produced a mixture of different repeat lengths, even after extensive efforts to optimise the PCR using a range of DNA polymerases and the use of 7-deazaguanasine. Further, commercial companies were not able to synthesise these repeats. However, other groups had previously cloned interrupted repeats to model repeat expansions, and this strategy was adopted for this project.

Interrupted (CTG)n repeat expansion constructs had been cloned in *Drosophila* to model the (CTG)n repeat expansion that causes Myotonic Dystrophy 1 (de Haro et al., 2006). This methodological approach was used here to clone large interrupted (G4C2)n repeat constructs by ligating together smaller (G4C2)10 repeat fragments (see section 3.3.2), and (G4C2)10, (G4C2)51 and (G4C2)102 repeats were successfully generated in the

pcDNA6.2 vector. These (G4C2)n repeat constructs were also challenging to sub-clone via restriction digest and ligation, but after several rounds of optimisation, the repeats were sub-cloned into the pcDNA5/FRT/TO-HIS vectors. The size and sequence of the (G4C2)n repeats in the pcDNA5/FRT/TO-HIS vectors were confirmed by electrophoresis and sequencing respectively. Interestingly, the (G4C2)51 and (G4C2)102 repeat constructs do not contain (G4C2)10 repeat tracts followed by the TCGAG interruptions, as predicted. Instead, the larger (G4C2)n repeat constructs actually contain (G4C2)n repeat tracts of various lengths interspersed with the expected interruptions (Appendices 2 and 3). Further, the (G4C2)n repeats are always conserved as a full (G4C2) unit. This suggests that the (G4C2)n repeat constructs generated underwent expansion, contraction and/or recombination, even though recombination deficient β-10 *E.coli* were used to clone the (G4C2)n repeat constructs.

3.4.2. Generating the Flp-In™ T-REx™ NSC34 Host Cell Line

To model how (G4C2)n repeat expression may cause ALS in a reductionist manner, the initial aim was to generate isogenic NSC34 cells with tetracycline-inducible (G4C2)n expression. The NSC34 cell line was used because it displays many properties of motor neurons (Cashman et al., 1992), and the aim was to study how (G4C2)n expression may reveal disease mechanisms in ALS. Ideally, the NSC34 cells would have no or, more realistically, low (G4C2)n expression at basal levels because it was unknown how toxic the (G4C2)n repeat constructs would be in the NSC34 cell context. Equally, the NSC34 cells would have high (G4C2)n expression upon tetracycline induction, such that the early biochemical effects of the (G4C2)n expression could be studied. In addition, the inducible nature of the (G4C2)n expression would potentially minimise the cell lines adapting to the (G4C2)n expression.

Also, the isogenic nature of the NSC34 (G4C2)n cells was desirable as there would be minimal noise between cell lines in experiments. Flp-In™ T-REx™ NSC34 clonal line B10-2 was generated and contained only one genomic FRT site, had low basal expression levels from the FRT site, and good tetracycline-inducible expression from the FRT site. After stable (G4C2)n construct insertion, the NSC34 (G4C2)n cells had the desirable characteristics described above. The ability to switch on (G4C2)n expression and the

isogenic nature of the cells were two advantages that the NSC34 (G4C2)n cells had over other models such as transiently (G4C2)n transfected cells or C9ORF72-ALS patient iPSC-derived motor neurons.

3.4.3. The HEK293 (G4C2)n and NSC34 (G4C2)n Cells Have Tetracycline-Inducible (G4C2)n RNA Expression

The (G4C2)n repeat is transcribed in both sense and antisense directions, and forms characteristic sense (G4C2)n and antisense (C4G2)n RNA foci in the CNS of C9ORF72-ALS/FTD patients (DeJesus-Hernandez et al., 2011; Gendron et al., 2013; Renton et al., 2011). Similarly to the (G4C2)n DNA, the sense (G4C2)n RNA forms G-quadruplex secondary structures (Fratta et al., 2012; Haeusler et al., 2014; Reddy et al., 2013a). The antisense (C4G2)n RNA is suggested to form i-motif secondary structures however, although the actual secondary structure remains unresolved (Kovanda et al., 2015). The characteristic RNA foci are probably formed by the RNA secondary structure, and/or the binding with RNA binding proteins.

RNA FISH was used to characterise the HEK293 (G4C2)n and NSC34 (G4C2)n cells for tetracycline-inducible (G4C2)n expression at the RNA level. As described previously, the (G4C2)n repeat cannot be amplified by PCR (section 3.4.1). Equally, there were no suitable regions flanking the (G4C2)n for qRT-PCR. RNA FISH, followed by confocal microscopy imaging and RNA foci counting was the only reliable method for (G4C2)n and (C4G2)n RNA detection in the NSC34 (G4C2)n and HEK293 (G4C2)n cells. Although this method is not strictly quantitative of total (G4C2)n or (C4G2)n levels, it did provide a relative measure of (G4C2)n and (C4G2)n expression in the cells. Also, the LNA probes bound specifically, but they did produce a diffuse background stain. Staining was classified as an RNA focus when the signal was strong, spherical in shape, and appeared in consecutive planes in the confocal z-stack images. RNase A treatment ablated these intense, spherical foci structures in both the NSC34 (G4C2)102 and HEK293 (G4C2)102 cells, and confirmed that they were RNA foci.

Using RNA FISH, the number of sense (G4C2)n RNA foci was dependent on the (G4C2)n repeat size. Intriguingly however, the number of RNA foci does not follow a linear trend with the (G4C2)n repeat size, as NSC34 (G4C2)51 cells induced with tetracycline contain

only 10.8±7.40% of the RNA foci number found in the NSC34 (G4C2)102 cells induced with tetracycline. Equally, the NSC34 (G4C2)10 cells induced with tetracycline contain 46.6±59.2% of the RNA foci number found in the NSC34 (G4C2)51 cells induced with tetracycline. This is unexpected because the (G4C2)n repeats are all expressed from the same CMV/TO promoter from the same isogenic genomic location in the NSC34 (G4C2)n cells, suggesting there are the same amount of (G4C2)10, (G4C2)51, and (G4C2)102 RNA molecules transcribed. In absolute (G4C2)n repeat terms, the NSC34 (G4C2)102 should therefore express double the number of (G4C2)n RNA repeats as the NSC34 (G4C2)51, and was predicted to form double the number of RNA foci.

Only two previous studies have looked at the correlation between RNA foci and (G4C2)n repeat length. SHSY5Y cells transfected with (G4C2)8 did not express any RNA foci, whilst cells transfected with (G4C2)38, and (G4C2)72 plasmids expressed 6 and 12 RNA foci per cell (Lee et al., 2013). This suggests a minimum number of repeats is necessary for RNA foci formation and/or detection using RNA FISH. In our hands, the cut off could be between 51 and 102 repeats, whilst it may be between 8 and 38 in this previous study. Above this number of (G4C2)n repeats there is a roughly linear correlation between RNA foci and (G4C2)n repeat length in cell lines however. Conversely, the number of RNA foci did not correlate with (G4C2)n repeat expansion length in C9ORF72-ALS iPSC-derived neurons (Almeida et al., 2013). Other genetic factors that arise from the great genetic diversity in these non-isogenic cells could affect the number of RNA foci however.

Importantly, the size of the (G4C2)n RNA foci varies greatly even in the same cell line, and this is clearly demonstrated in the NSC34 (G4C2)102 (Figure 3.10 inset). Only RNA foci above a certain size must be detectable using RNA FISH and confocal microscopy, which means only the largest RNA foci formed in the NSC34 (G4C2)10 and NSC34 (G4C2)51 cells are actually detectable, whilst the majority of RNA may exist as undetectable small RNA foci or soluble RNA molecules. Also, the (G4C2)102 RNA may aggregate more frequently, resulting in more detectable RNA foci. This could explain the lower RNA foci counts in the NSC34 (G4C2)10 and NSC34 (G4C2)51 cells. In addition, this suggests that soluble (G4C2)n RNA is not toxic, because (G4C2)10 and (G4C2)51

expression does not result in RNA foci and did not affect NSC34 cell viability (section 3.3.4.5). In contrast, RNA foci and toxcitiy are observed in the the NSC34 (G4C2)102 cells.

The NSC34 (G4C2)n cells do have tetracycline-inducible (G4C2)n RNA expression however. Staining was detected very rarely in the NSC34 sham, but the few foci that were detected are most likely non-specific staining. Also, the NSC34 sham cells have the same number of RNA foci with and without tetracycline, indicating that tetracycline treatment in itself is not sufficient to produce RNA foci. There are more RNA foci in the NSC34 (G4C2)10, NSC34 (G4C2)51, and NSC34 (G4C2)102 cells when treated with tetracycline compared to the respective non-induced controls, however the increases in RNA foci were only significant in the NSC34 (G4C2)102 cells.

The HEK293 (G4C2)102 cells also have tetracycline inducible (G4C2)n RNA expression. RNA foci were detected at very low frequency in the HEK293 sham cells, and similarly to the NSC34 sham cells, this staining is most likely to be non-specific. RNA foci were abundant in the HEK293 (G4C2)102 cells, and more were detected in the tetracycline induced cells than the non-induced cells.

3.4.4. The NSC34 (G4C2)n Cells Do Not Express Antisense (C4G2)n RNA Foci

In C9ORF72-ALS/FTD patients, the (G4C2)n repeat is transcribed in the antisense direction from a cryptic promoter, and forms characteristic antisense (C4G2)n RNA foci (Gendron et al., 2013). There were no antisense (C4G2)n RNA foci detected in the NSC34 (G4C2)102 cells. This is expected as the (G4C2)n repeats were only engineered under control of a sense CMV/TO promoter in the NSC34 (G4C2)n cells, and the antisense cryptic promoter found in the C9ORF72 gene context is absent. The antisense (C4G2)n RNA FISH probe did detect antisense (C4G2) RNA foci in HEK293 cells transiently transfected with (C4G2)102 constructs however, proving that the probe and assay do work. This means the NSC34 (G4C2)n cells specifically model sense (G4C2)n repeat expression independent of the C9ORF72 gene context.

3.4.5. The (G4C2)n Constructs Undergo RAN Translation in the HEK293 (G4C2)102 and NSC34 (G4C2)102 Cells

The (G4C2)n undergoes non-canonical RAN translation in cells throughout the CNS of C9ORF72-ALS/FTD patients, producing DPRs (Ash et al. 2013; Gendron et al. 2013; Mori et al. 2013a; Mori et al. 2013b). There is also strong evidence that the (G4C2)n repeat is RAN translated in the NSC34 (G4C2)102 cells, and interestingly, this shows that the (G4C2)102 RNA can form the secondary structure necessary to drive RAN translation despite containing interruptions. Firstly, DPR expression plasmids were transfected into HEK293 cells and immunoblotted using DPR antibodies. This showed the antibodies recognise the relevant DPR proteins. The DPR antibodies that detect DPR translated from the sense (G4C2)n RNA (anti-GA, anti-GR, and anti-GP) detect proteins at 24 and 27kDa specifically in the NSC34 (G4C2)102 cells treated with tetracycline, but not in the NSC34 sham cells. Secondly, tetracycline induction increases the amount of these proteins. Taken together with the RNA FISH result, which showed that NSC34 (G4C2)102 cells have tetracycline-inducible (G4C2)n RNA expression (measured by number of RNA foci), it confirms that the expression of the (G4C2)102 translated proteins is dependent on the expression of the (G4C2)102 RNA. Thirdly, there are no detectable proteins containing the antisense-specific DPR (AP and PR) that are specifically expressed in the NSC34 (G4C2)102 cells, and not the NSC34 sham cells. This supports the lack of antisense (C4G2)n RNA foci, and by extension, the lack of (G4C2)n transcription in the antisense direction. Lastly, when the (G4C2)n repeats were cloned into the pcDNA5/FRT/TO-HIS plasmids, the HIS tag and associated ATG codon were removed. This was confirmed by restriction digest of the plasmid and sequencing. Therefore, translation of the (G4C2)102 repeats is via RAN translation in the absence of an ATG start codon.

Several of the anti-DPR antibodies also detect non-DPR proteins, which is relatively unsurprising since the simple dipeptide motif that the antibodies recognise is likely to be contained in other proteins as well. Importantly, initial DPR immunoblot characterisation in the NSC34 (G4C2)n cells was performed solely using the anti-GA antibody, which failed to detect specific proteins in the NSC34 (G4C2)102 cells. This was because the anti-GA recognises a different protein at 40kDa that is much more abundant

than the (G4C2)102 translation products (Figure 3.14A). This made (G4C2)102 translation product detection difficult. Only after interrogating the immunoblots using an increased exposure time were the (G4C2)102 translation products detected.

The proteins translated from the interrupted (G4C2)102 repeats contain each of the three sense DPR motifs (Figure 3.13). This is because during translation of the interrupted (G4C2)n repeats, the 5 bp TCGAC interruptions cause a 'frame shift' to the next DPR motif. Therefore a different DPR motif is translated from each (G4C2)n repeat tract, but these different DPR motifs all exist in the same polypeptide. This means the protein bands at 24 and 27kDa detected by the anti-GA, anti-GR, and anti-GP are probably the same polypeptides, although this would require mass spectrometry to prove unequivocally. Unexpectedly however, the maximum predicted molecular weight for (G4C2)102 translation products is approximately 20.5kDa. The extra mass could be explained by post-translational modifications such as ubiquitination.

The evidence for RAN translation in the NSC34 (G4C2)102 cells is also true in the HEK293 (G4C2)102 cells. This shows the (G4C2)102 repeat construct is the necessary variable needed for RAN translation, in either the NSC34 or HEK293 cells. However, there are more detectable bands in the HEK293 (G4C2)102 cells treated with tetracycline using the anti-GA and anti-GP antibodies. This is possibly due to higher expression levels of the (G4C2)102 RAN proteins in the HEK293 (G4C2)102 cells compared to the NSC34 (G4C2)102, because the RAN protein signal is greater relative to the non-specific bands on the blot. This is most apparent in the anti-GA immunoblot (Figure 3.14B). This also means that the (G4C2)102 may produce multiple different molecular weight RAN proteins in the NSC34 (G4C2)102 cells, but the signal is too weak to be detected by the imaging system. Also, unexpectedly, there are high molecular weight bands detected by the anti-AP antibody specifically in the HEK293 (G4C2)102 cells, but not the HEK293 sham cells. This suggests that a promoter does control transcription of the (G4C2)102 repeat in the antisense direction, producing a polypeptide with the poly(AP) motif. Similarly to the (AP)100 used as the positive control, this polypeptide appears to have been trapped in the loading well at the top of the gel. Due to the huge difference in

molecular weights however, the AP containing polypeptide is different to the sense RAN peptides containing GA, GR and GP motifs.

3.4.6. (G4C2)102 Expression Reduces NSC34 Cell Growth Rate

The (G4C2)102 reduces NSC34 cell growth rate, but does not cause cell death. During preliminary cell viability assays it was discovered that 10µg/mL tetracycline is actually toxic to NSC34 cells. Optimisation suggested 0.5µg/mL tetracycline was a better dosage, as it did not reduce NSC34 sham cell viability, but did still induce (G4C2)n RNA foci expression in NSC34 (G4C2)102 cells. The MTT cell viability assay showed that expression of the (G4C2)10 and (G4C2)51 do not affect NSC34 cell viability. However, expression of the (G4C2)102 significantly reduced NSC34 cell viability after 7 days tetracycline induction, and appears to be above a threshold repeat length necessary to cause toxicity in the NSC34 cells. The EthD1 cell death assays showed that the reduced NSC34 cell viability was not due to an increase in NSC34 cell death, because there was no significant difference in % dead cells between NSC34 (G4C2)102 induced with tetracycline for 7 days, and non-induced NSC34 (G4C2)102. However, although not statistically significant, there was a greater % of dead cells in the NSC34 (G4C2)102 compared to the NSC34 sham. This could suggest that long term basal (G4C2)102 expression does contribute to a greater rate of cell death. The best explanation for the reduced NSC34 cell viability measured in the MTT assay is a reduction in growth rate however. In the growth curve, NSC34 (G4C2)102 cells induced with tetracycline grew slower than NSC34 sham ± tetracycline, and NSC34 (G4C2)102 cells that were not induced with tetracycline. This suggests expression of the (G4C2)102 dysregulates the NSC34 cell metabolism at some level, and causes a reduction in growth rate.

The toxic effects measured in the NSC34 (G4C2)102 cell model are similar to those described in previous (G4C2)n cell model studies (section 1.8.2.1), although the effects are more subtle in the NSC34 (G4C2)102 cells. Transient transfection of plasmids expressing (G4C2)30, but not (G4C2)3, reduce Neuro2a cell viability (Xu et al., 2013). Also, transient transfection of plasmids expressing (G4C2)38 and (G4C2)72, but not (G4C2)8, cause apoptosis in SHSY5Y neuronal cells (Lee et al., 2013). These plasmids do not induce apoptosis in HEK293 cells however (Lee et al., 2013). Further, several studies

using iPSC-derived neuronal cells and motor neuronal cells, derived from C9ORF72-ALS fibroblasts do not report reduced cell viability or increased cell death compared to control derived cells (Almeida et al., 2013; Devlin et al., 2015; Donnelly et al., 2013; Sareen et al., 2013). The expression level of C9ORF72 and the associated (G4C2)n repeats is likely to be much lower in iPSC derived neuronal cells than in the transient transfection models and the NSC34 (G4C2)n cells described here, which utilise strong promoters such as the CMV promoter. Therefore, the level of toxicity is probably a combination of the (G4C2)n repeat length, expression level, cell type, and ability of the (G4C2)n RNA to form secondary structure and/or DPR.

Crucially, ALS disease onset is age related, and therefore C9ORF72-ALS pathogenesis requires a 'second hit' (such as reduced mitochondrial efficiency or neuroinflammation) in combination with the (G4C2)n repeat derived toxicity. Therefore, the (G4C2)n repeat expansion is likely to produce low, or at least well tolerated, toxicity in the motor neurons. With this hypothesis in mind, the subtler toxic phenotype described in the NSC34 (G4C2)102 cells likely represents a more relevant model of C9ORF72-ALS than those cellular models that display more severe toxicity.

3.4.7. Summary

In summary, stable, isogenic, motor neuron-like NSC34 cell lines with tetracycline-inducible (G4C2)n expression have been successfully generated. The (G4C2)n is expressed in the NSC34 (G4C2)n cells in a tetracycline-inducible manner, and forms characteristic RNA foci. A unique advantage of the tetracycline-inducible (G4C2)n expression over other C9ORF72-ALS models, is that early biochemical effects can be studied. Additionally, the (G4C2)n RNA undergoes RAN translation to produce polypeptides containing all three sense DPR motifs (GA, GR, and GP). The NSC34 (G4C2)n cells express the (G4C2)n in the sense orientation only. Finally, tetracycline-induction of the (G4C2)102 RNA and/or RAN proteins reduces NSC34 cell growth rate.

Chapter 4

Chapter 4. Biochemical Analysis of the NSC34 (G4C2)n Cell Lines

4.1. Introduction

Initial characterisation of the NSC34 (G4C2)n cell lines showed that they have tetracycline-inducible expression of the (G4C2)n repeat, the (G4C2)n RNA forms RNA foci and is RAN translated, the (G4C2)n repeat is expressed in the (G4C2)n sense orientation only, and prolonged (G4C2)n expression reduced NSC34 (G4C2)102 growth rate. Next, biochemical analysis was performed to establish whether the NSC34 (G4C2)n cells recapitulate key pathological hallmarks of ALS in general, and specifically of C9ORF72-ALS. Importantly, NSC34 viability was only affected by (G4C2)102 expression, and not by (G4C2)51 or (G4C2)10 expression in the NSC34 (G4C2)n cell lines, and therefore, the early biochemical effects of (G4C2)n repeat expression would be most pronounced in the NSC34 (G4C2)102 cells. For that reason, and to conserve resources, only NSC34 sham and NSC34 (G4C2)102 cells were used in the biochemical characterisation experiments.

4.2. Aims and Objectives for Biochemical Analysis of NSC34 (G4C2)n Cells

1) Characterise NSC34 (G4C2)102 cells for TDP-43 aggregation and/or mis-localisation, and oxidative stress.
2) Characterise NSC34 (G4C2)102 cells for RNA foci and RNA binding protein co-localisation.
3) Interrogate potential biochemical effects of RNA binding protein sequestration using functional assays.

4.3. Results

4.3.1. NSC34 (G4C2)102 Cells Do Not Display TDP-43 Mislocalisation or Aggregation

TDP-43 is predominantly a nuclear protein that shuttles between the nucleus and cytoplasm of healthy cells. However, TDP-43 becomes mislocalised in the cytoplasm and forms aggregates in the neuronal and glial cells of most ALS patients (excluding SOD1 and FUS related ALS cases, but including C9ORF72-ALS cases) (Neumann et al., 2006). These cytoplasmic TDP-43 aggregates are a pathological hallmark of ALS. ICC was performed on the NSC34 (G4C2)102 and NSC34 sham ± tetracycline to assess whether (G4C2)102 expression in this model system would cause TDP-43 aggregation or mislocalisation. In both the NSC34 sham and NSC34 (G4C2)102 cells ± tetracycline, TDP-43 is predominantly nuclear with some cytoplasmic staining (Figure 4.1). There is no difference in nuclear and cytoplasmic TDP-43 distribution between the NSC34 sham and NSC34 (G4C2)102 cells (Figure 4.1). Also, TDP-43 is mainly diffuse in both the nucleus and cytoplasm, but there are also distinct small puncta structures in the nucleus and cytoplasm (Figure 4.1). There are no measurable changes in TDP-43 staining in the NSC34 (G4C2)102 compared to the NSC34 sham (Figure 4.1). In addition, tetracycline induction did not affect TDP-43 nuclear cytoplasmic distribution or staining in either the NSC34 sham or NSC34 (G4C2)102 cells (Figure 4.1). The TDP-43 ICC stained NSC34 sham and (G4C2)102 cells shown here were induced with tetracycline for 9 days, but showed the same results as NSC34 cells that were induced with tetracycline for shorter periods of time. The rationale was to induce (G4C2)102 for longer than 6 days (after which toxicity is observed in the NSC34 (G4C2)102 cells). If the (G4C2)102 expression did have an effect of TDP-43 localisation, it was predicted to have occurred by 9 days tetracycline induction.

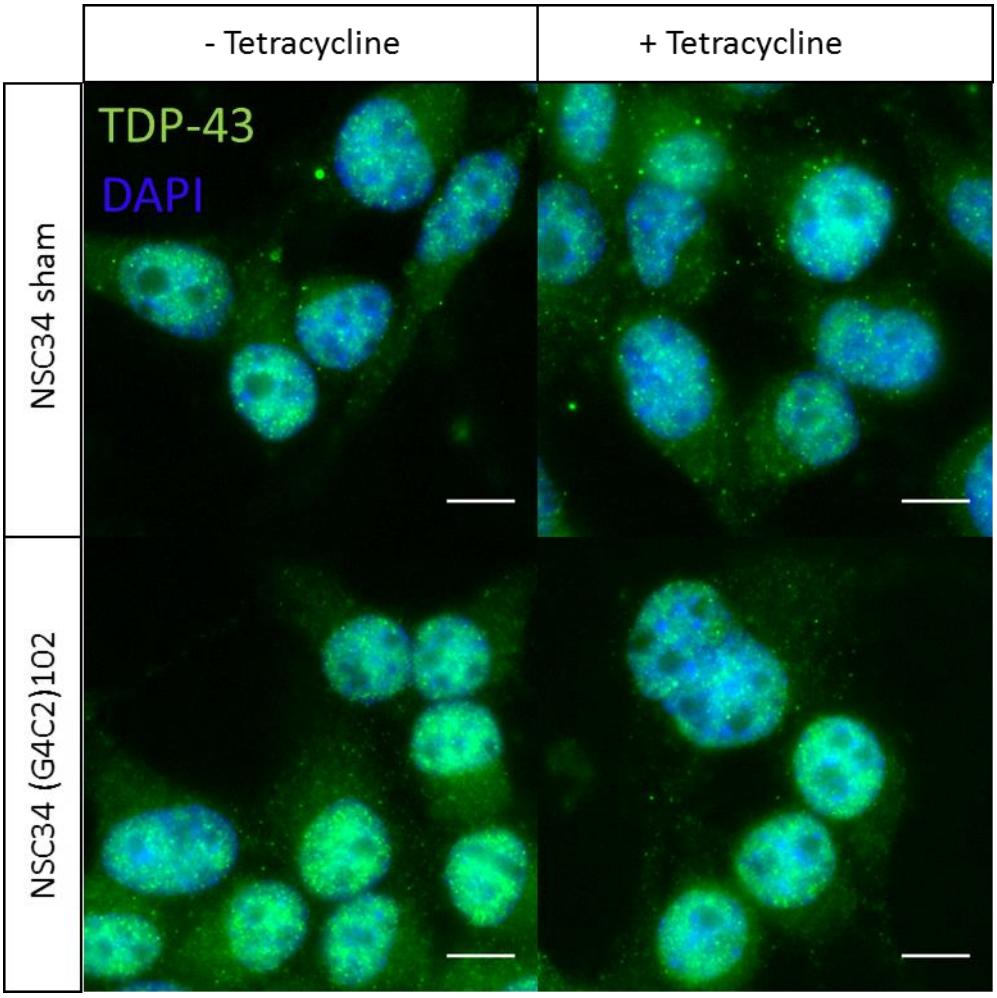


Figure 4.1 (G4C2)102 expression does not cause TDP-43 mislocalisation or aggregation in NSC34 cells. NSC34 sham and (G4C2)102 cells were cultured for 9 days ±0.5µg/mL tetracycline. Cells were stained for TDP-43 (Green) and Dapi (Blue). Scale bar = 10µm.

4.3.2. NSC34 (G4C2)102 Cells Do Not Have Increased Cellular Levels of Hydroxyl, Peroxyl, or Other Reactive Oxygen Species

Oxidative stress occurs when the production of reactive oxygen species (ROS) and their removal becomes imbalanced, and/or the ability of the biological system to repair oxidative damage caused by ROS becomes impaired. ROS disrupt redox sensitive cellular signalling and also damage DNA, RNA, proteins and lipids. Post-mortem tissue from ALS patients is widely reported to show increased levels of oxidative damage (Chang et al., 2008; Ferrante et al., 1997; Fitzmaurice et al., 1996; Shaw et al., 1995; Shibata et al., 2001).

Hydroxyl, peroxyl and other ROS levels were assessed in the NSC34 sham and NSC34 (G4C2)102 cells using the DCF assay. DCFDA is a cell permeant, fluorogenic dye. Once diffused through the plasma membrane, DCFDA is deacetylated by cellular esterases to produce a non-fluorescent, lipid membrane impermeant compound. Importantly, the removal of the acetyl group traps the compound inside the cell but also allows the DCF compound to be activated by hydroxyl, peroxyl and other cellular ROS, to produce fluorescent DCF. Therefore, DCF fluorescence directly measures the levels of these particular ROS within the cells. DCF data was normalised to total cell number, which was measured using EthD1 after freeze-thawing the cells. Firstly, as a positive control for the DCF assay, NSC34 cells were treated with menadione, which induces cellular ROS generation. The DCF fluorescence signal was significantly increased in the NSC34 sham and NSC34 (G4C2)102 cells treated with menadione, compared to the respective untreated control cells (Figure 4.2A). There was no significant difference in DCF signal between the untreated NSC34 sham and NSC34 (G4C2)102 cells either (Figure 4.2A).

The significant increase in DCF signal caused by the menadione showed the assay could detect differences in cellular ROS levels. The NSC34 sham and NSC34 (G4C2)102 cells were then cultured for 5 days, and induced with tetracycline for various lengths of time. The hypothesis being, that ROS levels would increase with time if the (G4C2)102 expression induced oxidative stress. There was no significant difference in DCF-detectable ROS levels between NSC34 sham and NSC34 (G4C2)102 cells after any of the tetracycline inductions (Figure 4.2B).

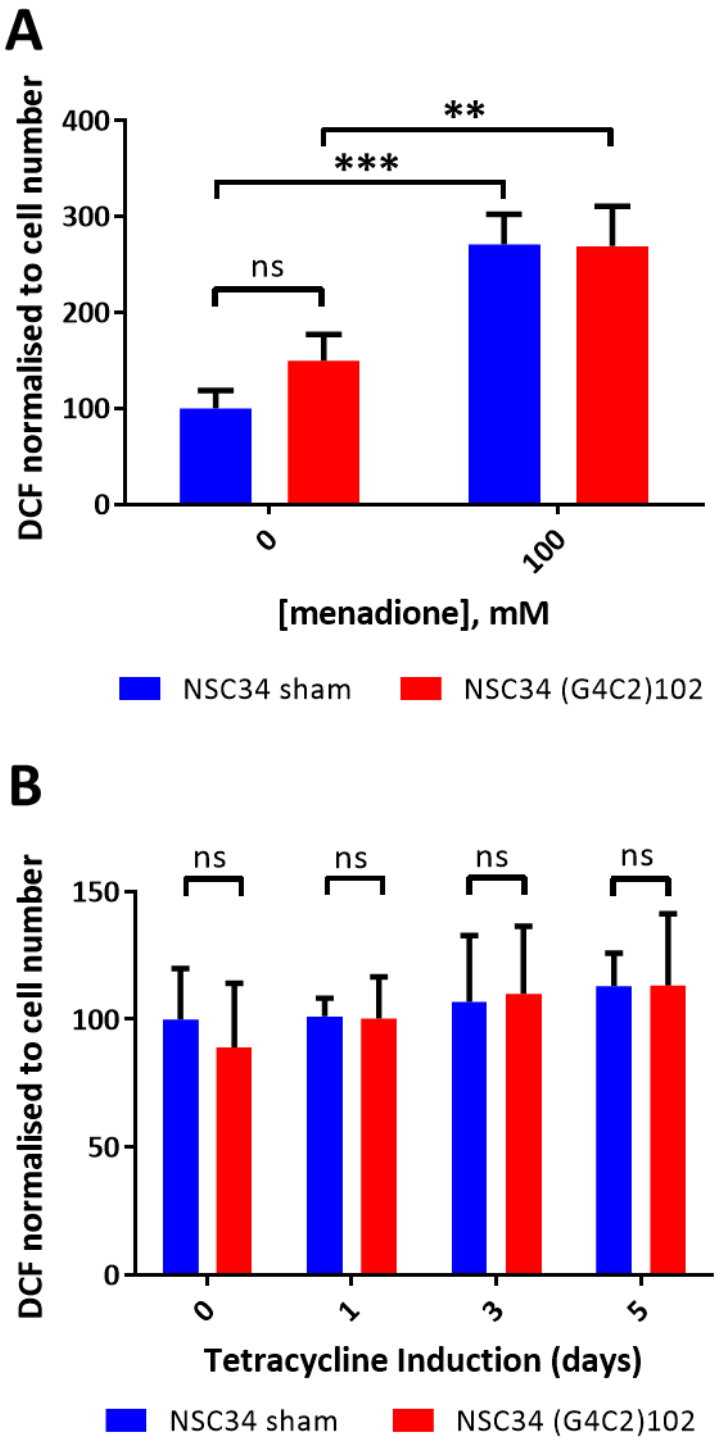


Figure 4.2 (G4C2)102 expression does not cause oxidative stress in NSC34 cells. The level of Reactive Oxygen Species (ROS) in the NSC34 cells was measured using the DCF assay, and normalised to cell number (measured after cell lysis using EthD1 fluorescence assay). **A)** NSC34 sham and NSC34 (G4C2)102 cells were treated with 100mM menadione for 24 h as a positive control for the DCF assay. **B)** NSC34 sham and (G4C2)102 cells were cultured for 5 days, and were induced for various lengths of time with 0.5µg/mL tetracycline. (**P<0.01; ***P<0.001; Two-way ANOVA with Tukey's post hoc test; Data are means ± SD; n=3).

4.3.3. RNA Foci Co-Localise with some RNA Binding Proteins in NSC34 (G4C2)102 Cells

The RNA sequestration hypothesis was one of the first proposed to explain how the (G4C2)n repeat expansion causes C9ORF72-ALS. The (G4C2)n and/or (C4G2)n RNA is suggested to bind and sequester RNA binding proteins (RBP), resulting in disrupted RNA metabolism. Therefore, various groups (including our own) performed *in vitro* (G4C2)n RNA pull downs in conjunction with mass spectrometry and/or western blotting to identify candidate binding proteins (Cooper-Knock et al., 2014b; Donnelly et al., 2013; Haeusler et al., 2014; Lee et al., 2013; Mori et al., 2013b; Rossi et al., 2015; Xu et al., 2013). Subsequently, many of these RBP were shown to co-localise with the (G4C2)n RNA foci in either cell models or C9ORF72-ALS post mortem tissue, suggesting *in vivo* sequestration (Cooper-Knock et al., 2014b; Donnelly et al., 2013; Haeusler et al., 2014; Lee et al., 2013; Mori et al., 2013b; Rossi et al., 2015; Xu et al., 2013).

The NSC34 (G4C2)102 cells were co-stained for sense (G4C2)n RNA foci as well as a selection of these RBP that have previously been shown to bind and/or localise with the RNA foci. The aim was firstly to see whether the interrupted (G4C2)102 repeat RNA would model the (G4C2)n repeat expansion of C9ORF72-ALS, and secondly, to see whether the reduced growth rate caused by the (G4C2)102 expression in the NSC34 (G4C2)102 could be caused by RBP sequestration. 50 cells were analysed for each RNA foci-RBP co-stain, and RNA foci-RBP co-localisation was performed manually.

During the experimental work up it was discovered that following RNA FISH, the subsequent ICC procedure was washing away RNA FISH staining. This was possibly because the PBS used in antibody staining incubations and washes altered the salt concentration which caused the probe to wash off. Therefore, the protocol was optimised to include a crosslinking step (using UV or PFA) after the RNA FISH procedure, and before the ICC procedure. The crosslinking step using UV improved the RNA FISH stain, and was therefore incorporated into the co-stain method. UV crosslinking causes molecules to form new covalent bonds with other molecules in close proximity, and therefore, it was predicted to strengthen the binding between the RNA FISH probe and the (G4C2)n RNA.

4.3.3.1. (G4C2)n RNA Foci Co-Localise with SRSF1 in the NSC34 (G4C2)102 Cells

Serine/Arginine-Rich Splicing Factor 1 (SRSF1/SF2/ASF) is an RBP that is involved in multiple gene expression processes including mRNA splicing (Ge and Manley, 1990; Krainer et al., 1990), nonsense-mediated mRNA decay (Sato et al., 2008; Zhang and Krainer, 2004), nuclear export of mRNA (Huang et al., 2003; Lai and Tarn, 2004; Tintaru et al., 2007), translation (Michlewski et al., 2008; Sanford et al., 2004), and miRNA processing (Wu et al., 2010). 19.5% of RNA foci in the 50 counted NSC34 (G4C2)102 cells co-localised with SRSF1 puncta (Figure 4.3).

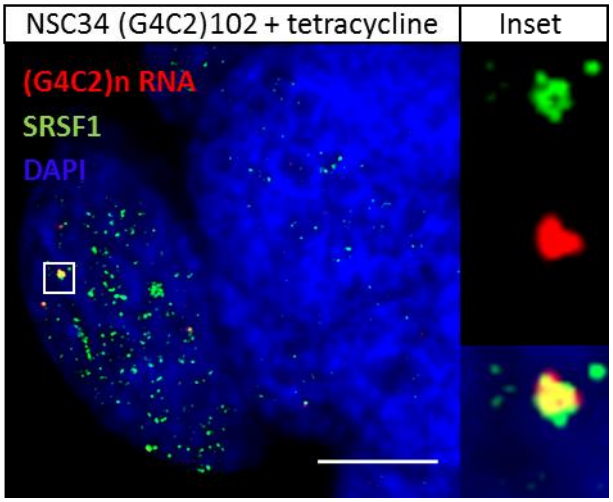


Figure 4.3 (G4C2)n RNA foci co-localise with SRSF1 in NSC34 (G4C2)102 cells. NSC34 (G4C2)102 cells were induced with 0.5µg/mL tetracycline for 5 days. Cells were then stained with a Locked Nucleic Acid (C4G2)3 sense probe (Red), anti-SRSF1 (Green), and Dapi (Blue). 50 nuclei were imaged using a confocal microscope, and the RNA foci-SRSF1 puncta colocalisation was quantified. The image shown is one z-plane imaged using the confocal microscope. Scale bar = 10µm, Inset = 5X magnification.

4.3.3.2. (G4C2)n RNA Foci Co-Localise with SRSF2 in the NSC34 (G4C2)102 Cells

Serine/Arginine-Rich Splicing Factor 2 (SRSF2/SC35) is another Serine/Arginine-Rich protein, and similarly to SRSF1, is involved in multiple gene expression processes including mRNA splicing (Fu and Maniatis, 1990; Fu et al., 1992), transcription elongation (Lin et al., 2008), nonsense-mediated mRNA decay (Zhang and Krainer, 2004). 11.9% of RNA foci in the 50 counted NSC34 (G4C2)102 cells co-localised with SRSF2 puncta (Figure 4.4).

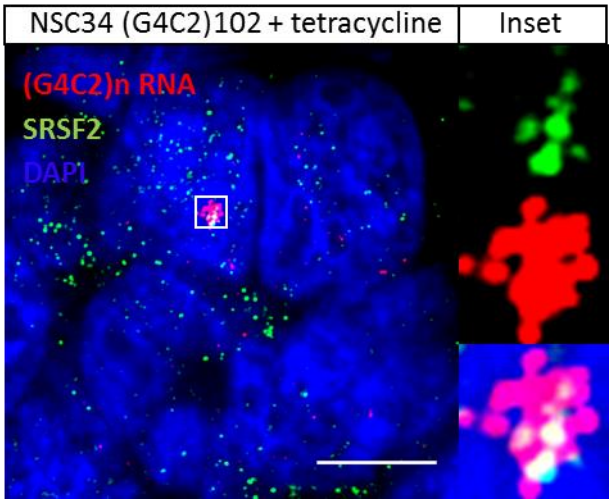


Figure 4.4 (G4C2)n RNA foci co-localise with SRSF2 in NSC34 (G4C2)102 cells. NSC34 (G4C2)102 cells were induced with 0.5µg/mL tetracycline for 5 days. Cells were then stained with a Locked Nucleic Acid (C4G2)3 sense probe (Red), anti-SRSF2 (Green), and DAPI (Blue). 50 nuclei were imaged using a confocal microscope, and the RNA foci-SRSF2 puncta colocalisation was quantified. The image shown is one z-plane imaged using the confocal microscope. Scale bar = 10µm, Inset = 5X magnification.

4.3.3.3. (G4C2)n RNA Foci Do Not Co-Localise with PURA in the NSC34 (G4C2)102 Cells

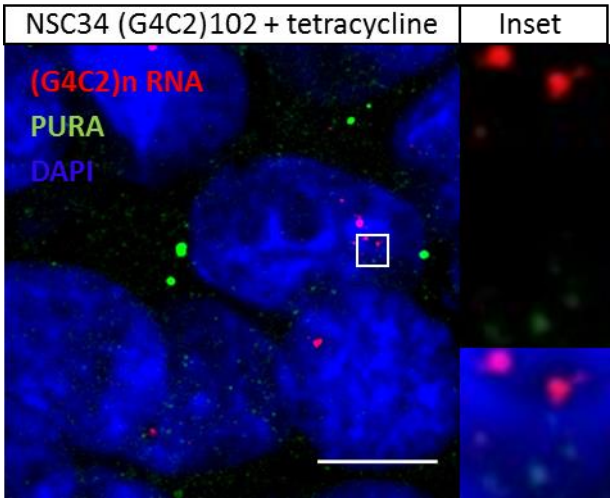


Figure 4.5 (G4C2)n RNA foci do not co-localise with PURA in NSC34 (G4C2)102 cells. NSC34 (G4C2)102 cells were induced with 0.5µg/mL tetracycline for 5 days. Cells were then stained with a Locked Nucleic Acid (C4G2)3 sense probe (Red), anti-PURA (Green), and DAPI (Blue). 50 nuclei were imaged using a confocal microscope, and the RNA foci-PURA puncta co-localisation was quantified. The image shown is one z-plane imaged using the confocal microscope. Scale bar = 10µm, Inset = 5X magnification.

Purine-Rich Element Binding Protein A (PURA) is a multifunctional protein that binds single-stranded DNA and RNA, and is involved in transcription (Haas et al., 1993; Haas

et al., 1995; White et al., 2009), mRNA transport and translation (Ohashi et al., 2000; Ohashi et al., 2002), DNA replication (Chang et al., 1996; Jurk et al., 1996), and DNA repair (Wang et al., 2007). The PURA staining was fairly diffuse throughout the cytoplasm and nuclei, but occasional large PURA aggregates were also detected (Figure 4.5). There was no evidence of RNA foci and PURA co-localisation in the NSC34 (G4C2)102 cells (Figure 4.5).

4.3.3.4. (G4C2)n RNA Foci Do Not Co-Localise with ALYREF in the NSC34 (G4C2)102 Cells

ALY/REF Export Factor (ALYREF) is involved in mRNA nuclear export (Rodrigues et al., 2001; Zhou et al., 2000). ALYREF staining was diffuse throughout the nuclei in the NSC34 (G4C2)102 cells (Figure 4.6). There was no evidence of RNA foci and ALYREF co-localisation in the NSC34 (G4C2)102 cells (Figure 4.6).

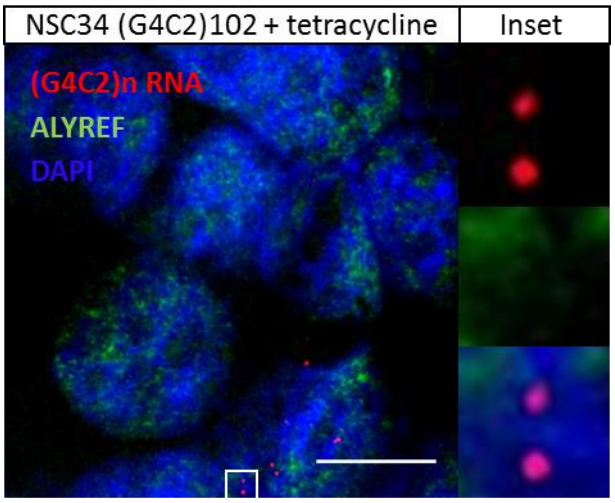


Figure 4.6 (G4C2)n RNA foci do not co-localise with ALYREF in NSC34 (G4C2)102 cells. NSC34 (G4C2)102 cells were induced with 0.5µg/mL tetracycline for 5 days. Cells were then stained with a Locked Nucleic Acid (C4G2)3 sense probe (Red), anti-ALYREF (Green), and DAPI (Blue). 50 nuclei were imaged using a confocal microscope. The image shown is one z-plane imaged using the confocal microscope. Scale bar = 10µm, Inset = 5X magnification.

4.3.3.5. (G4C2)n RNA Foci Co-Localise with Nucleolar NCL in the NSC34 (G4C2)102 Cells

Nucleolin (NCL) is another multifunctional RNA binding protein that is most abundant in the nucleolus. NCL is involved in multiple RNA and DNA processing events including ribosomal RNA (rRNA) transcription, rRNA maturation, and ribosome biogenesis (Ginisty

et al., 1998; Ginisty et al., 2000; Roger et al., 2003), mRNA transcription (Uribe et al., 2011), chromatin remodelling (Angelov et al., 2006; Yang et al., 1994), DNA replication (Seinsoth et al., 2003), telomere maintenance (Khurts et al., 2004), and DNA repair (Yang et al., 2009). 22.6% of RNA foci co-localised with the nucleolar NCL in the 50 counted NSC34 (G4C2)102 cells (Figure 4.7A). The RNA FISH-NCL co-stain was also performed on C9ORF72-ALS CNS tissue, and NCL co-localised with RNA foci in both cerebellar granule and Purkinje neurons (Figure 4.7B).

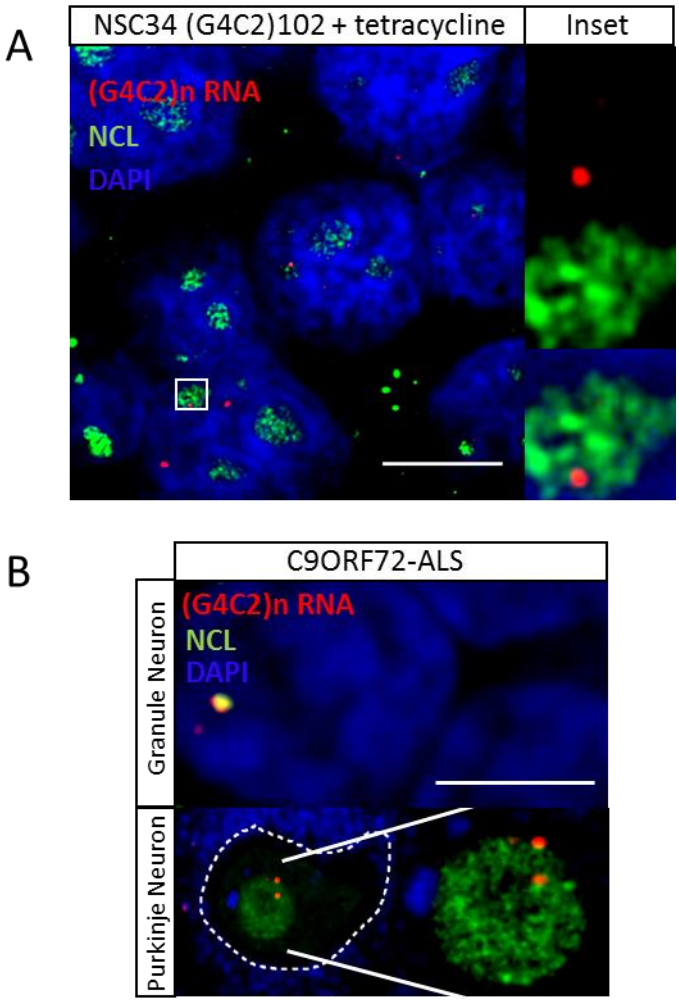


Figure 4.7 (G4C2)n RNA foci co-localise with NCL in NSC34 (G4C2)102 cells and C9ORF72-ALS CNS tissue. A) NSC34 (G4C2)102 cells were induced with 0.5µg/mL tetracycline for 5 days. Cells were then stained with a Locked Nucleic Acid (C4G2)3 sense probe (Red), anti-NCL (green), and Dapi (Blue). 50 nuclei were imaged using a confocal microscope, and the RNA foci-nucleolar NCL co-localisation was quantified. The image shown is one z-plane imaged using the confocal microscope. Scale bar = 10µm, Inset = 5X magnification. **B)** Cerebellar slices from C9ORF72-ALS cases were stained with a Locked Nucleic Acid (C4G2)3 sense probe (Red), anti-NCL (Green) and Dapi (Blue). The image shown is one z-plane imaged using the confocal microscope. Scale bar = 3µm.

4.3.4. (G4C2)102 Expression Does Not Cause Functional Nucleolar Stress in NSC34 (G4C2)102 Cells

4.3.4.1. (G4C2)102 Expression Disrupts Nucleolar Morphology in NSC34 (G4C2)102 Cells

NCL is a major functional protein of the nucleolus and sequestration by (G4C2)n RNA was hypothesised to cause nucleolar stress. Fragmented and/or enlarged nucleoli would indicate nucleolar stress in the NSC34 (G4C2)102 cells. NSC34 sham and NSC34 (G4C2)102 cells were cultured ± tetracycline for 5 days, and then fixed and stained for NCL. NCL staining in the NSC34 cells was visualised using a confocal microscope. A z-stack made up of images at 0.5µm intervals through the entire nuclear volume of the cells under consideration was imaged. The images shown in Figure 4.8A are a maximum projection of the z-stack, such that the total nucleolar area in each cell can be seen and quantified. NCL was localised abundantly in multiple compact globular nucleoli, but was also less abundantly localised diffusely throughout the nucleus in both the NSC34 sham and NSC34 (G4C2)102 ± tetracycline (Figure 4.8A). Extra-nuclear NCL puncta were also common in the NSC34 cells sham and (G4C2)102 ± tetracycline (Figure 4.8A).

A previously published method was used to quantify the nucleolar area in the NSC34 cells (Haeusler et al., 2014). A threshold of 50-100 was set in FIJI to measure the nucleolar NCL area, and exclude the diffuse nuclear NCL staining. The nucleolar NCL area was then normalised to the nuclear area, which was quantified by measuring the DAPI stained area. There was no significant difference in the nucleolar area (as a percentage of the nucleus) between NSC34 sham cells treated and untreated with tetracycline, where the nucleolar area was 23.7±1.8% and 22.1±3.3% respectively (Figure 4.8B). The nucleolar area was increased in the NSC34 (G4C2)102 cells, where the nucleolar area was 24.8 ± 1.3% and 30.6±2.2% in NSC34 (G4C2)102 cells untreated and treated with tetracycline respectively (Figure 4.8B). The nucleolar area was significantly increased in the NSC34 (G4C2)102 induced with tetracycline compared to NSC34 sham cells treated with tetracycline ($P<0.01$) (Figure 4.8B).

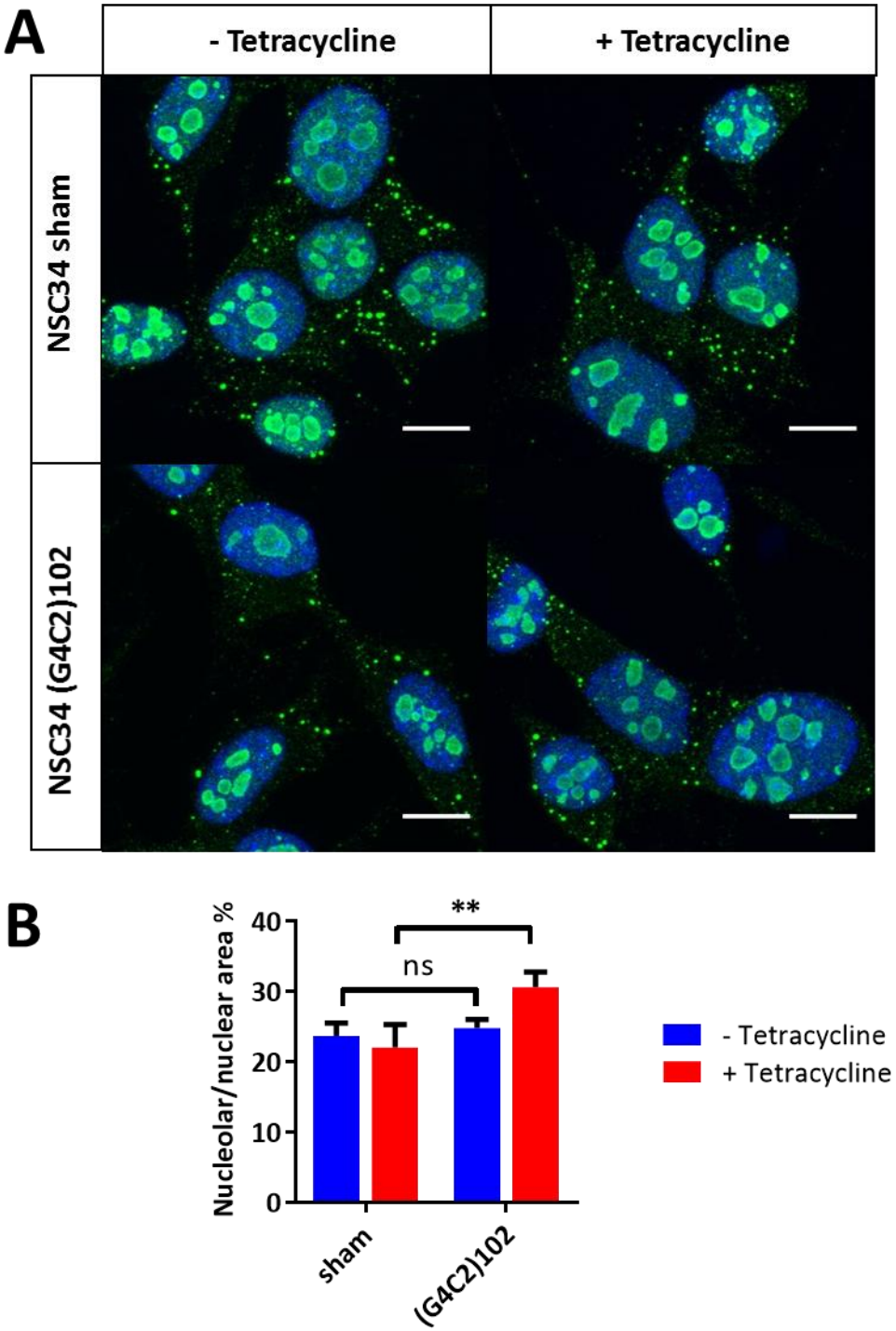


Figure 4.8 (G4C2)102 expression causes nucleolar fragmentation and dispersion in NSC34 (G4C2)102 cells. A) NSC34 (G4C2)102 cells were cultured with 0.5µg/mL tetracycline for 5 days. Cells were then stained with anti-NCL (Green), and Dapi (Blue) and imaged using a confocal microscope within a z-stack through the entire nuclear volume. The images shown are projections of the z-stack to show the maximum nucleolar area in the cells. The area of the nucleoli was measured as a percentage of total nuclear area. Scale bar = 10µm. **B)** Quantification of the nucleolar area as a percentage of nuclear area (**P<0.01; Two-way ANOVA with Tukey's multiple comparisons post hoc test; Data are means ± SD; n=3).

4.3.4.2. (G4C2)102 Expression Does Not Affect rRNA Levels

To assess nucleolar function, the processing and maturation of 45S pre-rRNA, and the levels of mature rRNA species were measured. In the nucleolus, the precursor 45S pre-rRNA is transcribed from rDNA genes, and is then processed to form the mature 18S, 5.8S, and 28S rRNA species. The 45S pre-rRNA molecule is very short lived, and is considered a sign of new rRNA transcription (Uemura et al., 2012). In addition, the levels of mature rRNA species present would indicate the amount of rRNA available for ribosome biogenesis, and also the maturation from pre-rRNA to mature rRNA (when rRNA is normalised to pre-rRNA). Therefore, reduction in 45S pre-RNA and/or mature rRNA levels would indicate impaired nucleolar function.

Total RNA was isolated from the NSC34 sham and NSC34 (G4C2)102 cells (both treated with tetracycline for 5 days), and qRT-PCR was performed to quantify the precursor 45S pre-rRNA, and the mature 18S, 5.8S and 28S rRNA species. There was no significant difference in 45s pre-rRNA, or any of the mature rRNA species in the NSC34 (G4C2)102 compared to NSC34 sham cells (Figure 4.9A), indicating pre-rRNA transcription is unaffected. In addition, there was no significant difference in any of the mature rRNA species when normalised to the 45S pre-rRNA in the NSC34 (G4C2)102 compared to the NSC34 sham cells (Figure 4.9B), indicating pre-rRNA maturation is unaffected.

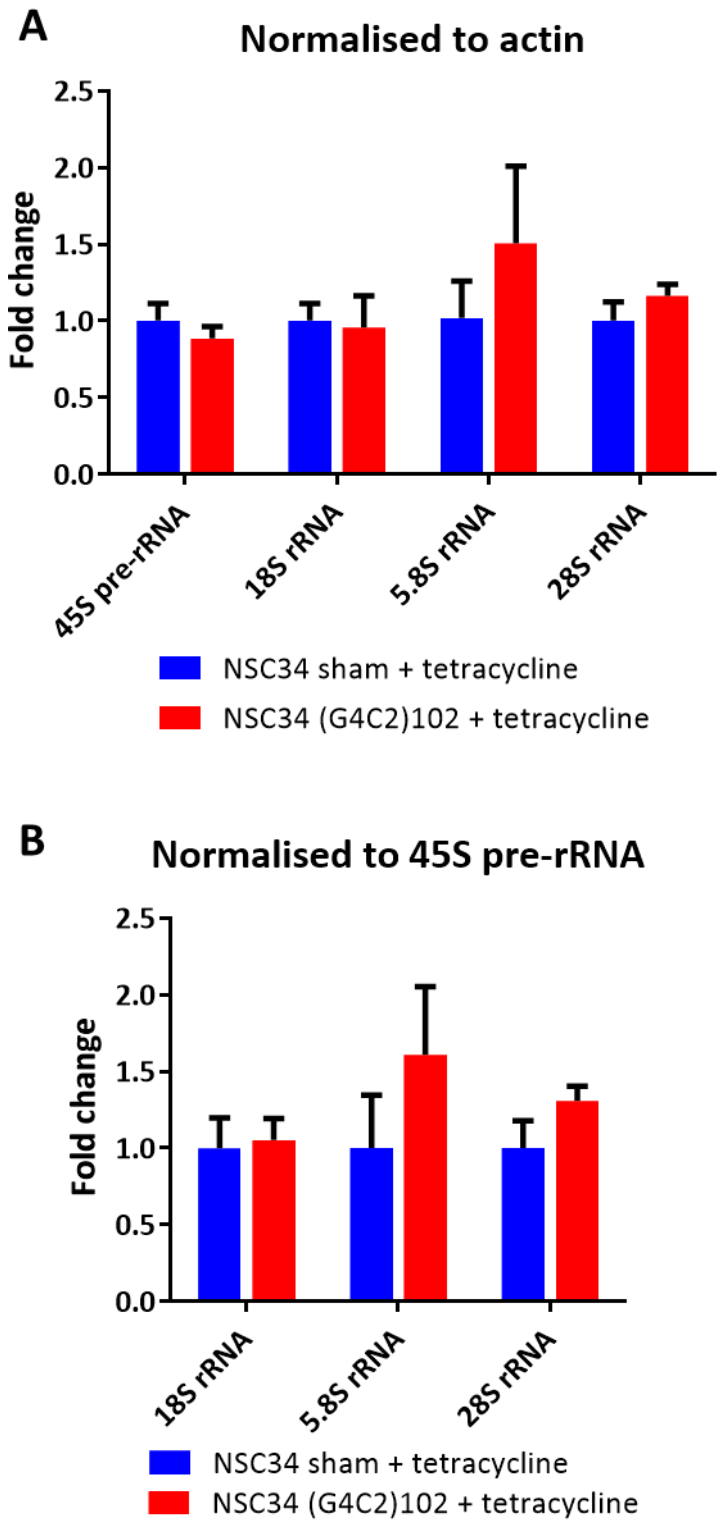


Figure 4.9 Ribosomal RNA maturation is not affected by (G4C2)102 expression in NSC34 cells. NSC34 sham and NSC34 (G4C2)102 cells were grown for 5 days with 0.5µg/mL tetracycline. qRT-PCR was performed on 45S pre-rRNA, and mature 18S rRNA, 5.8S rRNA, and 28S rRNA species, and normalised to β-actin **(A)** or 45S pre-rRNA **(B)** (Multiple t-tests; Data are means ± SD; n=3).

4.3.4.3. NSC34 (G4C2)102 Show Mild Translation Defects

Mature rRNA is assembled into the 60S and 40S ribosomal subunits in the nucleolus (Kressler et al., 1999; Venema and Tollervey, 1999). The 60S and 40S ribosomal subunits are then exported to the cytoplasm, where they can assemble onto mRNA forming the 80S translation initiation complex (Jackson et al., 2010). Once the 80S ribosome starts translating the mRNA, subsequent ribosomes can bind and translate the same mRNA molecule forming a polyribosome (polysome) complex (Jackson et al., 2010). Polysome profiling was performed to assess the translation efficiency in the NSC34 (G4C2)102 cells. It was predicted that nucleolar stress would lead to reduced ribosomal biogenesis, and translation defects.

NSC34 (G4C2)102 and NSC34 sham cells were cultured for 6 days with tetracycline, and then translation was stalled, cells were lysed, and ribosome fractionation was performed using sucrose density centrifugation. Ribosomes were separated into the soluble proteins, 40S and 60S ribosomal subunits, 80S initiation complex, and polysomes. Once the lysate was separated, the sucrose gradient protein content was measured by UV absorbance. NSC34 sham and NSC34 (G4C2)102 cells produced similar profiles in each of three experiments, and a representative profile is shown in Figure 4.10A. There were slight differences between the polysome profiles from the NSC34 (G4C2)102 and sham, but these differences were inconsistent across the three experimental repeats. The 40S peak was undetectable, but is likely in fraction 11 and 12 and hidden in the large soluble shoulder peak (Figure 4.10A). The 60S ribosomal subunit is in fractions 11-13 (Figure 4.10A), and corresponds with the increased 60S Ribosomal Protein L26 (RPL26) detected by immunoblot of the sucrose fractions (Figure 4.10B). The 80S initiation complex is in fractions 13-15 (Figure 4.10A), and corresponds with the abundant RPL26 bands detected by immunoblot (Figure 4.10B). The multiple peaks in fractions 16-21 (Figure 4.10A) also contain RPL26 (Figure 4.10B), confirming that they are the polysomes. The NSC34 (G4C2)102 cells appear to have slightly less RPL26 in fractions 11-21 compared to the NSC34 sham cells (Figure 4.10B), but this finding was weak and inconsistent across the three experimental repeats. There was consistently more RPL26 in fractions 2-7, which correspond to the soluble protein fractions (Figure

4.10B). The amount of soluble RPL26 was normalised to α-tubulin in those fractions, and there was 100±111% more soluble RPL26 in the NSC34 (G4C2)102 compared to NSC34 sham cells (Figure 4.10C).

4.3.4.4. (G4C2)102 Expression Increases p53 Protein Levels in NSC34 (G4C2)102 Cells

Ribosome biogenesis consumes a huge amount of cellular energy, and therefore many metabolic and signalling pathways regulate or are affected by the nucleolus (James et al., 2014). The nucleolus acts as a principle stress sensor, and initiates p53-dependent cell cycle arrest, which can lead to senescence or apoptosis, under cellular stress (Rubbi and Milner, 2003). NCL itself is involved in p53 regulation, and binds the 5'UTR of p53 mRNA inhibiting p53 translation (Takagi et al., 2005). NCL sequestration by the (G4C2)n RNA could therefore lead to increased p53 translation and cell cycle arrest or apoptosis. In addition, RPL26 also binds the 5'UTR of p53 mRNA, but unlike NCL binding, RPL26 activates p53 translation (Takagi et al., 2005).

p53 protein levels were increased by 191±95.8% ($P<0.01$) in NSC34 (G4C2)102 cells induced with tetracycline compared to NSC34 sham induced with tetracycline (Figure 4.11A-B). p53 protein levels were not affected by tetracycline in the NSC34 sham cells (Figure 4.11A-B). To identify whether the increased p53 protein levels could be caused by an increase in p53 translation, RNA was extracted from sucrose fractions containing translating ribosomes in the polysome profiling experiment (Figure 4.10A-B), and qRT-PCR was performed for p53 and actin. There was a 61±147% increase in p53 mRNA in the translating ribosomes in the NSC34 (G4C2)102 + tet compared to NSC34 sham + tet, although this difference was not statistically significant (Figure 4.11C).

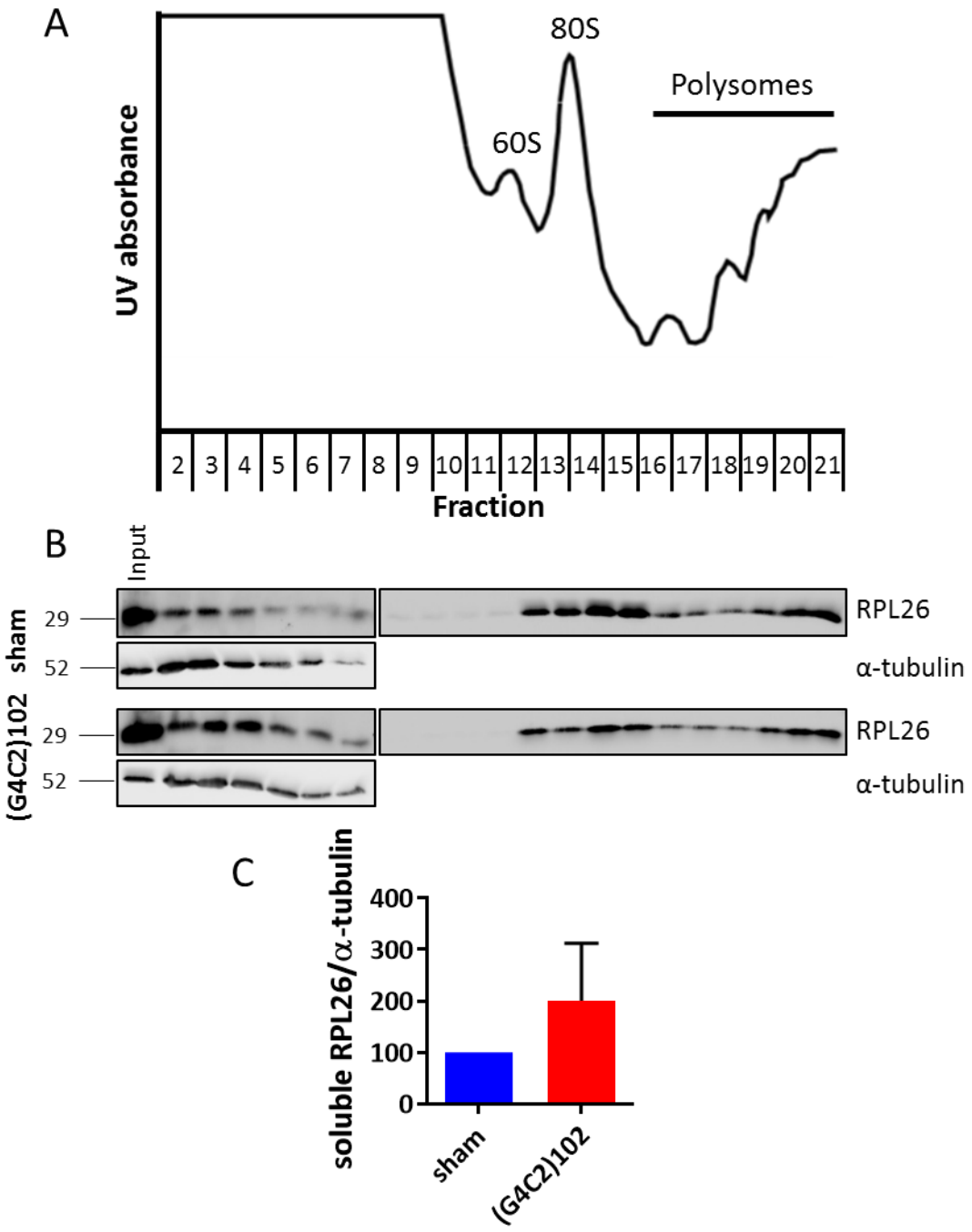


Figure 4.10 (G4C2)102 expression increases soluble RPL26 in NSC34 cells. NSC34 sham and NSC34 (G4C2)102 cell were cultured for 6 days with 0.5µg/mL tetracycline. The translating ribosomes (polysomes) were stalled, cells were lysed, and the ribosomal subunits and complexes were separated using sucrose density gradient centrifugation. The polysome profiles were measured using a FPLC machine, and sucrose fractions were collected. **A)** A representative polysome profile from the NSC34 cells, showing ribosomal subunits, 80S initiation complex, and polysomes. **B)** Fractions were immunoblotted for RPL26 and α-tubulin. **C)** Quantification of soluble RPL26 (fractions 2-7) normalized to α-tubulin. (Data are means ± SD; n=3).

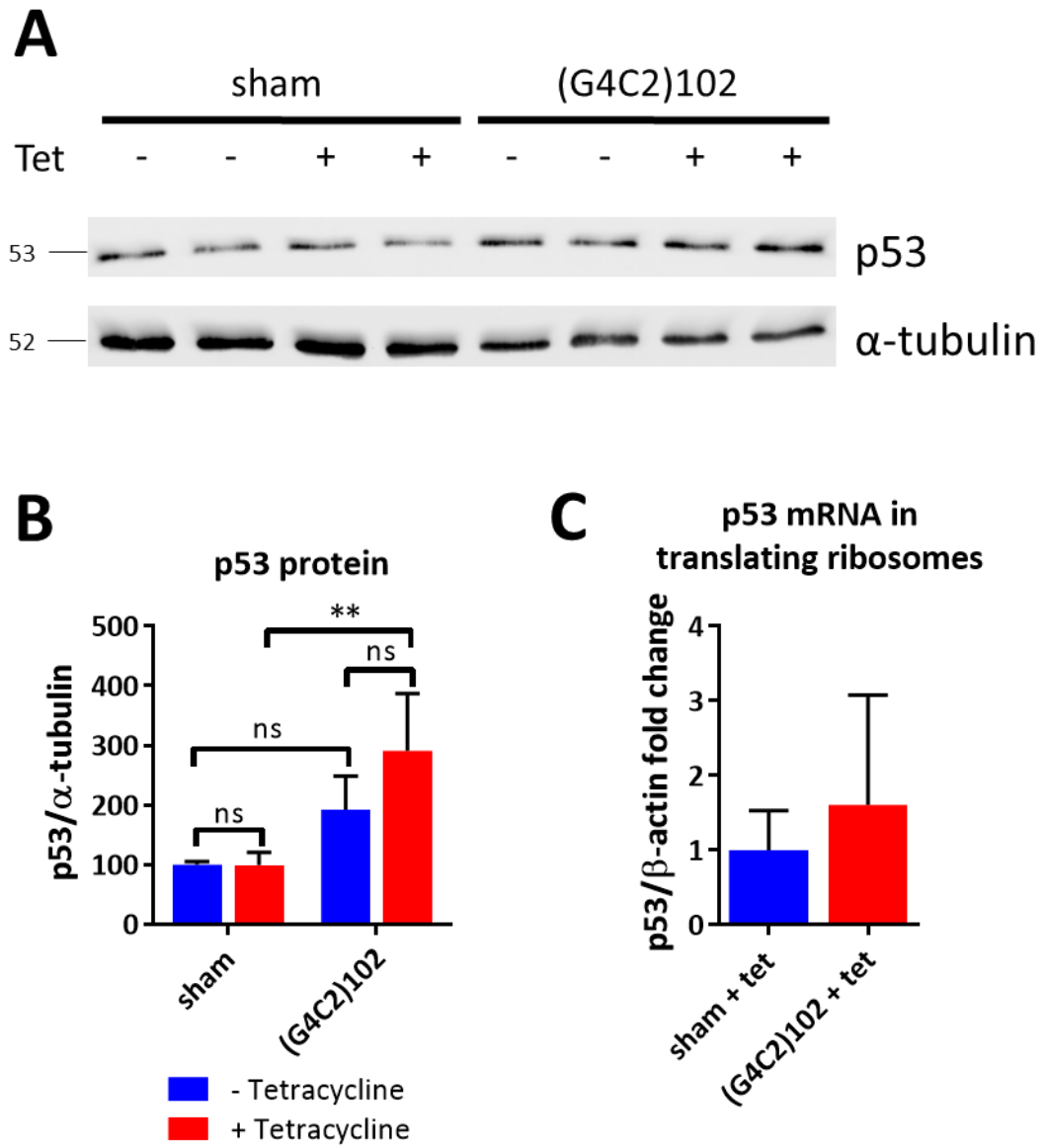


Figure 4.11 (G4C2)102 expression increases p53 protein levels. NSC34 sham and NSC34 (G4C2)102 cells were cultured for 6 days with or without 0.5µg/mL tetracycline. **A)** Cells were lysed and immunoblotted with anti-p53 and anti-α-tubulin. Molecular weight markers are indicated (kDa). **B)** Quantification of p53 protein normalised to α-tubulin (**P<0.01; Two-way ANOVA with Tukey's multiple comparisons post hoc test; Data are means ± SD; n = 3). **C)** After polysome profiling, RNA was extracted from fractions 20 and 21 (containing the translating polysomes) and qRT-PCR was performed on p53 and normalised to β-actin (t-test; Data are means ± SD; n=3).

4.4. Discussion

4.4.1. There Is No TDP-43 Mislocalisation or Aggregation

TDP-43 aggregates and mislocalisation are the pathological hallmarks of most genetic subtypes of ALS, including C9ORF72-ALS. However, it is unknown whether (G4C2)n expression, and/or C9ORF72 haploinsufficiency causes TDP-43 pathology in C9ORF72-ALS. The NSC34 (G4C2)102 cells are an inducible model of C9ORF72-ALS, and any effects in this model are early biochemical effects. There was no evidence of TDP-43 mislocalisation or aggregation in the NSC34 (G4C2)102 cells, and this is in agreement with the majority of other C9ORF72-ALS models (section 1.8.2). This is unsurprising however because TDP-43 mislocalisation and aggregation is not an early event in ALS pathogenesis.

Important comparisons about whether (G4C2)n expression causes TDP-43 pathology can be drawn from several recently published C9ORF72-ALS mouse models. One of the models expresses high levels of (G4C2)66 throughout the CNS, and is the only model of C9ORF72-ALS that does show TDP-43 pathology (Chew et al., 2015). The (G4C2)66 repeat is highly expressed from a chicken beta actin (CBA) promoter in the sense orientation only, and independently of the *C9ORF72* gene context, throughout the CNS (Chew et al., 2015). Nuclear (G4C2)n RNA foci, sense DPR aggregates, and TDP-43 aggregates were detected in the CNS of the mice at 6 months (Chew et al., 2015). However, two different *C9ORF72* BAC transgenic mice express a much larger (G4C2)n repeat (approximately 500 or 800 repeats in size) from the *C9ORF72* gene context, and do not show any signs of TDP-43 pathology even by 20 months (O'Rourke et al., 2015; Peters et al., 2015). Sense and antisense RNA foci, as well as DPR aggregates were detected in the *C9ORF72* BAC transgenic mice, but the promoters associated with the *C9ORF72* gene are much weaker than the CBA promoter used in the first mouse model (O'Rourke et al., 2015; Peters et al., 2015). Therefore, the (G4C2)n expression is much lower in the *C9ORF72* BAC transgenic mouse compared to the (G4C2)66 mouse. Importantly, the models show that the sense (G4C2)n RNA and/or sense DPR proteins are sufficient to cause TDP-43 pathology, but only when over expressed, and over a significant length of time.

There are several reasons that potentially explain why the NSC34 (G4C2)102 cells, and other cellular models of C9ORF72-ALS, do not show TDP-43 aggregation and/or mislocalisation. Firstly, (G4C2)n expression may only cause TDP-43 aggregation in post-mitotic CNS cells. Secondly, the (G4C2)n expression is not high enough, or expressed for long enough in the NSC34 cells. Thirdly, one of the pure DPR species may be responsible for the TDP-43 pathology, which the NSC34 (G4C2)102 cells do not express. Lastly, a second hit, either genetic or age-related, may be required in addition to the (G4C2)n expression to cause TDP-43 pathology, and could be particularly relevant for C9ORF72-ALS pathogenesis. However, any effects that are detected in the NSC34 (G4C2)102 cells are early biochemical effects that could occur prior to TDP-43 pathology in C9ORF72-ALS.

4.4.2. There Is No Increase in Hydroxyl, Peroxyl, or other Cellular ROS Levels in the NSC34 (G4C2)102 Cells

Oxidative stress is thought to be key mechanism in ALS pathogenesis, and many studies report increased levels of oxidative damage in post-mortem tissue and biosamples from SALS patients. There were increased 8-oxo-7,8-dihydroguanosine (8-OHG) levels in the motor cortex and cerebellum of sporadic ALS patients, which indicates increased oxidative damage to mRNA species (Chang et al., 2008). Also, there was increased protein carbonylation, a marker of oxidised protein, in both the motor cortex and spinal cord from SALS patients (Ferrante et al., 1997; Shaw et al., 1995), and elevated levels of 8-hydroxy-2'-deoxyguanosine (OH8dG), a marker of oxidized DNA, in both the motor cortex and cervical spinal cord from SALS patients (Ferrante et al., 1997; Fitzmaurice et al., 1996). Additionally, increased levels of 4-hydroxy-2-nonenal (HNE) and crotonaldehyde (CRA) in the motor neurons and glial cells of spinal cord from SALS patients, indicate increased lipid peroxidation (Shibata et al., 2001). Finally, OH8dG levels were elevated in urine samples from SALS patients (Mitsumoto et al., 2008), HNE levels were elevated in serum, spinal fluid, and CSF samples from SALS patients (Simpson et al., 2004; Smith et al., 1998).

However, the DCF assay did not detect increased levels of hydroxyl, peroxyl or other cellular ROS in the NSC34 (G4C2)102 cells. As described above in section 4.4.1, any

effects detected in the NSC34 (G4C2)102 cells are likely to be early biochemical effects, and suggests (G4C2)102 expression does not induce oxidative stress as an early biochemical effect. In addition, the vast majority of other C9ORF72-ALS models either make no comment on, or show no oxidative stress, which could suggest oxidative stress is not the primary toxic effect of (G4C2)n expression and/or C9ORF72 haploinsufficiency. In one study however, catalase transcript levels were significantly higher in C9ORF72-ALS iPSC-derived motor neurons, which indicates oxidative stress (Kiskinis et al., 2014).

There is stronger evidence for oxidative stress in other genetic models of ALS. Increased protein carbonylation, hydroxyl radicals, hydrogen peroxide, lipid peroxidation, and oxidative damage to DNA are described in SOD1(G93A) mouse models (Andrus et al., 1998; Liu et al., 1999; Liu et al., 1998; Poon et al., 2005). Increased lipid peroxidation and oxidative stress were also measured in NSC34 cells stably transfected with SOD1(G93A) (Wang et al., 2014). In addition to the SOD1(G93A) models, NSC34 cells expressing mutant TDP-43 have increased nuclear NRF2, which indicates increased oxidative stress (Duan et al., 2010). Glutathione S transferase D1 (GstD1) expression and protein carbonylation were elevated (both markers of oxidative stress) in a *Drosophila* model that expresses TDP-43 in the motor neurons (Zhan et al., 2015). In addition to biochemical studies, transcriptome analysis of NSC34 cells expressing SOD1(G93A) showed dysregulation in antioxidant response genes, suggesting oxidative stress (Kirby et al., 2005). Therefore, transcriptomic analysis of the NSC34 (G4C2)102 cells may detect whether the (G4C2)102 expression does cause oxidative stress or the NSC34 cells' ability to cope with oxidative stress.

4.4.3. (G4C2)n RNA Foci Co-Localise with Splicing Factors SRSF1 and SRSF2 in the NSC34 (G4C2)102 Cells

The hypothesis was that the (G4C2)n RNA foci would bind and co-localise with RBP in the NSC34 (G4C2)102 cells, which would lead to dysregulated RNA metabolism. SRSF1 and SRSF2 are both multifunctional RBP that co-localised with 19.5% and 11.9% of RNA foci counted in the NSC34 (G4C2)102 cells respectively, consistent with several *in vitro* studies. SRSF1 and SRSF2 were both pulled down from SHSY5Y whole-cell extract, SHSY5Y nuclear extract and dissected human cerebellum whole extract *in vitro* by

biotinylated-(G4C2)5 RNA (Cooper-Knock et al., 2014b). These interactions were confirmed to be direct, using UV crosslinking RNA pull down assays (Cooper-Knock et al., 2014b). SRSF1 was also pulled down from NSC34 whole cell extract *in vitro* by biotinylated-(G4C2)31 RNA in a different study (Rossi et al., 2015). In addition, SH-SY5Y cells transfected with (G4C2)72 constructs displayed colocalisation between RNA foci and both SRSF1 and SRSF2 (Lee et al., 2013). There was evidence of (G4C2)n RNA foci and SRSF2 colocalisation, in both cerebellar granule cells and ventral horn motor neurons from C9ORF72-ALS patients (Cooper-Knock et al., 2014b). However, there was no evidence of (G4C2)n RNA foci and SRSF1 colocalisation in the C9ORF72-ALS CNS tissue in the same study. In a different study, (G4C2)n RNA foci and both SRSF1 and SRSF2 co-localisation was much rarer in the cerebellum of C9ORF72-ALS patients (Lee et al., 2013). In summary, it is far from clear whether these protein-RNA interactions are physiologically relevant, but also possibly highlights the dynamic nature of these interactions.

SRSF1 and SRSF2 are both involved in mRNA splicing and several previous transcriptomic studies have shown splicing defects in C9ORF72-ALS cases. Splicing consistency is lower in C9ORF72-ALS lymphoblastoid cells compared to controls and non-C9ORF72-ALS cases (Cooper-Knock et al., 2015a). Extensive splicing and polyadenylation defects were found in C9ORF72-ALS patient cerebellum and frontal cortex (Prudencio et al., 2015). Further, amongst the spliced exons and introns with the splicing defects, binding motifs recognised by HnRNP H and/or SRSF1 were enriched (Prudencio et al., 2015). This certainly suggests that SRSF1 sequestration by (G4C2)n RNA could play a role in splicing dysregulation and neurodegeneration in C9ORF72-ALS. Differential splicing analysis of the NSC34 (G4C2)102 transcriptome also shows increased splicing defects caused by (G4C2)102 expression in the cellular model as well (section 5.3.4), suggesting splicing dysregulation is an early biochemical effect of (G4C2)n expression. Taken together, SRSF1 sequestration by the (G4C2)n RNA (either in RNA foci or soluble RNA molecules) could be a critical early pathological function of the (G4C2)n repeat expansion that contributes to C9ORF72-ALS.

4.4.4. (G4C2)n RNA Foci Do Not Co-Localise with either PURA or ALYREF in the NSC34 (G4C2)102 Cells

Unlike SRSF1 and SRSF2, there was no co-localisation between the (G4C2)n RNA foci and either PURA or ALYREF. This is inconsistent with several *in vitro* studies that show PURA and ALYREF both bound (G4C2)n RNA. PURA was pulled down from mouse spinal cord whole-cell lysate, dissected human cerebellum whole extract, and NSC34 whole-cell lysate *in vitro* by biotinylated-(G4C2)n RNA (Cooper-Knock et al., 2014b; Rossi et al., 2015; Xu et al., 2013). PURA also co-localises with (G4C2)n RNA foci in C9ORF72-ALS patient iPSC-derived motor neurons (Sareen et al., 2013). ALYREF was also pulled down from SHSY5Y whole cell extract *in vitro* by biotinylated-(G4C2)5 RNA, and the interaction was confirmed to be direct (Cooper-Knock et al., 2014b). ALYREF also co-localised with RNA foci in both cerebellar granule cells and ventral horn motor neurons from C9ORF72-ALS patients (Cooper-Knock et al., 2014b).

There are several different studies that have used *in vitro* RNA pull down techniques to identify RBP that bind the (G4C2)n RNA. However, there are inconsistencies across these studies, as certain RBP are shown to bind the (G4C2)n RNA in one (or more) studies, but not in others. This is probably caused by several different variables: the cellular origin of the cellular lysates used for the RNA pull downs (and the differential enrichment of certain RBP in the respective lysate), the stringency of the pull down, and the secondary structure of the (G4C2)n RNA used in the pull down. For example, NCL and HnRNP U preferentially bound (G4C2)n RNA with a G-quadruplex structure compared to (G4C2)n RNA with a hairpin structure (Haeusler et al., 2014). Whilst the RNA pull downs and co-localisation experiments are useful for identifying candidate RBP that may be sequestered by the (G4C2)n RNA, functional studies must be used to validate the importance of the RBP. For example, rescue assays, and detailed splicing analysis have been used to strengthen the case for PURA, and HnRNP F/H and SRSF1 sequestration by the (G4C2)n RNA respectively in C9ORF72-ALS pathogenesis (Prudencio et al., 2015; Xu et al., 2013).

In addition, (G4C2)n expression may also affect PURA localisation, in a mechanism separate to sequestration in RNA foci. PURA was diffusely distributed in the cytoplasm

and nucleus of iPSC-derived motor neurons (Sareen et al., 2013), and untransfected NSC34 and HeLa cells (Rossi et al., 2015). Although PURA was much more abundant in the cytoplasm in the NSC34 and HeLa cells (Rossi et al., 2015). However, PURA formed cytoplasmic and nuclear granules in the NSC34 and HeLa cells that were transfected with (G4C2)31 repeat constructs (Rossi et al., 2015). Large, intensely stained PURA aggregates were also detected in the NSC34 (G4C2)102 cells. Interestingly, a recent study also suggests PURA is essential for stress granule formation (Daigle et al., 2016). Taken together, it suggests the (G4C2)n expression could drive stress granule formation, and is important because altered stress granule dynamics are hypothesised to play an important role in ALS pathogenesis (Li et al., 2013; Ramaswami et al., 2013).

4.4.5. (G4C2)n RNA Foci Co-Localise with NCL, but There Is No Evidence for Functional Nucleolar Stress

RNA foci co-localised with the nucleolar NCL in the NSC34 (G4C2)102 cells and also in cerebellar granule and Purkinje neurons from C9ORF72-ALS patients, consistent with several other studies. NCL was pulled down from HEK293T whole-cell extract *in vitro* by biotinylated-(G4C2)4 RNA with both a hairpin and G-quadruplex secondary structure, although NCL preferentially bound the G-quadruplex motif (Haeusler et al., 2014). NCL was also the most significantly pulled down protein from SHSY5Y whole cell extract, SHSY5Y nuclear extract and dissected human cerebellum whole extract *in vitro* by biotinylated-(G4C2)5 RNA (Cooper-Knock et al., 2014b). In addition, (G4C2)n RNA foci co-localised with the nucleolar NCL in the motor cortex of C9ORF72-ALS patient post-mortem tissue (Haeusler et al., 2014). Taken together, the interaction between NCL and (G4C2)n RNA could be important in pathophysiology of C9ORF72-ALS.

NCL is a major functional protein of the nucleolus, and therefore, NCL sequestration by (G4C2)n RNA could cause nucleolar stress. The (G4C2)102 expression in the NSC34 (G4C2)102 cells increased nucleolar area compared to the NSC34 sham cells, indicating nucleolar stress. This finding is consistent with a previous study that showed nucleoli were more fragmented, dispersed and were larger in C9ORF72-ALS B lymphocytes, fibroblasts and iPSC-derived motor neurons compared to controls (Haeusler et al., 2014). Additionally, the increased nucleolar size in the NSC34 (G4C2)102 cells induced

with tetracycline is comparable to the nucleolar size increase in the C9ORF72-ALS fibroblasts (Haeusler et al., 2014). However, (G4C2)102 expression does not affect rRNA levels or maturation in the NSC34 (G4C2)102 cells. Interestingly, rRNA maturation is significantly reduced in C9ORF72-ALS motor cortex tissue, but not in C9ORF72-ALS B lymphocytes (Haeusler et al., 2014). Nucleolar stress was also hypothesised to affect translation efficiency in the NSC34 (G4C2)102 cells. However, there were no consistent differences between polysome profiles from the NSC34 (G4C2)102 and sham cells, suggesting no gross translation defect. The amount of RPL26 in the soluble fraction was increased in the NSC34 (G4C2)102 cells however, which indicates a defect in ribosome subunit biogenesis, and nucleolar dysfunction.

The (G4C2)n repeat expansion could also affect nucleolar function via the DPR proteins, in addition to potential NCL sequestration by the (G4C2)n RNA. Synthetic (GR)20 and (PR)20 peptides localise to the nucleolus in U2OS and cultured human astrocytes (Kwon et al., 2014). (GR)n and (PR)n DPR proteins also localised to the nucleolus in HEK293 cells transfected with DPR expression vectors (Tao et al., 2015). Further, the (GR)n and (PR)n induced signs of nucleolar stress in both these studies. The cultured human astrocytes treated with either (PR)20 or (GR)20 showed dysregulated rRNA metabolism, although it was unclear whether the defect was in the rRNA transcription and/or maturation process (Kwon et al., 2014). Also, the (GR)n and (PR)n caused an increase in nucleolar size in the transfected HEK293 cells (Tao et al., 2015). These nucleolar stress results are both consistent with the findings of Haeusler et al., in the C9ORF72-ALS tissue and actually suggests the DPRs could be contributing to some (if not all) of the nucleolar stress. This may also explain why the nucleolar defects are weak and inconsistent in the NSC34 (G4C2)102 cells, because no pure (GR)n or (PR)n DPRs are translated from the interrupted (G4C2)102 repeat (section 3.3.4.3). Although, this does not exclude the (G4C2)102 RAN proteins - which do contain stretches of the (GR)n motif - from causing subtle nucleolar stress.

4.4.6. Increased p53 Protein in the NSC34 (G4C2)102 Cells Indicates Cellular Stress

(G4C2)102 expression caused an increase in p53 protein levels in the NSC34 (G4C2)102 cells, and this was hypothesised to be caused by nucleolar stress. The nucleolus acts as

a principal stress sensor, and initiates p53-dependent cell cycle arrest, which can lead to senescence or apoptosis, under cellular stress (Rubbi and Milner, 2003). The increased p53 levels in the induced NSC34 (G4C2)102 cells suggests that (G4C2)102 expression causes cellular stress and leads to cell cycle arrest, and is consistent with earlier results that showed (G4C2)102 expression reduced NSC34 (G4C2)102 growth rate (section 3.3.4.7).

NCL binds the 5'UTR of p53 mRNA inhibiting p53 translation (Takagi et al., 2005), and therefore, it was reasoned that NCL sequestration by the (G4C2)n RNA could lead to increased p53 translation. Likewise, free RPL26 (not in the ribosomal subunits) also binds the 5'UTR of p53 mRNA, but activates p53 translation (Takagi et al., 2005), and therefore the increased soluble RPL26 in the NSC34 (G4C2)102 cells was also predicted to increase p53 translation. However, qRT-PCR of the translating polysomes did not show any significant increase in p53 mRNA, and does not support either the NCL sequestration hypothesis or increased soluble RPL26 result. There are many other molecular pathways that lead to increased p53 protein levels, and the (G4C2)102 expression could lead to increased p53 protein via one of these rather than via NCL sequestration.

4.4.7. Summary

The NSC34 (G4C2)102 cells did not recapitulate TDP-43 mislocalisation or aggregation, or an increase in oxidative stress. However, the NSC34 (G4C2)102 cells did recapitulate some phenotypes specific to C9ORF72-ALS. Firstly, the (G4C2)n RNA foci colocalised with the splicing factors SRSF1 and SRSF2, and also with nucleolar NCL. Secondly, the nucleolar area is enlarged. However, there was no functional evidence of nucleolar stress. Reassuringly, the (G4C2)102 expression causes subtle biochemical effects in the NSC34 (G4C2)102 cell model, which is arguably more relevant to an age of onset disease. The model therefore warrants further study using transcriptomic analysis, which may provide more clues as to the early biochemical effects caused by the (G4C2)n expression in the NSC34 (G4C2)102 cells.

Chapter 5. Gene Expression Profiling of NSC34 (G4C2)n Cell Lines

5.1. Introduction

Gene expression profiling (GEP) allows genome wide gene expression to be compared between disease and control patient tissue samples, animal models, or cellular models. GEP has been used to identify dysregulated biological pathways in various diseases, which could yield potential therapeutic targets. Microarray is a widely used and robust technique for GEP. Briefly, RNA is extracted from the cells or tissues of interest, amplified, and used to produce single-stranded complementary DNA (ss-cDNA), which is then fluorescently labelled. The labelled ss-cDNA is then hybridised to the microarray chip, which consists of an array of multiple ss-DNA probes fixed to a glass slide. These ss-DNA probes have a known target identity, and therefore the relative amount of labelled RNA in the starting material is proportional to the fluorescent signal intensity on the microarray chip. Thousands of copies of each probe allows for the quantification of many specific RNA transcripts simultaneously.

GeneChip® Mouse Transcriptome Arrays (MTA) 1.0 were used to perform GEP on the NSC34 (G4C2)n cells. MTA 1.0 have probe sets that recognise >23,000 protein coding genes and additionally >55,000 non-coding genes, and there are approximately 10 probe sets that recognise each exon, and 4 that recognise each exon-exon splice junction. This allows robust analysis of differential gene expression at both the whole transcript level and the exon level.

NSC34 sham, NSC34 (G4C2)10, NSC34 (G4C2)51, and NSC34 (G4C2)102 cells were induced with 0.5µg/mL tetracycline for 0, 1 or 5 days. However, as described previously, NSC34 viability was only affected by (G4C2)102 expression, and therefore the early transcriptomic effects of (G4C2)n repeat expression would be most pronounced in the NSC34 (G4C2)102 cells. Therefore, the main array comparisons were made between NSC34 (G4C2)102 and NSC34 sham cells. For brevity, the array names contain two numbers that correspond to the (G4C2)n repeat size (where 0 is sham), and the number

of days that the cells were induced with tetracycline for (0, 1, or 5) respectively, and a letter that denotes the biological repeat (A, B, or C). For example, the array for the first biological replicate of NSC34 (G4C2)102 cells that were induced with tetracycline for 5 days is named 102_5_A.

5.2. Aims and Objectives for Gene Expression Profiling of NSC34 (G4C2)n Cells

1) Identify changes in gene expression at the gene level in the NSC34 (G4C2)102 cells compared to NSC34 sham cells.
2) Perform enrichment analysis to identify which biological pathways are most dysregulated in the NSC34 (G4C2)102 cells compared to NSC34 sham cells.
3) Validate gene expression changes of key genes using qRT-PCR.
4) Identify changes in differential splicing in the NSC34 (G4C2)102 cells compared to NSC34 sham cells.

5.3. Results

5.3.1. RNA Extraction Quality Control

The RNA quality and yield were assessed using the NanoDrop™ 1000 spectrophotometer, following RNA extraction from the NSC34 cell lines. An average yield of 21.71±8.78µg total RNA was obtained, and all RNA samples had A280/260 ratios close to 2.00 (1.99±0.033) (Table 5.1) indicating high quality RNA. An Agilent 2100 Bioanalyzer was used to assess the RNA integrity, which is critical for successful microarray experiments. The RNA Integrity Number (RIN) describes the RNA integrity, where 0 is completely degraded and 10 is completely intact. However, the Bioanalyzer could not compute RIN values for most of the samples, and was displayed as N/A (Table 5.1). This was likely due to the extraction of small RNA species using the RNA extraction columns, which produced an additional peak between 50 and 150 nt (Figure 5.1). This peak is unexpected by the Bioanalyzer RIN computation software. The electropherograms from samples with RIN values of N/A and ≥8.0 were very similar, and all samples produced distinct 18S and 28S rRNA peaks on the electropherograms (Figure

5.1). Also, all RNA samples had 28S/18S rRNA ratios close to 1.50 (1.50±0.12) (Table 5.1), indicating the RNA has high integrity. Additionally, very little degradation products were present between the small RNA, 18 S rRNA and 28 S rRNA peaks (Figure 5.1).

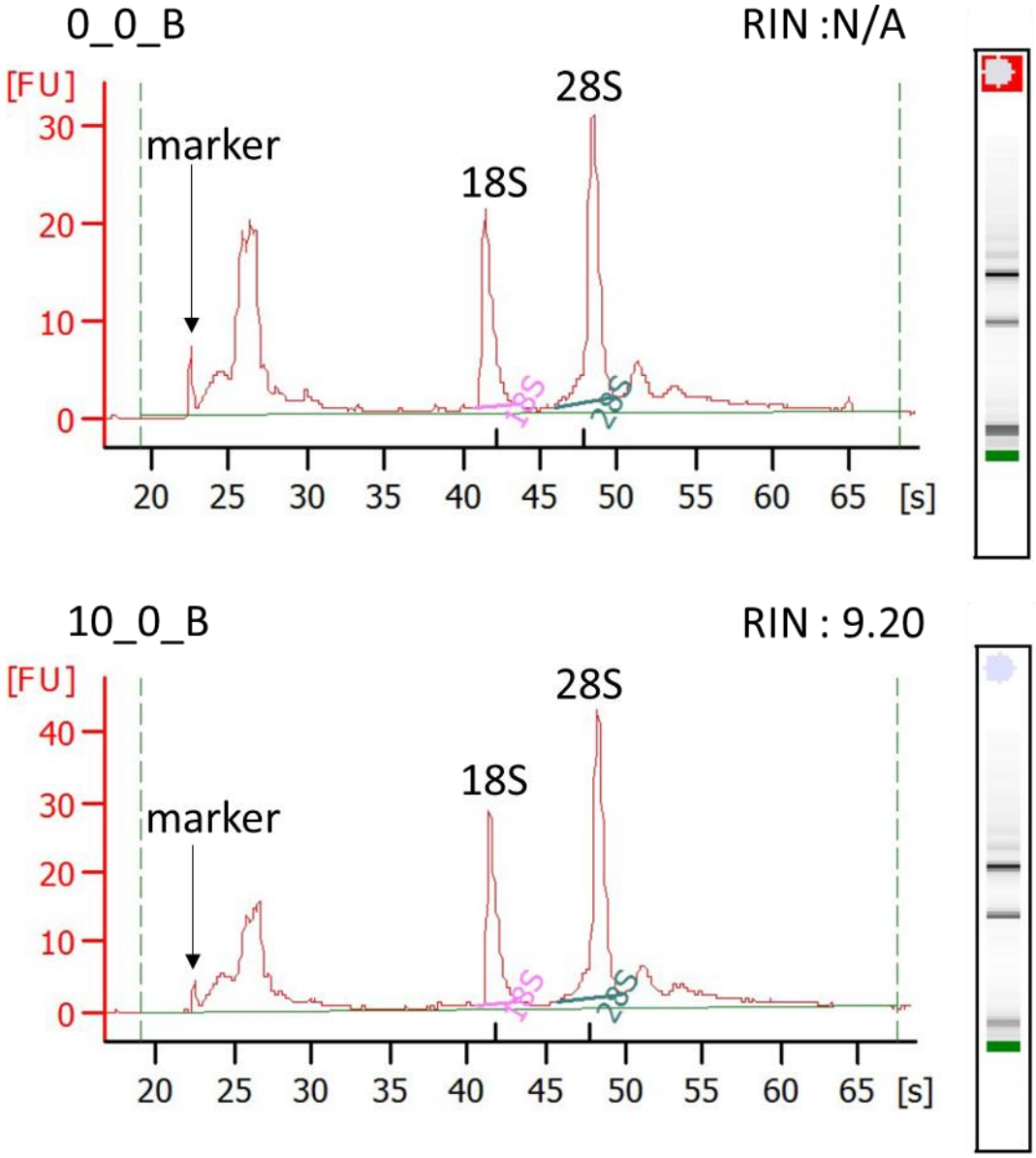


Figure 5.1 Agilent 2100 Bioanalyzer assessment of extracted RNA integrity from NSC34 cells. Example electropherograms and corresponding electrophoresis gels produced by the Agilent 2100 Bioanalyzer used to assess the RNA integrity of extracted RNA samples from NSC34 cells. Peaks corresponding to the 18S and 28S rRNA are labelled. The electropherogram from sample 0_0_B (top) is used here as a representative example to illustrate the integrity of extracted RNA samples with RIN values of N/A, whilst the electropherogram from sample 10_0_B (bottom) is used as a representative example to illustrate the integrity of the extracted RNA samples with RIN values ≥8.00.

Table 5.1 Quality and yield of RNA extracted from NSC34 cells. A NanoDrop™ 1000 spectrophotometer and Agilent 2100 Bioanalyzer were used to measure the concentration and quality of RNA extracted from NSC34 cells. The RNA yield was calculated using the total volume collected multiplied by the RNA concentration. The RNA quality was defined using the ratio of A280/260. The RNA integrity was measured using an Agilent 2100 Bioanalyzer, and described using RNA Integrity Number (RIN) on a scale of 0-10 (0 being completely degraded, and 10 being completely intact). The sample name corresponds to the (G4C2)n repeat size (where 0 is sham), the number of days that the cells were induced with tetracycline for (0, 1, or 5) respectively, and a letter that denotes the biological repeat (A, B, or C). For example, sample 102_5_A corresponds to the first biological replicate of NSC34 (G4C2)102 cells that were induced with tetracycline for 5 days.

Sample	RNA yield (µg)	A260/280	RIN	28S/18S rRNA ratio
0_0_ A	14.40	1.95	N/A	1.3
0_0_B	18.91	2.03	N/A	1.3
0_0_C	13.89	1.97	N/A	1.4
0_1_A	24.87	2.01	N/A	1.4
0_1_B	23.94	2.01	N/A	1.4
0_1_C	19.00	2.00	N/A	1.5
0_5_ A	15.91	2.00	N/A	1.6
0_5_B	24.47	2.02	8.7	1.7
0_5_C	25.14	2.01	8.5	1.7
10_0_A	18.91	1.99	N/A	1.6
10_0_B	30.84	2.02	9.2	1.5
10_0_C	26.24	2.02	8.7	1.3
10_5_A	22.96	2.07	8.7	1.6
10_5_B	15.52	2.02	N/A	1.5
10_5_C	16.79	2.02	N/A	1.4
51_0_A	46.16	2.00	8.7	1.5
51_0_B	47.25	2.02	9.1	1.6
51_0_C	38.38	1.99	8.4	1.7
51_5_A	19.49	1.98	N/A	1.6
51_5_B	22.91	1.98	N/A	1.6
51_5_C	15.89	2.01	N/A	1.7
102_0_A	19.69	1.99	N/A	1.4
102_0_B	20.80	1.99	N/A	1.4
102_0_C	15.66	1.97	N/A	1.4
102_1_A	16.09	1.95	N/A	1.4
102_1_B	17.51	1.97	N/A	1.5
102_1_C	13.66	1.92	N/A	1.4
102_5_A	18.15	1.99	N/A	1.5
102_5_B	14.03	1.96	N/A	1.5
102_5_C	13.70	1.91	N/A	1.5
Mean ± SD	**21.71 ± 8.78**	**1.99 ± 0.033**	**N/A**	**1.50 ± 0.12**

5.3.1.1. RNA Amplification and Labelling

500ng total RNA was reverse transcribed using a reverse transcriptase and primers containing the T7 promoter sequence at the 5' end. This was followed by simultaneous RNAse H and DNA polymerase treatment to degrade the starting RNA template, and synthesise the second-strand of the cDNA, to produce double-stranded cDNA (ds-cDNA). Complimentary RNA (cRNA) was then synthesised and amplified from the ds-cDNA by *in vitro* transcription (IVT) using T7 RNA polymerase. cRNA was then purified and the yield and size distribution of the cRNA were assessed using the NanoDrop™ 1000 Spectrophotometer and Agilent 2100 Bioanalyzer respectively. The mean cRNA yield was 62.18±21.66µg (Table 5.2). The cRNA was generated in batches A, B, and C (containing samples named respectively), to make sample handling more manageable. The cRNA yield was much higher in batch A compared to batches B and C, which accounts for the variability in cRNA yield between samples (Table 5.3). The cRNA profiles assessed by the Agilent 2100 Bioanalyzer showed the expected size distribution, with cRNAs ranging from 50 to 4500 nt in size, and most cRNAs ranging from 50 to 2000 nt in size (Figure 5.2). The distribution was also jagged as expected (Figure 5.2).

5.3.2. GeneChip® Mouse Transcriptome Arrays 1.0

Sense-strand cDNA was then synthesised from 15µg of cRNA in an *in vitro* reverse transcription reaction. dUTP was also incorporated into the ss-cDNA. RNAse H treatment then removed cRNA template leaving ss-cDNA only. After purification, the ss-cDNA yield and size distribution were assessed using the NanoDrop™ 1000 Spectrophotometer and the Agilent 2100 Bioanalyzer. The mean ss-cDNA yield was 24.28±2.61µg (Table 5.3). The ss-cDNA profiles assessed by the Agilent 2100 Bioanalyzer showed the expected sloped profile for the ss-cDNA size distribution, with most ss-cDNAs approximately 100 nt in size, and the median ss-cDNA size approximately 400 nt (Figure 5.3).

5.5µg of ss-cDNA was then fragmented by uracil-DNA glycosylase (UDG) and apurinic/apyrimidinic endonuclease 1 (APE 1) at the unnatural dUDP residues. The fragmented cDNA was then labelled by terminal deoxynucleotidyl transferase (TdT) using the Affymetrix® proprietary DNA Labelling Reagent that is covalently liked to

biotin. A gel-shift assay was used to assess the biotin labelling of the cDNA, and showed all samples had been labelled.

Table 5.2 Yield of cRNA post-amplification.

Sample	cRNA yield (μg)	A260/280
0_0_A	97.27	1.85
0_0_B	30.87	2.2
0_0_C	29.64	2.02
0_1_A	91.17	1.96
0_1_B	33.21	2.18
0_1_C	38.47	2.16
0_5_A	89.91	1.96
0_5_B	67.93	2.13
0_5_C	45.09	2.17
10_0_A	70.14	2.1
10_0_B	60.26	2.15
10_0_C	65.15	2.14
10_5_A	90.39	1.95
10_5_B	58.43	2.14
10_5_C	51.87	2.17
51_0_A	93.88	1.88
51_0_B	47.48	2.16
51_0_C	35.1	2.16
51_5_A	80.38	2.07
51_5_B	48.62	2.16
51_5_C	51.37	2.14
102_0_A	82.49	2.07
102_0_B	56.46	2.17
102_0_C	53.16	2.15
102_1_A	96.2	1.87
102_1_B	50.71	2.17
102_1_C	46.36	2.16
102_5_A	82.15	2.06
102_5_B	82.56	2.06
102_5_C	38.79	2.15
Mean ± SD	**62.18 ± 21.67**	**2.09 ± 0.10**

Table 5.3 Yield of ss-cDNA.

Sample	cDNA yield (µg)	A260/280
0_0_A	28.53	2
0_0_B	22.69	2.06
0_0_C	31.72	2.1
0_1_A	24.53	2.08
0_1_B	22.2	1.79
0_1_C	22.9	2.08
0_5_A	30.3	2.08
0_5_B	25.86	2.1
0_5_C	24.67	2.11
10_0_A	25.43	2.08
10_0_B	23.31	2.1
10_0_C	26.71	2.09
10_5_A	26.35	2.1
10_5_B	23.04	2.11
10_5_C	21.98	2.09
51_0_A	26.55	2.08
51_0_B	24.95	2.09
51_0_C	23.27	2.09
51_5_A	23.11	2.11
51_5_B	21.29	1.74
51_5_C	22.79	2.08
102_0_A	23.65	2.11
102_0_B	21.19	2.13
102_0_C	22.65	2.11
102_1_A	25.59	2.09
102_1_B	22.13	2.12
102_1_C	21.11	2.1
102_5_A	24.43	2.1
102_5_B	24.18	2.12
102_5_C	21.21	2.09
Mean	**24.28 ± 2.61**	**2.07 ± 0.087**

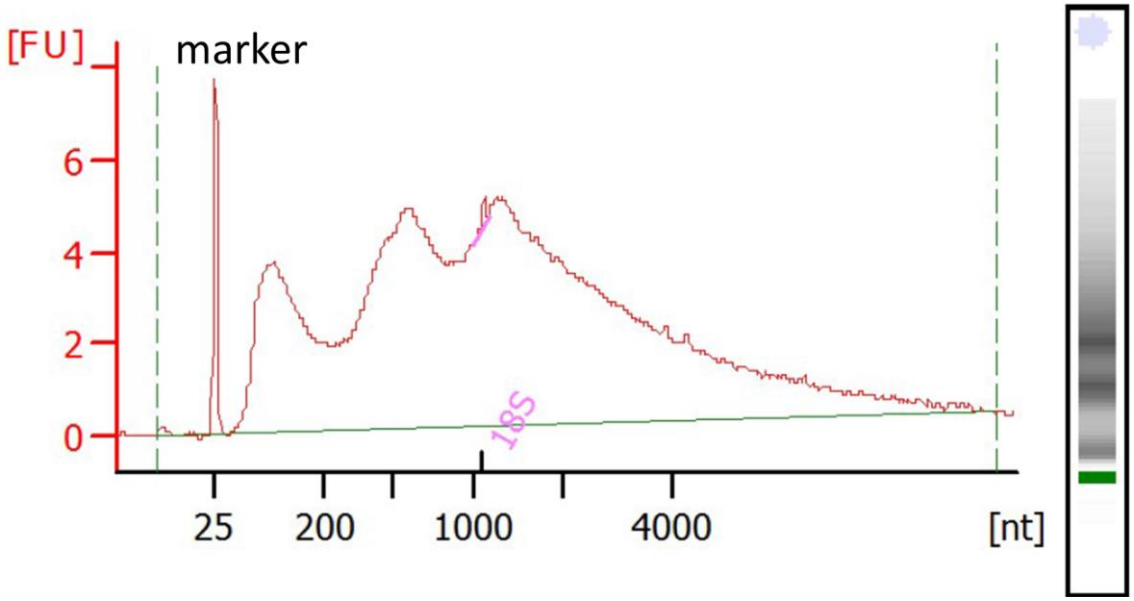


Figure 5.2 Agilent 2100 Bioanalyzer assessment of cRNA post-amplification. An example electropherogram and corresponding electrophoresis gel produced by the Agilent 2100 Bioanalyzer used to assess the cRNA amplification. The electropherogram from sample 0_0_A is used here as a representative example to illustrate the RNA amplification to cRNA.

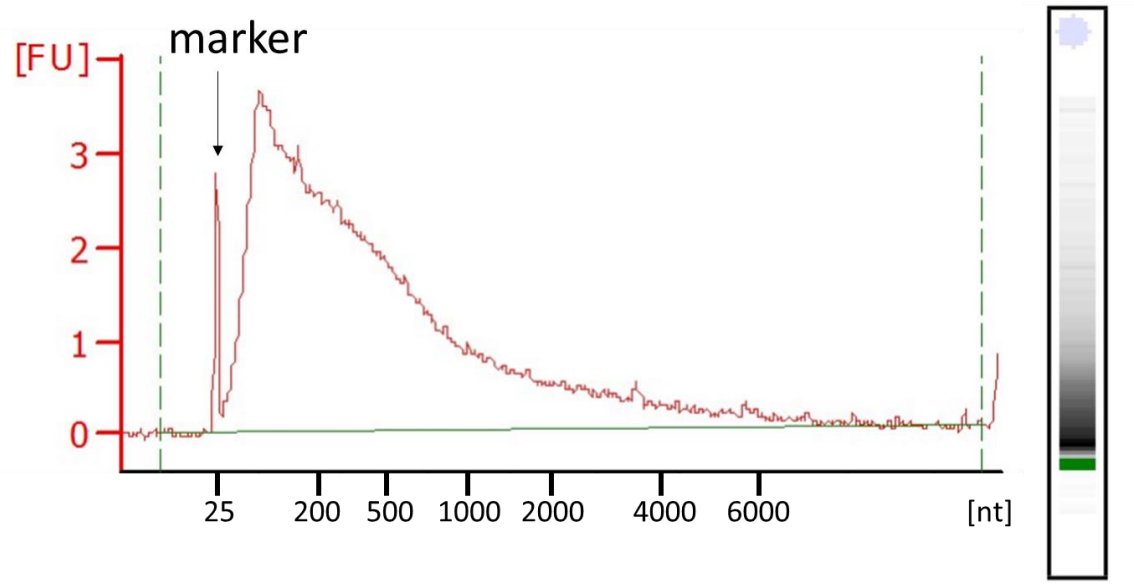


Figure 5.3 Agilent 2100 Bioanalyzer assessment of second-cycle ss-cDNA post-amplification. An example electropherogram and corresponding electrophoresis gel produced by the Agilent 2100 Bioanalyzer used to assess the ss-cDNA amplification. The 0_0_A sample is used here as a representative example to illustrate the amplification of the cRNA to ss-cDNA samples.

5.3.2.1. Affymetrix® Expression Console™ Quality Control Metrics

To ensure the arrays in the experiment are reliable, there are several quality control metrics that each array must pass. For gene and exon expression arrays, Affymetrix® recommends monitoring of hybridisation, as well as interrogating the positive vs negative area under the curve (pos vs neg AUC), all probe set mean, all probe set Relative Log Expression (RLE) mean, RLE plot, and Percentage Presence Call (%P) metrics, to determine any outlier arrays. Outlier arrays may be excluded from further downstream statistical analysis. As a general rule, arrays that have metric values of more than 2 standard deviations away from the mean of all array values are flagged as potential outliers. Arrays that are consistently flagged as outliers in two or more quality control parameters must be excluded from the study.

5.3.2.1.1. Hybridisation Controls

The Affymetrix arrays include four spike-in controls BioB, BioC, BioD (genes from the biotin synthesis pathway of *E.coli*,) and CreX (the recombinase gene from P1 bacteriophage). The spike in controls were added with the fragmented and labelled ss-cDNA in the hybridisation master mix at increasing concentrations (1.5, 5, 25, and 100pM respectively) prior to the hybridisation step. The spike in controls are not predicted to cross-react with eukaryotic ss-cDNA, and therefore the signal intensities for the spike-in controls should follow the trend BioB<BioC<BioD<CreX. Any other pattern would indicate poor hybridisation. All of the arrays in this experiment show the correct trend (Figure 5.4). In addition, BioB (at 1.5pM) is at the probe detection limit on the GeneChip® MTA 1.0, and the detection of BioB on all of the arrays indicates good overall sensitivity.

5.3.2.1.2. Positive vs Negative Area Under the Curve (AUC)

Positive vs negative area under the curve (AUC) is a robust metric for assessing overall array data quality. The AUC value is a measure of the detection of positive controls against the false detection of negative controls, which effectively means the ability to distinguish the true signal from noise in the array data. Values of 1 indicate perfect distinction between true signal and noise, whilst values of 0.5 indicate no distinction

between positive and negative controls. Arrays with AUC values <0.8 should be flagged as potential outliers. All the arrays in this experiment have an AUC value of ≥0.88 (Table 5.4), indicating that there are no significant differences in array data quality.

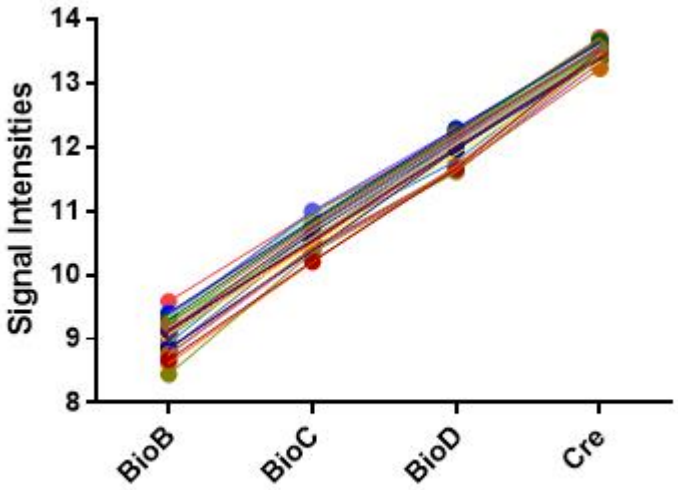


Figure 5.4 Eukaryotic Hybridisation controls for NSC34 arrays. Signal intensities of the eukaryotic hybridisation controls BioB, BioC, BioD, and CreX (at 1.5, 5, 25, 100pM respectively) on NSC34 GeneChip® Mouse Transcriptome Arrays 1.0 using Affymetrix Expression Console.

5.3.2.1.3. All Probe Set Mean

The all probe set mean is the mean of the signal of all the probe sets in the analysis, and allows detection of bright or dim arrays. The average of the all probe set means is 5.82 ± 0.042 for the gene level analysis, and 6.29 ± 0.0089 for the exon level analysis (Table 5.4). For the gene level analysis, the 102_5_A and 102_5_B arrays both have all probe set mean values slightly outside 2 standard deviations of the mean (both at 5.91) indicating potential outliers (Table 5.4). However, for the exon level analysis, all arrays have all probe set mean values within 2 standard deviations of the mean value (Table 5.4).

Table 5.4 Summarised Quality control metrics for NSC34 gene and exon level arrays. The positive vs negative AUC, all probe set mean and all probe set RLE (for both gene level and exon level arrays), and percentage exon level probe present (%P) quality control metrics are displayed for all arrays, with the mean and SD for each metric included. Outlier values are highlighted in yellow.

Sample	Both	Gene level analysis		Exon level analysis		
	Pos vs Neg AUC	All probe set mean	All probe set RLE mean	All probe set mean	All probe set RLE mean	%P
0_0_A	0.92	5.75	0.21	6.30	0.28	64.11
0_0_B	0.89	5.83	0.15	6.29	0.25	77.50
0_0_C	0.91	5.75	0.17	6.31	0.26	69.52
0_1_A	0.90	5.82	0.13	6.29	0.22	73.72
0_1_B	0.90	5.80	0.13	6.30	0.21	74.03
0_1_C	0.90	5.79	0.14	6.29	0.22	70.79
0_5_A	0.90	5.82	0.12	6.28	0.21	74.20
0_5_B	0.90	5.81	0.16	6.28	0.26	78.83
0_5_C	0.89	5.83	0.13	6.29	0.22	75.36
10_0_A	0.92	5.77	0.23	6.29	0.31	61.32
10_0_B	0.89	5.83	0.15	6.29	0.25	74.38
10_0_C	0.89	5.82	0.18	6.28	0.29	76.33
10_5_A	0.90	5.82	0.15	6.28	0.25	73.05
10_5_B	0.91	5.81	0.17	6.29	0.26	70.69
10_5_C	0.90	5.84	0.18	6.29	0.27	70.56
51_0_A	0.90	5.80	0.12	6.29	0.21	71.28
51_0_B	0.90	5.75	0.21	6.30	0.30	74.15
51_0_C	0.88	5.85	0.18	6.29	0.27	75.82
51_5_A	0.91	5.78	0.16	6.29	0.24	67.85
51_5_B	0.91	5.79	0.19	6.30	0.29	64.78
51_5_C	0.92	5.79	0.17	6.29	0.25	67.24
102_0_A	0.89	5.85	0.15	6.29	0.25	76.25
102_0_B	0.89	5.85	0.15	6.29	0.26	72.94
102_0_C	0.89	5.87	0.19	6.29	0.31	73.91
102_1_A	0.90	5.85	0.15	6.28	0.24	71.42
102_1_B	0.89	5.83	0.14	6.28	0.24	77.72
102_1_C	0.89	5.88	0.23	6.31	0.37	66.71
102_5_A	0.88	5.91	0.23	6.27	0.33	75.52
102_5_B	0.88	5.91	0.24	6.29	0.34	72.65
102_5_C	0.89	5.87	0.24	6.30	0.38	66.97
Mean ± SD	**0.90 ± 0.011**	**5.82 ± 0.041**	**0.17 ± 0.037**	**6.29 ± 0.0089**	**0.27 ± 0.045**	**71.99 ± 4.34**

5.3.2.1.4. Relative Log Expression (RLE) Signal

The Relative Log Expression (RLE) signal is derived by comparing the signal of each probe set to the median signal value for that probe set across all arrays in the experiment. The deviation of the RLE for all probe sets on the array can be plotted, and provides a robust method for assessing data quality. Arrays of poorer quality will have a larger spread and should be flagged as potential outliers. For the gene level RLE plot, all arrays have a similar spread (Figure 5.5). Equally, in the exon level RLE plot, all arrays have a similar spread (Figure 5.6). Therefore, the RLE plots do not indicate any obvious outlier arrays.

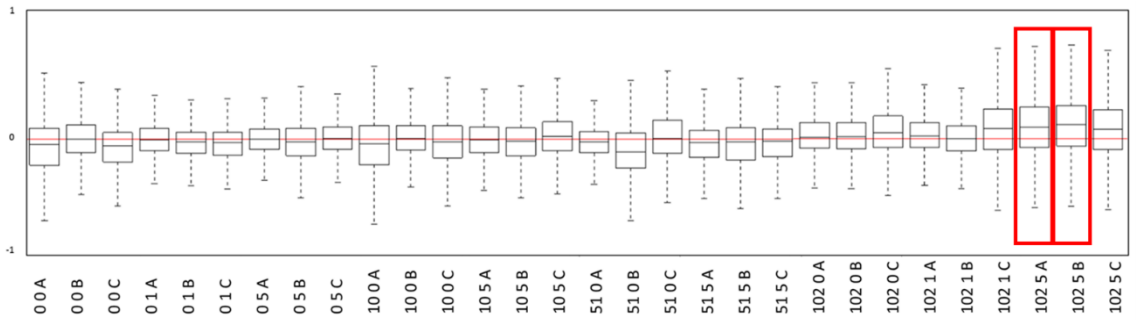


Figure 5.5 Relative log expression (RLE) box plots for GENE level arrays. The distribution of RLE values for the GENE level arrays are displayed as box plots. Potential outliers are highlighted in red boxes.

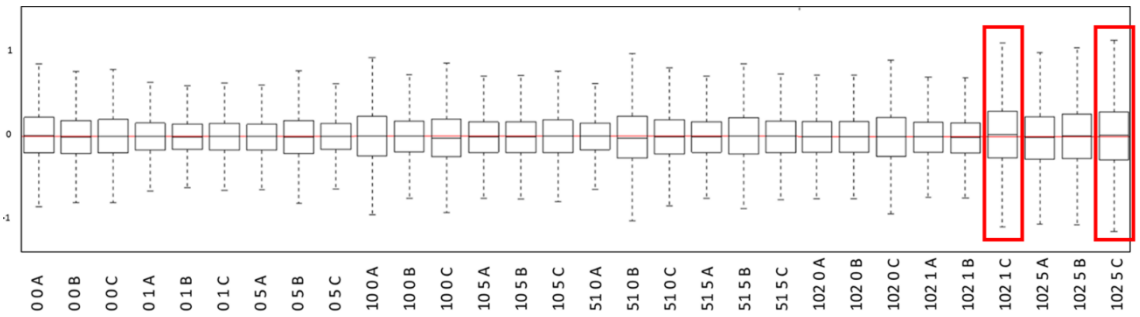


Figure 5.6 Relative log expression (RLE) box plots for EXON level. The distribution of RLE values for the EXON level arrays are displayed as box plots. Potential outliers are highlighted in red boxes.

In addition to the RLE plot, the all probe set RLE mean is the mean average of this deviation. Unusually high values may indicate the signals on the array are very different from the others in the experiment. The average of the all probe set RLE means is 0.17 ± 0.037 for the gene level analysis, and 0.27 ± 0.045 for the exon level analysis (Table 5.4).

For the gene level analysis, all arrays have all probe set RLE mean values within 2 standard deviations of the mean value (Table 5.4). For the exon level analysis, the 102_1_C and 102_5_C arrays both have all probe set RLE mean values slightly outside 2 standard deviations of the mean (0.37 and 0.38 respectively) indicating potential outliers (Table 5.4).

5.3.2.1.5. Percentage of Exon Level Probe Sets Present (%P)

Another metric specifically used to assess the quality of the exon level probe sets is the %P value, which measures the percentage of exon level probe sets detected based on the detection above background (DABG) algorithm. The average %P value for all arrays is 71.99±4.34, and only the 10_0_A array falls outside 2 standard deviations from the mean value at 61.32%P (Table 5.4).

5.3.2.1.6. Quality Control Summary

There are a few arrays with outlier metric values. However, there are no arrays which are consistently flagged up with outlier values in more than one quality control metric. Equally, all outlier values are only very slightly outside the two SD threshold. Therefore, all arrays passed the quality control assessment and were taken forward for statistical analysis.

5.3.3. Gene Level Data Analysis

5.3.3.1. Qlucore Omics Explorer to Define Transcriptomic Changes in NSC34 (G4C2)102 Cells

5.3.3.1.1. Selecting Filtering Stringency to Define Differentially Expressed Transcripts in NSC34 Arrays

The filtering stringency used to define differentially expressed (DE) transcripts between arrays first had to be selected before gene level analysis. When describing which transcripts are significantly DE, it is customary to filter using fold change (FC) and P value (using Student't T-test). FC describes the magnitude of the difference in expression levels, and a higher FC is important for future validation in further experiments. P value describes the statistical significance of the differential expression being real, and is used to determine the number of false positives in the list of DE transcripts. The a priori

criteria for the filtering stringency were to define around 3,000 DE transcripts between the NSC34 102 + 5 d tet and NSC34 0 + 5 d tet arrays, with the lowest possible P value. This is because the DAVID functional enrichment analysis was going to be used downstream to identify enriched biological functions and pathways in the DE transcripts, and DAVID can analyse up to 3,000 transcripts. Also, since the NSC34 (G4C2)n cells were isogenic, there should be very little genetic background noise between the arrays, and a lower P value was desired to minimise the number of false positives.

All Affymetrix® CHP files were opened in Qlucore Omics explorer (version 3.0), and normalised using the RMA-sketch method. Differentially expressed (DE) transcripts were defined for the NSC34 102 + 5 d tet vs NSC34 sham + 5 d tet array comparison using various combinations of filtering stringencies. P values of <0.05, <0.01, and <0.001, and FCs of ≥1.2, ≥ 1.5, and ≥2.0 were used, and the number of DE transcripts defined for the different filtering stringencies are displayed in Table 5.5. For the NSC34 102 + 5 d tet vs NSC34 sham + 5 d tet analysis, a ≥1.2 FC was selected, because ≥1.5 FC did not classify enough transcripts as DE for robust downstream enrichment analysis, even when P<0.05. At FC ≥1.2, 5572, 3069 and 788 transcripts were defined as DE using P values <0.05, 0.01, and 0.001 respectively (Table 5.5). Therefore, filtering criteria FC≥1.2 and P<0.01 were selected, as around 3000 transcripts were classified as DE in the NSC34 102 + 5 d tet vs NSC34 sham + 5 d tet analysis.

Table 5.5 Number of transcripts classified as differentially expressed between NSC34 (G4C2)102 + 5 tet vs NSC34 sham + 5 d tet using various P value and fold change filtering stringencies.

		Fold change		
		1.2	**1.5**	**2**
P value	**0.05**	5572	1028	135
	0.01	3069	835	129
	0.001	788	349	94

5.3.3.1.2. Clustering Analysis of All NSC34 Arrays

Clustering analysis was performed on all NSC34 arrays using multi-group analysis (two-way ANOVA at P<0.01), and displayed as a PCA plot (Figure 5.7A). There are three distinct clusters: one cluster contains all NSC34 sham and NSC34 (G4C2)51 arrays, the

second cluster contains all NSC34 (G4C2)10 arrays, and the last cluster contains all the NSC34 (G4C2)102 arrays. Reassuringly, the NSC34 (G4C2)102 + 5 day tet arrays are the most different of all NSC34 (G4C2)102 arrays along axis 3 of the PCA plot compared to the other arrays, indicating that increasing (G4C2)102 expression further alters gene expression. In addition, the NSC34 (G4C2)102 + 5 day tet arrays were compared to all other NSC34 sham and NSC34 (G4C2)102 arrays using Hierarchical Clustering analysis (Student's T-test at P<0.01, FC≥1.2) and displayed as a heat map (Figure 5.7B). The Hierarchical Clustering analysis showed there is a strong transcriptomic signal that distinguishes the NSC34 (G4C2)102 + 5 d tet arrays from the NSC34 (G4C2)102 at shorter tet induction timepoints and the NSC34 sham arrays (Figure 5.7B).

5.3.3.2. NSC34 Sham + 5 d Tet vs NSC34 Sham + 0 d Tet Direct Array Analysis

Firstly, the effect of tetracycline on the NSC34 cell transcriptome was assessed, because tetracycline had previously shown toxicity to the NSC34 cells, although at a higher concentration (section 3.3.4.4). Even at lower doses the tetracycline may affect gene expression. Therefore, NSC34 sham + 5 d tet arrays were compared to NSC34 sham 0 d tet arrays. There were only 28 DE transcripts (FC≥1.2 and P<0.01), where 10 (35.7%) were down-regulated and 18 (64.3%) were up-regulated. These 28 DE transcripts were compared to the 3089 DE transcripts in the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison in section 5.3.3.3, and there were only 3 transcripts in common.

5.3.3.2.1. DAVID Functional Enrichment Analysis of NSC34 Sham + 5 d Tet vs NSC34 Sham + 0 d Tet

The Entrez Gene IDs of the DE transcripts from the NSC34 sham + 5 d tet vs NSC34 sham + 0 d tet comparison were analysed using the functional annotation tool in DAVID bioinformatics resources 6.7. Total DE transcripts, as well as the down-regulated and up-regulated transcripts from the gene lists were analysed separately. Functional annotation clustering was subsequently performed using the GOTERM_BP_FAT gene ontology and KEGG_PATHWAY terms, applying a *Mus musculus* background, and filtering using medium stringency. Functional clusters with DAVID enrichment scores >1.30 (equivalent to a P value <0.05) were considered statistically significant.

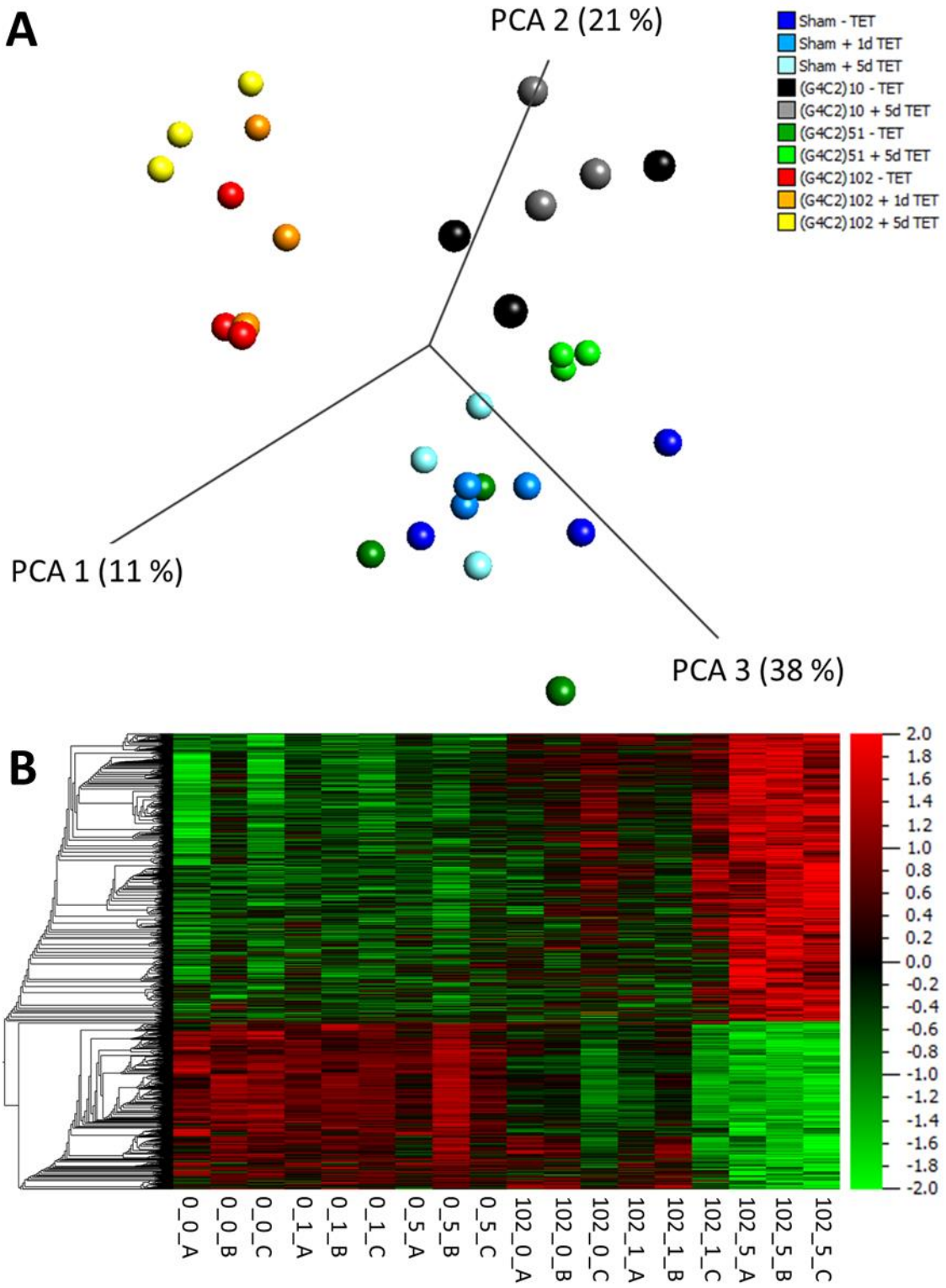


Figure 5.7 Clustering analysis of NSC34 arrays. A) Multiple group analysis PCA plot using two-way ANOVA at P<0.01. **B)** Hierarchical clustering heat map of NSC34 (G4C2)102 + 5 d tet compared to other NSC34 sham and NSC34 (G4C2)102 arrays using Student's T-test at FC≥1.2 and P<0.01.

Olfactory transduction was the only enriched functional annotation cluster, with a DAVID enrichment score of 2.37 in the total DE transcripts and 2.53 in the up-regulated transcripts. There were 8 (28.57%) transcripts in the largest cluster term for total DE transcripts and 7 (38.89%) transcripts in the largest cluster term for up-regulated transcripts. The most significant cluster term was GOTERM 'G-protein coupled receptor protein signalling pathway' for both total DE and up-regulated transcripts (P value = 0.00076; Benjamini value = 0.026; and P value = 0.00015; Benjamini value = 0.0023 respectively). Therefore, any transcriptional changes in genes from the 'olfactory transduction' biological process are likely to be caused by tetracycline, and would be excluded from further NSC34 (G4C2)102 +5 d tet vs NSC34 sham + 5 d tet analysis.

5.3.3.3. NSC34 (G4C2)102 vs NSC34 Sham Direct Array Analysis

The aim of the GEP was to determine which biological pathways are affected by the (G4C2)n expression, and to find any potential therapeutic targets for C9ORF72-ALS. The (G4C2)10 and (G4C2)51 expression caused no observable toxicity in the NSC34 cells, whereas expression of the (G4C2)102 did cause toxicity (section 3.3.4.5). Therefore, the strongest and most relevant transcriptomic changes related to (G4C2)n induced toxicity are most likely to be detected in the tetracycline induced NSC34 (G4C2)102 cells. Therefore the key array comparisons carried out were between the NSC34 (G4C2)102 and NSC34 sham cells, particularly at the 5 day tetracycline induction time point.

NSC34 (G4C2)102 arrays were compared to NSC34 sham arrays at each tetracycline induction time point to generate lists of DE transcripts (FC≥1.2 and P<0.01). In the NSC34 (G4C2)102 + 0 d tet vs NSC34 sham + 0 d tet comparison, there were 1576 DE transcripts, where 1128 (71.6%) were down-regulated, and 448 (28.4%) were up-regulated (Figure 5.8). In the NSC34 (G4C2)102 + 1 d tet vs NSC34 sham + 1 d tet comparison, there were 1725 DE transcripts, where 1146 (66.4%) were down-regulated, and 579 (33.6%) were up-regulated (Figure 5.8). In the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison, there were 3069 DE transcripts, where 1991 (64.9%) were down-regulated, and 1078 (35.1%) were up-regulated (Figure 5.8). Any potential overlap in the DE transcripts at the 0 and 5 day tet timepoints from the NSC34 (G4C2)102 vs NSC34 arrays was compared using a Venn diagram (Figure 5.9). 976 (61.9%) of 1576 DE transcripts in

the NSC34 (G4C2)102 0 d tet vs NSC34 sham 0 d tet comparison were also DE in the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison.

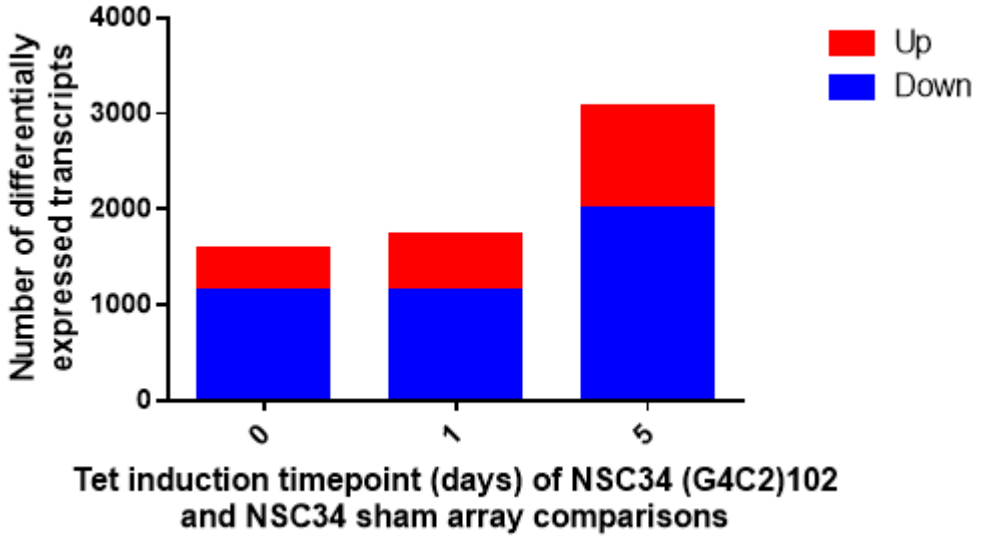


Figure 5.8 The numbers of differentially expressed transcripts with FC≥1.2 at a significance P<0.01 between NSC34 (G4C2)102 and NSC34 sham at 0, 1 and 5 days of tetracycline induction. Up regulated (red) and down regulated (blue) transcripts are displayed as a fraction of the total number of differentially expressed transcripts.

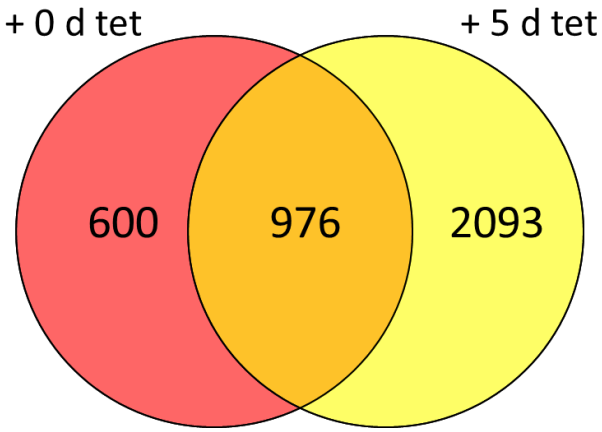


Figure 5.9 Venn diagram comparing the number of differentially expressed transcripts between the NSC34 (G4C2)102 vs NSC34 sham array comparisons at + 0 d and + 5 d tet induction timepoints. Differentially expressed transcripts with FC≥1.2 at a significance P<0.01 from the NSC34 (G4C2)102 + 0 d tet and NSC34 sham + 0 d tet analysis were compared to differentially expressed transcripts from the NSC34 (G4C2)102 + 5 d tet and NSC34 sham + 5 d tet analysis on GeneVenn.

The 28 DE transcripts from the NSC34 sham + 5 d tet vs NSC34 sham + 0 d tet comparsion in section 5.3.3.2 were compared to the 3069 DE transcripts from the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison. Only 3 transcripts (Olfr303, Olfr1396, and Atp5c1-ps) were found in both lists (Figure 5.10), and were removed from further analysis in the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison.

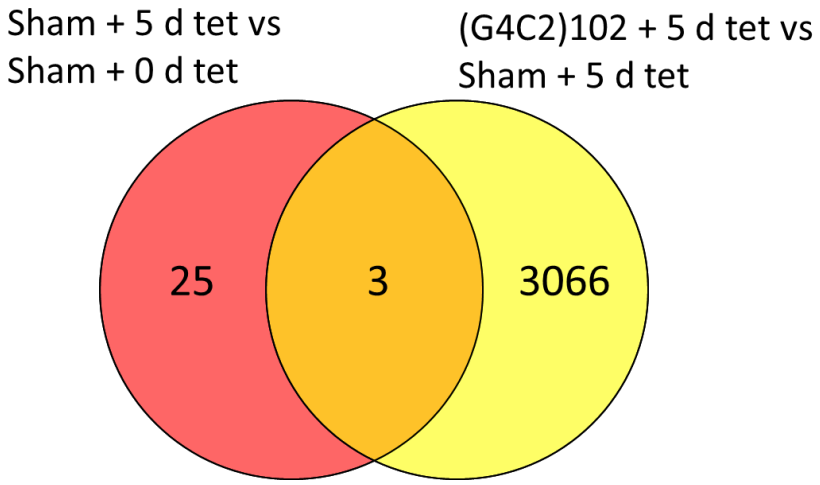


Figure 5.10 Venn diagram comparing the number of shared differentially expressed transcripts in the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison and the NSC34 sham + 5 d tet vs NSC34 sham + 0 d tet comparison. Differentially expressed transcripts with FC≥1.2 at a significance P<0.01 from the NSC34 (G4C2)102 + 5 d tet and NSC34 sham + 5 d tet analysis were compared to differentially expressed transcripts from the NSC34 sham + 5 d tet and NSC34 sham + 0 d tet analysis on GeneVenn.

5.3.3.3.1. DAVID Functional Enrichment Analysis of NSC34 (G4C2)102 + 5 d Tet vs NSC34 Sham + 5 d Tet

The Entrez Gene IDs of the DE transcripts from the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison were imported to the functional annotation tool in DAVID bioinformatics resources 6.7 (as described in section 5.3.3.2.1). The functional enrichment analysis of all 3069 DE transcripts is displayed in Table 5.6, the 1991 down-regulated transcripts only in Table 5.7, and the 1078 up-regulated transcripts only in Table 5.8. Functional clusters that are enriched include protein transport, phosphorylation, cytoskeletal organisation, vesicle-mediated transport, RNA processing, mRNA transport, and GTPase signalling. 63.4% of genes were annotated to

a biological function using the GOTERM_BP_FAT Gene Ontology term, and 26.2% of genes were annotated to a biological pathway using the KEGG_PATHWAY Pathway term.

5.3.3.3.2. IMPaLA Pathway Enrichment Analysis of NSC34 (G4C2)102 + 5 d Tet vs NSC34 Sham + 5 d Tet

The gene symbols of the DE transcripts from the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison were imported to the Integrated Pathway-level Analysis (IMPaLA; http://impala.molgen.mpg.de) and analysed using the pathway over-representation (enrichment) analysis tool (Kamburov et al., 2011). Where there were pathways with high similarity in the IMPaLA output table (for example 'EGF-EGFR signalling pathway' and 'signalling by EGFR'), only the most significantly enriched pathway was retained whilst others were excluded from the list. The pathway enrichment analysis of all 3069 DE transcripts is displayed in Table 5.9, and the 1991 down-regulated transcripts only in Table 5.10. There were no significantly enriched pathways (FDR<0.25) in the 1078 up-regulated transcripts however, and therefore no table is displayed. The most enriched pathways in both the total DE and down-regulated transcripts included EGFR signalling, TOR signalling, insulin signalling, and FOXO signalling. Strikingly, the aforementioned pathways consistently mapped to a subset of pathways in the PI3K/Akt signalling super-pathway (Figure 5.11). In addition, pathways including axon guidance, membrane trafficking, lysosome, developmental biology, protein processing in the endoplasmic reticulum, and regulation of the microtubule cytoskeleton were highly enriched in both total DE and down-regulated transcripts. Processing of capped intron-containing pre-mRNA was also enriched in the down-regulated transcripts. 47.4% of genes were annotated to a biological pathway using the pathway over-representation analysis tool.

Table 5.6 DAVID Functional Annotation Clustering Analysis for 3069 differentially expressed transcripts in NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison with FC≥1.2 and P<0.01.

	Cluster Term	DAVID Enrichment score	Number of genes	% of gene list	Most significant cluster term	P value (most significant term)	Benjamini (most significant)
1	Protein transport	5.31	178	5.99	Protein localisation	7.00E-08	1.30E-04
2	Phosphorylation	4.78	170	5.72	Phosphate metabolic process	1.20E-06	8.80E-04
3	Cytoskeletal organisation	4.34	75	2.52	Cytoskeletal organisation	8.50E-06	5.20E-03
4	Vesicle-mediated transport	3.58	120	4.04	Vesicle-mediated transport	2.40E-08	8.80E-05
5	RNA metabolism	3.51	85	2.86	mRNA metabolic process	8.70E-05	3.20E-02
6	GTPase mediated signalling	2.92	52	1.75	Regulation of small GTPase mediated signal transduction	3.00E-04	6.60E-02
7	Apoptosis	2.71	96	3.23	Cell death	1.60E-03	1.70E-01
8	mRNA transport	2.16	19	0.64	RNA localisation	3.60E-03	2.50E-01
9	Organelle organisation	2.1	34	1.14	regulation of protein complex disassembly	4.20E-04	7.80E-02
10	Glucose metabolism	1.99	40	1.35	Hexose metabolic process	5.10E-03	3.00E-01

Table 5.7 DAVID Functional Annotation Clustering Analysis for 1991 down-regulated transcripts in NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison with FC≥1.2 and P<0.01.

	Cluster Term	DAVID Enrichment score	Number of genes	% of gene list	Most significant cluster term	P value (most significant term)	Benjamini (most significant)
1	Protein transport	8.71	155	7.79	Protein localisation	4.60E-12	1.50E-08
2	RNA metabolism	6.29	80	4.02	mRNA metabolic process	8.30E-08	3.70E-05
3	Phosphorylation	5.25	133	6.69	Phosphate metabolic process	5.60E-07	1.50E-04
4	Vesicle-mediated transport	4.72	104	5.22	Vesicle-mediated transport	2.60E-11	2.10E-08
5	Cytoskeletal organisation	4.71	62	3.12	Cytoskeletal organisation	1.30E-06	3.00E-04
6	Proteolysis	4.47	122	6.13	Ubiquitin-dependent protein catabolic process	1.00E-06	2.50E-04
7	mRNA transport	3.7	19	0.96	RNA localisation	7.90E-05	8.90E-03
8	GTPase mediated signalling	3.23	41	2.06	Regulation of Ras GTPase activity	9.70E-05	9.90E-03
9	Chromosome organisation	2.48	62	3.12	Chromatin modification	8.80E-05	9.60E-03
10	Response to abiotic stimulus	2.38	40	2.01	Response to radiation	9.80E-04	6.30E-02

Table 5.8 DAVID Functional Annotation Clustering Analysis for 1078 up-regulated transcripts in NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison with FC≥1.2 and P<0.01.

	Cluster Term	DAVID Enrichment score	Number of genes	% of gene list	Most significant cluster term	P value (most significant term)	Benjamini (most significant)
1	Immune response	2.79	36	3.65	Immune effector process	5.30E-06	1.10E-02
2	Development	1.71	25	2.54	skeletal system development	5.40E-04	1.70E-01
3	Cell proliferation	1.42	18	1.83	B cell proliferation	9.20E-03	4.60E-01

Table 5.9 IMPaLA Pathway Enrichment Analysis for 3069 differentially expressed transcripts in NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison with FC≥1.2 and P<0.01. Pathways highlighted in grey are part of the PI3K/Akt super-pathway.

	Pathway Name	Pathway Source	Number of DE genes in pathway	Total number of genes in pathway	% overlap	P Value (most significant term)	FDR (most significant term)
1	EGFR1	NetPath	106	447	23.7	1.06E-09	4.12E-06
2	Axon guidance	Reactome	77	310	24.8	2.72E-08	3.54E-05
3	Membrane Trafficking	Reactome	45	153	29.4	1.40E-07	1.20E-04
4	Lysosome - Homo sapiens (human)	KEGG	38	120	31.7	1.61E-07	1.20E-04
5	TOR Signaling	Wikipathways	17	33	51.5	1.81E-07	1.20E-04
6	Developmental Biology	Reactome	93	426	21.8	7.06E-07	3.50E-04
7	Insulin signaling pathway - Homo sapiens (human)	KEGG	39	136	28.7	1.90E-06	8.30E-04
8	Protein processing in endoplasmic reticulum - Homo sapiens (human)	KEGG	44	167	26.3	5.17E-06	1.35E-03
9	Regulation of Microtubule Cytoskeleton	Wikipathways	17	43	39.5	1.74E-05	2.71E-03
10	FoxO signaling pathway - Homo sapiens (human)	KEGG	34	127	26.8	4.16E-05	5.08E-03

Table 5.10 IMPaLA Pathway Enrichment Analysis for 1991 down-regulated transcripts in NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet comparison with FC≥1.2 and P<0.01. Pathways highlighted in grey are part of the PI3K/Akt super-pathway.

	Pathway Name	Pathway Source	Number of DE genes in pathway	Total number of genes in pathway	% overlap	P Value (most significant term)	FDR (most significant term)
1	EGFR1	NetPath	91	447	20.4	3.20E-10	1.23E-06
2	Membrane Trafficking	Reactome	42	153	27.5	3.46E-09	4.51E-06
3	TOR Signalling	Wikipathways	17	33	51.5	5.09E-09	4.97E-06
4	Lysosome - Homo sapiens (human)	KEGG	35	120	29.2	1.28E-08	1.00E-05
5	Protein processing in endoplasmic reticulum - Homo sapiens (human)	KEGG	40	167	24.0	4.65E-07	1.70E-04
6	Processing of capped intron-containing pre-mRNA	Reactome	38	155	24.5	4.80E-07	1.70E-04
7	Axon guidance	Reactome	61	310	19.7	9.88E-07	3.20E-04
8	Regulation of Microtubule Cytoskeleton	Wikipathways	16	43	37.2	3.61E-06	7.80E-04
9	Developmental Biology	Reactome	75	426	17.6	5.00E-06	9.80E-04
10	Insulin signaling pathway - Homo sapiens (human)	KEGG	34	135	21.9	2.51E-05	2.59E-03

5.3.3.4. Biological Pathways of Interest

The IMPaLA enrichment analysis showed various pathways within the PI3K/Akt signalling super-pathway were highly enriched in both the total DE and down-regulated transcripts (Table 5.9). In addition, previous work showed PTEN and the PI3K/Akt signalling pathway are potential therapeutic targets for ALS (Kirby et al., 2011). Therefore, the PI3K/Akt signalling super-pathway was investigated further, with particular interest directed towards genes from the mTOR signalling, MAPK signalling, and insulin signalling pathways.

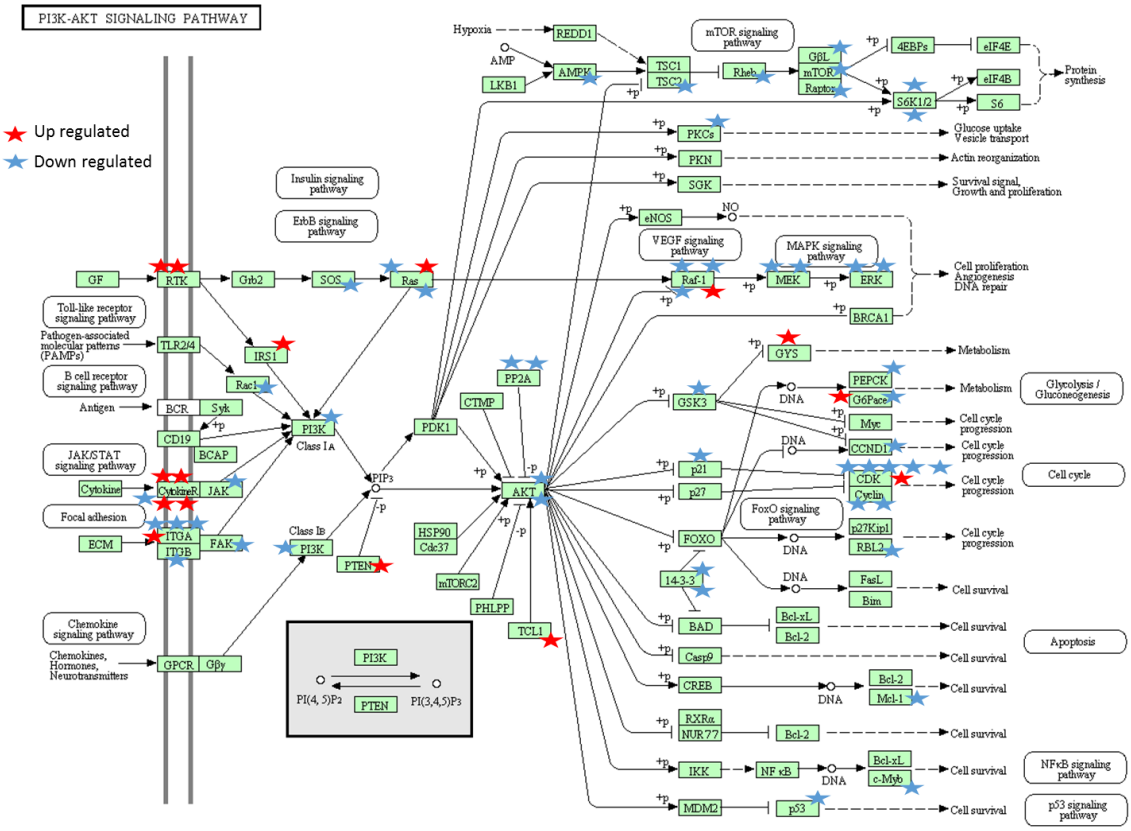


Figure 5.11 PI3K/Akt signalling pathway (KEGG). Square boxes indicate gene products such as proteins; rounded boxes indicate additional pathways. Arrows indicate molecular interactions. The PI3K/Akt signalling pathway regulates many fundamental cellular functions such as proliferation, growth and survival. PI3K phosphorylates PI[4,5]P2 to produce PI[3,4,5]P3, which then activates Akt. Active Akt promotes cell survival. Pten removes a phosphate group from PI[3,4,5]P3 to produce PI[4,5]P2, and therefore inhibits the PI3K/Akt signalling pathway. Transcripts that are DE in NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet are marked with stars (red for up regulation and blue for down regulation in a transcript that encodes the marked protein).

RNA metabolism and mRNA transport functional categories were highly enriched in the DAVID analysis, whilst the processing of capped intron-containing pre-mRNA pathway was also enriched in the down-regulated transcripts. RNA metabolism, and specifically mRNA splicing and mRNA export are hypothesised to be affected in C9ORF72-ALS (section 1.7.2). Therefore, the RNA metabolism and mRNA transport categories were also taken forward for further investigation.

Finally, protein localisation and vesicle-mediated transport were the first and fourth most highly enriched functional categories respectively in the DAVID analysis, and have significant overlap. Equally, membrane trafficking, lysosome, and protein processing in the ER pathways were highly enriched in the IMPaLA pathway enrichment analysis. As mentioned above the DE transcripts in the mTOR signalling pathway are enriched, and the mTOR signalling critically regulates autophagy. Taken together the transcriptomic data suggests there could be a fault in protein homeostasis at the processing and/or degradation stages.

5.3.3.4.1. PI3K/Akt Signalling Pathway

The genes in the 'PI3K/Akt signalling pathway' provided on gene cards (http://pathcards.genecards.org/pathway/29) were used to define the genes in the PI3K/Akt signalling super-pathway, and Table 5.11 lists the 75 of these 345 transcripts in the PI3K/Akt signalling pathway that are DE (FC≥1.2; P<0.01). Where possible, the DE transcripts were mapped onto a diagram of the PI3K/Akt signalling pathway, provided by KEGG (Figure 5.11). The insulin signalling, mTOR signalling, MAPK signalling, and FOXO signalling pathways, as well as the direct regulation of PI3K/Akt contain many DE transcripts (Figure 5.11).

The serine-threonine kinase Akt regulates a diverse set of cellular processes including cell survival, growth, proliferation, metabolism, transcription and protein synthesis (Figure 5.11). There are three highly related isoforms of Akt, of which Akt1 (FC=-1.22; P=0.0024) and Akt2 (FC=-1.21; P=0.00163) are DE in NSC34 (G4C2)102 + 5 d tet. The PI3K/Akt signalling pathway is activated by cell surface receptors. Signalling from the receptors activates phosphatidylinositol 3-kinases (PI3Ks), which phosphorylate phosphatidylinositol-4,5-P2 (PI[4,5]P2) to produce phosphatidylinositol-3,4,5-P3

(PI[3,4,5]P3). Akt is then recruited to the lipid membrane by PI[3,4,5]P3, where Akt is phosphorylated at Thr308 by phosphatidyl-dependent kinase (PDK), and Ser473 by various kinases including mTOR complex 2 (mTORC2) (Sarbassov et al., 2005). The phosphorylation of these two residues is critical for full activation of Akt (Bhaskar and Hay, 2007). Phosphatase and tensin homolog and deleted on chromosome 10 (Pten) however, inhibits the PI3K/Akt signalling pathway by dephosphorylating PI[3,4,5]P3 to produce PI[4,5]P2 (Stambolic et al., 2001) (Figure 5.11). Class I PI3K catalytic subunits PIK3CA (FC=-1.42; P=0.00977) and PIK3CB (FC=-1.32; P=0.00266) are downregulated, whilst Pten (FC=+1.67; P=0.00248) is upregulated in NSC34 (G4C2)102 + 5 d tet.

The mTOR Complexes 1 and 2 are structurally related serine/threonine kinases that are both involved in the PI3K/Akt signalling pathway (Bhaskar and Hay, 2007). Whilst Akt is activated by mTORC2, Akt indirectly activates mTORC1 via TSC1/2 complex inhibition. Active mTORC1 inhibits autophagy and activates protein synthesis. mTORC1 and 2 both contain the catalytic subunit mammalian target of Rapamycin (mTOR) (FC=-1.51; P=0.000155), and mTOR associated protein, LST8 homolog (Mlst8) (FC=-1.47; P=0.000496) subunit. mTORC1 additionally contains regulatory associated protein of mTOR, complex 1 (Rptor) (FC=-1.27; P=0.00396), whilst mTORC2 contains Rptor independent companion of mTOR, complex 2 (Rictor) and Mitogen-Activated protein kinase associated protein 1 (Mapkap1) (FC=-1.35; P=0.00257).

Akt1, mTOR, and Pten were taken forward for qRT-PCR validation. There was less Akt1 in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells at 0 (FC=-1.41; P=0.180), 1 (FC=-1.11; P=0.928) and 5 (FC=-1.79; P=0.117) days + tet, although none of these changes were significant (Figure 5.12). There was significantly less mTOR in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells at either 0 (FC=-3.29; P=0.0139), 1 (FC=-3.56; P=0.00382), or 5 (FC=-4.09; P = 0.0309) days + tet (Figure 5.13). Finally, there was more Pten in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells at 0 (FC=+1.20; P=0.267), 1 (FC=+1.42; P=0.264), and 5 (FC=+1.29; P=0.529) days + tet, although none of these changes were significant (Figure 5.14).

Table 5.11 Transcripts from the PI3K/Akt signalling pathway dysregulated in NSC34 (G4C2)102 + 5 d tet cells. Genes involved in PI3K/Akt signalling pathway which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells. Transcripts highlighted in grey were taken forward for qRT-PCR validation.

Gene Symbol	Gene Name	Fold Change	P-value
PI3K/Akt signalling pathway			
Ago1	argonaute RISC catalytic subunit 1	-1.33	3.42E-03
Akt1	thymoma viral proto-oncogene 1	-1.22	2.40E-03
Akt2	thymoma viral proto-oncogene 2	-1.21	1.63E-03
Calm1	calmodulin 1	-1.92	1.00E-03
Ccnd1	cyclin D1	-1.48	5.36E-05
Ccne2	cyclin E2	-1.66	7.50E-03
Cdkn1a/P21	cyclin-dependent kinase inhibitor 1A	-1.65	7.07E-03
Col4a1	collagen, type IV, alpha 1	+1.24	8.99E-03
Col6a3	collagen, type VI, alpha 3	-1.38	3.64E-04
Cxcl12	chemokine (C-X-C motif) ligand 12	+1.26	5.37E-04
Egfr	epidermal growth factor receptor	+1.20	2.46E-04
Fgf6	fibroblast growth factor 6	+1.25	6.81E-03
Fyn	Fyn proto-oncogene	-1.62	4.11E-04
G6pc	glucose-6-phosphatase, catalytic	+1.22	9.81E-03
G6pc3	glucose 6 phosphatase, catalytic, 3	-1.55	3.65E-04
Ghr	growth hormone receptor	+1.28	9.71E-03
Gnb1	guanine nucleotide binding protein (G protein), beta 1	-1.64	4.04E-03
Gnb4	guanine nucleotide binding protein (G protein), beta 4	-1.35	3.21E-03
Gng5	guanine nucleotide binding protein (G protein), gamma 5	-1.38	1.34E-03
Gsk3a	glycogen synthase kinase 3 alpha	-1.54	6.44E-03
Gys2	glycogen synthase 2	+1.23	2.52E-03
Igf1r	insulin-like growth factor I receptor	-1.45	8.04E-04
Il2ra	interleukin 2 receptor, alpha chain	+1.23	3.83E-03
Il6	interleukin 6	-2.23	4.70E-03
Il7r	interleukin 7 receptor	+1.26	7.45E-03
Irs2	insulin receptor substrate 2	+1.98	4.36E-03
Itch	itchy, E3 ubiquitin protein ligase	-1.57	7.04E-03
Itga1	integrin alpha 1	+1.34	1.09E-03
Itga3	integrin alpha 3	-1.29	2.42E-04
Itga5	integrin alpha 5	-1.48	1.42E-05
Itga8	integrin alpha 8	-1.96	7.99E-04
Itgb5	integrin beta 5	-1.30	2.79E-03
Jak1	Janus kinase 1	-1.71	1.55E-03
Lama4	laminin, alpha 4	+1.25	5.13E-04
Lamc1	laminin, gamma 1	-1.32	5.79E-03
Lcp2	lymphocyte cytosolic protein 2	+1.26	1.60E-03
Lpar3	lysophosphatidic acid receptor 3	+1.22	3.07E-03
Lpar5	lysophosphatidic acid receptor 5	+1.28	7.72E-04
Mapk8	mitogen-activated protein kinase 8	-1.54	1.66E-03
Map2k1	mitogen-activated protein kinase kinase 1	-1.34	3.22E-03

Map3k7	mitogen-activated protein kinase kinase kinase 7	-1.53	2.86E-03
Mapkap1	mitogen-activated protein kinase associated protein 1	-1.35	2.57E-03
Mcl1	myeloid cell leukemia sequence 1	-1.61	7.59E-05
Mlst8	MTOR associated protein, LST8 homolog (S. cerevisiae)	-1.47	4.96E-04
Mtor	mechanistic target of rapamycin	-1.51	1.55E-04
Myb	myeloblastosis oncogene	-1.65	2.82E-03
Ncstn	Nicastrin	-1.55	4.00E-03
Nedd4	neural precursor cell expressed, developmentally down-regulated 4	-1.78	5.36E-03
Nras	neuroblastoma ras oncogene	-1.68	1.68E-03
Pck2	phosphoenolpyruvate carboxykinase 2	-1.77	9.51E-04
Pdgfra	platelet derived growth factor receptor, alpha polypeptide	+1.31	3.36E-03
Pik3ca	phosphatidylinositol 3-kinase, catalytic, alpha polypeptide	-1.42	9.77E-03
Pik3cb	phosphatidylinositol 3-kinase, catalytic, beta polypeptide	-1.32	2.66E-03
Plcg1	phospholipase C, gamma 1	-1.52	1.63E-05
Ppp2r2d	protein phosphatase 2, regulatory subunit B, delta isoform	-1.35	1.57E-03
Ppp2r3c	protein phosphatase 2, regulatory subunit B'', gamma	-1.47	1.41E-03
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform	-1.40	7.03E-03
Prkca	protein kinase C, alpha	-1.24	4.26E-03
Psenen	presenilin enhancer 2 homolog	-1.49	5.23E-03
Pten	phosphatase and tensin homolog	+1.67	2.48E-03
Ptk2	protein tyrosine kinase 2	-1.45	9.11E-03
Rac1	RAS-related C3 botulinum substrate 1	-1.40	3.54E-03
Rbl2	retinoblastoma-like 2	-1.37	5.40E-03
Rheb	Ras homolog enriched in brain	-1.82	7.24E-04
Rhoa	ras homolog gene family, member A	-1.74	4.92E-03
Rptor	regulatory associated protein of MTOR, complex 1	-1.27	3.96E-03
Sos1	son of sevenless homolog 1 (*Drosophila*)	-1.54	4.12E-03
Stat3	signal transducer and activator of transcription 3	-1.22	1.06E-03
Tmem189	transmembrane protein 189	-1.24	6.41E-03
Tnrc6a	trinucleotide repeat containing 6a	-1.28	7.44E-03
Trp53	transformation related protein 53	-1.66	5.73E-03
Tsc2	tuberous sclerosis 2	-1.44	2.07E-04
Yes1	Yamaguchi sarcoma viral (v-yes) oncogene homolog 1	-1.45	9.89E-03
Ywhab	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	-1.51	3.67E-03
Ywhah	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	-1.38	1.03E-03

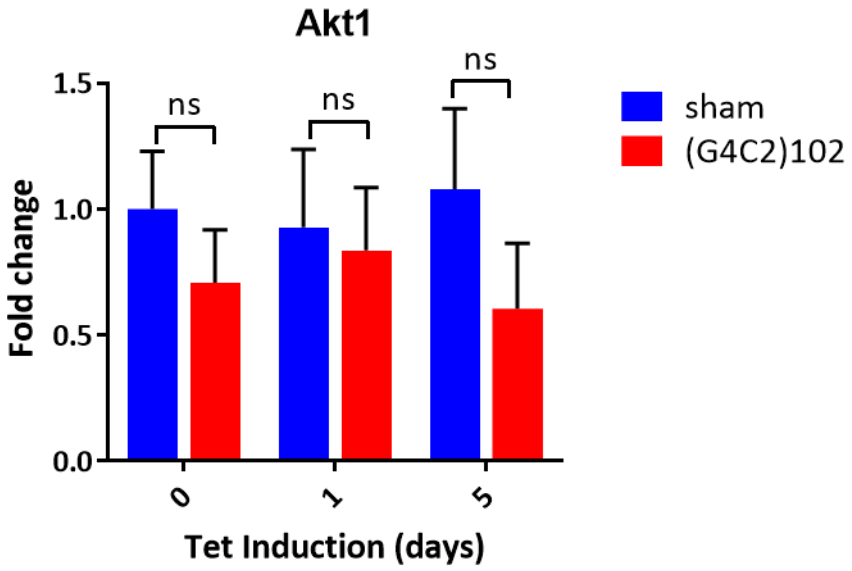


Figure 5.12 qRT-PCR Validation of Akt1 levels. NSC34 sham and NSC34 (G4C2)102 cells were grown for 0, 1, or 5 days with 0.5µg/mL tetracycline. qRT-PCR was performed on Akt1, and normalised to GAPDH. (Multiple t-tests; Data are means ± SD; n=3).

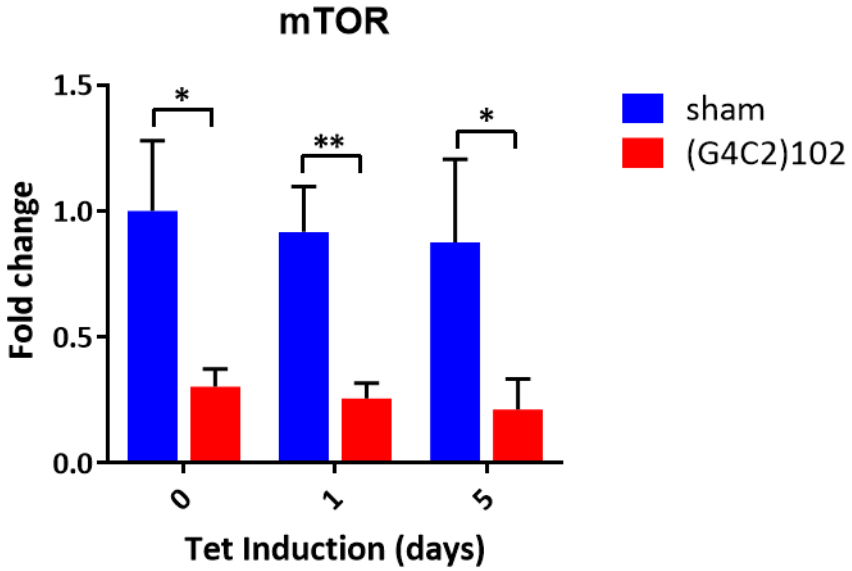


Figure 5.13 qRT-PCR Validation of mTOR levels. NSC34 sham and NSC34 (G4C2)102 cells were grown for 0, 1, or 5 days with 0.5µg/mL tetracycline. qRT-PCR was performed on mTOR, and normalised to GAPDH. (*P<0.05; Multiple t-tests; Data are means ± SD; n=3).

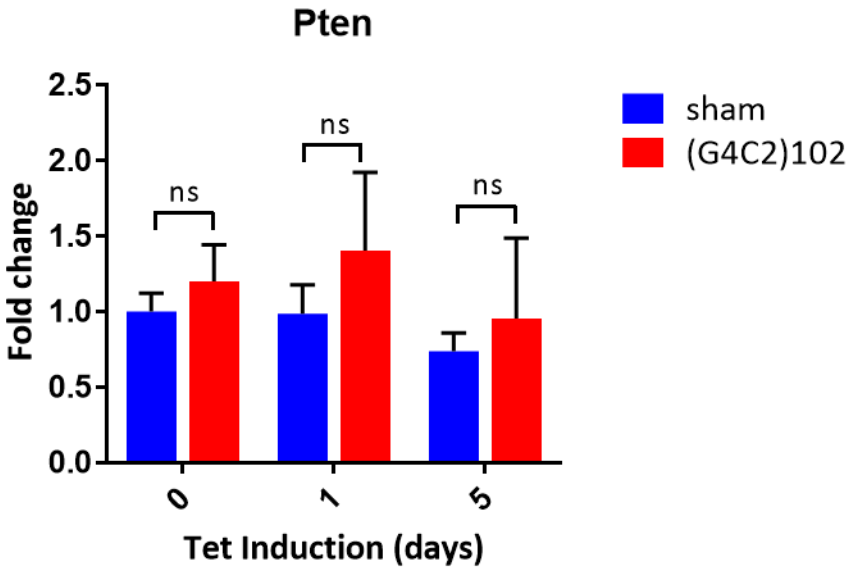


Figure 5.14 qRT-PCR Validation of Pten levels. NSC34 sham and NSC34 (G4C2)102 cells were grown for 0, 1, or 5 days with 0.5µg/mL tetracycline. qRT-PCR was performed on Pten, and normalised to GAPDH. (Multiple t-tests; Data are means ± SD; n=3).

Finally, our group had previously performed GEP on laser captured microdissected (LCM) motor neurons from the spinal cord of C9ORF72-ALS patients (Cooper-Knock et al., 2015a). Dr Johnathan Cooper-Knock investigated the PI3K/Akt signalling pathway (KEGG) in this dataset, and careful statistical analysis revealed that this pathway was significantly dysregulated in LCM motor neurons from C9ORF72-ALS patients compared to controls (rank-product, P=0.01). Further, *PTEN* had the highest fold change of all DE transcripts (FC=+11.3, P=0.00001) within this dataset. Transcripts in the PI3K/Akt signalling pathway (KEGG) are listed with transcript ID, gene ID, fold change and P value for the C9ORF72-ALS LCM motor neurons in Appendix 5.

5.3.3.4.2. RNA Metabolism and mRNA Transport

Aberrant RNA metabolism is described in a wide range of neurodegenerative diseases including ALS (sections 1.5.5 and 1.6.5). However, RNA metabolism includes a diverse set of biological processes. There were 111 DE transcripts (FC≥1.2; P<0.01) related to RNA metabolism and/or mRNA transport in the NSC34 (G4C2)102 cells. Transcripts annotated for RNA metabolism only (Table 5.12), mRNA transport only (Table 5.13), and both RNA metabolism and mRNA transport (Table 5.14) were listed separately. The 111

DE transcripts were functionally annotated again using the DAVID functional annotation tool (section 5.3.3.2.1) and IMPaLA pathway enrichment analysis (section 5.3.3.3.2) to identify which RNA metabolic processes and/or pathways are mainly affected by the (G4C2)102 expression. The top three annotated sub-categories were RNA splicing and/or spliceosome (52), ncRNA processing (23), and mRNA transport (18) using the DAVID annotation tool. These 111 DE transcripts were also enriched for the Spliceosome (30), Processing of Capped Intron-containing Pre-mRNA (26), mRNA Splicing (19), and RNA Transport (15) using the IMPaLA pathway enrichment analysis.

There are several splicing factors that bind (G4C2)n RNA (Cooper-Knock et al., 2014b) that are downregulated in the NSC34 (G4C2)102 + 5 d tet cells, including serine/arginine-rich splicing factor 1 (SRSF1) (FC=-1.54; P=0.00299), SRSF2 (FC=-1.68; P=0.00488), heterogeneous nuclear ribonucleoprotein H1 (HnRNPH1) (FC=-1.29; P=0.00758), nudix (nucleoside diphosphate linked moiety X)-type motif 21 (Nudt21) (FC=-1.27; P=0.00467), SRSF3 (FC=-1.42; P=0.00483) and SRSF6 (FC=-1.79; P=0.000240). Other factors involved in splicing regulation that bind (G4C2)n RNA (Cooper-Knock et al., 2014b) that are also downregulated in the NSC34 (G4C2)102 + 5 d tet cells include RNA binding motif protein 3 (RBM3) (FC=-1.74; P=0.00127), and serine/arginine-rich protein specific kinase 2 (SRPK2) (FC=-1.31; P=0.00603). SRSF1 and SRSF2 were taken forward for qRT-PCR validation. There was significantly less SRSF1 in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells at 0 (FC=-1.99; P=0.00363), 1 (FC=-1.78; P=0.0390) and 5 (FC=-2.28; P=0.00442) days + tet (Figure 5.15). There was significantly less SRSF2 in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells at 0 (FC=-2.63; P=0.0175), 1 (FC=-1.95; P=0.0344) and 5 (FC=-3.60; P=0.000395) days + tet (Figure 5.16).

Table 5.12 RNA Metabolism transcripts dysregulated in NSC34 (G4C2)102 + 5 d tet cells. Genes involved in RNA Metabolism which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells. Genes that are also functionally annotated as mRNA Transport in addition to RNA Metabolism are listed in **Table 5.14**. Transcripts highlighted in grey were taken forward for qRT-PCR validation.

Gene Symbol	Gene Title	Fold change	p-value
RNA Metabolism			
9530036M11Rik	RIKEN cDNA 9530036M11 gene	+1.30	6.17E-03
Adar	adenosine deaminase, RNA-specific	-1.28	1.62E-03
Adat1	adenosine deaminase, tRNA-specific 1	-1.22	6.79E-03
Ago1	argonaute RISC catalytic subunit 1	-1.33	3.42E-03
Aplp1	amyloid beta (A4) precursor-like protein 1	-1.53	6.54E-03
App	amyloid beta (A4) precursor protein	-1.31	3.77E-04
Aqr	Aquarius	-1.60	2.26E-03
Auh	AU RNA binding protein/enoyl-coenzyme A hydratase	-1.37	5.61E-05
Celf1	CUGBP, Elav-like family member 1	-1.24	1.74E-03
Celf3	CUGBP, Elav-like family member 3	-1.45	8.16E-03
Cherp	calcium homeostasis endoplasmic reticulum protein	-1.39	4.86E-03
Cpsf3l	cleavage and polyadenylation specific factor 3-like	-1.45	9.55E-04
Crnkl1	Crn, crooked neck-like 1 (*Drosophila*)	-1.43	7.02E-03
Cstf1	cleavage stimulation factor, 3' pre-RNA, subunit 1	-1.31	5.44E-04
Ctnnbl1	catenin, beta like 1	-1.33	3.79E-03
Ddx20	DEAD (Asp-Glu-Ala-Asp) box polypeptide 20	-1.37	7.27E-03
Ddx46	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46	-1.66	2.58E-03
Ddx5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	-1.24	9.92E-03
Ddx51	DEAD (Asp-Glu-Ala-Asp) box polypeptide 51	-1.25	8.60E-03
Dgcr8	DiGeorge syndrome critical region gene 8	-1.35	3.89E-03
Dhx15	DEAH (Asp-Glu-Ala-His) box polypeptide 15	-1.35	5.94E-03
Eftud2	elongation factor Tu GTP binding domain containing 2	+1.23	6.30E-03
Exosc10	exosome component 10	-1.67	6.31E-04
Hnrnph1	heterogeneous nuclear ribonucleoprotein H1	-1.29	7.58E-03
Hnrnpk	heterogeneous nuclear ribonucleoprotein K	-1.33	4.75E-03
Hnrnpll	heterogeneous nuclear ribonucleoprotein L-like	-1.34	4.89E-03
Ints3	integrator complex subunit 3	-1.57	1.89E-04
Ints6	integrator complex subunit 6	-1.60	3.74E-03
Ints7	integrator complex subunit 7	-1.35	5.68E-03
Ints8	integrator complex subunit 8	-1.65	2.68E-03
Mapkapk2	MAP kinase-activated protein kinase 2	-1.30	5.22E-03
Nol3	nucleolar protein 3 (apoptosis repressor with CARD domain)	-1.32	7.97E-04
Nono	non-POU-domain-containing, octamer binding protein	-1.51	6.77E-03
Nop14	NOP14 nucleolar protein homolog (yeast)	-1.73	2.29E-03
Nudt21	nudix (nucleoside diphosphate linked moiety X)-type motif 21	-1.27	4.67E-03
Plrg1	pleiotropic regulator 1, PRL1 homolog (Arabidopsis)	-1.37	9.40E-04
Pnrc2	proline-rich nuclear receptor coactivator 2	-2.43	3.28E-04

Pop4	processing of precursor 4, ribonuclease P/MRP family, (S. cerevisiae)	-2.64	6.98E-03
Ppp1r8	protein phosphatase 1, regulatory (inhibitor) subunit 8	-1.65	4.47E-03
Ppp4r2	protein phosphatase 4, regulatory subunit 2	-1.27	9.43E-03
Pqbp1	polyglutamine binding protein 1	-1.62	4.91E-03
Prkra	protein kinase, interferon inducible double stranded RNA dependent activator	-1.36	4.08E-03
Prmt5	protein arginine N-methyltransferase 5	-2.13	1.65E-03
Prpf18	PRP18 pre-mRNA processing factor 18 homolog (yeast)	-1.31	9.07E-03
Prpf31	PRP31 pre-mRNA processing factor 31 homolog (yeast)	-1.37	1.80E-03
Prpf40b	PRP40 pre-mRNA processing factor 40 homolog B (yeast)	+1.23	1.94E-03
Prpf8	pre-mRNA processing factor 8	-1.41	5.92E-03
Ptbp1	polypyrimidine tract binding protein 1	-1.35	5.54E-03
Ptbp2	polypyrimidine tract binding protein 2	-1.50	7.93E-03
Pus10	pseudouridylate synthase 10	-1.32	7.36E-03
Qtrtd1	queuine tRNA-ribosyltransferase domain containing 1	-1.33	2.32E-03
Rbm22	RNA binding motif protein 22	-1.38	6.71E-03
Rbm25	RNA binding motif protein 25	-1.34	6.14E-03
Rbm3	RNA binding motif protein 3	-1.74	1.27E-03
Rbm5	RNA binding motif protein 5	-1.27	2.80E-03
Rbmx	RNA binding motif protein, X chromosome	-1.42	1.85E-03
Rnasel	ribonuclease L (2', 5'-oligoisoadenylate synthetase-dependent)	+1.31	7.61E-04
Rnps1	ribonucleic acid binding protein S1	-1.44	9.39E-03
Rpf1	ribosome production factor 1 homolog (S. cerevisiae)	-1.45	8.11E-03
Rrp1b	ribosomal RNA processing 1 homolog B (S. cerevisiae)	-1.21	2.48E-03
Scgb1a1	secretoglobin, family 1A, member 1	+1.34	5.94E-03
Sf3a3	splicing factor 3a, subunit 3	-1.36	5.85E-03
Sf3b3	splicing factor 3b, subunit 3	-1.86	6.92E-05
Sfswap	splicing factor, suppressor of white-apricot homolog (*Drosophila*)	-1.28	1.22E-03
Slbp	stem-loop binding protein	-1.23	7.18E-03
Snrnp200	small nuclear ribonucleoprotein 200 (U5)	-1.73	5.32E-04
Snrnp25	small nuclear ribonucleoprotein 25 (U11/U12)	+1.21	1.01E-03
Snrpb	small nuclear ribonucleoprotein B	-1.29	2.42E-03
Srpk1	serine/arginine-rich protein specific kinase 1	-1.39	3.98E-04
Srpk2	serine/arginine-rich protein specific kinase 2	-1.31	6.03E-03
Srrm1	serine/arginine repetitive matrix 1	-1.54	3.08E-03
Srsf1/SF2	serine/arginine-rich splicing factor 1 (SF2)	-1.54	2.99E-03
Srsf2/SC35	serine/arginine-rich splicing factor 2 (SC35)	-1.68	4.88E-03
Srsf3	serine/arginine-rich splicing factor 3	-1.42	4.83E-03
Srsf6	serine/arginine-rich splicing factor 6	-1.79	2.40E-04
Ssu72	Ssu72 RNA polymerase II CTD phosphatase homolog (yeast)	-1.59	1.59E-03
Syf2	SYF2 homolog, RNA splicing factor (S. cerevisiae)	-1.42	7.23E-03

Syncrip	synaptotagmin binding, cytoplasmic RNA interacting protein	-1.34	6.48E-03
Tcp11	t-complex protein 11	+1.29	6.00E-03
Tfb2m	transcription factor B2, mitochondrial	-1.95	3.72E-03
Trmt44	tRNA methyltransferase 44	-1.26	5.11E-03
Tyw1	tRNA-yW synthesizing protein 1 homolog (S. cerevisiae)	-1.33	2.40E-03
U2af2	U2 small nuclear ribonucleoprotein auxiliary factor (U2AF) 2	-1.37	8.54E-05
U2surp	U2 snRNP-associated SURP domain containing	-1.61	7.29E-03
Upf1	UPF1 regulator of nonsense transcripts homolog (yeast)	-1.45	9.83E-04
Wdr3	WD repeat domain 3	-1.32	2.96E-04
Wdr55	WD repeat domain 55	-1.71	4.36E-03
Zcchc11	zinc finger, CCHC domain containing 11	-1.68	5.15E-03
Zcchc6	zinc finger, CCHC domain containing 6	-1.32	8.14E-03
Zcrb1	zinc finger CCHC-type and RNA binding motif 1	-1.37	5.17E-03
Zfc3h1	zinc finger, C3H1-type containing	+1.28	6.57E-03
Zfp36l2	zinc finger protein 36, C3H type-like 2	-1.86	2.19E-03

Table 5.13 mRNA Transport transcripts dysregulated in NSC34 (G4C2)102 + 5 d tet cells. Genes involved in mRNA Transport which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells. Genes that are also functionally annotated as RNA Metabolism in addition to mRNA Export are listed in **Table 5.14**.

Gene Symbol	Gene Title	Fold change	p-value
mRNA Transport			
Eif5a	eukaryotic translation initiation factor 5A	-1.26	5.90E-03
Eny2	enhancer of yellow 2 homolog (*Drosophila*)	-1.46	8.00E-03
G3bp2	GTPase activating protein (SH3 domain) binding protein 2	-1.58	1.71E-03
Gle1	GLE1 RNA export mediator (yeast)	-1.20	7.78E-03
Nup133	nucleoporin 133	-1.50	3.38E-03
Nup160	nucleoporin 160	-1.54	7.85E-03
Nup214	nucleoporin 214	-1.53	9.19E-04
Nup35	nucleoporin 35	-1.43	8.62E-03
Nup93	nucleoporin 93	-1.54	4.35E-04
Nupl1	nucleoporin like 1	-1.31	7.31E-03
Nupl2	nucleoporin like 2	-1.43	1.93E-03
Stau1	staufen (RNA binding protein) homolog 1 (*Drosophila*)	-1.38	1.14E-03
Xpo7	exportin 7	-1.44	3.79E-04

Table 5.14 RNA Metabolism and mRNA Transport transcripts dysregulated in NSC34 (G4C2)102 + 5 d tet cells. Genes involved in both RNA Metabolism and mRNA Transport which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells.

Gene Symbol	Gene Title	Fold change	p-value
RNA Metabolism and mRNA Transport			
Ddx39b	DEAD (Asp-Glu-Ala-Asp) box polypeptide 39B	-1.52	4.75E-05
Eif4a3	eukaryotic translation initiation factor 4A3	-1.35	4.79E-03
Khsrp	KH-type splicing regulatory protein	-1.30	1.10E-03
Qk	Quaking	-1.34	8.58E-03
Srsf10	serine/arginine-rich splicing factor 10	-1.54	3.33E-03
Thoc1	THO complex 1	-1.80	3.31E-03

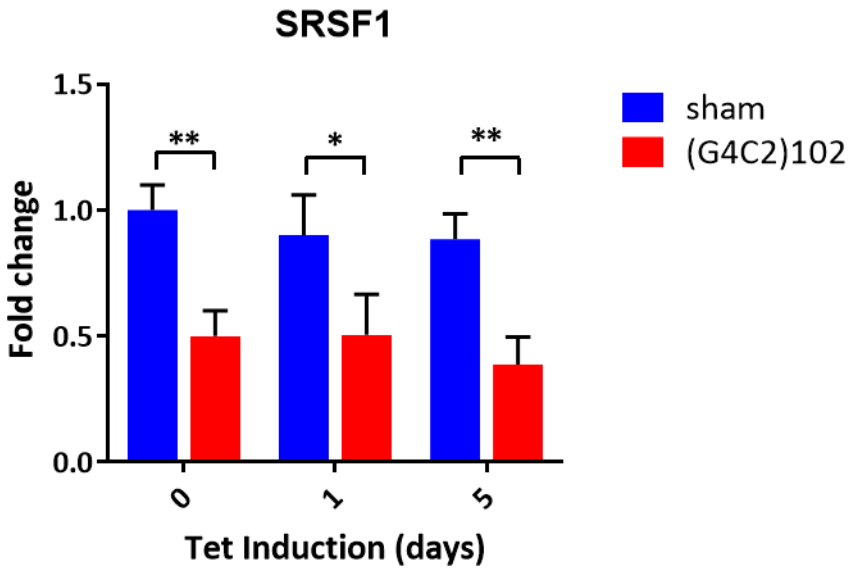


Figure 5.15 qRT-PCR Validation of SRSF1 levels. NSC34 sham and NSC34 (G4C2)102 cells were grown for 0, 1, or 5 days with 0.5µg/mL tetracycline. qRT-PCR was performed on SRSF1, and normalised to GAPDH. (*P<0.05, **P<0.01; Multiple t-tests; Data are means ± SD; n=3).

In addition, other factors involved in RNA and mRNA transport that are mutated in ALS and other motor neuron diseases are also DE in the NSC34 (G4C2)102 + 5 d tet. GLE1 RNA export mediator (Gle1) (FC=-1.20; P=0.00778) is an essential multifunctional protein involved in both mRNA export and translation (Bolger et al., 2008; Murphy and Wente, 1996). Mutations in GLE1 cause the autosomal recessive foetal motor neuron diseases lethal congenital contracture syndrome (LCCS1) and lethal arthrogryposis with anterior horn cell disease (LAAHD) (Nousiainen et al., 2008). In addition, three rare GLE1

mutations are found only in SALS patients, and are absent in controls (Kaneb et al., 2015). All three mutations cause loss of GLE1 function, indicating the importance of GLE1 in motor neuron survival and function.

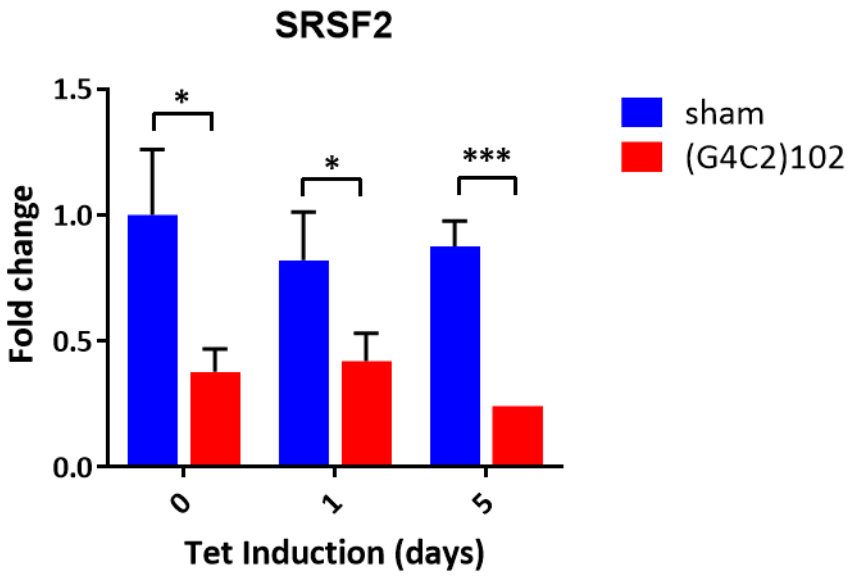


Figure 5.16 qRT-PCR Validation of SRSF2 levels. NSC34 sham and NSC34 (G4C2)102 cells were grown for 0, 1, or 5 days with 0.5µg/mL tetracycline. qRT-PCR was performed on SRSF2, and normalised to GAPDH. (**P<0.01; Multiple t-tests; Data are means ± SD; n=3).

Adenosine deaminase, RNA-specific (Adar) (FC=-1.28; P=0.00162) is an important paralog of Adar2, which is downregulated in LCM motor neurons from sporadic ALS patients (Hideyama et al., 2012). The Adar2 downregulation disrupts A-to-I RNA editing of GluA2 (a subunit of the AMPA receptor) producing Ca2+ permeable AMPA receptors, and excitotoxicity as a result (Hideyama et al., 2012). Aquarius (Aqr) (FC=-1.60; P=0.00226) is an important paralog of Senataxin (SETX), and links pre-mRNA splicing and snoRNA biogenesis (Hirose et al., 2006). Missense mutations in SETX cause juvenile onset ALS (Chen et al., 2004). SETX was also found to be down regulated (FC=-1.38; P=0.00721), although it was not annotated in the RNA metabolism biological function category, and was identified through manual interrogation of the DE transcripts list.

5.3.3.4.3. Protein Transport and Vesicle-Mediated Transport

Similarly to RNA metabolism, aberrant protein homeostasis is also described in a wide range of neurodegenerative diseases including ALS (section 1.6.6). There were 225 (178 and 120) DE trancripts (FC≥1.2; P<0.01) related to protein transport and/or vesicle-mediated transport in the NSC34 (G4C2)102 cells. Transcripts annotated for protein transport only (Table 5.15), vesicle-mediated transport only (Table 5.16), and both protein transport and vesicle-mediated transport (Table 5.17) were listed separately. The 225 DE transcripts were functionally annotated again using the DAVID functional annotation tool (section 5.3.3.3.1) and IMPaLA pathway enrichment analysis (section 5.3.3.3.2) to identify which protein and vesicle-mediated transport processes and/or pathways are mainly affected by (G4C2)102 expression. The top three annotated sub-categories were intracellular transport (91), membrane organisation (54), and secretion by cell (29) using the DAVID annotation tool. These 225 DE transcripts were also enriched for membrane trafficking (30), endocytosis (20), and trans-golgi network vesicle budding (10) using the IMPaLA pathway enrichment analysis.

There are several genes involved in protein transport and/or vesicle mediated transport that are mutated in ALS or other neurodegenerative diseases that are downregulated in the NSC34 (G4C2)102 + 5 d tet cells. Mutations in valosin containing protein (Vcp) (FC=-1.54; P=0.00148) cause ALS (Johnson et al., 2010; Koppers et al., 2012) and inclusion body myopathy with early-onset Paget disease and frontotemporal dementia (IBMPFD) (Watts et al., 2004). Vcp regulates a variety of cellular processes including ubiquitin-dependent protein degradation, nuclear envelope construction, Golgi and ER assembly, and autophagosome maturation (Ju et al., 2009; Tresse et al., 2010). Mutations in charged multivesicular body protein 2B (Chmp2b) cause both ALS and FTD (Parkinson et al., 2006; Skibinski et al., 2005). Chmp2a (FC=-1.32; P=0.00398) is an important paralog of Chmp2b, and both proteins are components of the Endosomal Sorting Complex Required for Transport III (ESCRT-III) which is involved in protein cargo trafficking in the endolysosomal pathway and autophagy (Henne et al., 2011). A (CAG)n repeat expansion in Huntingtin (Htt) (FC=-1.63; P=0.000522) causes Huntington's disease (HD). Htt interacts with hundreds of different proteins and is involved in many different cellular

functions (Marques Sousa and Humbert, 2013). Mutations in mitofusin-2 (Mfn2) (FC=-1.43; P=0.000971) cause Charchot-Marie-Tooth disease (CMT), and have also been linked to ALS (Marchesi et al., 2011). Mfn2 is a transmembrabe GTPase involved in mitochondrial fusion and mitochondria-ER membrane interactions (Züchner et al., 2004).

Table 5.15 Protein Transport transcripts dysregulated in NSC34 (G4C2)102 cells. Genes involved in Protein Transport which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells. Genes that are also functionally annotated as Vesicle-mediated transport in addition to Protein Transport are listed in **Table 5.17**.

Gene Symbol	Gene Title	Fold Change	p-value
Protein Localisation			
Actn4	actinin alpha 4	- 1.52	2.63E-03
Adam10	a disintegrin and metallopeptidase domain 10	-1.36	6.89E-03
Ank2	ankyrin 2, brain	-1.22	1.58E-03
Arfip1	ADP-ribosylation factor interacting protein 1	+1.21	8.94E-03
Arntl	aryl hydrocarbon receptor nuclear translocator-like	-1.47	9.63E-03
Atp7b	ATPase, Cu++ transporting, beta polypeptide	+1.30	1.87E-03
Bbs4	Bardet-Biedl syndrome 4 (human)	-1.78	9.29E-03
Bet1l	blocked early in transport 1 homolog (S. cerevisiae)-like	-1.25	3.02E-03
Bin3	bridging integrator 3	-1.25	2.69E-03
Cacnb1	calcium channel, voltage-dependent, beta 1 subunit	+1.23	1.06E-03
Cd27	CD27 antigen	+1.21	4.45E-03
Cdk5	cyclin-dependent kinase 5	-1.62	3.22E-05
Cep290	centrosomal protein 290	+1.44	7.01E-03
Chmp2a	charged multivesicular body protein 2A	-1.32	3.98E-03
Cog1	component of oligomeric golgi complex 1	-1.30	6.15E-03
Cog4	component of oligomeric golgi complex 4	-1.46	7.79E-03
Dnajc14	DnaJ (Hsp40) homolog, subfamily C, member 14	+1.20	3.87E-03
Dopey1	dopey family member 1	-1.38	5.13E-03
Eif5a	eukaryotic translation initiation factor 5A	-1.26	5.90E-03
Eps15	epidermal growth factor receptor pathway substrate 15	-1.21	1.52E-04
Faf1	Fas-associated factor 1	-1.46	5.70E-03
Fam125a	family with sequence similarity 125, member A	+1.36	2.78E-03
Folr1	folate receptor 1 (adult)	-1.27	5.13E-03
Gabarap	gamma-aminobutyric acid receptor associated protein	-1.29	1.77E-03
Gle1	GLE1 RNA export mediator (yeast)	-1.20	7.78E-03
Hgs	HGF-regulated tyrosine kinase substrate	-1.20	1.83E-03
Hspg2	perlecan (heparan sulfate proteoglycan 2)	-1.27	1.19E-03
Ift46	intraflagellar transport 46	-1.53	1.33E-03
Ipo4	importin 4	-1.40	1.65E-05
Kif1b	kinesin family member 1B	-1.83	1.57E-03

Lax1	lymphocyte transmembrane adaptor 1	+1.26	2.93E-03
Lcp2	lymphocyte cytosolic protein 2	+1.26	1.60E-03
Lman2	lectin, mannose-binding 2	-1.37	1.72E-03
Mecp2	methyl CpG binding protein 2	-1.29	7.89E-03
Mlh3	mutL homolog 3 (E coli)	+1.33	2.36E-03
Mlph	Melanophilin	+1.21	4.39E-03
Mtx2	metaxin 2	-1.80	5.76E-03
Nckap1	NCK-associated protein 1	-1.63	2.32E-03
Nefm	neurofilament, medium polypeptide	-1.53	3.69E-04
Nup133	nucleoporin 133	-1.50	3.38E-03
Nup160	nucleoporin 160	-1.54	7.85E-03
Nup214	nucleoporin 214	-1.53	9.19E-04
Nup35	nucleoporin 35	-1.43	8.62E-03
Nup93	nucleoporin 93	-1.54	4.35E-04
Nupl1	nucleoporin like 1	-1.31	7.31E-03
Nupl2	nucleoporin like 2	-1.43	1.93E-03
Nutf2	predicted gene 10349; predicted pseudogene 9386; nuclear transport factor 2; nuclear transport factor 2, pseudogene 1	-1.26	4.05E-03
Pcna	proliferating cell nuclear antigen	-1.49	9.91E-03
Pdcd6ip	programmed cell death 6 interacting protein	-1.76	5.24E-03
Pex1	peroxisomal biogenesis factor 1	-1.47	6.96E-04
Pex14	peroxisomal biogenesis factor 14	-1.23	6.44E-03
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform	-1.40	7.03E-03
Rab11b	RAB11B, member RAS oncogene family	-2.11	3.15E-04
Rab14	RAB14, member RAS oncogene family	-1.31	3.21E-03
Rab22a	RAB22A, member RAS oncogene family	-1.60	4.05E-04
Rab31	RAB31, member RAS oncogene family	-1.58	2.74E-03
Rab35	RAB35, member RAS oncogene family	-1.27	5.47E-04
Rab39b	RAB39B, member RAS oncogene family	-1.54	4.15E-03
Rab3c	RAB3C, member RAS oncogene family	-1.32	1.57E-03
Rab4a	RAB4A, member RAS oncogene family	-1.37	1.21E-03
Rab8a	RAB8A, member RAS oncogene family	-1.36	3.04E-03
Ranbp3l	RAN binding protein 3-like	+1.27	4.13E-03
Rrbp1	ribosome binding protein 1	-1.59	4.26E-04
Scamp2	secretory carrier membrane protein 2	-1.51	6.37E-05
Scamp3	secretory carrier membrane protein 3	-1.30	5.56E-03
Sec61a1	Sec61 alpha 1 subunit (S. cerevisiae)	-1.35	1.98E-03
Sec61a2	Sec61, alpha subunit 2 (S. cerevisiae)	-1.44	9.17E-03
Selenbp2	selenium binding protein 2	+1.30	9.95E-03
Sft2d2	SFT2 domain containing 2	-2.32	2.62E-03
Shroom3	shroom family member 3	+1.27	7.77E-03
Slc15a3	solute carrier family 15, member 3	+1.23	9.89E-03
Slc15a5	solute carrier family 15, member 5	+1.25	4.86E-03
Slc37a2	solute carrier family 37 (glycerol-3-phosphate transporter), member 2	+1.22	2.73E-03
Snx1	sorting nexin 1	-1.25	5.23E-04
Snx12	sorting nexin 12	-2.01	3.34E-04

Snx14	sorting nexin 14	-1.75	2.77E-03
Snx18	sorting nexin 18	-1.25	2.82E-03
Snx25	sorting nexin 25	-1.36	3.12E-03
Snx3	sorting nexin 3	-1.26	1.09E-03
Snx9	sorting nexin 9	-1.69	1.04E-03
Srpr	signal recognition particle receptor (docking protein)	-1.76	2.82E-03
Srsf10	serine/arginine-rich splicing factor 10	-1.54	3.33E-03
Stam2	signal transducing adaptor molecule (SH3 domain and ITAM motif) 2	-1.62	6.55E-03
Stau1	staufen (RNA binding protein) homolog 1 (*Drosophila*)	-1.38	1.14E-03
Syngr2	synaptogyrin 2	-1.30	7.86E-04
Timm17b	translocase of inner mitochondrial membrane 17b	-1.22	1.56E-04
Timm23	translocase of inner mitochondrial membrane 23	-1.41	4.66E-03
Tln1	talin 1	-1.27	3.25E-04
Tmed3	transmembrane emp24 domain containing 3	-1.20	2.64E-03
Tnfsf13b	tumor necrosis factor (ligand) superfamily, member 13b	+1.35	1.22E-04
Tnrc6a	trinucleotide repeat containing 6a	-1.28	7.44E-03
Trpc4ap	transient receptor potential cation channel, subfamily C, member 4 associated protein	-1.52	1.22E-03
Tsc2	tuberous sclerosis 2	-1.44	2.07E-04
Uchl1	ubiquitin carboxy-terminal hydrolase L1	-1.86	2.61E-04
Vcp	valosin containing protein	-1.54	1.48E-03
Vps11	vacuolar protein sorting 11 (yeast)	-1.61	1.18E-04
Vps13b	vacuolar protein sorting 13B (yeast)	-1.41	4.90E-03
Vps13d	vacuolar protein sorting 13D (yeast)	-1.35	1.56E-03
Vps16	vacuolar protein sorting 16 (yeast)	-1.31	3.49E-03
Vps25	vacuolar protein sorting 25 (yeast)	-1.53	6.33E-03
Vti1a	vesicle transport through interaction with t-SNAREs 1A (yeast)	-1.39	1.26E-03
Xpo6	exportin 6	-1.24	3.85E-03
Xpo7	exportin 7	-1.44	3.79E-04
Ywhab	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	-1.51	3.67E-03
Ywhah	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta polypeptide	-1.38	1.03E-03

Table 5.16 Vesicle-mediated Transport transcripts dysregulated in NSC34 (G4C2)102 cells. Genes involved in Vesicle-mediated Transport which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells. Genes that are also functionally annotated as Protein Transport in addition to Vesicle-mediated Transport are listed in **Table 5.17**.

Gene Symbol	Gene Title	Fold change	p-value
Vesicle-mediated Transport			
Abca1	ATP-binding cassette, sub-family A (ABC1), member 1	-2.18	1.31E-04
Agrn	Agrin	-1.44	2.53E-04
Aplp1	amyloid beta (A4) precursor-like protein 1	-1.53	6.54E-03

Arhgap17	Rho GTPase activating protein 17	-1.31	2.19E-03
Cap1	CAP, adenylate cyclase-associated protein 1 (yeast)	-1.65	8.84E-03
Chrna7	cholinergic receptor, nicotinic, alpha polypeptide 7	-1.48	3.91E-04
Cnih	cornichon homolog (*Drosophila*)	-1.38	3.98E-03
Coro1c	coronin, actin binding protein 1C	-1.55	1.27E-03
Cplx1	complexin 1	-1.42	5.45E-04
Cstad	CSA-conditional, T cell activation-dependent protein	+1.25	3.91E-03
Cttn	Cortactin	-1.30	7.99E-04
Dennd1a	DENN/MADD domain containing 1A	-1.38	1.09E-03
Dnm1	dynamin 1	-1.28	1.15E-03
Dnm2	dynamin 2	-1.24	4.39E-03
Dock1	dedicator of cytokinesis 1	-1.59	9.69E-05
Ehd2	EH-domain containing 2	-1.55	5.24E-04
Elmo2	engulfment and cell motility 2	-1.39	1.07E-03
Eps15l1	epidermal growth factor receptor pathway substrate 15-like 1	-1.53	4.71E-03
Fcgr3	Fc receptor, IgG, low affinity III	+1.27	4.82E-03
Fnbp1l	formin binding protein 1-like	-1.61	1.63E-03
Gapvd1	GTPase activating protein and VPS9 domains 1	-1.59	3.43E-03
Ghr	growth hormone receptor	+1.28	9.71E-03
Glra1	glycine receptor, alpha 1 subunit	+1.20	5.68E-03
Gria1	glutamate receptor, ionotropic, AMPA1 (alpha 1)	+1.26	6.17E-03
Gsn	Gelsolin	-1.21	5.95E-03
Gulp1	GULP, engulfment adaptor PTB domain containing 1	+1.21	6.32E-03
Itga3	integrin alpha 3	-1.29	2.42E-04
Mfge8	milk fat globule-EGF factor 8 protein	-1.36	3.11E-04
Nploc4	nuclear protein localization 4 homolog (S. cerevisiae)	-1.50	2.11E-03
Nras	neuroblastoma ras oncogene	-1.68	1.68E-03
Pacsin2	protein kinase C and casein kinase substrate in neurons 2	-1.28	1.23E-03
Plcd4	phospholipase C, delta 4	+1.29	9.06E-05
Rac1	RAS-related C3 botulinum substrate 1	-1.40	3.54E-03
Rims3	regulating synaptic membrane exocytosis 3	-1.40	7.99E-06
Rims4	regulating synaptic membrane exocytosis 4	-1.31	2.79E-03
Rin2	Ras and Rab interactor 2	+1.22	3.40E-04
Sgca	sarcoglycan, alpha (dystrophin-associated glycoprotein)	+1.25	2.61E-03
Sgcb	sarcoglycan, beta (dystrophin-associated glycoprotein)	-1.72	4.15E-03
Sirpa	signal-regulatory protein alpha	+1.24	8.33E-03
Sort1	sortilin 1	-1.45	1.58E-03
Syp	Synaptophysin	-1.20	8.45E-04
Tac4	tachykinin 4	+1.32	6.65E-03
Tfrc	transferrin receptor	-1.59	5.95E-04
Trappc10	trafficking protein particle complex 10	-1.49	2.78E-04
Trappc6b	trafficking protein particle complex 6B	-1.34	5.69E-03
Vamp3	vesicle-associated membrane protein 3	-1.82	3.19E-03
Zmpste24	zinc metallopeptidase, STE24	-1.70	2.59E-03

Table 5.17 Protein Transport and Vesicle-mediated Transport transcripts dysregulated in NSC34 (G4C2)102 cells. Genes involved in both Protein Transport and Vesicle-mediated Transport which are differentially expressed in NSC34 (G4C2)102 + 5 d tet cells compared to NSC34 sham + 5 d tet cells.

Gene Symbol	Gene Title	Fold change	p-value
Protein Localisation and Vesicle-mediated Transport			
Ap1g2	adaptor protein complex AP-1, gamma 2 subunit	-1.37	4.14E-04
Ap1m1	adaptor-related protein complex AP-1, mu subunit 1	-1.67	1.41E-03
Ap1s2	adaptor-related protein complex 1, sigma 2 subunit	-1.50	1.91E-03
Ap2a2	adaptor protein complex AP-2, alpha 2 subunit	-1.94	2.84E-04
Ap3s2	adaptor-related protein complex 3, sigma 2 subunit	-1.70	3.07E-03
Ap4b1	adaptor-related protein complex AP-4, beta 1	-1.33	5.90E-03
Ap4e1	adaptor-related protein complex AP-4, epsilon 1	-1.39	6.86E-03
App	amyloid beta (A4) precursor protein	-1.31	3.77E-04
Arcn1	archain 1	-1.65	9.16E-03
Arfgap2	ADP-ribosylation factor GTPase activating protein 2	-1.33	1.18E-03
Cadps	Ca2+-dependent secretion activator	-1.32	4.54E-03
Cdc42	cell division cycle 42	-1.85	2.49E-03
Chic2	cysteine-rich hydrophobic domain 2	-1.46	1.95E-03
Chmp7	charged multivesicular body protein 7	-1.30	1.52E-03
Cltb	clathrin, light polypeptide	-1.39	2.72E-03
Cog3	component of oligomeric golgi complex 3	-1.29	8.10E-03
Copb2	coatomer protein complex, subunit beta 2 (beta prime)	-1.72	6.06E-03
Copg1	coatomer protein complex, subunit gamma 1	-1.46	6.82E-03
Cox16	COX16 cytochrome c oxidase assembly homolog (S. cerevisiae)	-1.24	8.48E-03
D230025D16Rik	RIKEN cDNA D230025D16 gene	-1.27	9.37E-03
Dhcr24	24-dehydrocholesterol reductase	-1.46	3.96E-03
Ergic3	ERGIC and golgi 3	-1.65	4.70E-04
Exoc1	exocyst complex component 1	-1.53	3.03E-03
Exoc2	exocyst complex component 2	-1.53	6.12E-03
Flna	filamin, alpha	-2.57	1.01E-05
Golga5	golgi autoantigen, golgin subfamily a, 5	-1.33	3.38E-03
Gria2	glutamate receptor, ionotropic, AMPA2 (alpha 2)	+1.42	3.42E-03
Htt	Huntingtin	-1.63	5.22E-04
Kdelr1	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 1	-1.21	9.00E-03
Kdelr2	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2	-1.33	8.67E-04
Kdelr3	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3	-1.54	3.31E-03
Lin7a	lin-7 homolog A (C. elegans)	+1.22	1.63E-03
Mapk8ip3	mitogen-activated protein kinase 8 interacting protein 3	-1.20	5.06E-04
Mfn2	mitofusin 2	-1.43	9.71E-04
Napa	N-ethylmaleimide sensitive fusion protein attachment protein alpha	-1.35	2.73E-03
Necap2	NECAP endocytosis associated 2	-1.65	6.17E-04
Npc1	Niemann Pick type C1	-1.50	9.08E-03

Nrbp1	nuclear receptor binding protein 1	-1.74	3.47E-04
Oxa1l	oxidase assembly 1-like	-1.64	4.64E-03
Pldn	Pallidin	-1.46	6.57E-05
Preb	prolactin regulatory element binding	-1.26	2.58E-03
Rab27a	RAB27A, member RAS oncogene family	+1.25	3.12E-03
Rab2a	RAB2A, member RAS oncogene family	-1.60	4.04E-03
Rab6a	RAB6A, member RAS oncogene family	-1.63	8.00E-03
Rhob	ras homolog gene family, member B	-1.87	1.02E-04
Rims1	regulating synaptic membrane exocytosis 1	+1.22	8.36E-03
Rint1	RAD50 interactor 1	-1.30	6.46E-03
Scfd2	Sec1 family domain containing 2	-1.37	1.62E-03
Sec13	SEC13 homolog (S. cerevisiae)	-1.24	7.44E-03
Sec23b	SEC23B (S. cerevisiae)	-1.45	7.08E-03
Sec24c	Sec24 related gene family, member C (S. cerevisiae)	-1.24	1.72E-03
Sec24d	Sec24 related gene family, member D (S. cerevisiae)	-1.31	2.49E-03
Sec31a	Sec31 homolog A (S. cerevisiae)	-1.68	2.86E-03
Slc30a6	solute carrier family 30 (zinc transporter), member 6	-2.08	2.06E-03
Snap23	synaptosomal-associated protein 23	-1.52	5.48E-03
Snx17	sorting nexin 17	-1.36	2.88E-03
Spast	Spastin	-1.76	8.24E-04
Steap2	six transmembrane epithelial antigen of prostate 2	-1.26	1.75E-03
Stx12	syntaxin 12	-2.07	1.25E-04
Stx16	syntaxin 16	-1.31	2.08E-03
Stx18	syntaxin 18	-1.40	2.67E-04
Stxbp1	syntaxin binding protein 1	-1.44	5.05E-04
Stxbp2	syntaxin binding protein 2	+1.22	3.12E-03
Stxbp3a	syntaxin binding protein 3A	-1.66	7.13E-04
Trappc1	trafficking protein particle complex 1	-1.28	2.29E-03
Trappc4	trafficking protein particle complex 4	-1.43	6.50E-03
Trp53	transformation related protein 53	-1.66	5.73E-03
Ulk1	Unc-51 like kinase 1 (C. elegans)	-1.32	1.31E-04
Uso1	USO1 homolog, vesicle docking protein (yeast)	-1.78	3.24E-03
Vps29	vacuolar protein sorting 29 (S. pombe)	-1.36	9.28E-03
Vps33b	vacuolar protein sorting 33B (yeast)	-1.36	7.45E-04
Wasf2	WAS protein family, member 2	-1.48	9.50E-04
Zw10	ZW10 homolog (*Drosophila*), centromere/kinetochore protein	-1.24	4.23E-04

5.3.4. Differential Splicing Analysis

CEL files were normalised using RMA alt splice analysis in Expression Console, and saved as CHP files using Expression Console. The CHP files were then opened in TAC for alternative (differential) splicing analysis. To identify true differential splicing events between array conditions, the Splicing Index (SI) was used. The SI is the relative difference between a splicing event (such as exon level or an exon skipping event)

between two array conditions and is summarised in Figure 5.17. The SI can be likened to Fold Change, whereas it is the relative level of a splicing event being compared instead of the level of a particular transcript. Several criteria must be met for an exon to be considered for SI analysis however. The gene must be expressed in both conditions and the exon or junction must be expressed in at least one condition. Once these criteria are met, the signal intensity for each gene, and subsequently each exon and junction, is normalised. The normalised intensities are then compared between the two conditions using One Way Between-Subject ANOVA (unpaired), giving an SI value and FDR corrected P value. Differential spicing events were filtered to include only events with SI≥2.0 or ≤-2.0, and ANOVA FDR<0.05.

$$\text{Splicing Index (SI)} = \frac{\left\{\dfrac{\text{Exon 1 condition 1 Intensity}}{\text{Gene 1 condition 1 Intensity}}\right\}}{\left\{\dfrac{\text{Exon 1 condition 2 Intensity}}{\text{Gene 1 condition 2 Intensity}}\right\}}$$

Figure 5.17 Algorithm describing Splicing Index used to assess differential splicing events between two biological conditions.

Differential splicing analysis was performed on the NSC34 sham and NSC34 (G4C2)102 cells. There were 3029 differentially spliced transcripts in the NSC34 (G4C2)102 + 0 d tet compared to NSC34 sham + 0 d tet, 2455 differentially spliced transcripts in the NSC34 (G4C2)102 + 1 d tet compared to NSC34 sham + 1 d tet, and 5834 differentially spliced transcripts in the NSC34 (G4C2)102 + 5 d tet compared to NSC34 sham + 5 d tet (Figure 5.18). Thus the low level (G4C2)102 expression at basal conditions modestly disrupts splicing, but the increased (G4C2)102 expression at 5 days tetracycline induction causes more extensive splicing disruption.

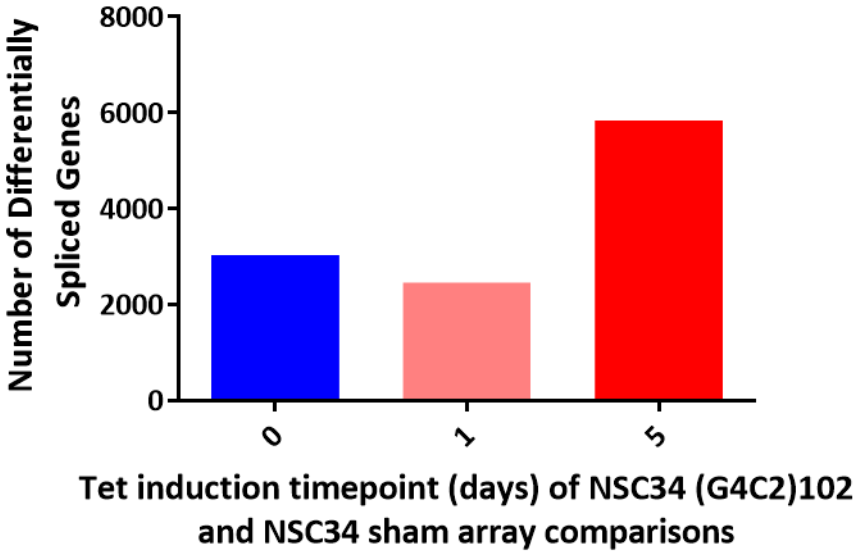


Figure 5.18 Summary of the differentially spliced genes between NSC34 (G4C2)102 and NSC34 sham cells. The numbers of genes with at least one differential splice event with a splicing index >2 and with One-Way ANOVA (unpaired) P<0.05 between NSC34 (G4C2)102 and NSC34 sham at 0, 1 and 5 days of tetracycline induction.

5.4. Discussion

5.4.1. Tetracycline Does Not Affect the NSC34 Transcriptome

Previous experiments had shown 10µg/mL tetracycline was toxic to NSC34 sham cells (section 3.3.4.4). Therefore the NSC34 sham arrays with and without 5 days tetracycline induction were compared. There were only a few DE transcripts between the arrays which shows that 0.5µg/mL tetracycline has only a very small effect on the NSC34 transcriptome. Functional enrichment analysis showed olfactory transduction was the only enriched functional annotation cluster. The 3 DE transcripts from this tetracycline analysis that overlapped with the DE transcripts in the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet analysis were removed from further analysis. Therefore, the DE transcripts, and functionally enriched biological processes and pathways identified in the NSC34 (G4C2)102 + 5 d tet vs NSC34 sham + 5 d tet analysis are not caused by tetracycline, but are caused by the (G4C2)102.

5.4.2. (G4C2)n Expression Alters the NSC34 Transcriptome

GEP of the NSC34 cells shows that the (G4C2)102 expression alters the NSC34 transcriptome. Firstly, the NSC34 (G4C2)102 arrays are clustered together, separate from the NSC34 sham, (G4C2)10 and (G4C2)51 arrays on the PCA plot. Secondly, the NSC34 (G4C2)102 + 5 day tet arrays are the most different of the NSC34 (G4C2)102 arrays on the PCA plot. Thirdly, when compared to NSC34 sham cells, the number of DE transcripts increases from 1576 to 3069 in the NSC34 (G4C2)102 cells when induced with tetracycline for 0 and 5 days respectively. Importantly, The NSC34 (G4C2)102 still express the (G4C2)102 repeat at lower levels when not induced with tetracycline because RNA foci and RAN translation products are still detected under basal conditions (sections 3.4.3 and 3.4.5 respectively). The low basal expression of (G4C2)102 probably accounts for the transcriptomic differences, and could explain why the NSC34 (G4C2)102 + 0 d tet arrays cluster with the NSC34 (G4C2)102 + 1 d and + 5 d arrays. Additionally, the (G4C2)102 DNA repeat in the genome could contribute to the DE transcripts and altered gene expression. However, the greater (G4C2)102 expression further alters the NSC34 cell gene expression and there is a high degree of overlap in the DE transcripts in the 0 and 5 day tet comparisons, which shows the (G4C2)102 expression does alter gene expression.

5.4.3. Enrichment Analysis

Enrichment analysis identifies which biological functions and pathways are most altered at the transcript level in disease. Enrichment analysis was first performed here using DAVID to assess how (G4C2)102 expression affected the NSC34 cell transcriptome. Using DAVID, 63.4% of the DE transcripts were annotated to a biological function, whilst only 26.2% of the DE genes were annotated to a KEGG pathway. The percentage of genes annotated to a biological pathway using DAVID seemed very low, and the actual biological pathway annotations are more informative and specific than the vaguer biological function annotations. Therefore, an alternative pathway enrichment analysis tool with greater annotation was used. IMPaLA annotates genes from 11 different public databases (including KEGG), and was used to repeat the pathway enrichment analysis. Using IMPaLA, 47.4% of genes were annotated to a biological pathway, which was nearly

double the coverage that DAVID provided, and was much more acceptable for pathway analysis. The DAVID and IMPaLA enrichment analyses identified several interesting biological functions and pathways including braches of the PI3K/Akt signalling super-pathway, RNA metabolism, and protein transport which are discussed in greater detail below.

5.4.4. Pathways of interest

5.4.4.1. PI3K/Akt signalling

The PI3K/Akt signalling pathway regulates a range of cellular functions including metabolism, cell growth, proliferation, survival, and protein translation. GEP of the NSC34 (G4C2)102 cells identified transcriptomic dysregulation in this pathway. More specifically, Akt1 and Akt2 are both downregulated at transcript level, suggesting the Akt levels and signalling may be reduced. Also, PI3K activates Akt, whilst Pten antagonises this activation, and the class I PI3K catalytic subunits PIK3CA and PIK3CB are downregulated at transcript level, whilst Pten is upregulated at transcript level in NSC34 (G4C2)102, further suggesting reduced Akt activation and signalling. Finally, mTOR, which is a member of the Akt activating complex MTORC2, is also downregulated at transcript level again suggesting reduced Akt activation. Reduced Akt levels and activity would potentially cause a reduced growth rate, since Akt promotes cell growth and proliferation. The (G4C2)102 expression reduced NSC34 cell growth rate (section 3.4.6), and supports the reduced Akt signalling hypothesis. Further, the PI3K/Akt signalling pathway is also dysregulated at transcriptomic level in C9ORF72-ALS LCM motor neurons, and corroborates the GEP in the NSC34 (G4C2)102 cells. However, qRT-PCR analysis only validated the DE of mTOR, but not Akt1 or Pten in the NSC34 (G4C2)102 cells. Further investigation of the PI3K/Akt signalling pathway at protein level is required to strengthen the hypothesis that (G4C2)102 expression in a motor neuronal context affects PI3K/Akt signalling.

Previously, GEP on LCM motor neurons from spinal cord of SOD1-ALS patients identified transcriptomic changes in the PI3K/Akt signalling pathway (Kirby et al., 2011). Specifically, the transcriptomic data suggested the PI3K/Akt signalling pathway is more active in the surviving motor neurons, and actually contributes to their survival (Kirby et

al., 2011). This supports the finding that the PI3K/Akt signalling pathway is dysregulated at transcript level in both the NSC34 (G4C2)102 cells and the C9ORF72-ALS LCM motor neurons. However, the transcriptomic data and cellular assay data in the NSC34 (G4C2)102 cells suggests (G4C2)102 expression results in reduced Akt signalling, rather than in a prosurvival manner. To understand how (G4C2)102 expression affects The PI3K/Akt signalling pathway, further biochemical investigation in the NSC34 (G4C2)102 cells was required.

Interestingly, Pten, which is a key negative regulator of the PI3K/Akt signalling pathway, has been identified as a potential therapeutic target in ALS and other motor neuron diseases (Kirby et al., 2011; Little et al., 2015; Ning et al., 2010). PTEN knock-down via siRNA promoted survival in SOD1-G93A NSC34 cells and motor neuron survival (Kirby et al., 2011) and axon growth in an SMA mouse model (Ning et al., 2010). Further, Pten knock-down by siRNA in an SMA mouse model tripled survival time, improved motor function, and motor neuron survival (Little et al., 2015). Interestingly, in the C9ORF72-ALS LCM motor neuron transcriptomic dataset, *PTEN* had the highest fold change of all DE genes. As described above, Pten is upregulated at transcript level in the NSC34 (G4C2)102 cells as well. This suggests that PTEN upregulation in C9ORF72-ALS may contribute to motor neuronal cell death. More specifically, the NSC34 (G4C2)102 data suggests the (G4C2)n expression may cause this PTEN upregulation. Therefore, Pten was taken forward as a potential therapeutic target in chapter 6.

5.4.4.2. RNA Metabolism and mRNA Transport

RNA metabolism is an important biological process involved in many levels of gene expression (Walsh et al., 2015). Dysregulated RNA metabolism is implicated in a wide range of neurodegenerative diseases including ALS (sections 1.5.5 and 1.6.5). GEP of the NSC34 (G4C2)102 cells identified transcriptomic dysregulation in RNA metabolism, but more specifically in RNA splicing, ncRNA processing and mRNA transport. This is in agreement with a previous transcriptomic study that showed RNA splicing is dysregulated in C9ORF72-ALS lymphoblastoid cells and LCM motor neurons (Cooper-Knock et al., 2015a).

The (G4C2)n repeat expansion is hypothesised to cause neurotoxicity and neurodegeneration in C9ORF72-ALS via an RNA toxicity mechanism (section 1.7.2). Briefly, the (G4C2)n RNA transcribed from the repeat expansion is thought to bind and sequester RNA binding proteins. Therefore, several studies used *in vitro* (G4C2)n RNA pull downs to identify candidate binding proteins (Cooper-Knock et al., 2014b; Donnelly et al., 2013; Haeusler et al., 2014; Lee et al., 2013; Mori et al., 2013b; Rossi et al., 2015; Xu et al., 2013). Interestingly, several of the (G4C2)n RNA binding candidates identified by our group, including many splicing factors, were downregulated in the NSC34 (G4C2)102 cells. These transcripts included SRSF1 and SRSF2, and the downregulation was validated by qRT-PCR. SRSF1 and SRSF2 had previously been shown to co-localise with the (G4C2)n RNA foci in the NSC34 (G4C2)102 cells (section 4.3.3) as well as in C9ORF72-ALS pathological tissue (Cooper-Knock et al., 2014b; Lee et al., 2013).

The downregulation of (G4C2)n RNA binding candidates is contrary to a seemingly more obvious hypothesis: sequestration of RNA binding proteins by the (G4C2)n RNA would result in loss of activity of those specific RNA binding proteins, and the cell would therefore compensate by increasing expression of those RNA binding proteins. This is in fact what our group found in C9ORF72-ALS patient tissue. However, 20% and 10% of the identified binding candidates from the aforementioned study (Cooper-Knock et al., 2014b) were DE in C9ORF72-ALS lymphoblastoid cells and LCM motor neurons respectively, where 89% and 77% of those DE transcripts were upregulated respectively (Cooper-Knock et al., 2015a). This casts some doubt over the physiological relevance of the transcriptomic changes identified in the NSC34 (G4C2)102 cells.

The differential splicing analysis provides a functional readout of RNA splicing dysregulation in the NSC34 (G4C2)102 cells. The number of differentially spliced transcripts between the NSC34 (G4C2)102 and NSC34 sham increased from 3029 in basal conditions (no tetracycline) to 5834 when the cells were induced with tetracycline for 5 days, and suggests that the (G4C2)102 expression alters splicing. The (G4C2)102 is expressed at low levels at basal conditions (section 3.4.3), which could cause some differential splicing events. However, the greater (G4C2)102 expression after 5 days tetracycline induction further disrupts splicing in the NSC34 cells. This importantly

suggests that RNA splicing defects are an early biochemical effect of (G4C2)n expression. This corroborates previous transcriptomic analysis that showed there are extensive splicing defects in C9ORF72-ALS frontal cortex and cerebellum (Prudencio et al., 2015). In addition, splicing is less consistent in C9ORF72-ALS lymphoblastoid cells compared to non-C9ORF72-ALS and control cells, although the actual number of differential splicing events was not altered (Cooper-Knock et al., 2015a). It is not clear whether the (G4C2)n RNA and/or the RAN translation products cause the RNA splicing defects, although the RNA sequestration hypothesis makes it more logical to attribute these effects to the (G4C2)n RNA.

In addition to RNA splicing defects, dysregulated mRNA export is also thought to play an important role in C9ORF72-ALS pathogenesis. The (G4C2)n RNA binds RanGAP1 in transgenic fly models, C9ORF72-ALS iPSC-derived neurons and C9ORF72-ALS patient brain tissue (Zhang et al., 2015). The (G4C2)n RNA also disrupts nuclear import in transgenic fly models and C9ORF72-ALS iPSC-derived neurons (Zhang et al., 2015), and nuclear export of RNA in transfected NSC34 and HeLa cells, transgenic fly models, and C9ORF72-ALS iPSC-derived neurons (Freibaum et al., 2015; Rossi et al., 2015). In addition, genetic screening in transgenic fly models show proteins involved in nucleocytoplasmic transport modify toxicity caused by the (G4C2)n RNA (Freibaum et al., 2015), whilst a screen performed in *S. cerevisiae* identified components of nucleocytoplasmic transport as modifiers of DPR-induced toxicity (Jovičić et al., 2015).

5.4.4.3. Protein Transport and Vesicle-Mediated Transport

Protein transport and vesicle-mediated transport, in addition to protein degradation, are thought to be important biological processes in ALS pathogenesis (sections 1.5.6 and 1.6.6). This is due to the number of ALS associated genes that are involved in protein homeostasis, and also because protein aggregates are pathological hallmarks of ALS, suggesting cellular protein homeostasis is inefficient or overwhelmed. GEP of the NSC34 (G4C2)102 cells identified transcriptomic dysregulation in protein transport and vesicle-mediated transport. Also, previous studies have suggested the (G4C2)n repeat and/or the DPR cause nucleocytoplasmic transport dysregulation (Freibaum et al., 2015; Jovičić et al., 2015; Zhang et al., 2015).

The (G4C2)102 repeat is transcribed to form RNA foci (section 3.3.4.1), but also undergoes RAN translation producing proteins that contain the (GA)n, (GR)n and (GP)n sense DPR motifs (section 3.3.4.3). The DPR proteins aggregate to form insoluble inclusions *in vitro* and also in C9ORF72-ALS CNS tissue (sections 1.4, 1.7.3, and 1.8.2.3). Whilst there is no direct evidence to show the RAN proteins in the NSC34 (G4C2)102 cells aggregate, it is certainly possible. Therefore, the transcriptomic alterations in protein transport and vesicle-mediated transport may suggest a cellular response to potentially misfolded and aggregated RAN proteins within the NSC34 (G4C2)102 cells. Additionally, in times of stress, protein synthesis and cellular growth are inhibited, whilst autophagy is upregulated (Sarbassov et al., 2005). Critically, mTOR signalling favours growth and inhibits autophagy, and therefore the downregulation of mTOR in the NSC34 (G4C2)102 cells fits with the reduced growth rate previously described in the NSC34 (G4C2)102 cells (sections 3.3.4.7 and 3.4.6), but also suggests increased autophagy.

5.4.5. Summary

Early transcriptomic changes caused by the (G4C2)102 expression were identified in the NSC34 cells using microarray analysis. Importantly, the transcriptomic changes corroborate with previous transcriptomic studies in C9ORF72-ALS patient derived cells and post-mortem tissue. Specifically, RNA metabolism, and RNA splicing and mRNA transport in particular, were dysregulated at transcript level. Further, (G4C2)102 expression increased the number of differentially spliced transcripts, suggesting functional splicing dysregulation. In addition, significant transcriptomic dysregulation was identified in the PI3K/Akt signalling pathway in the NSC34 (G4C2)102 cells. This was validated in LCM motor neurons from C9ORF72-ALS patients, and suggests this is an early transcriptomic change in C9ORF72-ALS pathogenesis. Lastly, since the PI3K/Akt signalling pathway regulates cellular survival it may represent a potent therapeutic target, and was taken forward for *in vitro* manipulation and rescue assays.

Chapter 6. Modulation of Pten and PI3K/Akt Signalling Pathway in NSC34 (G4C2)102 Cells

6.1. Introduction

PTEN and the PI3K/AKT signalling pathway have previously been described as potential therapeutic targets in ALS and other motor neuron diseases (see section 5.4.4.1). Pten was elevated at the transcript level in the NSC34 (G4C2)102 cells, as well as C9ORF72-ALS LCM motor neurons. Further, the PI3K/AKT signalling pathway (or branches of the super-pathway) were also dysregulated at transcript level in the NSC34 (G4C2)102 and C9ORF72-ALS LCM motor neurons. Therefore, it was hypothesised that PTEN inhibition would activate the PI3K/Akt signalling pathway rescuing the NSC34 cells from (G4C2)102 induced toxicity. Two different approaches were taken to inhibit PTEN activity: PTEN knock-down using commercial lentiviral Pten shRNA vectors, and pharmacological inhibition using the specific PTEN inhibitor bisperoxo(1,10-phenanthroline)oxovanadate (BpV(phen)).

The PI3K/AKT signalling pathway is highly conserved, and regulates various cellular functions including cellular survival, cell proliferation, cell growth, metabolism, transcription and protein synthesis (Figure 5.11) (Hers et al., 2011). Phosphoinositide 3-kinases (PI3K) are central to the PI3K/AKT signalling cascade, and are activated by a wide range of receptors such as receptor tyrosine kinases and G protein coupled receptors (Figure 5.11). Active PI3Ks phosphorylate phosphatidylinositol-4,5-P2 (PI[4,5]P2) to produce phosphatidylinositol-3,4,5-P3 (PI[3,4,5]P3) (Figure 6.1), which acts as a secondary messenger. Importantly, PTEN negatively regulates this signalling cascade via it's intrinsic phosphatase activity, which converts PI[3,4,5]P3 back to PI[4,5]P2 thereby antagonising the PI3K signalling (Figure 6.1). PTEN activity can also be reduced by phosphorylation at Ser380 (Vazquez et al., 2000).

AKT is activated downstream in the PI3K/AKT signalling cascade by the PI[3,4,5]P3. The three genes *AKT1*, *AKT2*, and *AKT3* encode highly similar isoforms of AKT. The PH domain of AKT allows docking to PI[3,4,5]P3, and recruitment to the plasma membrane. Once

docked to PI[3,4,5]P3, AKT changes conformation allowing Thr308 and Ser473 residues to be phosphorylated by various phosphatidyl-dependent kinases (PDK). The phosphorylation of these two residues is critical for full activation of AKT (Bhaskar and Hay, 2007). Therefore, quantification of p-Akt(Ser473) levels is widely used to measure AKT signalling activity.

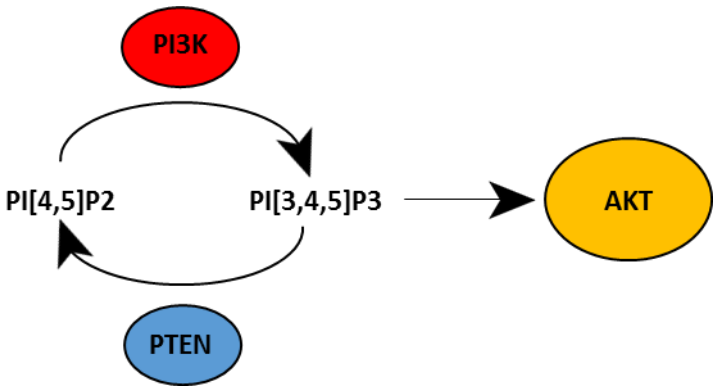


Figure 6.1 PTEN antagonises PI3K/AKT signalling. PI3K phosphorylates phosphatidylinositol-4,5-P2 (PI[4,5]P2) to produce phosphatidylinositol-3,4,5-P3 (PI[3,4,5]P3). PI[3,4,5]P3 recruits AKT at the plasma membrane, leading to AKT activation. PTEN antagonises this cascade by phosphatase activity, which converts PI[3,4,5]P3 to PI[4,5]P2.

BpV(phen) is a competitive and reversible inhibitor of PTEN (Schmid et al., 2004), and is thought to reversibly oxidise the catalytic cysteine residue in the phosphatase active site of PTEN (Lee et al., 2015). BpV(phen), and other structurally related BpV compounds, increase pAKT(Ser473) levels *in vitro* in various cell lines, (Dimchev et al., 2013; Schmid et al., 2004; Zhao et al., 2013) and in *in vivo* models (Ding et al., 2013; Guo et al., 2013; Mao et al., 2013) indicating increased AKT signalling. However, although the BpV compounds inhibit PTEN, they do not increase p-PTEN(Ser380) levels (relative to total PTEN) in any models tested. Additionally, the BpV compounds have also been shown to provide neuroprotective effects in a range of *in vivo* stroke models (Chen et al., 2015; Ding et al., 2013; Guo et al., 2013; Mao et al., 2015).

Vanadium compounds, such as sodium orthovanadate have been recognized as inhibitors of several classes of phosphatase enzymes since the 1970s, in some cases with reasonable potency (eg. human liver alkaline phosphatase Ki <1µM) (Seargeant and Stinson, 1979; VanEtten et al., 1974). Although this broad spectrum phosphatase

inhibition by vanadate appears to be mediated by simple reversible competitive inhibition, a more selective irreversible inhibition of several members of the protein tyrosine phosphatase family appears to be achieved by aqueous peroxovanadium compounds due to oxidation of the active site cysteine thiol (Bevan et al., 1995; Huyer et al., 1997). Subsequently, peroxovanadium compounds such as bisperoxovanadium 1,10 phenanthroline (bpV(phen)) and bisperoxovanadium 5-hydroxypiridine-2-carboxyl (bpV(HOpic)) were studied due to their increased biological potency and evidence that these vanadium complexes have greater target selectivity than the simple vanadate compounds. For example, bpV(phen) and bisperoxovanadium 2-carboxypyridine (bpV(pic)) were shown to inhibit Cdc25A with some selectivity, displaying IC50 determined *in vitro* in the presence of 1mM DTT in the 10–50nM range (Scrivens et al., 2003).

6.2. Aims and Objectives for PI3K/Akt Signalling Pathway in NSC34 (G4C2)102 Cells

1) Validate changes in total Pten and total Akt protein levels in NSC34 (G4C2)102 cells.
2) Validate changes in PI3K/Akt signalling in NSC34 (G4C2)102 cells.
3) Use lentiviral Pten shRNA vectors to knock-down Pten and rescue (G4C2)102 induced toxicity.
4) Use Pten inhibitor BpV(phen) to activate PI3K/Akt signalling pathway and rescue (G4C2)102 induced toxicity.

6.3. Results

6.3.1. Akt, p-Akt(Ser473), Pten and p-Pten(Ser380) Levels are Unaffected by (G4C2)102 Expression in NSC34 Cells

Immunoblotting was used to validate the DE of Akt and Pten transcripts after 5 days tetracycline induction (section 5.3.3.4.1), and also to investigate whether the PI3K/Akt signalling pathway was dysregulated at a biochemical level in the NSC34 (G4C2)102 cells.

NSC34 sham and NSC34 (G4C2)102 cells were grown for 5 days with or without tetracycline and then immunoblotted for total Akt, and Pten, as well as p-Akt(Ser473) and p-Pten(Ser380) phosphor-isomers. The anti-Akt and anti-p-Akt(Ser473) antibodies detected a protein band at 60kDa (Figure 6.2A), which is the expected molecular weight. There was no significant difference in total Akt levels between tetracycline treated and untreated NSC34 sham cells (Figure 6.2). Neither was there any significant difference in total Akt levels between tetracycline treated NSC34 (G4C2)102 cells and either untreated NSC34 (G4C2)102 or treated NSC34 sham cells (Figure 6.2). Also, there was no significant difference in p-Akt(Ser473) levels between tetracycline treated and untreated NSC34 sham cells, when normalised to either β-actin or total Akt levels (Figure 6.2). Finally, there were slight increases in p-Akt(Ser473) levels in NSC34 (G4C2)102 cells compared to NSC34 sham cells, although these comparisons were not statistically significant. p-Akt(Ser473) levels were increased by 15.4 ± 9.5% (ns) and 29.3 ± 21.1% (ns) in untreated NSC34 (G4C2)102 cells compared to untreated NSC34 sham cells, when normalised to β-actin or total Akt levels respectively (Figure 6.2). p-Akt(Ser473) levels were also increased by 35.1 ± 41.0% (ns) and 30.0 ± 41.4% (ns) in tetracycline treated NSC34 (G4C2)102 cells compared to tetracycline treated NSC34 sham cells, when normalised to β-actin or total Akt levels respectively (Figure 6.2).

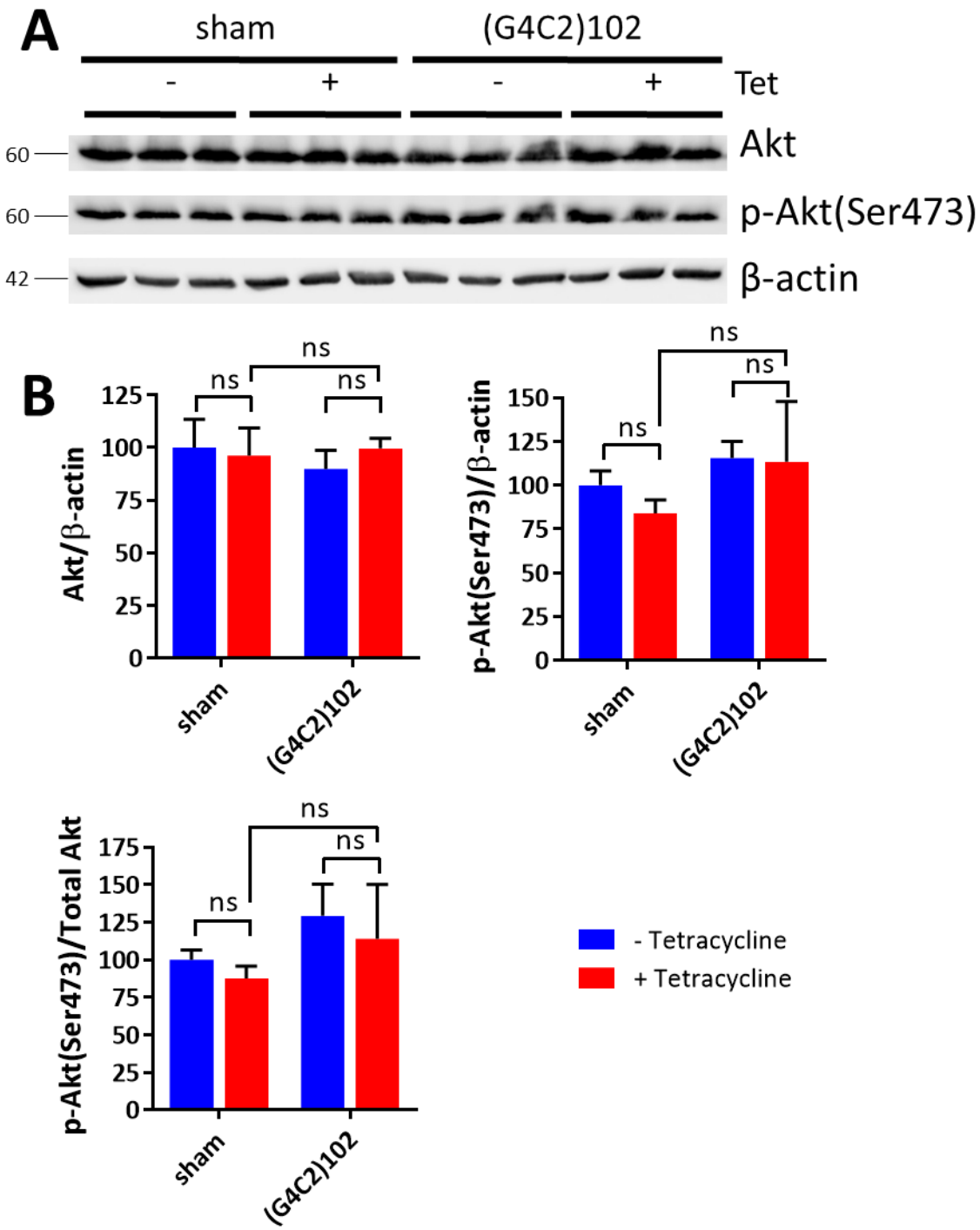


Figure 6.2 (G4C2)102 expression does not affect total Akt or phospho-Akt(Ser473) levels in NSC34 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 5 days with or without 0.5µg/mL tetracycline. Cells were lysed and immunoblotted with anti-Akt, anti-p-Akt(Ser473) and anti-β-actin. A representative immunoblot is shown. Molecular weight markers are indicated (kDa). **B)** Quantification of Akt normalised to β-actin, and p-Akt(Ser473) normalised to β-actin or Akt (Two-way ANOVA with Tukey's multiple comparisons post hoc test; Data are means ± SD; n = 3).

The anti-Pten and anti-p-Pten(Ser380) antibodies detected a protein band at 54 kDa (Figure 6.3A), which is the expected molecular weight. There was no significant difference in total Pten levels between tetracycline treated and untreated NSC34 sham cells (Figure 6.3). However, total Pten levels were increased by 23.1 ± 40.8% (ns) in the untreated NSC34 (G4C2)102 cells and 24.1 ± 33.6% (ns) in the tetracycline induced NSC34 (G4C2)102 cells when compared to untreated and tetracycline treated NSC34 sham cells respectively, although these increases were not statistically significant (Figure 6.3). Similarly, there was no significant difference in p-Pten(Ser380) levels between tetracycline treated and untreated NSC34 sham cells, when normalised to either β-actin or total Pten levels (Figure 6.3). However, p-Pten(Ser380) levels (normalised to β-actin) were increased by 15.1 ± 32.6% (ns) in the untreated NSC34 (G4C2)102 cells and 31.9 ± 26.0% (ns) in the tetracycline induced NSC34 (G4C2)102 cells when compared to untreated and tetracycline treated NSC34 sham cells respectively, although these increases were not statistically significant (Figure 6.3). Finally, the ratio of p-Pten(Ser380) to total Pten was not significantly different in either the untreated NSC34 (G4C2)102 cells or tetracycline treated NSC34 (G4C2)102 cells compared to the untreated and tetracycline treated NSC34 sham cells respectively (Figure 6.3).

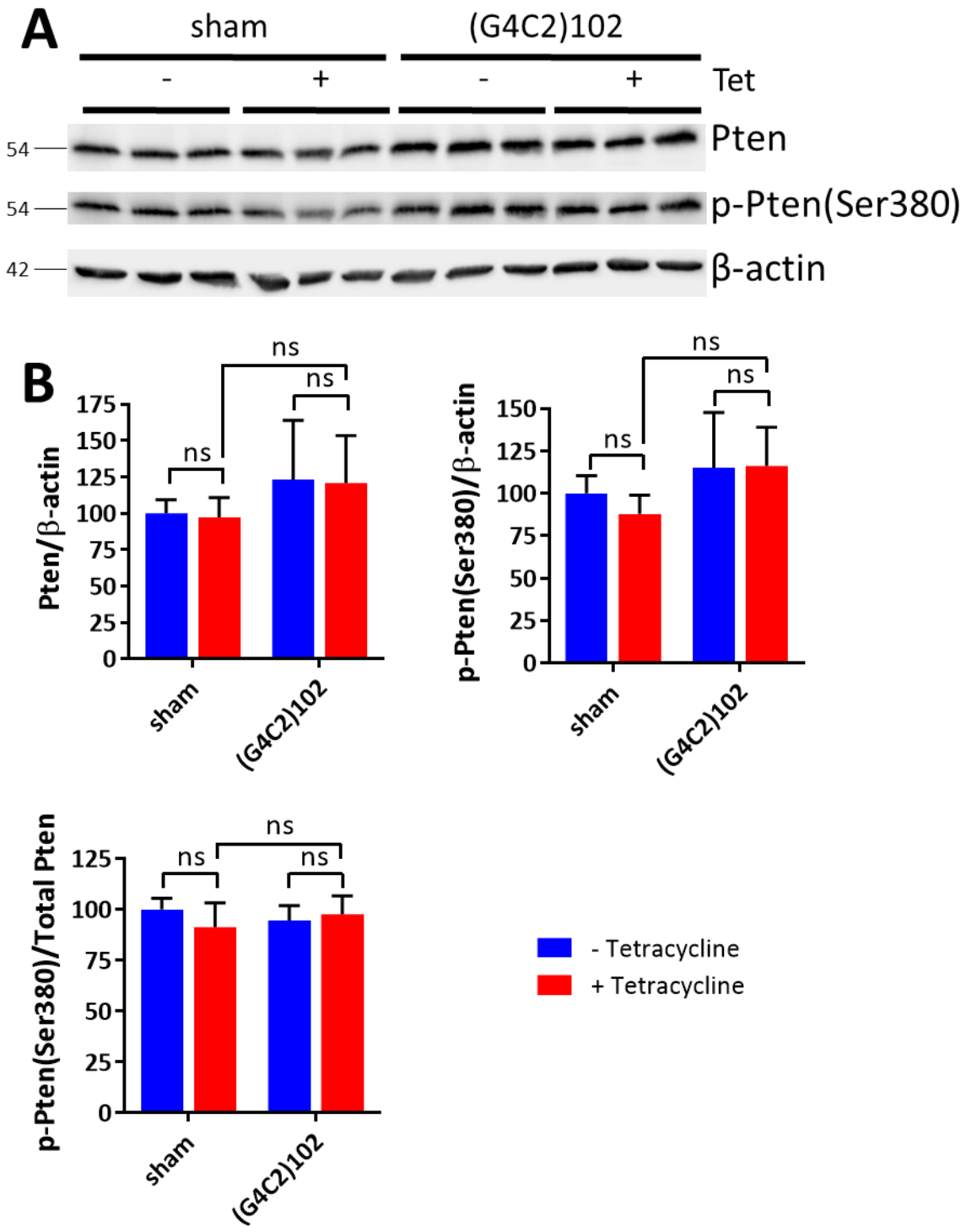


Figure 6.3 (G4C2)102 expression does not affect total Pten or phospho-Pten(Ser380) levels in NSC34 cells. A) NSC34 sham and NSC34 (G4C2)102 cells were cultured for 5 days with or without 0.5µg/mL tetracycline. Cells were lysed and immunoblotted with anti-Pten, anti-p-Pten(Ser380) and anti-β-actin. A representative immunoblot is shown. Molecular weight markers are indicated (kDa). **B)** Quantification of Pten normalised to β-actin, and p-Pten(Ser380) normalised to β-actin or Pten (Two-way ANOVA with Tukey's multiple comparisons post hoc test; Data are means ± SD; n = 3).

6.3.2. Pten Inhibition Using a Lentiviral Vector Encoding Pten shRNA

6.3.2.1. Puromycin Dose-Response in NSC34 Sham and (G4C2)102 Cells

Pten was knocked down by stably integrating shRNA into the NSC34 sham and (G4C2)102 cell genome by lentiviral delivery. The lentiviral vectors used all confer puromycin resistance, and therefore cells containing stably integrated lentivirus can be selected using puromycin selection. Therefore, the minimum concentration of puromycin that would kill NSC34 sham and NSC34 (G4C2)102 had to be identified. NSC34 sham cells were cultured for 7 days with various concentrations of puromycin, and then an MTT assay was performed to assess cell viability. ≤0.01µg/mL puromycin did not affect NSC34 sham or NSC34 (G4C2)102 cell viability (Figure 6.4). 0.05 and 0.1µg/mL puromycin partially reduced NSC34 sham and NSC34 (G4C2)102 cell viability, whilst ≥0.5µg/mL puromycin completely devastated NSC34 sham and NSC34 (G4C2)102 cells (Figure 6.4). Therefore, puromycin was used at 1µg/mL to select for stable lentiviral transduced NSC34 cells.

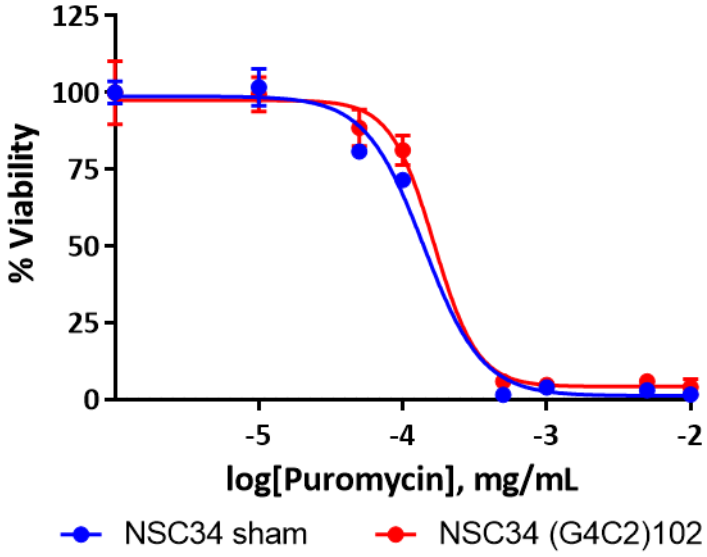


Figure 6.4 Dose-response of NSC34 sham and NSC34 (G4C2)102 cell viability against puromycin. NSC34 sham cells were cultured for 7 days with various concentrations of puromycin. Cell viability was measured using an MTT assay (Data shown are means and SD; n = 1).

6.3.2.2. Lentiviral ShRNA Vectors Reduced Total Pten Levels in Stably Transduced NSC34 Sham and (G4C2)102 Cells

NSC34 sham and NSC34 (G4C2)102 cells were transduced with either GFP control lentiviral particles or Pten shRNA (mouse) lentiviral particles. The media was removed and replenished 24 hours after transduction. Puromycin selection was added to the media to select for cells which had the lentivirus stably integrated into the genome. Over the next week there was a lot of cell death, but visible puromycin-resistant colonies could be seen growing. The stable transformants were expanded as mixed populations, and frozen down. Later, in a separate procedure, control shRNA lentiviral particles were also stably transduced into NSC34 sham and NSC34 (G4C2)102 cells, as described above. The second transduction using the control shRNA was performed because initial experiments using the GFP control produced encouraging results, and a better control was desired for the Pten shRNA. The GFP transduced cells only controlled for viral transduction, whilst the control shRNA transduced cells additionally controlled for shRNA expression, and therefore represent a more robust control for the Pten shRNA transduced cells.

The effect of the stably transduced Pten shRNA was assessed by immunoblot. Pten was knocked down by 63.0 ± 14.4% ($P<0.05$) in the NSC34 sham Pten shRNA compared to NSC34 sham GFP control cells (Figure 6.5). Whilst Pten was knocked down by 62.5 ± 22.2 ($P<0.01$) in the NSC34 (G4C2)102 Pten shRNA compared to NSC34 (G4C2)102 GFP control cells (Figure 6.5), there was no significant difference in the Pten levels between NSC34 sham GFP and NSC34 (G4C2)102 GFP control cells, or NSC34 sham Pten shRNA compared to NSC34 (G4C2)102 Pten shRNA cells (Figure 6.5).

In a separate set of experiments, the effect of Pten shRNA was also assessed compared to the control shRNA cell lines. Unexpectedly however, the NSC34 sham Pten shRNA cells no longer showed a significant Pten knock-down, and were removed from later experiments where the control shRNA cell lines were used as the control rather than the GFP lines. Pten was knocked down by 70.0 ± 4.5% ($P<0.01$) in the NSC34 (G4C2)102 Pten shRNA compared to NSC34 (G4C2)102 control shRNA cells (Figure 6.6).

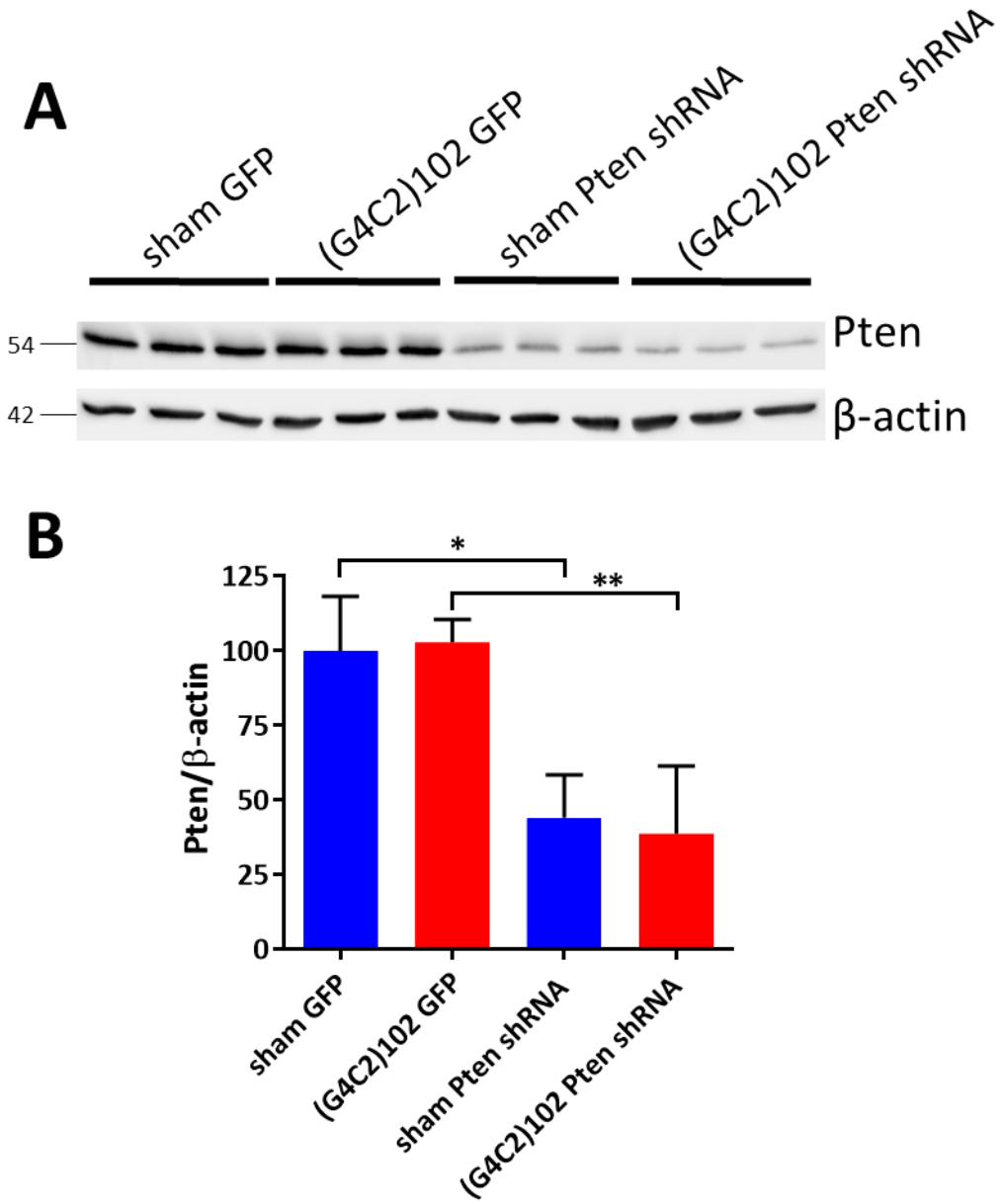


Figure 6.5 Pten shRNA reduces Pten levels in stably transduced NSC34 sham and (G4C2)102 cells. A) NSC34 sham and (G4C2)102 cells were stably transduced with either LV GFP or Pten shRNA LV. Cells were lysed and immunoblotted with anti-Pten and anti-β-actin. A representative immunoblot is shown. Molecular 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 are means ± SD; n=3).

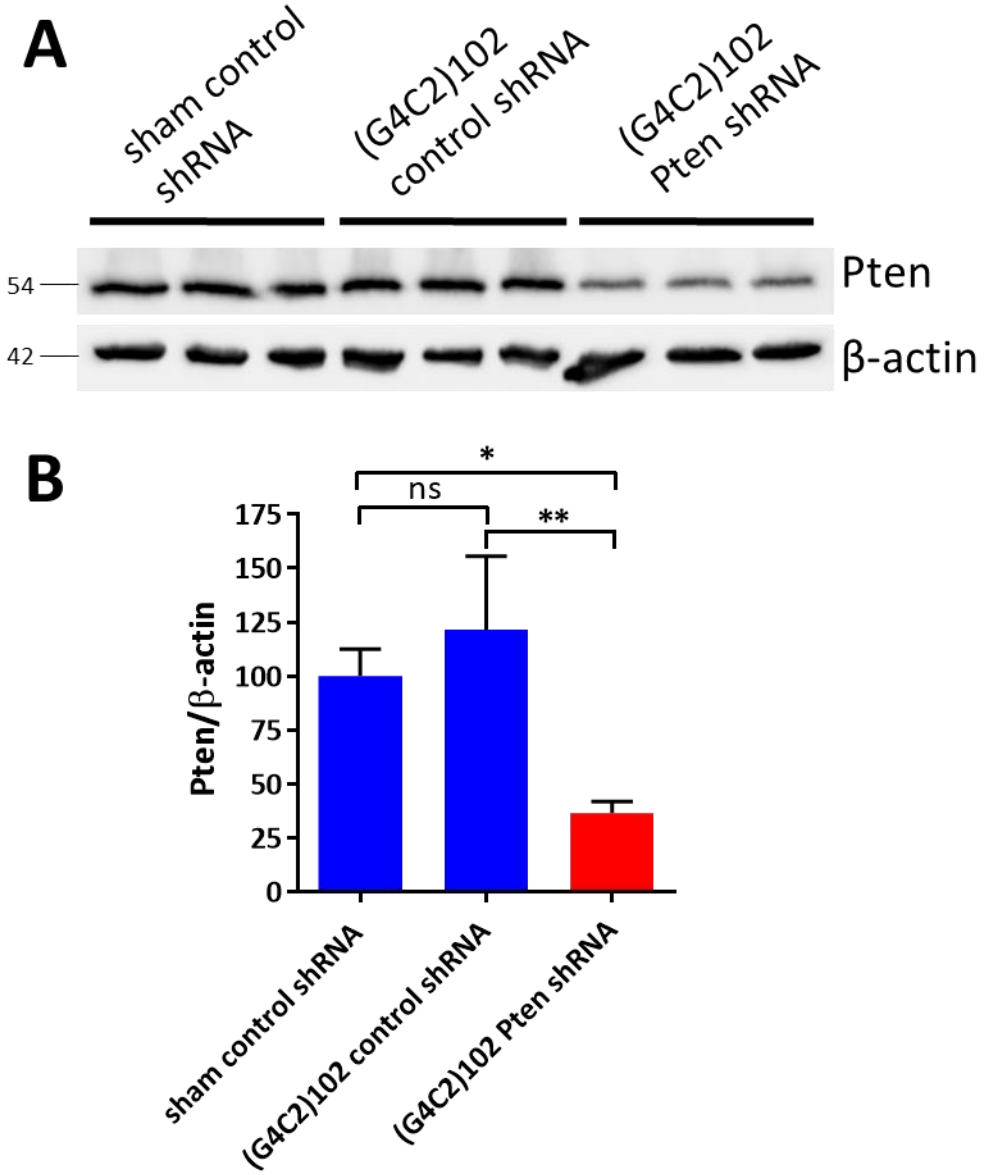


Figure 6.6 Pten shRNA reduces Pten levels in stably transduced NSC34 (G4C2)102 cells. **A)** NSC34 sham and (G4C2)102 cells were stably transduced with either LV control shRNA or Pten shRNA LV. Cells were lysed and immunoblotted with anti-Pten and anti-β-actin. A representative immunoblot is shown. Molecular 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 are means ± SD; n=3).

6.3.2.3. Pten Knock-Down Rescues NSC34 Cells from (G4C2)102 Induced Toxicity

The hypothesis was that Pten knock-down would rescue the NSC34 (G4C2)102 cells from the (G4C2)102 induced toxicity (which had previously been described in section 3.3.4.5).

A rescue assay was performed that was modelled on the MTT cell viability assay performed in section 3.3.4.5, which showed (G4C2)102 expression reduced NSC34 cell viability after 7 days of tetracycline induction. NSC34 sham GFP, NSC34 sham Pten shRNA, NSC34 (G4C2)102 GFP and NSC34 (G4C2)102 Pten shRNA cells were cultured for 7 days with or without tetracycline, and then cell viability was assessed using an MTT assay. Neither NSC34 sham GFP control cell nor NSC34 sham Pten shRNA cell viability was affected by tetracycline induction (Figure 6.7). As expected, NSC34 (G4C2)102 GFP cell viability was reduced by 16.9 ± 11.3% ($P<0.01$) when cells were induced with tetracycline compared to NSC34 (G4C2)102 GFP cell untreated with tetracycline (Figure 6.7). However, there was no significant reduction in NSC34 (G4C2)102 Pten shRNA cell viability when cells were induced with tetracycline compared to NSC34 (G4C2)102 Pten shRNA cells untreated with tetracycline (Figure 6.7), indicating Pten knock-down provides a rescue effect against the (G4C2)102 induced toxicity.

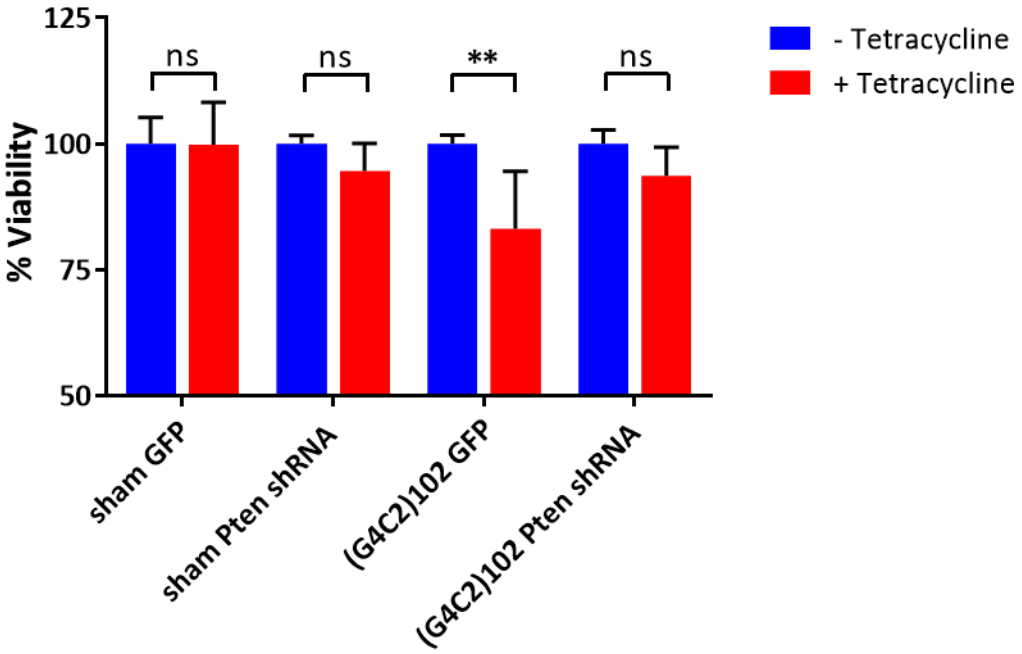


Figure 6.7 Pten knock-down rescues (G4C2)102 induced toxicity in NSC34 cells. NSC34 sham GFP, sham Pten shRNA, (G4C2)102 GFP and (G4C2)102 cells Pten shRNA cells were cultured for 7 days with or without 0.5µg/mL tetracycline. Cell viability was measured using an MTT assay (**P<0.01; Two-way ANOVA with Sidak's multiple comparisons post hoc test; Data are means ± SD; n=4).

In a similar set of experiments, NSC34 sham control shRNA, NSC34 (G4C2)102 control shRNA and NSC34 (G4C2)102 Pten shRNA cells were cultured for 7 days with or without tetracycline, and then cell viability was assessed using an MTT assay. NSC34 sham GFP control cell viability was not affected by tetracycline induction (Figure 6.8). NSC34 (G4C2)102 control shRNA cell viability was reduced by 15.0 ± 5.0% ($P<0.01$) when cells were induced with tetracycline compared to NSC34 (G4C2)102 control shRNA cells untreated with tetracycline (Figure 6.8). In a repeat of the above experiments, there was no significant reduction in NSC34 (G4C2)102 Pten shRNA cell viability when cells were induced with tetracycline compared to NSC34 (G4C2)102 Pten shRNA cells untreated with tetracycline (Figure 6.8), further demonstrating Pten knock-down provides a rescue effect against the (G4C2)102 induced toxicity.

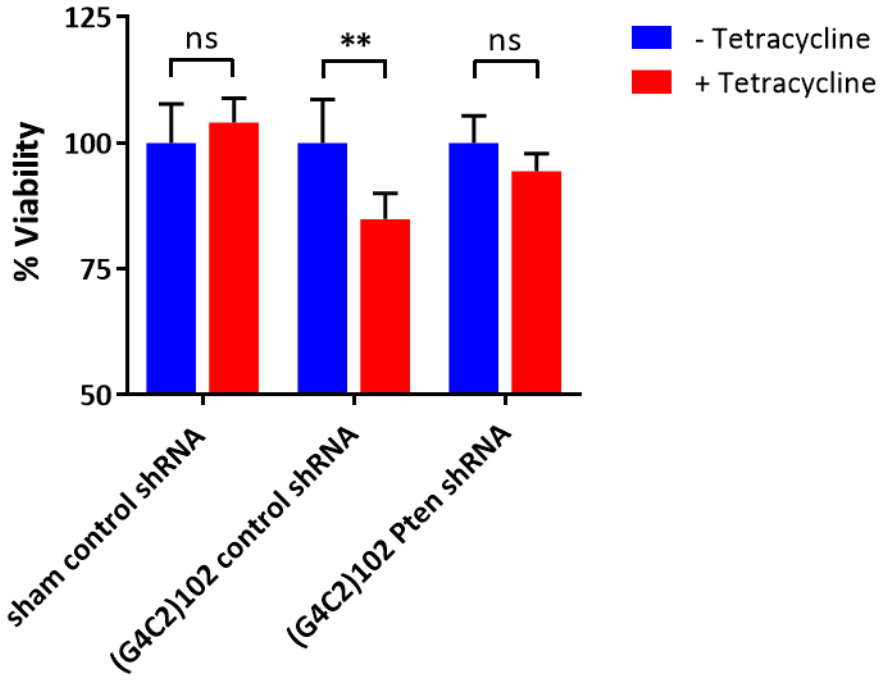


Figure 6.8 Pten knock-down stops (G4C2)102 induced toxicity in NSC34 cells. NSC34 sham GFP, sham Pten shRNA, (G4C2)102 GFP and (G4C2)102 cells Pten shRNA cells were cultured for 7 days with or without 0.5µg/mL tetracycline. Cell viability was measured using an MTT assay (**P<0.01; Two-way ANOVA with Sidak's multiple comparisons post hoc test; Data are means ± SD; n=4).

6.3.2.4. Akt Signalling Analysis in NSC34 Pten ShRNA Cells

Pten knock-down was hypothesised to increase Akt signalling in the NSC34 cells. NSC34 sham control shRNA, NSC34 (G4C2)102 control shRNA, and NSC34 (G4C2)102 Pten shRNA cells were cultured for 5 days with or without tetracycline, and then immunoblotted for total Akt and p-Akt(Ser473). There was no significant difference in total Akt levels between tetracycline induced NSC34 sham control shRNA, NSC34 (G4C2)102 control shRNA or NSC34 (G4C2)102 Pten shRNA compared to their respective control cells untreated with tetracycline (Figure 6.9). In addition, there was no significant difference in total Akt levels between either the NSC34 (G4C2)102 control shRNA cells or the NSC34(G4C2)102 Pten shRNA cells compared to the NSC34 sham control shRNA cells (Figure 6.9). Lastly, there was no significant difference in total Akt levels between the NSC34 (G4C2)102 control shRNA cells and the NSC34(G4C2)102 Pten shRNA cells (Figure 6.9).

There was no significant difference in p-Akt(Ser473) levels (normalised to both β-actin and total Akt) between tetracycline induced NSC34 sham control shRNA, NSC34 (G4C2)102 control shRNA or NSC34 (G4C2)102 Pten shRNA compared to their respective control cells untreated with tetracycline (Figure 6.9). However, there was significantly more p-Akt(Ser473) (normalised to both β-actin and total Akt) in NSC34 (G4C2)102 control shRNA cells compared to NSC34 sham control shRNA cells ($P<0.001$) (Figure 6.9). Also, there was significantly more p-Akt(Ser473) (normalised to both β-actin and total Akt) in NSC34 (G4C2)102 Pten shRNA cells compared to NSC34 sham control shRNA cells ($P<0.001$) (Figure 6.9). There was no significant difference in p-Akt(Ser473) (normalised to both β-actin and total Akt) in NSC34 (G4C2)102 Pten shRNA cells compared to NSC34 (G4C2)102 control shRNA cells however (Figure 6.9).

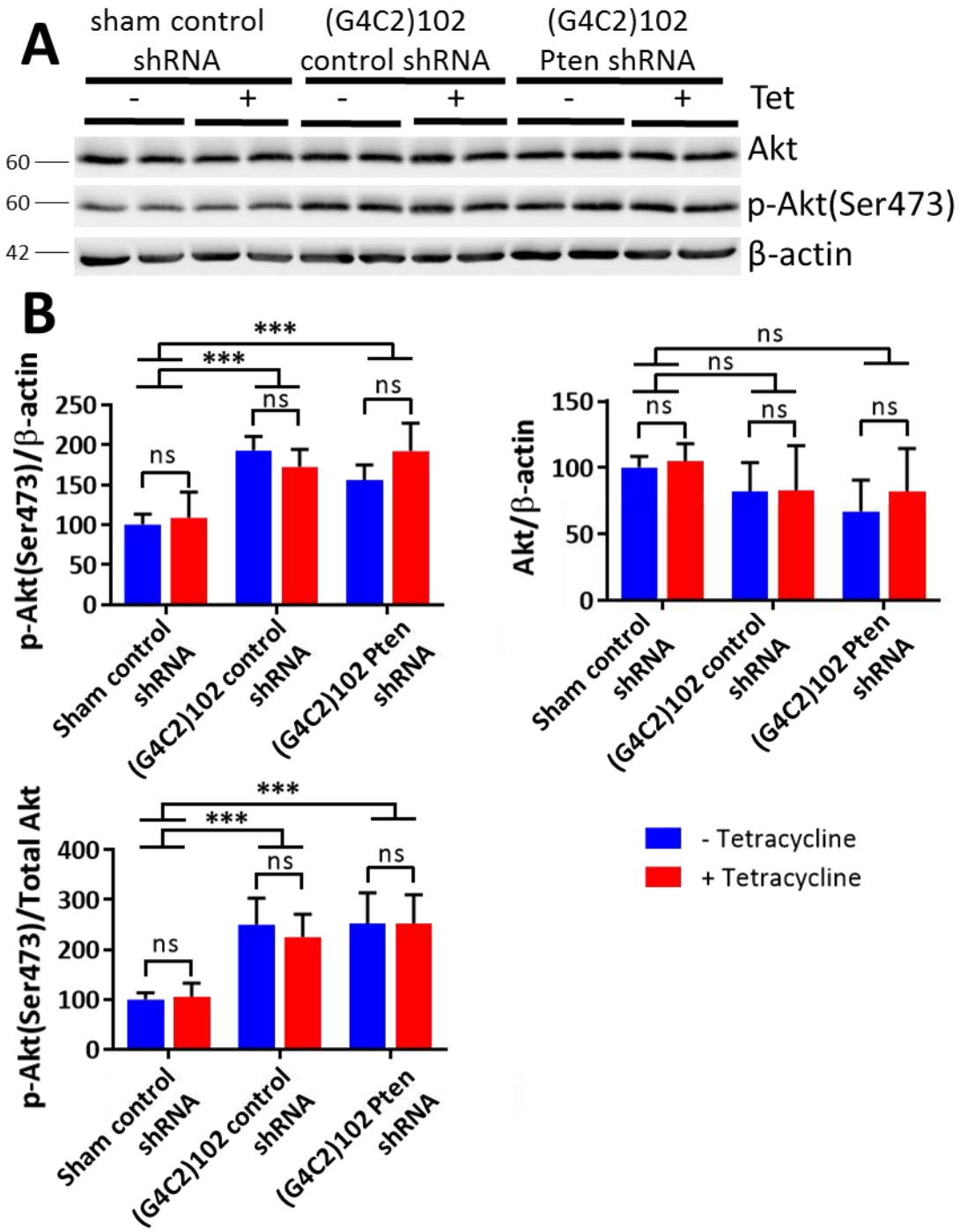


Figure 6.9 p-Akt levels are increased in NSC34 (G4C2)102 cells stably transduced with Lentivirus compared to NSC34 sham cells stably transduced with Lentivirus. A) NSC34 sham control shRNA, (G4C2)102 control shRNA and (G4C2)102 Pten shRNA cells were cultured for 5 days with or without 0.5µg/mL tetracycline. Cells were lysed and immunoblotted with anti-Akt, anti-p-Akt(Ser473) and anti-β-actin. A representative immunoblot is shown. Molecular weight markers are indicated (kDa). **B)** Quantification of Akt normalised to β-actin, and p-Akt(Ser473) normalised to β-actin or Akt (***P<0.001; Two-way ANOVA with Tukey's multiple comparisons post hoc test; Data are means ± SD; n=3).

From these results, the following can be concluded: 1) total Akt levels are unaffected by Pten knock-down; 2) total Akt levels are unaffected by tetracycline induction and/or increased (G4C2)102 expression; 3) p-Akt(Ser473) levels are increased in NSC34 cells that contain both the genomic (G4C2)102 repeat and a stably integrated lentiviral vector; 4) (G4C2)102 expression does not affect p-Akt(Ser473) levels in the lentiviral transduced NSC34 cells; 5) Pten knock-down does not measurably alter p-Akt(Ser473) levels at the 5 day timepoint.

6.3.3. Pten Inhibition Using BpV(phen)

6.3.3.1. BpV(phen) Viability Dose-Response in NSC34 Sham Cells

To validate the Pten knock-down rescue effect in the shRNA lentiviral transduced NSC34 (G4C2)102 cells, Pten inhibitors were used to pharmacologically inhibit Pten activity. The hypothesis was to treat NSC34 (G4C2)102 cells with the potent Pten inhibitor BpV(phen) (Schmid et al., 2004), to see whether Pten inhibition, and Akt signalling activation, would rescue the (G4C2)102 induced toxicity. Firstly however, a dose-response of BpV(phen) concentration and NSC34 sham cell viability was performed to identify a sub-lethal dose of BpV(phen) to use in these rescue assays. NSC34 sham cells were cultured for 6 days with various concentrations of BpV(phen), and then an MTT assay was performed to assess cell viability. ≤1µM BpV(phen) did not affect NSC34 sham cell viability, whilst ≥3µM BpV(phen) completely devastated NSC34 sham cells (Figure 6.10). This was unsurprising since the cells appeared dead upon visual observation. 1µM BpV(phen) was taken forward as the highest sub-lethal dose.

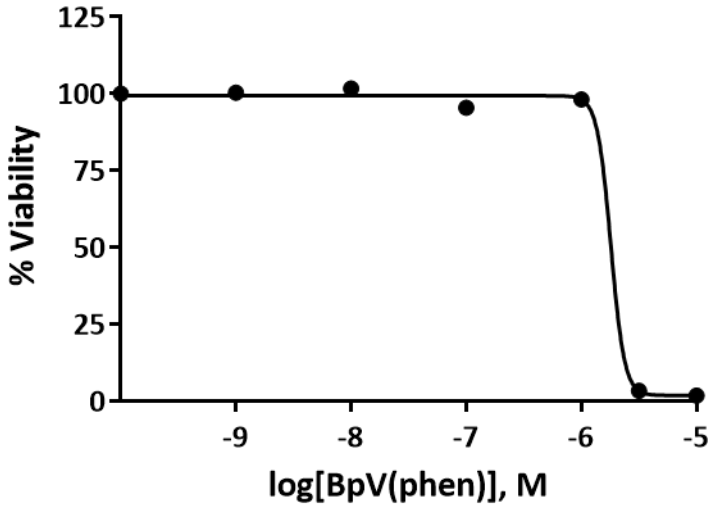


Figure 6.10 Dose-response of NSC34 sham cell viability against BpV(phen). NSC34 sham cells were cultured for 6 days with various concentrations of BpV(phen). Cell viability was measured using an MTT assay (data shown are mean and SD; n=1).

6.3.3.2. 1 µM BpV(phen) Effect on Pten and Akt Activity in NSC34 Sham Cells

1µM was the highest sublethal dose of BpV(phen) tested in the NSC34 sham cells, and this concentration was taken forward to establish whether BpV(phen) could inhibit Pten and activate Akt in the NSC34 sham cells. To do this, NSC34 sham cells were treated with 1µM BpV(phen) for various lengths of time, and then immunoblotted for total Akt, p-Akt(Ser473), total Pten and p-Pten(Ser380). There was no significant difference in p-Pten(Ser380) levels (normalised to either β-actin or total Pten) between the untreated NSC34 sham cells and the NSC34 sham cells treated with BpV(phen) for 12, 24 or 48 hours (Figure 6.11). There was no significant difference in p-Akt(Ser473) levels (normalised to either β-actin or total Akt) between the untreated NSC34 sham cells and the NSC34 sham cells treated with BpV(phen) for 12, 24 or 48 hours (Figure 6.12).

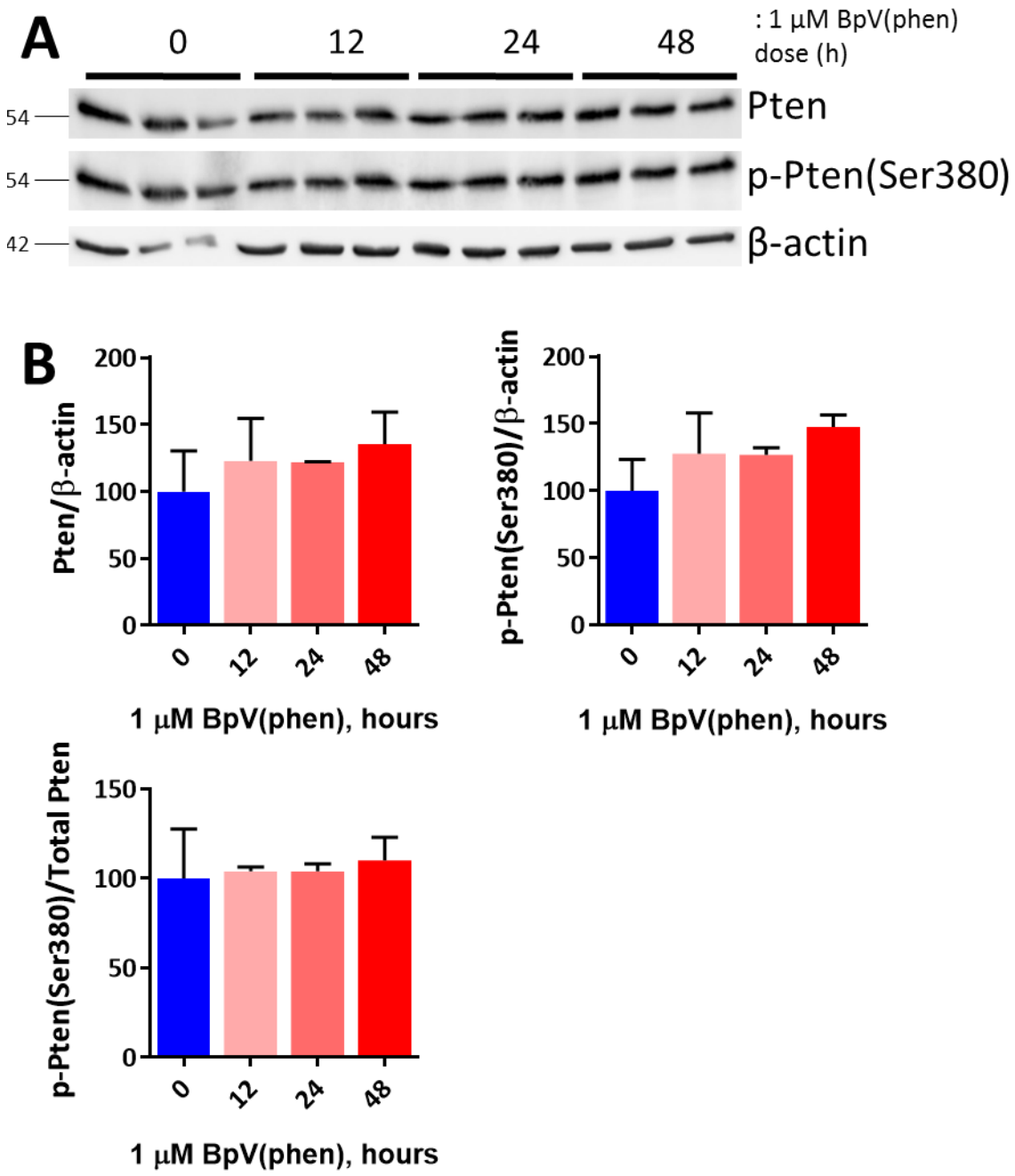


Figure 6.11 Time-response of Pten inhibition by 1µM BpV(phen) in NSC34 sham cells. A) NSC34 sham cells were cultured for various lengths of time with 1µM BpV(phen). Cells were lysed and immunoblotted with anti-Akt, anti-p-Akt(Ser473) and anti-β-actin. A representative immunoblot is shown. Molecular weight markers are indicated (kDa). **B)** Quantification of Pten normalised to β-actin, and p-Pten(Ser380) normalised to β-actin or Pten (One-way ANOVA with Sidak's multiple comparisons post hoc test; Data shown are mean and SD; n=3).

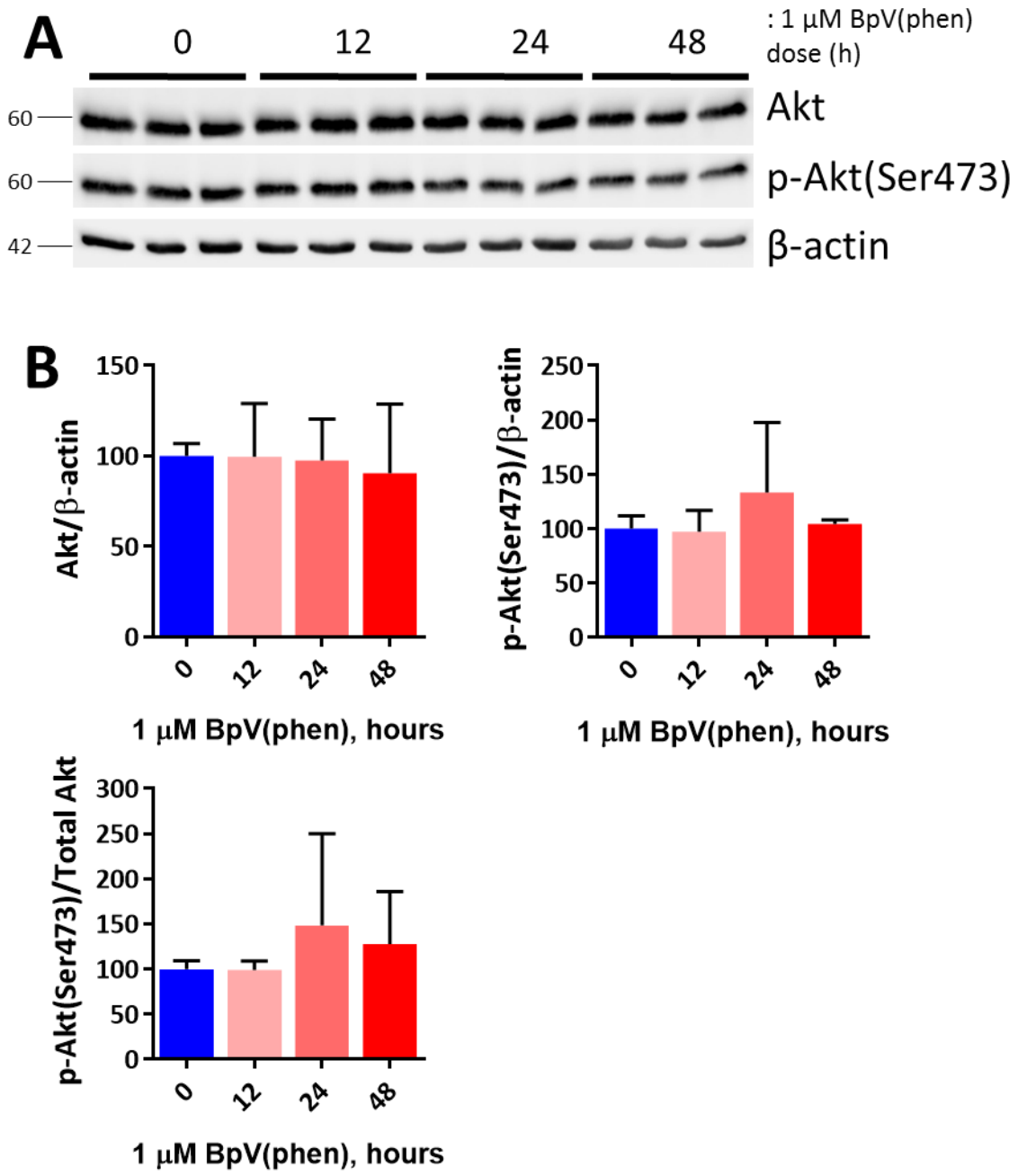


Figure 6.12 Time-response of Akt activation by 1µM BpV(phen) in NSC34 sham cells.NSC34 sham cells were cultured for various lengths of time with 1µM BpV(phen). Cells were lysed and immunoblotted with anti-Akt, anti-p-Akt(Ser473) and anti-β-actin. A representative immunoblot is shown. Molecular weight markers are indicated (kDa). **B)** Quantification of Akt normalised to β-actin, and p-Akt(Ser473) normalised to β-actin or Akt (One-way ANOVA with Sidak's multiple comparisons post hoc test; Data shown are means and SD; n=3).

6.3.3.3. Higher Dose (Short Time) Effect on Pten and Akt Activities in NSC34 Sham Cells

≥1µM doses of BpV(phen) were tested for Akt activation and Pten inhibition in the NSC34 sham cells, because 1µM BpV(phen) did not elicit an increase in p-Akt(Ser473) levels or p-Pten(Ser380) levels. NSC34 cells were cultured with various concentrations of BpV(phen) for 2 days, and then immunoblotted for total Akt, p-Akt(Ser473), total Pten and p-Pten(Ser380). The higher the BpV(phen) dose, the lower the p-Pten(Ser380) levels (normalised to either β-actin or total Pten) were in the NSC34 sham cells (Figure 6.13). However, p-Akt(Ser473) levels (normalised to either β-actin or total Akt) were far greater in the NSC34 sham cells treated with 3µM BpV(phen) compared to cells treated with ≤1µM BpV(phen) (Figure 6.14). This suggested that 3µM BpV(phen) did activate the Akt signalling pathway, and was therefore re-tested for toxicity in the NSC34 sham cells at shorter dose lengths. To note however, this experiment was only performed once.

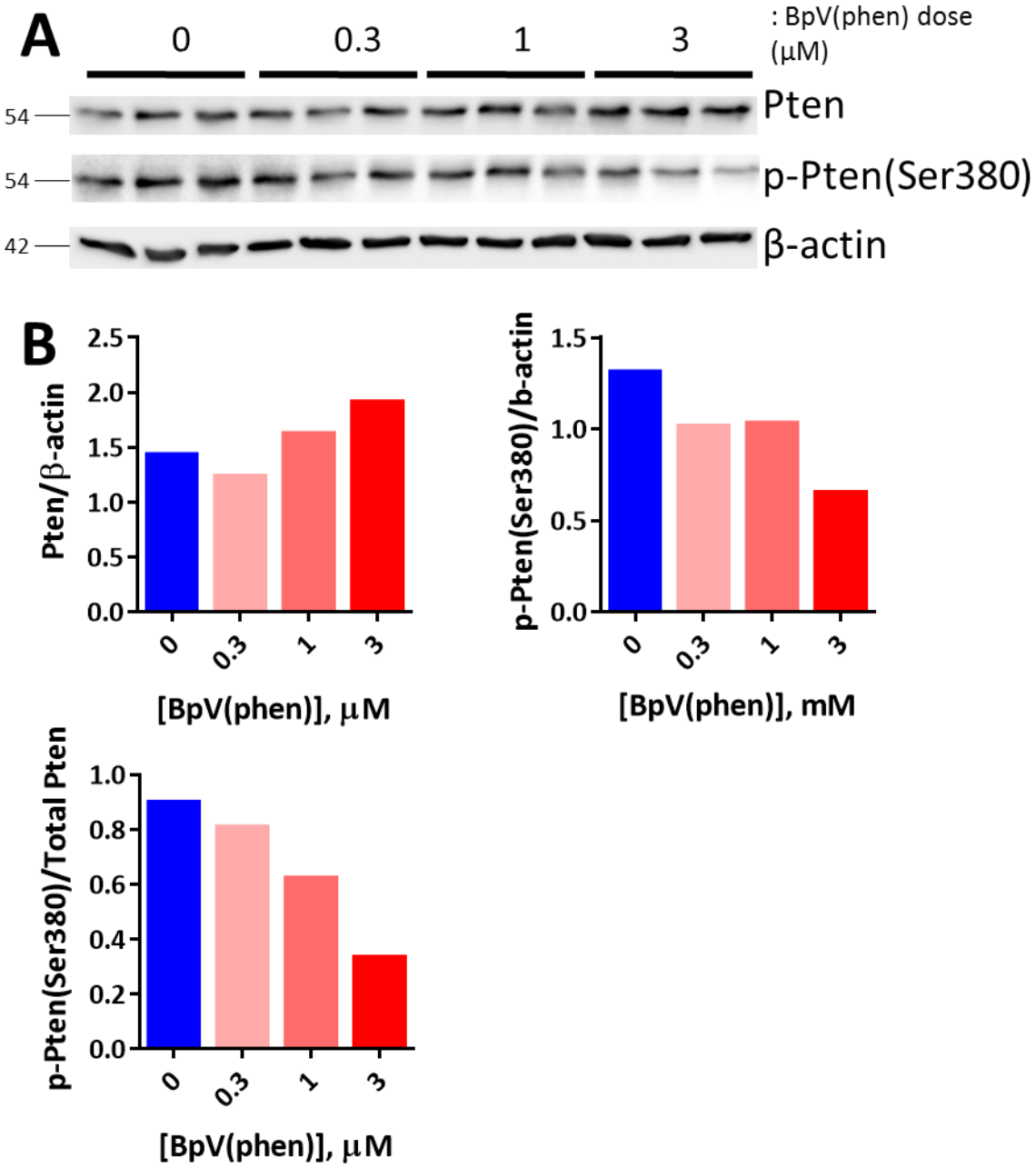


Figure 6.13 3μM BpV(phen) decreases p-Pten levels in NSC34 sham cells. A) NSC34 sham cells were cultured with various concentrations of BpV(phen). Cells were lysed and immunoblotted with anti-Pten, anti-p-Pten(Ser380) and anti-β-actin. Molecular weight markers are indicated (kDa). **B)** Quantification of Pten normalised to β-actin, and p-Pten(Ser380) normalised to β-actin or Pten (Data are means; n=1).

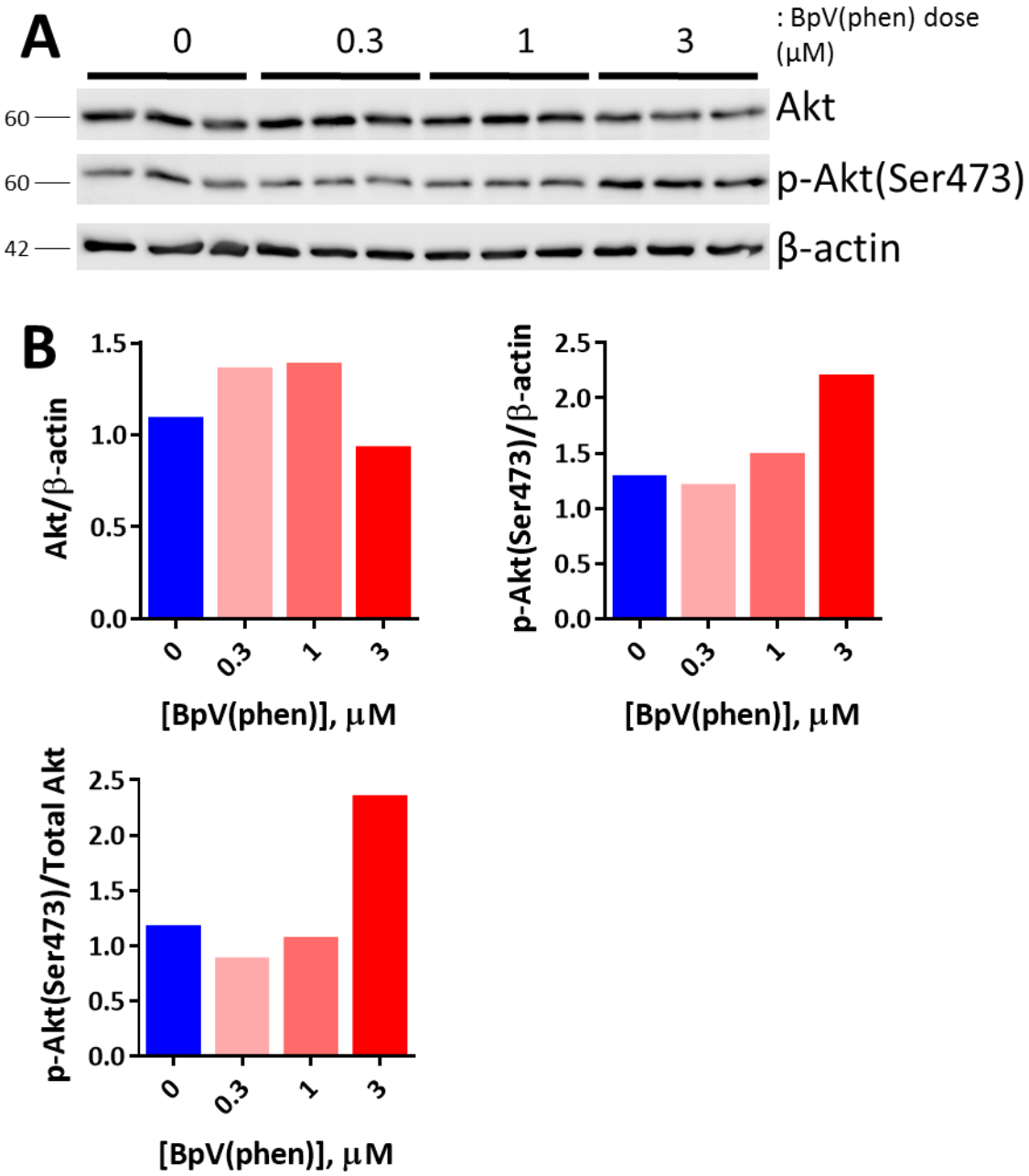


Figure 6.14 3μM BpV(phen) increases p-Akt levels in NSC34 sham cells. A) NSC34 sham cells were cultured with various concentrations of BpV(phen). Cells were lysed and immunoblotted with anti-Akt, anti-p-Akt(Ser473) and anti-β-actin. Molecular weight markers are indicated (kDa). **B)** Quantification of Akt normalised to β-actin, and p-Akt(Ser473) normalised to β-actin or Akt. n=1.

6.3.3.4. Time-Response of NSC34 Sham Cell Viability Against 3μM BpV(phen)

NSC34 sham cells were cultured for one week, and 3μM BpV(phen) was added to the media for various lengths of time. An MTT assay was then performed to assess the NSC34 sham cell viability. The NSC34 sham cell viability was not affected after 6 hours

of 3 µM BpV(Phen) treatment (Figure 6.15). However, NSC34 sham cell viability was reduced by 18, 73 and 87% after 24, 48 and 144 hours of 3µM BpV(Phen) treatment respectively (Figure 6.15). Therefore, NSC34 (G4C2)102 cells would be dosed with 3µM BpV(phen) for 6 hours on day 2 and 5 of a 7 day cell viability rescue assay.

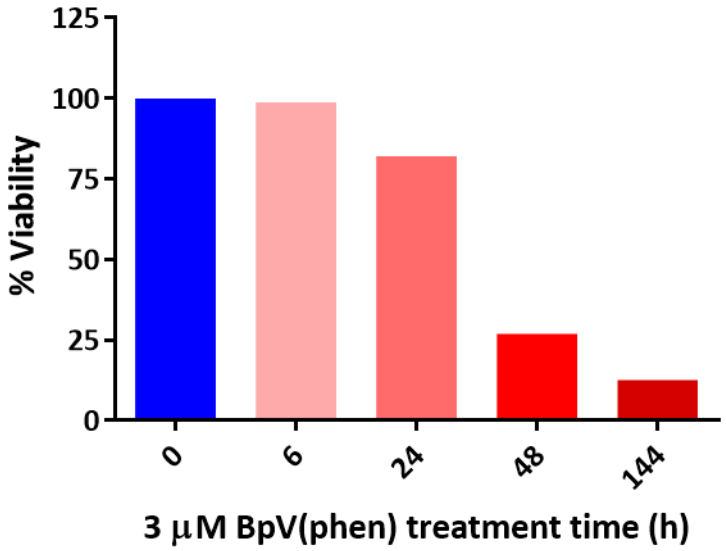


Figure 6.15 Time-response of NSC34 sham cell viability against 3µM BpV(phen). NSC34 sham cells were cultured for various lengths of time with 3 µM BpV(phen). Cell viability was measured using an MTT assay. n=1.

6.3.3.5. BpV(phen) Rescue Assay in NSC34 Sham and NSC34 (G4C2)102 cells

The hypothesis was that Pten inhibition, and Akt signalling activation, would rescue the NSC34 (G4C2)102 cells from the (G4C2)102 induced toxicity (which had previously been described in section 3.3.4.5). A rescue assay was performed that was modelled on the MTT cell viability assay performed in section 3.3.4.5, which showed (G4C2)102 expression reduced NSC34 cell viability after 7 days of tetracycline induction. NSC34 sham and NSC34 (G4C2)102 cells were cultured for 7 days with or without tetracycline, and cells were additionally dosed with either 3µM BpV(phen) or a vehicle control for 6 hours on days 2 and 5. NSC34 cells were given 24 hours to adhere to the plates prior to BpV(phen) treatment for two reasons. Firstly, because the cells could be more susceptible to any BpV(phen) induced toxicity when not adhered to the plate. Secondly, because the media was removed and replaced after 6 hours of BpV(phen) (or vehicle

control) dosing, an initial 24 hours was necessary to ensure all NSC34 had a chance to adhere before any media was removed.

NSC34 sham cell viability was not affected by tetracycline induction in cells that were dosed with the vehicle control (Figure 6.16). Neither was NSC34 sham cell viability affected by tetracycline induction in cells that were dosed with BpV(phen) (Figure 6.16). However, NSC34 sham cell viability was reduced by 20.7 ± 4.9% (P<0.0001) in cells treated with BpV(phen) but not tetracycline, compared to the vehicle control cells also untreated with tetracycline (Figure 6.16). In addition, NSC34 sham cell viability was reduced by 28.3 ± 2.8% (P<0.0001) in cells treated with BpV(phen) and tetracycline, compared to the vehicle control cells also treated with tetracycline (Figure 6.16).

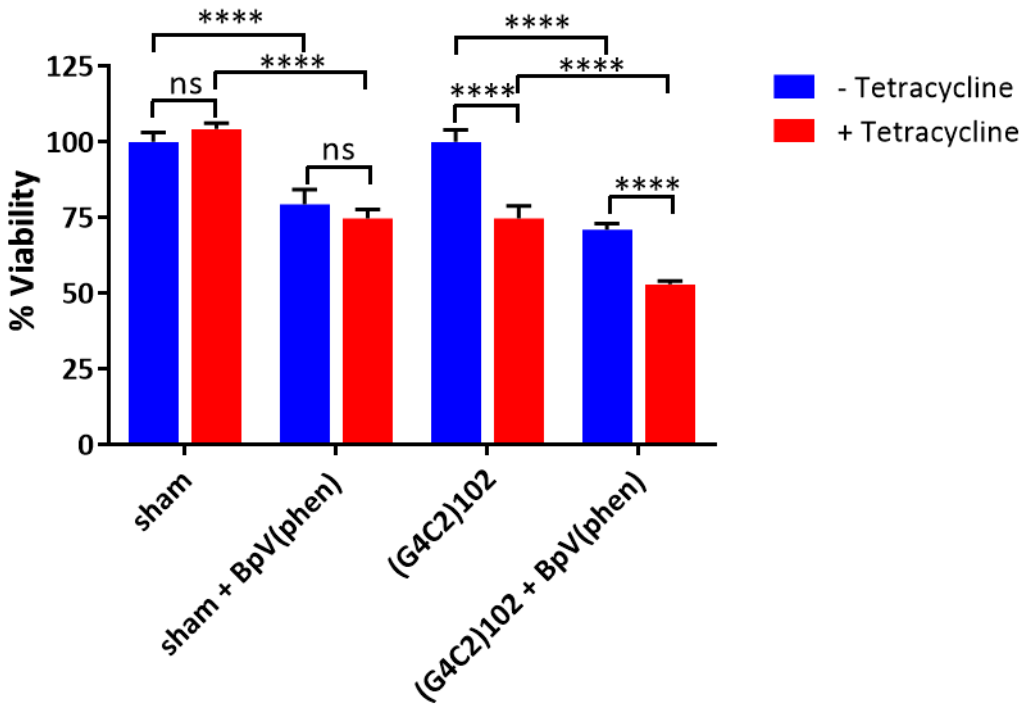


Figure 6.16 BpV(phen) rescue assay of (G4C2)102 toxicity in NSC34 cells. NSC34 sham and (G4C2)102 cells were cultured for 7 days with or without 0.5µg/mL tetracycline. NSC34 cells were dosed with 3µM BpV(phen), or a vehicle control, for 6 h twice (on day 2 and day 5). Cell viability was measured using an MTT assay (****P<0.0001; Two-way ANOVA with Sidak's multiple comparisons post hoc test; Data are means ± SD; n=3).

NSC34 (G4C2)102 cell viability was reduced by 25.3 ± 4.1% (P<0.0001) in tetracycline induced cells compared to NSC34 (G4C2)102 cells untreated with tetracycline, where

both cells were additionally dosed with the vehicle control (Figure 6.16). NSC34 (G4C2)102 cell viability was also reduced by 25.4 ± 1.4% (P<0.0001) in tetracycline induced cells compared to NSC34 (G4C2)102 cells untreated with tetracycline, where both cells were additionally dosed with BpV(phen) (Figure 6.16). Similarly to the NSC34 sham cells, NSC34 (G4C2)102 cell viability was reduced by 29.0 ± 2.0% (P<0.0001) in cells treated with BpV(phen) but not tetracycline, compared to the vehicle control cells also untreated with tetracycline (Figure 6.16). Finally, NSC34 (G4C2)102 cell viability was reduced by 29.1 ± 1.3% (P<0.0001) in cells treated with BpV(phen) and tetracycline, compared to the vehicle control cells also treated with tetracycline (Figure 6.16).

6.4. Discussion

6.4.1. There are no Significant Differences in Pten or Akt Protein levels in NSC34 (G4C2)102 Cells

There was no significant difference in total Akt protein levels in the NSC34 (G4C2)102 cells, suggesting that the (G4C2)102 expression does not affect total Akt expression. This protein result supports the Akt1 qRT-PCR result in section 5.3.3.4.1, which showed that there was no significant difference in Akt1 mRNA levels in the NSC34 (G4C2)102 cells. Neither of these results validate the GEP, which showed Akt1 and Akt2 mRNA levels were decreased by 22 and 21% respectively. Importantly however, there are three different *Akt* genes (*Akt1*, *Akt2*, and *Akt3*) that each express highly similar isoforms of Akt protein. AKT1 is ubiquitously expressed (Hers et al., 2011), AKT2 is highly expressed in muscles and adipose tissue (Cho et al., 2001), whilst AKT3 expression is restricted to the testes and brain (Yang et al., 2003). The polyclonal anti-AKT antibody used in this study recognises Akt protein expressed from all three *Akt* genes, due to the high similarity in the three Akt isoforms. Also, given that NSC34 cells are derived from embryonic spinal cord cells and neuroblastoma cells, Akt3 expression is likely to contribute substantially to total Akt protein levels in the NSC34 cell line. Slight reductions in Akt1 and Akt2 expression may be masked by the potentially unaffected Akt3 expression levels, and therefore have an insignificant effect on overall Akt protein levels, explaining why total Akt protein is the same in the NSC34 (G4C2)102 cells

compared to NSC34 sham cells. Akt1, Akt2, and Akt3 specific monoclonal antibodies could be used to determine the actual levels of each of the Akt isoforms in the NSC34 (G4C2)102 cells.

In addition, there was no significant difference in total Pten protein levels in NSC34 (G4C2)102 cells within the timeframe studied. This again supports the qRT-PCR result in section 5.3.3.4.1, which showed that there was no significant difference in Pten mRNA levels in the NSC34 (G4C2)102 cells. As with total Akt, the Pten immunoblotting and qRT-PCR results do not validate the GEP. However, whilst neither the qRT-PCR nor the immunoblots showed statistically significant alterations, the Pten protein and mRNA levels were both slightly elevated in the NSC34 (G4C2)102 cells compared to NSC34 sham cells, regardless of tetracycline induction time. This does suggest that Pten expression may be slightly elevated in the NSC34 (G4C2)102 cell line, but the increase in Pten is so small the qRT-PCR and immunoblot quantification failed to detect any significance.

The differences in sensitivity between the microarray technology and the qRT-PCR and immunoblot methods used to validate the results may account for the apparent lack of validation. The GeneChip® MTA 1.0 arrays use a median of 30 probes per gene which target regions throughout the mRNA, whilst the qRT-PCR uses only one probe (primer pair). Therefore, mRNA levels are quantified by measuring several or only one small defined region of the mRNA using the microarray and qRT-PCR respectively. This means differences in mRNA splice isoforms could cause differences in the measured mRNA levels using the different techniques. To minimise any differences, primers were selected for qRT-PCR that recognised all splice variants, however, differences in splice variant levels may still affect the microarray measurements. Also, the microarray and qRT-PCR experiments were performed using different experimental samples, which may have contributed to the differences between the microarray and qRT-PCR results. Lastly, for the protein levels, DE at mRNA level does not always result in DE at protein level due to post-transcriptional regulation of gene expression. However, despite the lack of validation using qRT-PCR and immunoblotting, the main strength of the GEP results

come from the fact that the PI3K/Akt pathway shows dysregulation not only in the NSC34 (G4C2)102 but also in the C9ORF72-ALS LCM motor neurons.

6.4.2. PI3K/Akt Signalling is not Altered at the Biochemical Level in NSC34 (G4C2)102 cells

Whilst the total Akt protein levels were unaffected by (G4C2)102 expression, they do not provide a good read out for PI3K/Akt signalling in the cells. Phosphorylation of AKT at Serine 473 is required for AKT activation, and therefore, the p-AKT(Ser473) to total AKT ratio is a better, and widely used, measure of AKT signalling. The p-Akt(Ser473)/total Akt levels were slightly elevated in the NSC34 (G4C2)102 cells compared to the NSC34 sham cells (with or without tetracycline treatment), although the increases were not significant. Therefore, in the experimental conditions used, our data provided no evidence to suggest the PI3K/Akt signalling pathway was affected by (G4C2)102 expression. The GEP analysis did detect alterations in the PI3K/Akt signalling pathway however, and the biochemical data does not disprove that the PI3K/Akt signalling pathway is altered, because the pathway is hugely complex and integrates many stimuli. Under different experimental conditions, for example in serum starvation or stimulation with specific growth factors or cytokines, the (G4C2)102 expression may result in altered Akt activation.

Previous studies looking at p-AKT(Ser473) levels in ALS models have mainly been performed in SOD1(G93A) models, and have produced conflicting results. There were no changes in p-Akt(Ser473) levels in the spinal cord motor neurons of SOD1(G93A) mice compared to controls, p-Akt(Ser473) levels did not change with disease progression and they were also consistent in healthy and degenerating motor neurons in the SOD1(G93A) mice (Peviani et al., 2007). However, p-Akt(Ser473) levels were elevated in astrocytes and microglia in these SOD1(G93A) mice (Peviani et al., 2007). In a different study, p-Akt(Ser473) staining was much stronger in the spinal cord motor neurons of SOD1(G93A) mice compared to controls, and p-Akt(Ser473) levels were increased in the spinal cord of a SOD1(G93A) mice (Ilieva et al., 2003). However, in a motor neuron-like *in vitro* model of SOD1(G93A)-ALS, p-Akt(Ser473) levels were decreased when compared to control cells (Koh et al., 2004). Therefore, because of the conflicting results, and the

use of models with a different genetic variant of ALS, it is difficult to say how the results in the NSC34 (G4C2)102 cells fit with previous studies. In addition, no studies have looked at pAKT(Ser473) levels in other C9ORF72-ALS models, such as C9ORF72-ALS iPSC derived motor neurons.

6.4.3. Pten Knock-Down Provides Rescue from (G4C2)102 Induced Toxicity in the NSC34 (G4C2)102 Cells Independently of Increased Akt Signalling

The NSC34 sham and NSC34 (G4C2)102 were successfully transduced with Pten shRNA and control lentiviral vectors. Importantly, NSC34 cells were first transduced with GFP control lentivirus to visually confirm using a fluorescence microscope that the cells had been stably transduced, and also used as a control for lentiviral transduction. Later, when rescue was observed, NSC34 cells were then transduced with a control shRNA lentivirus that served as a control for the Pten shRNA expression off target effects. NSC34 (G4C2)102 cells that were transduced with the Pten shRNA had significant Pten knock-down over several passages, indicating stable transduction. Unexpectedly however, the NSC34 sham cells that were transduced with Pten shRNA only had significant Pten knock-down for a few passages, after which the Pten levels returned to the same as control cells. This suggests there was potentially an outgrowth of clones that contained the lentivirus intergrated at a silent genomic locus (ie weak expression of the Pten shRNA, and poor knock-down). Additionally, a transient burst of lentiviral expression may have accounted for the Pten knock-down in the early passages. For that reason, the initial immunoblot and rescue experiments (that used the GFP control lines) include the NSC34 sham Pten shRNA line, whilst the later experiments that used the control shRNA lines, do not include the NSC34 sham Pten shRNA line. Ideally, clonal transformants would have been selected and characterised for Pten knock-down, but there was insufficient time for clonal selection and therefore the mixed population of transduced cells was taken forward.

Another unexpected result was that p-Akt(Ser473) levels were elevated in both the NSC34 (G4C2)102 Pten and control shRNA cells compared to the NSC34 sham control shRNA cells. Firstly, this suggests that Pten knock-down had no effect on Akt signalling. Secondly, the lack of elevated Akt signalling in the NSC34 sham control shRNA cells

suggests that a double hit from the genomic (G4C2)102 repeat and lentivirus caused the increase in Akt signalling. Increased (G4C2)102 expression has no effect on the Akt signalling in the NSC34 (G4C2)102 lentiviral transduced cells. Also, the presence of the (G4C2)102 repeat alone had no effect on Akt signalling in the NSC34 cells (as described in section 6.4.1). However, viral vectors alone can activate the PI3K/AKT signalling pathway (Philpott et al., 2004; Tan et al., 2005; Tan et al., 2006), and stable transduction of the lentivirus into the NSC34 cells may be the cause of the increased Akt signalling. Alternatively to the double hit hypothesis, the NSC34 sham cells may have only been transiently transduced, or kicked out the control shRNA lentiviral vector (as was suspected with the Pten shRNA in the NSC34 sham cells described above), explaining the lack of elevated p-Akt(Ser473) levels.

The hypothesis was that Pten knock-down in the NSC34 cells would increase Akt signalling, providing a protective effect to the NSC34 (G4C2)102 cells. However, the Pten knock-down did not elevate p-Akt(Ser473) levels. This is contrary to previous work using Pten siRNA in both cultured motor neurons and mice. Pten siRNA delivered by lentivirus into cultured murine motor neurons caused Pten knock-down and increased p-Akt(Ser473) levels compared to cells transduced with control siRNA lentivirus (Kirby et al., 2011; Ning et al., 2010). In addition, Pten siRNA delivered by adeno associated virus 9 (AAV9) into mice also knocked down Pten and increased p-Akt(Ser473) levels in the skeletal muscle compared to mice transduced with control siRNA lentivirus (Little et al., 2015). In all three studies, the level of Pten knock-down was similar to that in the NSC34 (G4C2)102 Pten shRNA cells.

Whilst the Pten knock-down did not activate Akt signalling to a measurable level, it did provide a protective effect against the (G4C2)102 induced toxicity in the NSC34 (G4C2)102 cells. This rescue effect was detected in two separate sets of experiments, whilst both the NSC34 (G4C2)102 GFP and NSC34 (G4C2)102 control shRNA cell lines both displayed reduced viability after 7 days of increased (G4C2)102 expression. This suggests that the rescue effect derived from Pten knock-down in the NSC34 (G4C2)102 cells was independent of PI[3,4,5]P3 metabolism and PI3K/Akt signalling modulation. Pten is suggested to have functions independent of PI[3,4,5]P3 metabolism, including

DNA repair and sensitivity to genotoxic stress, cell migration, and protein phosphatase activity (Bassi et al., 2013; Leslie et al., 2007; Shi et al., 2014; Song et al., 2011). Pten knock-down has previously been shown to provide neuroprotective effects in both ALS and SMA models, although these effects were attributed to increased PI3K/AKT signalling (Kirby et al., 2011; Little et al., 2015; Ning et al., 2010). However, loss of PI3K/AKT independent Pten functions cannot be ruled out as providing some neuroprotective effects in these studies.

6.4.4. BpV(phen) was not Suitable as a Pten Inhibitor in the NSC34 Cells

BpV(phen) was tested for toxicity and Pten inhibition in the NSC34 sham cells to establish whether it could be used as a Pten inhibitor in rescue assays. The aim was to perform a rescue assay using BpV(phen) on NSC34 (G4C2)102 cells that were induced with tetracycline for 7 days. This is because previous experiments had shown (G4C2)102 induced toxicity was detectable after 7 days tetracycline induction (section 3.3.4.5). Therefore, the toxicity of various BpV(phen) concentrations at 6 days treatment was tested in NSC34 sham cells. 1µM was the highest sub-lethal dose of BpV(phen) in the NSC34 sham cells tested in the initial 6 day BpV(phen) dose-response experiment. This sub-lethal BpV(phen) concentration is fairly consistent with other *in vitro* studies. For example BpV(phen) caused mild cytotoxicity after 48 hours treatment at ≥2µM in BEAS2B cells (Lai et al., 2007), 24 hours treatment at ≥20µM in H9c2 cardiomyoblasts (Tian et al., 2012), and 2 hours treatment at ≥100µM in NIH3T cells (Schmid et al., 2004).

However, Akt signalling was not increased in NSC34 sham cells treated with 1µM BpV(phen), suggesting Pten was not inhibited. This is inconsistent with other previous *in vitro* studies that have used ≤1µM BpV(phen) to inhibit Pten and activate AKT signalling. In one study, Pten was activated and p-AKT(ser473) levels were elevated in BEAS2B cells treated for 30 minutes with as little as 0.1µM BpV(phen) (Lai et al., 2007). BpV(phen) is also known to be relatively stable in cell culture conditions, and can be incubated in media over night without any reduction in its ability to inhibit Pten (Schmid et al., 2004). In addition, increased AKT signalling was maintained over several days in BEAS2B cells by BpV(phen) treatment (Lai et al., 2007). Therefore, alterations in Akt signalling should have been detectable between 12 and 48 hours BpV(phen) treatment.

1µM BpV(phen) was too low a concentration to use as a Pten inhibitor in the NSC34 cells.

When the NSC34 sham cells were treated with a higher dose of 3µM BpV(phen) for 48 hours, there was an increase in p-Akt(Ser473) levels, indicating Pten inhibition. This was important because no previous studies had shown that BpV(phen) could inhibit Pten in the NSC34 cell line. Rather contradictorily however, 3µM BpV(phen) caused cytotoxicity in the NSC34 sham cells after only 24 hours. This actually suggests the elevated p-Akt(Ser473) levels were in fact due to the NSC34 activating the prosurvival pathway in a last ditch attempt to avoid BpV(Phen) induced toxicity. A shorter 6 hour dose of 3µM BpV(phen) (which did not cause detectable cytotoxicity) was taken forward for the rescue assay, but did not rescue the NSC34 (G4C2)102 cells from (G4C2)102 induced toxicity, and actually caused slight cytotoxicity to the NSC34 cells. In conclusion, there was no suitable concentration of BpV(phen) that would inhibit Pten, activate Akt signalling and was not cytotoxic.

6.4.5. Summary

GEP identified transcriptomic dysregulation in the PI3K/AKT signalling pathway in NSC34 (G4C2)102 cells, and also in C9ORF72-ALS LCM motor neurons. This suggests that transcriptomic dysregulation in this pro-survival pathway is an early biochemical event in C9ORF72-ALS, and could therefore serve as a potent therapeutic target for protecting motor neurons in ALS pathogenesis. However, there was no measurable effect on Akt or Pten protein levels in the NSC34 (G4C2)102 cells, nor was the PI3K/Akt signalling pathway altered under basal conditions. Although, the reduced sensitivity of immunoblotting as a technique compared to the microarray technology used is the likely cause for the apparent lack of validation. Importantly, Pten knock-down did provide a rescue effect against the (G4C2)102 induced toxicity, independently of the Akt signalling pathway, and suggests Pten deserves further study in different models of C9ORF72-ALS as a potential therapeutic target. Lastly however, BpV(phen) was not a useful Pten inhibitor in the NSC34 cells, because there was no sub-lethal dose that activated the Akt signalling pathway.

Chapter 7. Conclusions and Future Work

ALS is an incurable neurodegenerative disease characterised by progressive loss of both upper and lower motor neurons. The most common genetic cause of ALS is a (G4C2)n repeat expansion in intron 1 of *C9ORF72*. This project aimed to understand how the (G4C2)n repeat expansion could cause motor neuron cell death, and to identify potential therapeutic targets for neuroprotection. To do this, motor neuron like cellular models were generated using NSC34 cells that expressed different lengths of the (G4C2)n in a tetracycline inducible manner. Initially, these NSC34 (G4C2)n cells were characterised for C9ORF72-ALS pathology using a range of biochemical techniques. Analysis of gene expression was then employed to assess how expression of the (G4C2)n affects the transcriptome in a motor neuron like context. Additionally, the transcriptomic analysis identified biological pathways that may hold therapeutic potential for C9ORF72-ALS patients. In the last part of the project, one of the potential therapeutic targets was assessed in the NSC34 (G4C2)n cells to see whether modulation would provide rescue effects and could therefore serve as a potential therapeutic target.

When the project was started, I attempted to use RNA tagging systems with the (G4C2)n constructs that would allow detection of (G4C2)n expression, but which would also allow the (G4C2)n RNA to be studied in live cells. The MS2 system and tRSA systems were employed (Bertrand et al., 1998; Iioka et al., 2011), to facilitate visualisation of the (G4C2)n RNA in live cells and direct (G4C2)n RNA pull down from live cells respectively. Whilst the MS2 and tRSA tags were each cloned upstream of the (G4C2)n repeat constructs, I was unable to optimise either system to detect (G4C2)n RNA expression in the live cells. Also, at the time, the locked nucleic acid (LNA) probe and FISH methodologies were developed, which allowed detection of the (G4C2)n RNA expressed from the repeat, and could be employed to measure (G4C2)n expression in the cellular model system. Given that the main aim of the project was to develop a cellular model that expressed the (G4C2)n repeat expansion, and to use the model to identify potential therapeutic targets, I decided to discontinue the RNA tag work. However, with improvements in future technology, these sorts of live cell RNA tagging tools would be

incredibly useful for further understanding the molecular biology of the (G4C2)n repeat. Particularly, visualisation of the (G4C2)n RNA in live cells could help elucidate how the (G4C2)n RNA is exported from the nucleus, how dynamic the (G4C2)n RNA foci and (G4C2)n RNA-protein interactions are, and also how and where the (G4C2)n RNA is RAN translated to produce DPR proteins.

At present, it is still not fully understood how the (G4C2)n repeat expansion in *C9ORF72* causes motor neuron cell death in ALS pathogenesis. Therefore, generating stable cellular models that expressed the (G4C2)n repeat in a motor neuron cell context was designed to better understand this problem. Many early studies that also aimed to elucidate the potentially toxic mechanisms associated with the (G4C2)n repeats either used transiently transfected cell lines, C9ORF72-ALS patient derived cells or post-mortem tissue (sections 1.7 and 1.8). However, I wanted to generate a reductionist model that would interrogate only the effect of the (G4C2)n expression independent of the *C9ORF72* gene context. Also, I specifically used the FRT system to generate isogenic control and disease model cell lines, because random insertion of the (G4C2)n repeats could have had off target effects unrelated to the (G4C2)n expression itself. Additionally, overexpression of a gene of interest can provide misleading results, and tetracycline-inducible expression had the advantage of allowing control of the (G4C2)n expression level. Another unique advantage of the tetracycline inducible expression is that the early biochemical effects of the (G4C2)n expression could be studied. Stable and isogenic NSC34 (G4C2)n cell lines were successfully generated that had tetracycline inducible (G4C2)n expression. The (G4C2)n repeat forms characteristic sense RNA foci, but not antisense RNA foci, in the NSC34 (G4C2)n cells. In addition, the interrupted (G4C2)102 repeat underwent RAN translation to produce proteins containing all sense DPR motifs ((GA)n, (GR)n and (GP)n). Therefore, the NSC34 (G4C2)n cells are sense only (G4C2)n gain of function models of C9ORF72-ALS, with tetracycline inducible expression allowing identification of early biochemical effects.

Expression of the (G4C2)102 reduced NSC34 cell growth rate, indicating that the RNA and/or RAN proteins were causing mild toxicity or posed a metabolic burden in the cells. This is in agreement with other (G4C2)n gain of function models of C9ORF72-ALS that

also describe (G4C2)n induced toxicity (section 1.8.2). However, the more subtle toxic effects described in the NSC34 (G4C2)102 cells may actually represent a more physiologically relevant model compared to the more severe phenotypes described in other cellular models. ALS onset generally occurs later in life, and therefore any toxicity associated with the (G4C2)n repeat expansion is likely to be well tolerated and fairly low. However, the (G4C2)n must provide some toxic burden that aged motor neurons and/or other cells in the CNS fail to handle. Although many cellular and animal models suggest that the arginine rich DPR species are the main toxic entity associated with the (G4C2)n repeat expansion, the NSC34 (G4C2)102 cells do not produce these pure DPR, which may also explain the mild toxic phenotype. Further, in house work with Guillaume Hautbergue suggests that pure (G4C2)n repeats produce far more RAN translation products compared to the interrupted (G4C2)102 repeats used in this project (unpublished data).

In addition, the NSC34 (G4C2)102 cells recapitulated some other aspects of C9ORF72-ALS pathology. Firstly, the sense RNA foci colocalised with SRSF1, SRSF2, and nucleolar NCL, but not with PURA or ALYREF. Secondly, the (G4C2)102 expression caused mild disruption to the nucleolar morphology in the NSC34 cells comparable to previously published data, although there was no clear evidence of functional nucleolar stress. However, (G4C2)102 expression did not cause TDP-43 aggregation or mislocalisation. This is in agreement with the majority of other (G4C2)n gain of function cellular and animal models of C9ORF72-ALS (section 1.8.2). In fact, TDP-43 pathology has only been described in one model – 6 month old mice that over express a (G4C2)66 repeat construct in the CNS (O'Rourke et al., 2015). Several other (G4C2)n gain of function animal models do not recapitulate TDP-43 pathology (section 1.8.2). In conclusion though, the subtle biochemical effects described in the NSC34 (G4C2)102 cells model early events that occur before TDP-43 aggregation and mislocalisation, and any therapeutic targets identified may prevent downstream events that culminate in TDP-43 pathology and/or ALS pathogenesis. Additionally, the subtle phenotype described in the NSC34 (G4C2)102 cells is arguably more relevant for disease with age-related onset.

A pipeline of work was previously used to identify potential therapeutic targets in SOD1-ALS. Firstly, transcriptomic analysis was performed on NSC34 cell models expressing G93A SOD1, and identified transcriptomic dysregulation in genes encoding antioxidant response proteins (Kirby et al., 2005). In particular, nuclear erythroid 2-related-factor 2 (NRF2) was identified as a potential therapeutic target because it promotes expression of the cytoprotective Phase II detoxification and antioxidant enzymes (McMahon et al., 2001; Nguyen et al., 2003). NRF2 was down regulated in the G93A SOD1 NSC34 cell models, and also in spinal motor neurons and motor cortex of SALS cases (Kirby et al., 2005; Sarlette et al., 2008). Secondly, small molecules were screened and S[+] Apomorphine was identified as a CNS penetrant NRF2 activator, which is a promising therapeutic candidate for ALS patients (Mead et al., 2013).

Here, we aimed to repeat the pipeline of work described above to identify therapeutic targets, and eventually drug candidates, for C9ORF72-ALS. Previous transcriptomic studies on C9ORF72-ALS have used either patient derived cell lines or post-mortem CNS tissue. These transcriptomic changes may therefore reflect the end stage of the disease, and the crucial early transcriptomic changes associated with (G4C2)n repeat toxicity may be lost or more difficult to identify. Also, these studies did not interrogate the (G4C2)n expression in a reductionist manner. However, transcriptomic analysis of the NSC34 (G4C2)102 cells identified dysregulated pathways that may be crucial in the early pathogenesis of C9ORF72-ALS. Specifically, transcriptomic dysregulation was identified in RNA metabolism, and particularly in RNA splicing. Further, differential splicing analysis of the NSC34 (G4C2)102 cells showed that (G4C2)102 expression functionally affected the splicing. This corroborates several other transcriptomic studies performed on C9ORF72-ALS patient derived cells and post-mortem tissue (Cooper-Knock et al., 2015a; Prudencio et al., 2015), and critically suggests that RNA metabolism and splicing dysregulation is an early event in C9ORF72-ALS pathogenesis caused by the sense (G4C2)n repeat.

Also using the GEP analysis, significant dysregulation in the PI3K/Akt signalling pathway was identified in the NSC34 (G4C2)102 cells. Again, this was validated in LCM motor neurons from C9ORF72-ALS patients, suggesting this is an early transcriptomic change

in C9ORF72-ALS. Given that the PI3K/AKT signalling pathway regulates cellular survivial (amongst other functions) it could represent a particularly potent therapeutic target in prolonguing motor neuron survival in C9ORF72-ALS. This also demonstrates that the NSC34 (G4C2)102 cells can successfully be used to identify disease relevant early transcriptomic changes in C9ORF72-ALS. Additionally, previous work had also identified the PI3K/Akt signalling pathway, and Pten in particular, as potential therapeutic targets in ALS, which therefore encouraged further interrogation.

Although the microarray analysis identified transcriptomic dysregulation in the NSC34 (G4C2)102 cells, and this was also found in C9ORFF72-ALS LCM motor neurons, qRT-PCR and immunoblot validation did not find any significant differences in Pten or Akt levels in the NSC34 (G4C2)102 cells. This is likely due to reduced sensitivity in validation methods compared to the microarray technology. Additionally, there was no measurable change in Akt signalling at basal levels in the NSC34 (G4C2)102 cells. However, Pten knock-down did provide a rescue effect against the (G4C2)102 induced toxicity described in the NSC34 (G4C2)102 cells. Further, the Pten knock-down did not affect Akt signalling, and suggests that the rescue effect associated with the Pten knock-down could be independent of the PI3K/Akt signalling pathway. In conclusion though, the transcriptomic validation in C9ORF72-ALS patient motor neurons, and the Pten knock-down rescue effect suggest Pten deserves further study in different models of C9ORF72-ALS.

7.1. Future Work

The relative contributions of C9ORF72 haploinsufficiency, RNA toxicity and DPR toxicity in ALS pathogenesis still remain far from clear, and therefore future work should be invested in this direction. In terms of C9ORF72 haploinsufficiency, there are two key problems: there are no reliable antibodies available for C9ORF72 detection, and also the function of C9ORF72 is still unknown. Therefore investment in developing reliable C9ORF72 antibodies will allow research into the function of C9ORF72, and also identify whether C9ORF72 protein levels are affected by the presence of the (G4C2)n repeat expansion.

Additionally, it is also unclear which cell types are affected directly and indirectly by the (G4C2)n repeat expansion. Elegant studies using the Cre-LoxP system in SOD1 mouse models of ALS, show that deleting the mutant SOD1 in astrocytes and microglia slows disease progression, whilst deleting mutant SOD1 in the motor neurons delays disease onset (Boillée et al., 2006b; Wang et al., 2011; Yamanaka et al., 2008a; Yamanaka et al., 2008b). Additionally, co-culture models show that FALS and SALS patient derived astrocytes (including C9ORF72-ALS patient derived astrocytes) and microglial cells confer *in vitro* toxicity when cultured with Hb9-GFP murine motor neurons (Di Giorgio et al., 2008; Haidet-Phillips et al., 2011; Meyer et al., 2014; Nagai et al., 2007). Therefore, further investigation into which cell types contribute to motor neuron death *in vivo* in C9ORF72-ALS is crucial for identifying effective therapeutics.

The first line of future work I would consider is to interrogate PTEN as a therapeutic target in different models of both C9ORF72-ALS, but also other genetic forms of FALS and SALS. Firstly, *in vitro* co-culture models (described above) would be used, and if PTEN modulation provided rescue effects, secondary *in vivo* mouse models would be used. The Hb9-GFP murine motor neurons would be treated with Pten shRNA, and then co-cultured with C9ORF72-ALS derived astrocytes or microglial cells to see whether Pten knock-down rescues the motor neurons from glial toxicity. Additionally, other pharmacological PTEN inhibitors such as BpV(Phen) would be trialled for efficacy in these motor neuron cells as well. Further, astrocytes and microglial cells would also be treated with PTEN shRNA, and co-cultured with PTEN shRNA treated and untreated Hb9-GFP murine motor neurons. This would identify whether PTEN knock-down in ALS patient derived glial cells would exacerbate their toxicity toward motor neurons, in which case, a motor neuron specific PTEN knock-down would be required for *in vivo* models as opposed to a pan-CNS PTEN shRNA delivery. Experiments using other genetic forms of FALS, and also SALS derived glial cells would also be performed alongside the C9ORF72-ALS work.

Dependent on the efficacy of the PTEN shRNA and/or PTEN inhibition in the co-culture system, this strategy would then be taken forward for *in vivo* work. Mice provide a robust and widely used model organism for studying human disease, and would

therefore be used as the *in vivo* model. Since the aim of the *in vivo* work would be to identify whether PTEN knock-down and/or inhibition can prevent motor neuron degeneration, a mouse model with a clear phenotype would be required. At present, only one C9ORF72-ALS mouse model exists that recapitulates TDP-43 pathology, and behavioural and motor defects similar to the clinical symptoms of C9ORF72-ALS/FTD patients (Chew et al., 2015). Therefore, this model would be selected, unless a more convincing mouse model was generated. In this model, a (G4C2)66 construct would be delivered to the CNS via adeno-associated virus, with the addition of either Pten shRNA or scrambled shRNA. The mice would then be assessed for TDP-43 inclusions in the CNS, astrogliosis, cortical and cerebellar degeneration, as well as cognitive and motor performance (all of which have been described in this mouse model previously (Chew et al., 2015)).

Finally, ALS is a complex multifactorial disease, and there are several known dysregulated biological pathways but also many different cell types involved in disease onset and progression. Therefore, effective treatment of the disease will likely require modulation of more than one target, and also multiple cell types. Also, successful early diagnosis of ALS is likely to critical in treating the disease effectively, and therefore the development of better biomarkers will be necessary.

References

Abhinav, K., Stanton, B., Johnston, C., Hardstaff, J., Orrell, R.W., Howard, R., Clarke, J., Sakel, M., Ampong, M.A., Shaw, C.E.*, et al.* (2007). Amyotrophic lateral sclerosis in South-East England: a population-based study. The South-East England register for amyotrophic lateral sclerosis (SEALS Registry). Neuroepidemiology 29, 44-48.

Abramzon, Y., Johnson, J.O., Scholz, S.W., Taylor, J.P., Brunetti, M., Calvo, A., Mandrioli, J., Benatar, M., Mora, G., Restagno, G.*, et al.* (2012). Valosin-containing protein (VCP) mutations in sporadic amyotrophic lateral sclerosis. Neurobiol Aging 33, 2231.e2231-2231.e2236.

Achi, E.Y., and Rudnicki, S.A. (2012). ALS and Frontotemporal Dysfunction: A Review. Neurol Res Int 2012, 806306.

Al-Chalabi, A., Fang, F., Hanby, M.F., Leigh, P.N., Shaw, C.E., Ye, W., and Rijsdijk, F. (2010). An estimate of amyotrophic lateral sclerosis heritability using twin data. J Neurol Neurosurg Psychiatry 81, 1324-1326.

Al-Chalabi, A., Jones, A., Troakes, C., King, A., Al-Sarraj, S., and van den Berg, L.H. (2012). The genetics and neuropathology of amyotrophic lateral sclerosis. Acta Neuropathol 124, 339-352.

Al-Saif, A., Al-Mohanna, F., and Bohlega, S. (2011). A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis. Ann Neurol 70, 913-919.

Al-Sarraj, S., King, A., Troakes, C., Smith, B., Maekawa, S., Bodi, I., Rogelj, B., Al-Chalabi, A., Hortobagyi, T., and Shaw, C.E. (2011). p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. Acta Neuropathol 122, 691-702.

Alami, N.H., Smith, R.B., Carrasco, M.A., Williams, L.A., Winborn, C.S., Han, S.S., Kiskinis, E., Winborn, B., Freibaum, B.D., Kanagaraj, A.*, et al.* (2014). Axonal transport of TDP-43 mRNA granules is impaired by ALS-causing mutations. Neuron 81, 536-543.

Alexianu, M.E., Kozovska, M., and Appel, S.H. (2001). Immune reactivity in a mouse model of familial ALS correlates with disease progression. Neurology 57, 1282-1289.

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.

Andersen, P.M. (2006). Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. Curr Neurol Neurosci Rep 6, 37-46.

Andersen, P.M., Forsgren, L., Binzer, M., Nilsson, P., Ala-Hurula, V., Keränen, M.L., Bergmark, L., Saarinen, A., Haltia, T., Tarvainen, I.*, et al.* (1996). Autosomal recessive adult-onset amyotrophic lateral sclerosis associated with homozygosity for Asp90Ala CuZn-superoxide dismutase mutation. A clinical and genealogical study of 36 patients. Brain 119 (Pt 4), 1153-1172.

Anderson, P., and Kedersha, N. (2006). RNA granules. J Cell Biol 172, 803-808.

Andreadou, E., Kapaki, E., Kokotis, P., Paraskevas, G.P., Katsaros, N., Libitaki, G., Zis, V., Sfagos, C., and Vassilopoulos, D. (2008). Plasma glutamate and glycine levels in patients with amyotrophic lateral sclerosis: the effect of riluzole treatment. Clin Neurol Neurosurg 110, 222-226.

Andrus, P.K., Fleck, T.J., Gurney, M.E., and Hall, E.D. (1998). Protein oxidative damage in a transgenic mouse model of familial amyotrophic lateral sclerosis. J Neurochem 71, 2041-2048.

Angelov, D., Bondarenko, V.A., Almagro, S., Menoni, H., Mongélard, F., Hans, F., Mietton, F., Studitsky, V.M., Hamiche, A., Dimitrov, S., and Bouvet, P. (2006). Nucleolin is a histone chaperone with FACT-like activity and assists remodeling of nucleosomes. EMBO J 25, 1669-1679.

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.

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.

Atkinson, R.A., Fernandez-Martos, C.M., Atkin, J.D., Vickers, J.C., and King, A.E. (2015). C9ORF72 expression and cellular localization over mouse development. Acta Neuropathol Commun 3, 59.

Ayala, Y.M., Zago, P., D'Ambrogio, A., Xu, Y.F., Petrucelli, L., Buratti, E., and Baralle, F.E. (2008). Structural determinants of the cellular localization and shuttling of TDP-43. J Cell Sci 121, 3778-3785.

Bannwarth, S., Ait-El-Mkadem, S., Chaussenot, A., Genin, E.C., Lacas-Gervais, S., Fragaki, K., Berg-Alonso, L., Kageyama, Y., Serre, V., Moore, D.G*., et al.* (2014). A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. Brain 137, 2329-2345.

Barber, S.C., and Shaw, P.J. (2010). Oxidative stress in ALS: key role in motor neuron injury and therapeutic target. Free Radic Biol Med 48, 629-641.

Bartolome, F., Wu, H.C., Burchell, V.S., Preza, E., Wray, S., Mahoney, C.J., Fox, N.C., Calvo, A., Canosa, A., Moglia, C*., et al.* (2013). Pathogenic VCP mutations induce mitochondrial uncoupling and reduced ATP levels. Neuron 78, 57-64.

Bassi, C., Ho, J., Srikumar, T., Dowling, R.J., Gorrini, C., Miller, S.J., Mak, T.W., Neel, B.G., Raught, B., and Stambolic, V. (2013). Nuclear PTEN controls DNA repair and sensitivity to genotoxic stress. Science 341, 395-399.

Bauer, P.O. (2016). Methylation of C9orf72 expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. Neurosci Lett 612, 204-209.

Beck, J., Poulter, M., Hensman, D., Rohrer, J.D., Mahoney, C.J., Adamson, G., Campbell, T., Uphill, J., Borg, A., Fratta, P*., et al.* (2013). Large C9orf72 Hexanucleotide Repeat Expansions Are Seen in Multiple Neurodegenerative Syndromes and Are More Frequent Than Expected in the UK Population. Am J Hum Genet 92, 345-353.

Beeldman, E., van der Kooi, A.J., de Visser, M., van Maarle, M.C., van Ruissen, F., and Baas, F. (2015). A Dutch family with autosomal recessively inherited lower motor neuron predominant motor neuron disease due to optineurin mutations. Amyotroph Lateral Scler Frontotemporal Degener 16, 410-411.

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., and Shaw, P.J. (2015). Intermediate length C9orf72 expansion in an ALS patient without classical C9orf72 neuropathology. Amyotroph Lateral Scler Frontotemporal Degener 16, 249-251.

Beers, D.R., Henkel, J.S., Xiao, Q., Zhao, W., Wang, J., Yen, A.A., Siklos, L., McKercher, S.R., and Appel, S.H. (2006). Wild-type microglia extend survival in PU.1 knockout mice with familial amyotrophic lateral sclerosis. Proc Natl Acad Sci U S A 103, 16021-16026.

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.

Bensimon, G., Lacomblez, L., and Meininger, V. (1994). A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group. N Engl J Med 330, 585-591.

Benussi, L., Rossi, G., Glionna, M., Tonoli, E., Piccoli, E., Fostinelli, S., Paterlini, A., Flocco, R., Albani, D., Pantieri, R.*, et al.* (2014). C9ORF72 hexanucleotide repeat number in frontotemporal lobar degeneration: a genotype-phenotype correlation study. J Alzheimers Dis 38, 799-808.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. Mol Cell 2, 437-445.

Bevan, A.P., Drake, P.G., Yale, J.F., Shaver, A., and Posner, B.I. (1995). Peroxovanadium compounds: biological actions and mechanism of insulin-mimesis. Mol Cell Biochem 153, 49-58.

Bhaskar, P.T., and Hay, N. (2007). The two TORCs and Akt. Dev Cell 12, 487-502.

Bjørkøy, G., Lamark, T., and Johansen, T. (2006). p62/SQSTM1: a missing link between protein aggregates and the autophagy machinery. Autophagy 2, 138-139.

Boillée, S., Vande Velde, C., and Cleveland, D.W. (2006a). ALS: a disease of motor neurons and their nonneuronal neighbors. Neuron 52, 39-59.

Boillée, S., Yamanaka, K., Lobsiger, C.S., Copeland, N.G., Jenkins, N.A., Kassiotis, G., Kollias, G., and Cleveland, D.W. (2006b). Onset and progression in inherited ALS determined by motor neurons and microglia. Science 312, 1389-1392.

Bolger, T.A., Folkmann, A.W., Tran, E.J., and Wente, S.R. (2008). The mRNA export factor Gle1 and inositol hexakisphosphate regulate distinct stages of translation. Cell 134, 624-633.

Boopathy, S., Silvas, T.V., Tischbein, M., Jansen, S., Shandilya, S.M., Zitzewitz, J.A., Landers, J.E., Goode, B.L., Schiffer, C.A., and Bosco, D.A. (2015). Structural basis for mutation-induced destabilization of profilin 1 in ALS. Proc Natl Acad Sci U S A 112, 7984-7989.

Bowling, A.C., Schulz, J.B., Brown, R.H., and Beal, M.F. (1993). Superoxide dismutase activity, oxidative damage, and mitochondrial energy metabolism in familial and sporadic amyotrophic lateral sclerosis. J Neurochem 61, 2322-2325.

Bruijn, L.I., Becher, M.W., Lee, M.K., Anderson, K.L., Jenkins, N.A., Copeland, N.G., Sisodia, S.S., Rothstein, J.D., Borchelt, D.R., Price, D.L., and Cleveland, D.W. (1997). ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. Neuron 18, 327-338.

Bruijn, L.I., Houseweart, M.K., Kato, S., Anderson, K.L., Anderson, S.D., Ohama, E., Reaume, A.G., Scott, R.W., and Cleveland, D.W. (1998). Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. Science 281, 1851-1854.

Buratti, E., and Baralle, F.E. (2008). Multiple roles of TDP-43 in gene expression, splicing regulation, and human disease. Frontiers in bioscience : a journal and virtual library 13, 867-878.

Buratti, E., and Baralle, F.E. (2012). TDP-43: gumming up neurons through protein-protein and protein-RNA interactions. Trends Biochem Sci 37, 237-247.

Buratti, E., Brindisi, A., Giombi, M., Tisminetzky, S., Ayala, Y.M., and Baralle, F.E. (2005). TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing. J Biol Chem 280, 37572-37584.

Burrell, J.R., Kiernan, M.C., Vucic, S., and Hodges, J.R. (2011). Motor neuron dysfunction in frontotemporal dementia. Brain 134, 2582-2594.

Bury, J.J., Highley, J.R., Cooper-Knock, J., Goodall, E.F., Higginbottom, A., McDermott, C.J., Ince, P.G., Shaw, P.J., and Kirby, J. (2015). Oligogenic inheritance of optineurin (OPTN) and C9ORF72 mutations in ALS highlights localisation of OPTN in the TDP-43-negative inclusions of C9ORF72-ALS. Neuropathology.

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.

Byrne, S., Walsh, C., Lynch, C., Bede, P., Elamin, M., Kenna, K., McLaughlin, R., and Hardiman, O. (2011). Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. J Neurol Neurosurg Psychiatry 82, 623-627.

Bäumer, D., Talbot, K., and Turner, M.R. (2014). Advances in motor neurone disease. J R Soc Med 107, 14-21.

Cashman, N.R., Durham, H.D., Blusztajn, J.K., Oda, K., Tabira, T., Shaw, I.T., Dahrouge, S., and Antel, J.P. (1992). Neuroblastoma x spinal cord (NSC) hybrid cell lines resemble developing motor neurons. Dev Dyn 194, 209-221.

Chang, C.F., Gallia, G.L., Muralidharan, V., Chen, N.N., Zoltick, P., Johnson, E., and Khalili, K. (1996). Evidence that replication of human neurotropic JC virus DNA in glial cells is regulated by the sequence-specific single-stranded DNA-binding protein Pur alpha. J Virol 70, 4150-4156.

Chang, L., and Monteiro, M.J. (2015). Defective Proteasome Delivery of Polyubiquitinated Proteins by Ubiquilin-2 Proteins Containing ALS Mutations. PLoS One 10, e0130162.

Chang, Y., Kong, Q., Shan, X., Tian, G., Ilieva, H., Cleveland, D.W., Rothstein, J.D., Borchelt, D.R., Wong, P.C., and Lin, C.L. (2008). Messenger RNA oxidation occurs early in disease pathogenesis and promotes motor neuron degeneration in ALS. PloS one 3, e2849.

Chang, Y.J., Jeng, U.S., Chiang, Y.L., Hwang, I.S., and Chen, Y.R. (2016). The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. J Biol Chem 291, 4903-4911.

Chen, H.J., Anagnostou, G., Chai, A., Withers, J., Morris, A., Adhikaree, J., Pennetta, G., and de Belleroche, J.S. (2010). Characterization of the properties of a novel mutation in VAPB in familial amyotrophic lateral sclerosis. J Biol Chem 285, 40266-40281.

Chen, Y., Luo, C., Zhao, M., Li, Q., Hu, R., Zhang, J.H., Liu, Z., and Feng, H. (2015). Administration of a PTEN inhibitor BPV(pic) attenuates early brain injury via modulating AMPA receptor subunits after subarachnoid hemorrhage in rats. Neurosci Lett 588, 131-136.

Chen, Y.Z., Bennett, C.L., Huynh, H.M., Blair, I.P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M.L., Rabin, B.A.*, et al.* (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). Am J Hum Genet 74, 1128-1135.

Cheng, R., and Banack, S.A. (2009). Previous studies underestimate BMAA concentrations in cycad flour. Amyotroph Lateral Scler 10 Suppl 2, 41-43.

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.

Chiò, A., Borghero, G., Restagno, G., Mora, G., Drepper, C., Traynor, B.J., Sendtner, M., Brunetti, M., Ossola, I., Calvo, A.*, et al.* (2012a). Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. Brain 135, 784-793.

Chiò, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., Traynor, B.G., and Consortium, E. (2009). Prognostic factors in ALS: A critical review. Amyotroph Lateral Scler 10, 310-323.

Chiò, A., Logroscino, G., Traynor, B.J., Collins, J., Simeone, J.C., Goldstein, L.A., and White, L.A. (2013). Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. Neuroepidemiology 41, 118-130.

Chiò, A., Restagno, G., Brunetti, M., Ossola, I., Calvo, A., Canosa, A., Moglia, C., Floris, G., Tacconi, P., Marrosu, F.*, et al.* (2012b). ALS/FTD phenotype in two Sardinian families carrying both C9ORF72 and TARDBP mutations. J Neurol Neurosurg Psychiatry 83, 730-733.

Cho, H., Mu, J., Kim, J.K., Thorvaldsen, J.L., Chu, Q., Crenshaw, E.B., Kaestner, K.H., Bartolomei, M.S., Shulman, G.I., and Birnbaum, M.J. (2001). Insulin resistance and a diabetes mellitus-like syndrome in mice lacking the protein kinase Akt2 (PKB beta). Science 292, 1728-1731.

Chow, C.Y., Landers, J.E., Bergren, S.K., Sapp, P.C., Grant, A.E., Jones, J.M., Everett, L., Lenk, G.M., McKenna-Yasek, D.M., Weisman, L.S*., et al.* (2009). Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. Am J Hum Genet 84, 85-88.

Chow, C.Y., Zhang, Y., Dowling, J.J., Jin, N., Adamska, M., Shiga, K., Szigeti, K., Shy, M.E., Li, J., Zhang, X*., et al.* (2007). Mutation of FIG4 causes neurodegeneration in the pale tremor mouse and patients with CMT4J. Nature 448, 68-72.

Cirulli, E.T., Lasseigne, B.N., Petrovski, S., Sapp, P.C., Dion, P.A., Leblond, C.S., Couthouis, J., Lu, Y.F., Wang, Q., Krueger, B.J*., et al.* (2015). Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. Science 347, 1436-1441.

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. Annals of Neurology 74, 180-187.

Clement, A.M., Nguyen, M.D., Roberts, E.A., Garcia, M.L., Boillée, S., Rule, M., McMahon, A.P., Doucette, W., Siwek, D., Ferrante, R.J*., et al.* (2003). Wild-type nonneuronal cells extend survival of SOD1 mutant motor neurons in ALS mice. Science 302, 113-117.

Cooper-Knock, J., Bury, J.J., Heath, P.R., Wyles, M., Higginbottom, A., Gelsthorpe, C., Highley, J.R., Hautbergue, G., Rattray, M., Kirby, J., and Shaw, P.J. (2015a). C9ORF72 GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis. PLoS One 10, e0127376.

Cooper-Knock, J., Hewitt, C., Highley, J.R., Brockington, A., Milano, A., Man, S., Martindale, J., Hartley, J., Walsh, T., Gelsthorpe, C*., et al.* (2012). Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. Brain 135, 751-764.

Cooper-Knock, J., Higginbottom, A., Connor-Robson, N., Bayatti, N., Bury, J.J., Kirby, J., Ninkina, N., Buchman, V.L., and Shaw, P.J. (2013). C9ORF72 transcription in a frontotemporal dementia case with two expanded alleles. Neurology 81, 1719-1721.

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. (2015b). Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. Acta Neuropathol 130, 63-75.

Cooper-Knock, J., Shaw, P.J., and Kirby, J. (2014a). The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. Acta Neuropathol 127, 333-345.

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.* (2014b). Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. Brain 137, 2040-2051.

Cox, L.E., Ferraiuolo, L., Goodall, E.F., Heath, P.R., Higginbottom, A., Mortiboys, H., Hollinger, H.C., Hartley, J.A., Brockington, A., Burness, C.E*., et al.* (2010). Mutations in CHMP2B in lower motor neuron predominant amyotrophic lateral sclerosis (ALS). PLoS One 5, e9872.

Cozzolino, M., and Carrì, M.T. (2012). Mitochondrial dysfunction in ALS. Prog Neurobiol 97, 54-66.

Cudkowicz, M.E., McKenna-Yasek, D., Sapp, P.E., Chin, W., Geller, B., Hayden, D.L., Schoenfeld, D.A., Hosler, B.A., Horvitz, H.R., and Brown, R.H. (1997). Epidemiology of mutations in superoxide dismutase in amyotrophic lateral sclerosis. Ann Neurol 41, 210-221.

Daigle, J.G., Krishnamurthy, K., Ramesh, N., Casci, I., Monaghan, J., McAvoy, K., Godfrey, E.W., Daniel, D.C., Johnson, E.M., Monahan, Z*., et al.* (2016). Pur-alpha regulates cytoplasmic stress granule dynamics and ameliorates FUS toxicity. Acta Neuropathol.

Dal Canto, M.C., and Gurney, M.E. (1994). Development of central nervous system pathology in a murine transgenic model of human amyotrophic lateral sclerosis. Am J Pathol 145, 1271-1279.

Daoud, H., Zhou, S., Noreau, A., Sabbagh, M., Belzil, V., Dionne-Laporte, A., Tranchant, C., Dion, P., and Rouleau, G.A. (2012). Exome sequencing reveals SPG11 mutations causing juvenile ALS. Neurobiology of aging 33, 839 e835-839.

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.

De Conti, L., Akinyi, M.V., Mendoza-Maldonado, R., Romano, M., Baralle, M., and Buratti, E. (2015). TDP-43 affects splicing profiles and isoform production of genes involved in the apoptotic and mitotic cellular pathways. Nucleic Acids Res 43, 8990-9005.

de Haro, M., Al-Ramahi, I., De Gouyon, B., Ukani, L., Rosa, A., Faustino, N.A., Ashizawa, T., Cooper, T.A., and Botas, J. (2006). MBNL1 and CUGBP1 modify expanded CUG-induced toxicity in a Drosophila model of myotonic dystrophy type 1. Hum Mol Genet 15, 2138-2145.

De Vos, K.J., Chapman, A.L., Tennant, M.E., Manser, C., Tudor, E.L., Lau, K.F., Brownlees, J., Ackerley, S., Shaw, P.J., McLoughlin, D.M*., et al.* (2007). Familial amyotrophic lateral sclerosis-linked SOD1 mutants perturb fast axonal transport to reduce axonal mitochondria content. Hum Mol Genet 16, 2720-2728.

De Vos, K.J., Grierson, A.J., Ackerley, S., and Miller, C.C. (2008). Role of axonal transport in neurodegenerative diseases. Annu Rev Neurosci 31, 151-173.

De Vos, K.J., Morotz, G.M., Stoica, R., Tudor, E.L., Lau, K.F., Ackerley, S., Warley, A., Shaw, C.E., and Miller, C.C. (2012). VAPB interacts with the mitochondrial protein PTPIP51 to regulate calcium homeostasis. Hum Mol Genet 21, 1299-1311.

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.

Deng, H., Gao, K., and Jankovic, J. (2014). The role of FUS gene variants in neurodegenerative diseases. Nat Rev Neurol 10, 337-348.

Deng, H.X., Chen, W., Hong, S.T., Boycott, K.M., Gorrie, G.H., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H*., et al.* (2011). Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature 477, 211-215.

Devlin, A.C., Burr, K., Borooah, S., Foster, J.D., Cleary, E.M., Geti, I., Vallier, L., Shaw, C.E., Chandran, S., and Miles, G.B. (2015). Human iPSC-derived motoneurons harbouring TARDBP or C9ORF72 ALS mutations are dysfunctional despite maintaining viability. Nat Commun 6, 5999.

Di Giorgio, F.P., Boulting, G.L., Bobrowicz, S., and Eggan, K.C. (2008). Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. Cell Stem Cell 3, 637-648.

Dimchev, G.A., Al-Shanti, N., and Stewart, C.E. (2013). Phospho-tyrosine phosphatase inhibitor Bpv(Hopic) enhances C2C12 myoblast migration in vitro. Requirement of PI3K/AKT and MAPK/ERK pathways. J Muscle Res Cell Motil 34, 125-136.

Ding, J., Guo, J., Yuan, Q., Yuan, F., Chen, H., and Tian, H. (2013). Inhibition of phosphatase and tensin homolog deleted on chromosome 10 decreases rat cortical neuron injury and blood-brain barrier permeability, and improves neurological functional recovery in traumatic brain injury model. PLoS One 8, e80429.

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.

Duan, W., Li, X., Shi, J., Guo, Y., Li, Z., and Li, C. (2010). Mutant TAR DNA-binding protein-43 induces oxidative injury in motor neuron-like cell. Neuroscience 169, 1621-1629.

Elden, A.C., Kim, H.J., Hart, M.P., Chen-Plotkin, A.S., Johnson, B.S., Fang, X., Armakola, M., Geser, F., Greene, R., Lu, M.M.*, et al.* (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 466, 1069-1075.

Engelhardt, J.I., and Appel, S.H. (1990). IgG reactivity in the spinal cord and motor cortex in amyotrophic lateral sclerosis. Arch Neurol 47, 1210-1216.

Engelhardt, J.I., Tajti, J., and Appel, S.H. (1993). Lymphocytic infiltrates in the spinal cord in amyotrophic lateral sclerosis. Arch Neurol 50, 30-36.

Esanov, R., Belle, K.C., van Blitterswijk, M., Belzil, V.V., Rademakers, R., Dickson, D.W., Petrucelli, L., Boylan, K.B., Dykxhoorn, D.M., Wuu, J.*, et al.* (2016). C9orf72 promoter hypermethylation is reduced while hydroxymethylation is acquired during reprogramming of ALS patient cells. Exp Neurol 277, 171-177.

Farg, M.A., Soo, K.Y., Warraich, S.T., Sundaramoorthy, V., Blair, I.P., and Atkin, J.D. (2013). Ataxin-2 interacts with FUS and intermediate-length polyglutamine expansions enhance FUS-related pathology in amyotrophic lateral sclerosis. Hum Mol Genet 22, 717-728.

Farg, M.A., Sundaramoorthy, V., Sultana, J.M., Yang, S., Atkinson, R.A., Levina, V., Halloran, M.A., Gleeson, P.A., Blair, I.P., Soo, K.Y.*, et al.* (2014). C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. Hum Mol Genet 23, 3579-3595.

Fecto, F., Yan, J., Vemula, S.P., Liu, E., Yang, Y., Chen, W., Zheng, J.G., Shi, Y., Siddique, N., Arrat, H.*, et al.* (2011). SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. Arch Neurol 68, 1440-1446.

Feiler, M.S., Strobel, B., Freischmidt, A., Helferich, A.M., Kappel, J., Brewer, B.M., Li, D., Thal, D.R., Walther, P., Ludolph, A.C.*, et al.* (2015). TDP-43 is intercellularly transmitted across axon terminals. J Cell Biol 211, 897-911.

Ferrante, R.J., Browne, S.E., Shinobu, L.A., Bowling, A.C., Baik, M.J., MacGarvey, U., Kowall, N.W., Brown, R.H., Jr., and Beal, M.F. (1997). Evidence of increased oxidative damage in both sporadic and familial amyotrophic lateral sclerosis. Journal of neurochemistry 69, 2064-2074.

Ferri, A., Cozzolino, M., Crosio, C., Nencini, M., Casciati, A., Gralla, E.B., Rotilio, G., Valentine, J.S., and Carrì, M.T. (2006). Familial ALS-superoxide dismutases associate with mitochondria and shift their redox potentials. Proc Natl Acad Sci U S A 103, 13860-13865.

Fiesel, F.C., and Kahle, P.J. (2011). TDP-43 and FUS/TLS: cellular functions and implications for neurodegeneration. FEBS J 278, 3550-3568.

Fifita, J.A., Williams, K.L., McCann, E.P., O'Brien, A., Bauer, D.C., Nicholson, G.A., and Blair, I.P. (2015). Mutation analysis of MATR3 in Australian familial amyotrophic lateral sclerosis. Neurobiol Aging 36, 1602.e1601-1602.

Figley, M.D., Bieri, G., Kolaitis, R.M., Taylor, J.P., and Gitler, A.D. (2014). Profilin 1 associates with stress granules and ALS-linked mutations alter stress granule dynamics. J Neurosci 34, 8083-8097.

Fitzmaurice, P.S., Shaw, I.C., Kleiner, H.E., Miller, R.T., Monks, T.J., Lau, S.S., Mitchell, J.D., and Lynch, P.G. (1996). Evidence for DNA damage in amyotrophic lateral sclerosis. Muscle & nerve 19, 797-798.

Fratta, P., Charnock, J., Collins, T., Devoy, A., Howard, R., Malaspina, A., Orrell, R., Sidle, K., Clarke, J., Shoai, M.*, et al.* (2014). Profilin1 E117G is a moderate risk factor for amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 85, 506-508.

Fratta, P., Mizielinska, S., Nicoll, A.J., Zloh, M., Fisher, E.M., Parkinson, G., and Isaacs, A.M. (2012). C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. Sci Rep 2, 1016.

Fratta, P., Poulter, M., Lashley, T., Rohrer, J.D., Polke, J.M., Beck, J., Ryan, N., Hensman, D., Mizielinska, S., Waite, A.J.*, et al.* (2013). Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. Acta Neuropathol 126, 401-409.

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

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Müller, K., Marroquin, N., Nordin, F., Hübers, A., Weydt, P.*, et al.* (2015). Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. Nat Neurosci 18, 631-636.

Fu, X.D., and Maniatis, T. (1990). Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature 343, 437-441.

Fu, X.D., Mayeda, A., Maniatis, T., and Krainer, A.R. (1992). General splicing factors SF2 and SC35 have equivalent activities in vitro, and both affect alternative 5' and 3' splice site selection. Proc Natl Acad Sci U S A 89, 11224-11228.

Fujita, K., Yamauchi, M., Shibayama, K., Ando, M., Honda, M., and Nagata, Y. (1996). Decreased cytochrome c oxidase activity but unchanged superoxide dismutase and glutathione peroxidase activities in the spinal cords of patients with amyotrophic lateral sclerosis. J Neurosci Res 45, 276-281.

Fukunaga, K., Shinoda, Y., and Tagashira, H. (2015). The role of SIGMAR1 gene mutation and mitochondrial dysfunction in amyotrophic lateral sclerosis. J Pharmacol Sci 127, 36-41.

Galimberti, D., Arosio, B., Fenoglio, C., Serpente, M., Cioffi, S.M., Bonsi, R., Rossi, P., Abbate, C., Mari, D., and Scarpini, E. (2014). Incomplete penetrance of the C9ORF72 hexanucleotide repeat expansions: frequency in a cohort of geriatric non-demented subjects. J Alzheimers Dis 39, 19-22.

Gautier, G., Verschueren, A., Monnier, A., Attarian, S., Salort-Campana, E., and Pouget, J. (2010). ALS with respiratory onset: clinical features and effects of non-invasive ventilation on the prognosis. Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases 11, 379-382.

Ge, H., and Manley, J.L. (1990). A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. Cell 62, 25-34.

Gellera, C., Tiloca, C., Del Bo, R., Corrado, L., Pensato, V., Agostini, J., Cereda, C., Ratti, A., Castellotti, B., Corti, S.*, et al.* (2013). Ubiquilin 2 mutations in Italian patients with amyotrophic lateral sclerosis and frontotemporal dementia. J Neurol Neurosurg Psychiatry 84, 183-187.

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.

Gijselinck, I., Van Langenhove, T., van der Zee, J., Sleegers, K., Philtjens, S., Kleinberger, G., Janssens, J., Bettens, K., Van Cauwenberghe, C., Pereson, S.*, et al.* (2012). A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. Lancet neurology 11, 54-65.

Gijselinck, I., Van Mossevelde, S., van der Zee, J., Sieben, A., Engelborghs, S., De Bleecker, J., Ivanoiu, A., Deryck, O., Edbauer, D., Zhang, M.*, et al.* (2015). The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. Mol Psychiatry.

Gilpin, K.M., Chang, L., and Monteiro, M.J. (2015). ALS-linked mutations in ubiquilin-2 or hnRNPA1 reduce interaction between ubiquilin-2 and hnRNPA1. Hum Mol Genet 24, 2565-2577.

Ginisty, H., Amalric, F., and Bouvet, P. (1998). Nucleolin functions in the first step of ribosomal RNA processing. EMBO J 17, 1476-1486.

Ginisty, H., Serin, G., Ghisolfi-Nieto, L., Roger, B., Libante, V., Amalric, F., and Bouvet, P. (2000). Interaction of nucleolin with an evolutionarily conserved pre-ribosomal RNA sequence is required for the assembly of the primary processing complex. J Biol Chem 275, 18845-18850.

Gitcho, M.A., Baloh, R.H., Chakraverty, S., Mayo, K., Norton, J.B., Levitch, D., Hatanpaa, K.J., White, C.L., Bigio, E.H., Caselli, R.*, et al.* (2008). TDP-43 A315T mutation in familial motor neuron disease. Ann Neurol 63, 535-538.

Goldstein, O., Nayshool, O., Nefussy, B., Traynor, B.J., Renton, A.E., Gana-Weisz, M., Drory, V.E., and Orr-Urtreger, A. (2016). OPTN 691_692insAG is a founder mutation causing recessive ALS and increased risk in heterozygotes. Neurology 86, 446-453.

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.

Gordon, P.H. (2013). Amyotrophic Lateral Sclerosis: An update for 2013 Clinical Features, Pathophysiology, Management and Therapeutic Trials. Aging Dis 4, 295-310.

Greenway, M.J., Andersen, P.M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., Patterson, V., Swingler, R., Kieran, D., Prehn, J.*, et al.* (2006). ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. Nat Genet 38, 411-413.

Grosskreutz, J., Van Den Bosch, L., and Keller, B.U. (2010). Calcium dysregulation in amyotrophic lateral sclerosis. Cell Calcium 47, 165-174.

Guo, J.Y., Ding, J., Yuan, F., Chen, H., Chen, S.W., and Tian, H.L. (2013). Dose-dependent protective effect of bisperoxovanadium against acute cerebral ischemia in a rat model of ischemia/reperfusion injury. Int J Mol Sci 14, 12013-12022.

Haas, S., Gordon, J., and Khalili, K. (1993). A developmentally regulated DNA-binding protein from mouse brain stimulates myelin basic protein gene expression. Mol Cell Biol 13, 3103-3112.

Haas, S., Thatikunta, P., Steplewski, A., Johnson, E.M., Khalili, K., and Amini, S. (1995). A 39-kD DNA-binding protein from mouse brain stimulates transcription of myelin basic protein gene in oligodendrocytic cells. J Cell Biol 130, 1171-1179.

Hadano, S., Hand, C.K., Osuga, H., Yanagisawa, Y., Otomo, A., Devon, R.S., Miyamoto, N., Showguchi-Miyata, J., Okada, Y., Singaraja, R.*, et al.* (2001). A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. Nat Genet 29, 166-173.

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.

Haidet-Phillips, A.M., Hester, M.E., Miranda, C.J., Meyer, K., Braun, L., Frakes, A., Song, S., Likhite, S., Murtha, M.J., Foust, K.D.*, et al.* (2011). Astrocytes from familial and sporadic ALS patients are toxic to motor neurons. Nat Biotechnol 29, 824-828.

Hall, E.D., Oostveen, J.A., and Gurney, M.E. (1998). Relationship of microglial and astrocytic activation to disease onset and progression in a transgenic model of familial ALS. Glia 23, 249-256.

Hand, C.K., Khoris, J., Salachas, F., Gros-Louis, F., Lopes, A.A., Mayeux-Portas, V., Brewer, C.G., Brown, R.H., Meininger, V., Camu, W., and Rouleau, G.A. (2002). A novel locus for familial amyotrophic lateral sclerosis, on chromosome 18q. Am J Hum Genet 70, 251-256.

Harms, M.B., Cady, J., Zaidman, C., Cooper, P., Bali, T., Allred, P., Cruchaga, C., Baughn, M., Libby, R.T., Pestronk, A.*, et al.* (2013). Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. Neurobiol Aging 34, 2234.e2213-2239.

Hayashi, T., and Su, T.P. (2007). Sigma-1 receptor chaperones at the ER-mitochondrion interface regulate Ca(2+) signaling and cell survival. Cell 131, 596-610.

Henne, W.M., Buchkovich, N.J., and Emr, S.D. (2011). The ESCRT pathway. Dev Cell 21, 77-91.

Hers, I., Vincent, E.E., and Tavaré, J.M. (2011). Akt signalling in health and disease. Cell Signal 23, 1515-1527.

Hideyama, T., Yamashita, T., Aizawa, H., Tsuji, S., Kakita, A., Takahashi, H., and Kwak, S. (2012). Profound downregulation of the RNA editing enzyme ADAR2 in ALS spinal motor neurons. Neurobiol Dis 45, 1121-1128.

Highley, J.R., Kirby, J., Jansweijer, J.A., Webb, P.S., Hewamadduma, C.A., Heath, P.R., Higginbottom, A., Raman, R., Ferraiuolo, L., Cooper-Knock, J.*, et al.* (2014). Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones. Neuropathol Appl Neurobiol 40, 670-685.

Hirano, M., Quinzii, C.M., Mitsumoto, H., Hays, A.P., Roberts, J.K., Richard, P., and Rowland, L.P. (2011). Senataxin mutations and amyotrophic lateral sclerosis. Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases 12, 223-227.

Hirose, T., Ideue, T., Nagai, M., Hagiwara, M., Shu, M.D., and Steitz, J.A. (2006). A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. Mol Cell 23, 673-684.

Huang, Y., Gattoni, R., Stévenin, J., and Steitz, J.A. (2003). SR splicing factors serve as adapter proteins for TAP-dependent mRNA export. Mol Cell 11, 837-843.

Hukema, R.K., Riemslagh, F.W., Melhem, S., van der Linde, H.C., Severijnen, L.A., Edbauer, D., Maas, A., Charlet-Berguerand, N., Willemsen, R., and van Swieten, J.C. (2014). A new inducible transgenic mouse model for C9orf72-associated GGGGCC repeat expansion supports a gain-of-function mechanism in C9orf72-associated ALS and FTD. Acta Neuropathol Commun 2, 166.

Huyer, G., Liu, S., Kelly, J., Moffat, J., Payette, P., Kennedy, B., Tsaprailis, G., Gresser, M.J., and Ramachandran, C. (1997). Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate. J Biol Chem 272, 843-851.

Iioka, H., Loiselle, D., Haystead, T.A., and Macara, I.G. (2011). Efficient detection of RNA-protein interactions using tethered RNAs. Nucleic Acids Res 39, e53.

Ilieva, H., Nagano, I., Murakami, T., Shiote, M., Shoji, M., and Abe, K. (2003). Sustained induction of survival p-AKT and p-ERK signals after transient hypoxia in mice spinal cord with G93A mutant human SOD1 protein. J Neurol Sci 215, 57-62.

Ince, P.G., Tomkins, J., Slade, J.Y., Thatcher, N.M., and Shaw, P.J. (1998). Amyotrophic lateral sclerosis associated with genetic abnormalities in the gene encoding Cu/Zn superoxide dismutase: molecular pathology of five new cases, and comparison with previous reports and 73 sporadic cases of ALS. Journal of neuropathology and experimental neurology 57, 895-904.

Ingre, C., Landers, J.E., Rizik, N., Volk, A.E., Akimoto, C., Birve, A., Hübers, A., Keagle, P.J., Piotrowska, K., Press, R.*, et al.* (2013a). A novel phosphorylation site mutation in profilin 1 revealed in a large screen of US, Nordic, and German amyotrophic lateral sclerosis/frontotemporal dementia cohorts. Neurobiol Aging 34, 1708.e1701-1706.

Ingre, C., Pinto, S., Birve, A., Press, R., Danielsson, O., de Carvalho, M., Guđmundsson, G., and Andersen, P.M. (2013b). No association between VAPB mutations and familial or sporadic

ALS in Sweden, Portugal and Iceland. Amyotroph Lateral Scler Frontotemporal Degener 14, 620-627.

Ivanov, P., O'Day, E., Emara, M.M., Wagner, G., Lieberman, J., and Anderson, P. (2014). G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. Proc Natl Acad Sci U S A 111, 18201-18206.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol 11, 113-127.

James, A., Wang, Y., Raje, H., Rosby, R., and DiMario, P. (2014). Nucleolar stress with and without p53. Nucleus 5, 402-426.

Johnson, J.O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V.M., Trojanowski, J.Q., Gibbs, J.R., Brunetti, M., Gronka, S., Wuu, J.*, et al.* (2010). Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 68, 857-864.

Johnson, J.O., Pioro, E.P., Boehringer, A., Chia, R., Feit, H., Renton, A.E., Pliner, H.A., Abramzon, Y., Marangi, G., Winborn, B.J.*, et al.* (2014). Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. Nat Neurosci 17, 664-666.

Johnston, C.A., Stanton, B.R., Turner, M.R., Gray, R., Blunt, A.H., Butt, D., Ampong, M.A., Shaw, C.E., Leigh, P.N., and Al-Chalabi, A. (2006). Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. J Neurol 253, 1642-1643.

Jovičić, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N., Yamada, S.B., Paul, J.W., Sun, S., Herdy, J.R., Bieri, G.*, et al.* (2015). Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. Nat Neurosci 18, 1226-1229.

Joyce, P.I., Fratta, P., Fisher, E.M., and Acevedo-Arozena, A. (2011). SOD1 and TDP-43 animal models of amyotrophic lateral sclerosis: recent advances in understanding disease toward the development of clinical treatments. Mamm Genome 22, 420-448.

Ju, J.S., Fuentealba, R.A., Miller, S.E., Jackson, E., Piwnica-Worms, D., Baloh, R.H., and Weihl, C.C. (2009). Valosin-containing protein (VCP) is required for autophagy and is disrupted in VCP disease. J Cell Biol 187, 875-888.

Jurk, M., Weissinger, F., Lottspeich, F., Schwarz, U., and Winnacker, E.L. (1996). Characterization of the single-strand-specific BPV-1 origin binding protein, SPSF I, as the HeLa Pur alpha factor. Nucleic Acids Res 24, 2799-2806.

Kaal, E.C., Vlug, A.S., Versleijen, M.W., Kuilman, M., Joosten, E.A., and Bär, P.R. (2000). Chronic mitochondrial inhibition induces selective motoneuron death in vitro: a new model for amyotrophic lateral sclerosis. J Neurochem 74, 1158-1165.

Kabashi, E., El Oussini, H., Bercier, V., Gros-Louis, F., Valdmanis, P.N., McDearmid, J., Mejier, I.A., Dion, P.A., Dupre, N., Hollinger, D.*, et al.* (2013). Investigating the contribution of VAPB/ALS8 loss of function in amyotrophic lateral sclerosis. Hum Mol Genet 22, 2350-2360.

Kabashi, E., Valdmanis, P.N., Dion, P., Spiegelman, D., McConkey, B.J., Vande Velde, C., Bouchard, J.P., Lacomblez, L., Pochigaeva, K., Salachas, F.*, et al.* (2008). TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet 40, 572-574.

Kamburov, A., Cavill, R., Ebbels, T.M., Herwig, R., and Keun, H.C. (2011). Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. Bioinformatics 27, 2917-2918.

Kaneb, H.M., Folkmann, A.W., Belzil, V.V., Jao, L.E., Leblond, C.S., Girard, S.L., Daoud, H., Noreau, A., Rochefort, D., Hince, P.*, et al.* (2015). Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. Hum Mol Genet 24, 1363-1373.

Kanekura, K., Nishimoto, I., Aiso, S., and Matsuoka, M. (2006). Characterization of amyotrophic lateral sclerosis-linked P56S mutation of vesicle-associated membrane protein-associated protein B (VAPB/ALS8). J Biol Chem 281, 30223-30233.

Karam, C., Scelsa, S.N., and Macgowan, D.J. (2010). The clinical course of progressive bulbar palsy. Amyotroph Lateral Scler 11, 364-368.

Khurts, S., Masutomi, K., Delgermaa, L., Arai, K., Oishi, N., Mizuno, H., Hayashi, N., Hahn, W.C., and Murakami, S. (2004). Nucleolin interacts with telomerase. J Biol Chem 279, 51508-51515.

Kihira, T., Yoshida, S., Kondo, T., Iwai, K., Wada, S., Morinaga, S., Kazimoto, Y., Okamoto, K., Kokubo, Y., and Kuzuhara, S. (2012). An increase in ALS incidence on the Kii Peninsula, 1960-2009: a possible link to change in drinking water source. Amyotroph Lateral Scler 13, 347-350.

Kim, H.J., Kim, N.C., Wang, Y.D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A.*, et al.* (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature 495, 467-473.

King, A., Maekawa, S., Bodi, I., Troakes, C., and Al-Sarraj, S. (2011). Ubiquitinated, p62 immunopositive cerebellar cortical neuronal inclusions are evident across the spectrum of TDP-43 proteinopathies but are only rarely additionally immunopositive for phosphorylation-dependent TDP-43. Neuropathology 31, 239-249.

King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res 1462, 61-80.

Kirby, J., Halligan, E., Baptista, M.J., Allen, S., Heath, P.R., Holden, H., Barber, S.C., Loynes, C.A., Wood-Allum, C.A., Lunec, J., and Shaw, P.J. (2005). Mutant SOD1 alters the motor neuronal transcriptome: implications for familial ALS. Brain 128, 1686-1706.

Kirby, J., Highley, J.R., Cox, L., Goodall, E.F., Hewitt, C., Hartley, J.A., Hollinger, H.C., Fox, M., Ince, P.G., McDermott, C.J., and Shaw, P.J. (2013). Lack of unique neuropathology in amyotrophic lateral sclerosis associated with p.K54E angiogenin (ANG) mutation. Neuropathol Appl Neurobiol 39, 562-571.

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.

Kiskinis, E., Sandoe, J., Williams, L.A., Boulting, G.L., Moccia, R., Wainger, B.J., Han, S., Peng, T., Thams, S., Mikkilineni, S.*, et al.* (2014). Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. Cell Stem Cell 14, 781-795.

Koh, S.H., Kwon, H., Kim, K.S., Kim, J., Kim, M.H., Yu, H.J., Kim, M., Lee, K.W., Do, B.R., Jung, H.K.*, et al.* (2004). Epigallocatechin gallate prevents oxidative-stress-induced death of mutant Cu/Zn-superoxide dismutase (G93A) motoneuron cells by alteration of cell survival and death signals. Toxicology 202, 213-225.

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.

Koppers, M., van Blitterswijk, M.M., Vlam, L., Rowicka, P.A., van Vught, P.W., Groen, E.J., Spliet, W.G., Engelen-Lee, J., Schelhaas, H.J., de Visser, M.*, et al.* (2012). VCP mutations in familial and sporadic amyotrophic lateral sclerosis. Neurobiol Aging 33, 837.e837-813.

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

Krainer, A.R., Conway, G.C., and Kozak, D. (1990). Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells. Genes Dev 4, 1158-1171.

Kressler, D., Linder, P., and de La Cruz, J. (1999). Protein trans-acting factors involved in ribosome biogenesis in Saccharomyces cerevisiae. Mol Cell Biol 19, 7897-7912.

Kwiatkowski, T.J., Jr., Bosco, D.A., Leclerc, A.L., Tamrazian, E., Vanderburg, C.R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E.J., Munsat, T.*, et al.* (2009). Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science 323, 1205-1208.

Kwok, C.T., Morris, A., and de Belleroche, J.S. (2014). Sequestosome-1 (SQSTM1) sequence variants in ALS cases in the UK: prevalence and coexistence of SQSTM1 mutations in ALS kindred with PDB. Eur J Hum Genet 22, 492-496.

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.

Kyotani, A., Azuma, Y., Yamamoto, I., Yoshida, H., Mizuta, I., Mizuno, T., Nakagawa, M., Tokuda, T., and Yamaguchi, M. (2016). Knockdown of the Drosophila FIG4 induces deficient locomotive behavior, shortening of motor neuron, axonal targeting aberration, reduction of life span and defects in eye development. Exp Neurol 277, 86-95.

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 U S A 110, E4530-4539.

Lai, C., Xie, C., Shim, H., Chandran, J., Howell, B.W., and Cai, H. (2009). Regulation of endosomal motility and degradation by amyotrophic lateral sclerosis 2/alsin. Mol Brain 2, 23.

Lai, J.P., Dalton, J.T., and Knoell, D.L. (2007). Phosphatase and tensin homologue deleted on chromosome ten (PTEN) as a molecular target in lung epithelial wound repair. Br J Pharmacol 152, 1172-1184.

Lai, M.C., and Tarn, W.Y. (2004). Hypophosphorylated ASF/SF2 binds TAP and is present in messenger ribonucleoproteins. J Biol Chem 279, 31745-31749.

Lattante, S., de Calbiac, H., Le Ber, I., Brice, A., Ciura, S., and Kabashi, E. (2015). Sqstm1 knock-down causes a locomotor phenotype ameliorated by rapamycin in a zebrafish model of ALS/FTLD. Hum Mol Genet 24, 1682-1690.

Leblond, C.S., Gan-Or, Z., Spiegelman, D., Laurent, S.B., Szuto, A., Hodgkinson, A., Dionne-Laporte, A., Provencher, P., de Carvalho, M., Orrù, S.*, et al.* (2016). Replication study of MATR3 in familial and sporadic amyotrophic lateral sclerosis. Neurobiol Aging 37, 209.e217-221.

Lee, C.U., Hahne, G., Hanske, J., Bange, T., Bier, D., Rademacher, C., Hennig, S., and Grossmann, T.N. (2015). Redox Modulation of PTEN Phosphatase Activity by Hydrogen Peroxide and Bisperoxidovanadium Complexes. Angew Chem Int Ed Engl 54, 13796-13800.

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.

Leigh, P.N., Whitwell, H., Garofalo, O., Buller, J., Swash, M., Martin, J.E., Gallo, J.M., Weller, R.O., and Anderton, B.H. (1991). Ubiquitin-immunoreactive intraneuronal inclusions in amyotrophic lateral sclerosis. Morphology, distribution, and specificity. Brain 114 (Pt 2), 775-788.

Leslie, N.R., Yang, X., Downes, C.P., and Weijer, C.J. (2007). PtdIns(3,4,5)P(3)-dependent and -independent roles for PTEN in the control of cell migration. Curr Biol 17, 115-125.

Lev, S., Ben Halevy, D., Peretti, D., and Dahan, N. (2008). The VAP protein family: from cellular functions to motor neuron disease. Trends Cell Biol 18, 282-290.

Levine, T.P., Daniels, R.D., Gatta, A.T., Wong, L.H., and Hayes, M.J. (2013). The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics 29, 499-503.

Li, Y.R., King, O.D., Shorter, J., and Gitler, A.D. (2013). Stress granules as crucibles of ALS pathogenesis. J Cell Biol 201, 361-372.

Lillo, P., and Hodges, J.R. (2009). Frontotemporal dementia and motor neurone disease: overlapping clinic-pathological disorders. J Clin Neurosci 16, 1131-1135.

Lin, K.P., Tsai, P.C., Liao, Y.C., Chen, W.T., Tsai, C.P., Soong, B.W., and Lee, Y.C. (2015). Mutational analysis of MATR3 in Taiwanese patients with amyotrophic lateral sclerosis. Neurobiol Aging 36, 2005.e2001-2004.

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X.D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. Nat Struct Mol Biol 15, 819-826.

Ling, J.P., Pletnikova, O., Troncoso, J.C., and Wong, P.C. (2015). TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. Science 349, 650-655.

Ling, S.C., Albuquerque, C.P., Han, J.S., Lagier-Tourenne, C., Tokunaga, S., Zhou, H., and Cleveland, D.W. (2010). ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. Proc Natl Acad Sci U S A 107, 13318-13323.

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.

Liu, D., Wen, J., Liu, J., and Li, L. (1999). The roles of free radicals in amyotrophic lateral sclerosis: reactive oxygen species and elevated oxidation of protein, DNA, and membrane phospholipids. FASEB J 13, 2318-2328.

Liu, R., Althaus, J.S., Ellerbrock, B.R., Becker, D.A., and Gurney, M.E. (1998). Enhanced oxygen radical production in a transgenic mouse model of familial amyotrophic lateral sclerosis. Ann Neurol 44, 763-770.

Logroscino, G., Traynor, B.J., Hardiman, O., Chio', A., Couratier, P., Mitchell, J.D., Swingler, R.J., Beghi, E., and EURALS (2008). Descriptive epidemiology of amyotrophic lateral sclerosis: new evidence and unsolved issues. J Neurol Neurosurg Psychiatry 79, 6-11.

Lomen-Hoerth, C., Anderson, T., and Miller, B. (2002). The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. Neurology 59, 1077-1079.

Lomen-Hoerth, C., Murphy, J., Langmore, S., Kramer, J.H., Olney, R.K., and Miller, B. (2003). Are amyotrophic lateral sclerosis patients cognitively normal? Neurology 60, 1094-1097.

Luty, A.A., Kwok, J.B., Dobson-Stone, C., Loy, C.T., Coupland, K.G., Karlström, H., Sobow, T., Tchorzewska, J., Maruszak, A., Barcikowska, M*., et al.* (2010). Sigma nonopioid intracellular receptor 1 mutations cause frontotemporal lobar degeneration-motor neuron disease. Ann Neurol 68, 639-649.

Mackenzie, I.R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S., Mori, K., Weng, S.M., Haass, C., Kretzschmar, H.A., Edbauer, D., and Neumann, M. (2013). Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. Acta Neuropathol 126, 859-879.

Mackenzie, I.R., Bigio, E.H., Ince, P.G., Geser, F., Neumann, M., Cairns, N.J., Kwong, L.K., Forman, M.S., Ravits, J., Stewart, H*., et al.* (2007). Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. Ann Neurol 61, 427-434.

Mackenzie, I.R., Frick, P., Grässer, F.A., Gendron, T.F., Petrucelli, L., Cashman, N.R., Edbauer, D., Kremmer, E., Prudlo, J., Troost, D., and Neumann, M. (2015). Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. Acta Neuropathol 130, 845-861.

Majounie, E., Renton, A.E., Mok, K., Dopper, E.G., Waite, A., Rollinson, S., Chio, A., Restagno, G., Nicolaou, N., Simon-Sanchez, J*., et al.* (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol 11, 323-330.

Mann, D.M., Rollinson, S., Robinson, A., Bennion Callister, J., Thompson, J.C., Snowden, J.S., Gendron, T., Petrucelli, L., Masuda-Suzukake, M., Hasegawa, M*., et al.* (2013). Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal

lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun 1, 68.

Mao, L., Jia, J., Zhou, X., Xiao, Y., Wang, Y., Mao, X., Zhen, X., Guan, Y., Alkayed, N.J., and Cheng, J. (2013). Delayed administration of a PTEN inhibitor BPV improves functional recovery after experimental stroke. Neuroscience 231, 272-281.

Mao, L.L., Hao, D.L., Mao, X.W., Xu, Y.F., Huang, T.T., Wu, B.N., and Wang, L.H. (2015). Neuroprotective effects of bisperoxovanadium on cerebral ischemia by inflammation inhibition. Neurosci Lett 602, 120-125.

Marchesi, C., Ciano, C., Salsano, E., Nanetti, L., Milani, M., Gellera, C., Taroni, F., Fabrizi, G.M., Uncini, A., and Pareyson, D. (2011). Co-occurrence of amyotrophic lateral sclerosis and Charcot-Marie-Tooth disease type 2A in a patient with a novel mutation in the mitofusin-2 gene. Neuromuscul Disord 21, 129-131.

Marin, B., Logroscino, G., Boumédiene, F., Labrunie, A., Couratier, P., Babron, M.C., Leutenegger, A.L., Preux, P.M., and Beghi, E. (2015). Clinical and demographic factors and outcome of amyotrophic lateral sclerosis in relation to population ancestral origin. Eur J Epidemiol.

Marques Sousa, C., and Humbert, S. (2013). Huntingtin: here, there, everywhere! J Huntingtons Dis 2, 395-403.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H*., et al.* (2010). Mutations of optineurin in amyotrophic lateral sclerosis. Nature 465, 223-226.

Matus, S., Valenzuela, V., Medinas, D.B., and Hetz, C. (2013). ER Dysfunction and Protein Folding Stress in ALS. Int J Cell Biol 2013, 674751.

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.

McKhann, G.M., Albert, M.S., Grossman, M., Miller, B., Dickson, D., Trojanowski, J.Q., and Disease, W.G.o.F.D.a.P.s. (2001). Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. Arch Neurol 58, 1803-1809.

McMahon, M., Itoh, K., Yamamoto, M., Chanas, S.A., Henderson, C.J., McLellan, L.I., Wolf, C.R., Cavin, C., and Hayes, J.D. (2001). The Cap'n'Collar basic leucine zipper transcription factor Nrf2 (NF-E2 p45-related factor 2) controls both constitutive and inducible expression of intestinal detoxification and glutathione biosynthetic enzymes. Cancer Res 61, 3299-3307.

Mead, R.J., Higginbottom, A., Allen, S.P., Kirby, J., Bennett, E., Barber, S.C., Heath, P.R., Coluccia, A., Patel, N., Gardner, I*., et al.* (2013). S[+] Apomorphine is a CNS penetrating activator of the Nrf2-ARE pathway with activity in mouse and patient fibroblast models of amyotrophic lateral sclerosis. Free Radic Biol Med 61, 438-452.

Menzies, F.M., Cookson, M.R., Taylor, R.W., Turnbull, D.M., Chrzanowska-Lightowlers, Z.M., Dong, L., Figlewicz, D.A., and Shaw, P.J. (2002a). Mitochondrial dysfunction in a cell culture model of familial amyotrophic lateral sclerosis. Brain 125, 1522-1533.

Menzies, F.M., Ince, P.G., and Shaw, P.J. (2002b). Mitochondrial involvement in amyotrophic lateral sclerosis. Neurochem Int 40, 543-551.

Meyer, H., and Weihl, C.C. (2014). The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. J Cell Sci 127, 3877-3883.

Meyer, K., Ferraiuolo, L., Miranda, C.J., Likhite, S., McElroy, S., Renusch, S., Ditsworth, D., Lagier-Tourenne, C., Smith, R.A., Ravits, J*., et al.* (2014). Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS. Proc Natl Acad Sci U S A 111, 829-832.

Michlewski, G., Sanford, J.R., and Cáceres, J.F. (2008). The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. Mol Cell 30, 179-189.

Millecamps, S., Boillée, S., Le Ber, I., Seilhean, D., Teyssou, E., Giraudeau, M., Moigneu, C., Vandenberghe, N., Danel-Brunaud, V., Corcia, P.*, et al.* (2012). Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. J Med Genet 49, 258-263.

Millecamps, S., De Septenville, A., Teyssou, E., Daniau, M., Camuzat, A., Albert, M., LeGuern, E., Galimberti, D., Brice, A., Marie, Y.*, et al.* (2014). Genetic analysis of matrin 3 gene in French amyotrophic lateral sclerosis patients and frontotemporal lobar degeneration with amyotrophic lateral sclerosis patients. Neurobiol Aging 35, 2882.e2813-2885.

Millecamps, S., Salachas, F., Cazeneuve, C., Gordon, P., Bricka, B., Camuzat, A., Guillot-Noël, L., Russaouen, O., Bruneteau, G., Pradat, P.F.*, et al.* (2010). SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations. J Med Genet 47, 554-560.

Mitchell, J.D., and Borasio, G.D. (2007). Amyotrophic lateral sclerosis. Lancet 369, 2031-2041.

Mitsumoto, H., Santella, R.M., Liu, X., Bogdanov, M., Zipprich, J., Wu, H.C., Mahata, J., Kilty, M., Bednarz, K., Bell, D.*, et al.* (2008). Oxidative stress biomarkers in sporadic ALS. Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases 9, 177-183.

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.

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

Mizuno, Y., Amari, M., Takatama, M., Aizawa, H., Mihara, B., and Okamoto, K. (2006). Transferrin localizes in Bunina bodies in amyotrophic lateral sclerosis. Acta Neuropathol 112, 597-603.

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

Mori, K., Lammich, S., Mackenzie, I.R.A., Forne, I., Zilow, S., Kretzschmar, H., Edbauer, D., Janssens, J., Kleinberger, G., Cruts, M.*, et al.* (2013b). hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. Acta Neuropathol 125, 413-423.

Mori, K., Weng, S.-M., Arzberger, T., May, S., Rentzsch, K., Kremmer, E., Schmid, B., Kretzschmar, H.A., Cruts, M., Van Broeckhoven, C.*, et al.* (2013c). The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. Science 339, 1335-1338.

Morton, S., Hesson, L., Peggie, M., and Cohen, P. (2008). Enhanced binding of TBK1 by an optineurin mutant that causes a familial form of primary open angle glaucoma. FEBS Lett 582, 997-1002.

Murphy, R., and Wente, S.R. (1996). An RNA-export mediator with an essential nuclear export signal. Nature 383, 357-360.

Murray, M.E., DeJesus-Hernandez, M., Rutherford, N.J., Baker, M., Duara, R., Graff-Radford, N.R., Wszolek, Z.K., Ferman, T.J., Josephs, K.A., Boylan, K.B.*, et al.* (2011). Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. Acta Neuropathol 122, 673-690.

Mórotz, G.M., De Vos, K.J., Vagnoni, A., Ackerley, S., Shaw, C.E., and Miller, C.C. (2012). Amyotrophic lateral sclerosis-associated mutant VAPBP56S perturbs calcium homeostasis to disrupt axonal transport of mitochondria. Hum Mol Genet 21, 1979-1988.

Münch, C., O'Brien, J., and Bertolotti, A. (2011). Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. Proc Natl Acad Sci U S A 108, 3548-3553.

Nagai, M., Re, D.B., Nagata, T., Chalazonitis, A., Jessell, T.M., Wichterle, H., and Przedborski, S. (2007). Astrocytes expressing ALS-linked mutated SOD1 release factors selectively toxic to motor neurons. Nat Neurosci 10, 615-622.

Nagy, D., Kato, T., and Kushner, P.D. (1994). Reactive astrocytes are widespread in the cortical gray matter of amyotrophic lateral sclerosis. J Neurosci Res 38, 336-347.

Neuenschwander, A.G., Thai, K.K., Figueroa, K.P., and Pulst, S.M. (2014). Amyotrophic lateral sclerosis risk for spinocerebellar ataxia type 2 ATXN2 CAG repeat alleles: a meta-analysis. JAMA Neurol 71, 1529-1534.

Neumann, M., Sampathu, D.M., Kwong, L.K., Truax, A.C., Micsenyi, M.C., Chou, T.T., Bruce, J., Schuck, T., Grossman, M., Clark, C.M.*, et al.* (2006). Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science 314, 130-133.

Nguyen, T., Sherratt, P.J., and Pickett, C.B. (2003). Regulatory mechanisms controlling gene expression mediated by the antioxidant response element. Annu Rev Pharmacol Toxicol 43, 233-260.

Niblock, M., Smith, B.N., Lee, Y.B., Sardone, V., Topp, S., Troakes, C., Al-Sarraj, S., Leblond, C.S., Dion, P.A., Rouleau, G.A.*, et al.* (2016). Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: implications for the pathogenesis of ALS/FTD. Acta Neuropathol Commun 4, 18.

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.

Nishimura, A.L., Mitne-Neto, M., Silva, H.C., Richieri-Costa, A., Middleton, S., Cascio, D., Kok, F., Oliveira, J.R., Gillingwater, T., Webb, J.*, et al.* (2004). A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. Am J Hum Genet 75, 822-831.

Nonis, D., Schmidt, M.H., van de Loo, S., Eich, F., Dikic, I., Nowock, J., and Auburger, G. (2008). Ataxin-2 associates with the endocytosis complex and affects EGF receptor trafficking. Cell Signal 20, 1725-1739.

Nordin, A., Akimoto, C., Wuolikainen, A., Alstermark, H., Jonsson, P., Birve, A., Marklund, S.L., Graffmo, K.S., Forsberg, K., Brännström, T., and Andersen, P.M. (2015). Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. Hum Mol Genet 24, 3133-3142.

Nousiainen, H.O., Kestilä, M., Pakkasjärvi, N., Honkala, H., Kuure, S., Tallila, J., Vuopala, K., Ignatius, J., Herva, R., and Peltonen, L. (2008). Mutations in mRNA export mediator GLE1 result in a fetal motoneuron disease. Nat Genet 40, 155-157.

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.

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.

Ohashi, S., Kobayashi, S., Omori, A., Ohara, S., Omae, A., Muramatsu, T., Li, Y., and Anzai, K. (2000). The single-stranded DNA- and RNA-binding proteins pur alpha and pur beta link BC1 RNA to microtubules through binding to the dendrite-targeting RNA motifs. J Neurochem 75, 1781-1790.

Ohashi, S., Koike, K., Omori, A., Ichinose, S., Ohara, S., Kobayashi, S., Sato, T.A., and Anzai, K. (2002). Identification of mRNA/protein (mRNP) complexes containing Puralpha,

mStaufen, fragile X protein, and myosin Va and their association with rough endoplasmic reticulum equipped with a kinesin motor. J Biol Chem 277, 37804-37810.

Okamoto, K., Hirai, S., Amari, M., Watanabe, M., and Sakurai, A. (1993). Bunina bodies in amyotrophic lateral sclerosis immunostained with rabbit anti-cystatin C serum. Neurosci Lett 162, 125-128.

Okamoto, K., Mizuno, Y., and Fujita, Y. (2008). Bunina bodies in amyotrophic lateral sclerosis. Neuropathology : official journal of the Japanese Society of Neuropathology 28, 109-115.

Orlacchio, A., Babalini, C., Borreca, A., Patrono, C., Massa, R., Basaran, S., Munhoz, R.P., Rogaeva, E.A., St George-Hyslop, P.H., Bernardi, G., and Kawarai, T. (2010). SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. Brain 133, 591-598.

Osaka, M., Ito, D., Yagi, T., Nihei, Y., and Suzuki, N. (2015). Evidence of a link between ubiquilin 2 and optineurin in amyotrophic lateral sclerosis. Hum Mol Genet 24, 1617-1629.

Pan, L., Deng, X., Ding, D., Leng, H., Zhu, X., and Wang, Z. (2015). Association between the Angiogenin (ANG) K17I variant and amyotrophic lateral sclerosis risk in Caucasian: a meta-analysis. Neurol Sci 36, 2163-2168.

Parkinson, N., Ince, P.G., Smith, M.O., Highley, R., Skibinski, G., Andersen, P.M., Morrison, K.E., Pall, H.S., Hardiman, O., Collinge, J.*, et al.* (2006). ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). Neurology 67, 1074-1077.

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.

Peviani, M., Cheroni, C., Troglio, F., Quarto, M., Pelicci, G., and Bendotti, C. (2007). Lack of changes in the PI3K/AKT survival pathway in the spinal cord motor neurons of a mouse model of familial amyotrophic lateral sclerosis. Mol Cell Neurosci 34, 592-602.

Philips, T., and Robberecht, W. (2011). Neuroinflammation in amyotrophic lateral sclerosis: role of glial activation in motor neuron disease. Lancet Neurol 10, 253-263.

Philpott, N.J., Nociari, M., Elkon, K.B., and Falck-Pedersen, E. (2004). Adenovirus-induced maturation of dendritic cells through a PI3 kinase-mediated TNF-alpha induction pathway. Proc Natl Acad Sci U S A 101, 6200-6205.

Piao, Y.S., Wakabayashi, K., Kakita, A., Yamada, M., Hayashi, S., Morita, T., Ikuta, F., Oyanagi, K., and Takahashi, H. (2003). Neuropathology with clinical correlations of sporadic amyotrophic lateral sclerosis: 102 autopsy cases examined between 1962 and 2000. Brain Pathol 13, 10-22.

Pilli, M., Arko-Mensah, J., Ponpuak, M., Roberts, E., Master, S., Mandell, M.A., Dupont, N., Ornatowski, W., Jiang, S., Bradfute, S.B.*, et al.* (2012). TBK-1 promotes autophagy-mediated antimicrobial defense by controlling autophagosome maturation. Immunity 37, 223-234.

Polymenidou, M., and Cleveland, D.W. (2011). The seeds of neurodegeneration: prion-like spreading in ALS. Cell 147, 498-508.

Poon, H.F., Hensley, K., Thongboonkerd, V., Merchant, M.L., Lynn, B.C., Pierce, W.M., Klein, J.B., Calabrese, V., and Butterfield, D.A. (2005). Redox proteomics analysis of oxidatively modified proteins in G93A-SOD1 transgenic mice--a model of familial amyotrophic lateral sclerosis. Free Radic Biol Med 39, 453-462.

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.

Pérez-Brangulí, F., Mishra, H.K., Prots, I., Havlicek, S., Kohl, Z., Saul, D., Rummel, C., Dorca-Arevalo, J., Regensburger, M., Graef, D.*, et al.* (2014). Dysfunction of spatacsin leads to

axonal pathology in SPG11-linked hereditary spastic paraplegia. Hum Mol Genet 23, 4859-4874.
Rakhit, R., and Chakrabartty, A. (2006). Structure, folding, and misfolding of Cu,Zn superoxide dismutase in amyotrophic lateral sclerosis. Biochim Biophys Acta 1762, 1025-1037.
Rakhit, R., Crow, J.P., Lepock, J.R., Kondejewski, L.H., Cashman, N.R., and Chakrabartty, A. (2004). Monomeric Cu,Zn-superoxide dismutase is a common misfolding intermediate in the oxidation models of sporadic and familial amyotrophic lateral sclerosis. J Biol Chem 279, 15499-15504.
Ralser, M., Nonhoff, U., Albrecht, M., Lengauer, T., Wanker, E.E., Lehrach, H., and Krobitsch, S. (2005). Ataxin-2 and huntingtin interact with endophilin-A complexes to function in plastin-associated pathways. Hum Mol Genet 14, 2893-2909.
Ramaswami, M., Taylor, J.P., and Parker, R. (2013). Altered ribostasis: RNA-protein granules in degenerative disorders. Cell 154, 727-736.
Ratnavalli, E., Brayne, C., Dawson, K., and Hodges, J.R. (2002). The prevalence of frontotemporal dementia. Neurology 58, 1615-1621.
Ravits, J.M., and La Spada, A.R. (2009). ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration. Neurology 73, 805-811.
Reddy, K., Schmidt, M.H., Geist, J.M., Thakkar, N.P., Panigrahi, G.B., Wang, Y.H., and Pearson, C.E. (2014). Processing of double-R-loops in (CAG)·(CTG) and C9orf72 (GGGGCC)·(GGCCCC) repeats causes instability. Nucleic Acids Res 42, 10473-10487.
Reddy, K., Zamiri, B., Stanley, S.Y., Macgregor, R.B., and Pearson, C.E. (2013a). The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. J Biol Chem 288, 9860-9866.
Reddy, K., Zamiri, B., Stanley, S.Y.R., Macgregor, R.B., Jr., and Pearson, C.E. (2013b). The Disease-associated r(GGGGCC)(n) Repeat from the C9orf72 Gene Forms Tract Length-dependent Uni- and Multimolecular RNA G-quadruplex Structures. Journal of Biological Chemistry 288, 9860-9866.
Renoux, A.J., and Todd, P.K. (2012). Neurodegeneration the RNA way. Prog Neurobiol 97, 173-189.
Renton, A.E., Chiò, A., and Traynor, B.J. (2014). State of play in amyotrophic lateral sclerosis genetics. Nat Neurosci 17, 17-23.
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.
Robberecht, W. (2000). Oxidative stress in amyotrophic lateral sclerosis. J Neurol 247 Suppl 1, I1-6.
Rodrigues, J.P., Rode, M., Gatfield, D., Blencowe, B.J., Carmo-Fonseca, M., and Izaurralde, E. (2001). REF proteins mediate the export of spliced and unspliced mRNAs from the nucleus. Proc Natl Acad Sci U S A 98, 1030-1035.
Roger, B., Moisand, A., Amalric, F., and Bouvet, P. (2003). Nucleolin provides a link between RNA polymerase I transcription and pre-ribosome assembly. Chromosoma 111, 399-407.
Rosen, D.R. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 364, 362.
Rossi, S., Serrano, A., Gerbino, V., Giorgi, A., Di Francesco, L., Nencini, M., Bozzo, F., Schininà, M.E., Bagni, C., Cestra, G.*, et al.* (2015). Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. J Cell Sci 128, 1787-1799.
Rothstein, J.D., Martin, L.J., and Kuncl, R.W. (1992). Decreased glutamate transport by the brain and spinal cord in amyotrophic lateral sclerosis. N Engl J Med 326, 1464-1468.

Rowland, L.P., and Shneider, N.A. (2001). Amyotrophic lateral sclerosis. N Engl J Med 344, 1688-1700.

Rubbi, C.P., and Milner, J. (2003). Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. EMBO J 22, 6068-6077.

Salton, M., Elkon, R., Borodina, T., Davydov, A., Yaspo, M.L., Halperin, E., and Shiloh, Y. (2011). Matrin 3 binds and stabilizes mRNA. PLoS One 6, e23882.

Sanford, J.R., Gray, N.K., Beckmann, K., and Cáceres, J.F. (2004). A novel role for shuttling SR proteins in mRNA translation. Genes Dev 18, 755-768.

Sapp, P.C., Hosler, B.A., McKenna-Yasek, D., Chin, W., Gann, A., Genise, H., Gorenstein, J., Huang, M., Sailer, W., Scheffler, M*., et al.* (2003). Identification of two novel loci for dominantly inherited familial amyotrophic lateral sclerosis. Am J Hum Genet 73, 397-403.

Sarbassov, D.D., Guertin, D.A., Ali, S.M., and Sabatini, D.M. (2005). Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. Science 307, 1098-1101.

Sareen, D., O'Rourke, J.G., Meera, P., Muhammad, A.K., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A*., et al.* (2013). Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. Sci Transl Med 5, 208ra149.

Sarlette, A., Krampfl, K., Grothe, C., Neuhoff, N., Dengler, R., and Petri, S. (2008). Nuclear erythroid 2-related factor 2-antioxidative response element signaling pathway in motor cortex and spinal cord in amyotrophic lateral sclerosis. J Neuropathol Exp Neurol 67, 1055-1062.

Sasaki, S., Horie, Y., and Iwata, M. (2007). Mitochondrial alterations in dorsal root ganglion cells in sporadic amyotrophic lateral sclerosis. Acta Neuropathol 114, 633-639.

Sato, H., Hosoda, N., and Maquat, L.E. (2008). Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. Mol Cell 29, 255-262.

Satterfield, T.F., and Pallanck, L.J. (2006). Ataxin-2 and its Drosophila homolog, ATX2, physically assemble with polyribosomes. Hum Mol Genet 15, 2523-2532.

Schludi, M.H., May, S., Grässer, F.A., Rentzsch, K., Kremmer, E., Küpper, C., Klopstock, T., Arzberger, T., Edbauer, D., Degeneration, G.C.f.F.L., and Alliance, B.B.B. (2015). Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. Acta Neuropathol 130, 537-555.

Schmid, A.C., Byrne, R.D., Vilar, R., and Woscholski, R. (2004). Bisperoxovanadium compounds are potent PTEN inhibitors. FEBS Lett 566, 35-38.

Scrivens, P.J., Alaoui-Jamali, M.A., Giannini, G., Wang, T., Loignon, M., Batist, G., and Sandor, V.A. (2003). Cdc25A-inhibitory properties and antineoplastic activity of bisperoxovanadium analogues. Mol Cancer Ther 2, 1053-1059.

Seargeant, L.E., and Stinson, R.A. (1979). Inhibition of human alkaline phosphatases by vanadate. Biochem J 181, 247-250.

Seibenhener, M.L., Babu, J.R., Geetha, T., Wong, H.C., Krishna, N.R., and Wooten, M.W. (2004). Sequestosome 1/p62 is a polyubiquitin chain binding protein involved in ubiquitin proteasome degradation. Mol Cell Biol 24, 8055-8068.

Seinsoth, S., Uhlmann-Schiffler, H., and Stahl, H. (2003). Bidirectional DNA unwinding by a ternary complex of T antigen, nucleolin and topoisomerase I. EMBO Rep 4, 263-268.

Shaw, P.J. (2005). Molecular and cellular pathways of neurodegeneration in motor neurone disease. J Neurol Neurosurg Psychiatry 76, 1046-1057.

Shaw, P.J., Ince, P.G., Falkous, G., and Mantle, D. (1995). Oxidative damage to protein in sporadic motor neuron disease spinal cord. Annals of neurology 38, 691-695.

Shi, P., Gal, J., Kwinter, D.M., Liu, X., and Zhu, H. (2010). Mitochondrial dysfunction in amyotrophic lateral sclerosis. Biochim Biophys Acta 1802, 45-51.

Shi, Y., Wang, J., Chandarlapaty, S., Cross, J., Thompson, C., Rosen, N., and Jiang, X. (2014). PTEN is a protein tyrosine phosphatase for IRS1. Nat Struct Mol Biol 21, 522-527.

Shibata, H., Huynh, D.P., and Pulst, S.M. (2000). A novel protein with RNA-binding motifs interacts with ataxin-2. Hum Mol Genet 9, 1303-1313.

Shibata, N., Nagai, R., Uchida, K., Horiuchi, S., Yamada, S., Hirano, A., Kawaguchi, M., Yamamoto, T., Sasaki, S., and Kobayashi, M. (2001). Morphological evidence for lipid peroxidation and protein glycoxidation in spinal cords from sporadic amyotrophic lateral sclerosis patients. Brain research 917, 97-104.

Simpson, E.P., Henry, Y.K., Henkel, J.S., Smith, R.G., and Appel, S.H. (2004). Increased lipid peroxidation in sera of ALS patients: a potential biomarker of disease burden. Neurology 62, 1758-1765.

Skibinski, G., Parkinson, N.J., Brown, J.M., Chakrabarti, L., Lloyd, S.L., Hummerich, H., Nielsen, J.E., Hodges, J.R., Spillantini, M.G., Thusgaard, T*., et al.* (2005). Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. Nat Genet 37, 806-808.

Skourti-Stathaki, K., Proudfoot, N.J., and Gromak, N. (2011). Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Molecular cell 42, 794-805.

Smith, B.N., Ticozzi, N., Fallini, C., Gkazi, A.S., Topp, S., Kenna, K.P., Scotter, E.L., Kost, J., Keagle, P., Miller, J.W*., et al.* (2014). Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. Neuron 84, 324-331.

Smith, B.N., Vance, C., Scotter, E.L., Troakes, C., Wong, C.H., Topp, S., Maekawa, S., King, A., Mitchell, J.C., Lund, K*., et al.* (2015). Novel mutations support a role for Profilin 1 in the pathogenesis of ALS. Neurobiol Aging 36, 1602.e1617-1627.

Smith, R.G., Henry, Y.K., Mattson, M.P., and Appel, S.H. (1998). Presence of 4-hydroxynonenal in cerebrospinal fluid of patients with sporadic amyotrophic lateral sclerosis. Annals of neurology 44, 696-699.

Snowden, J.S., Rollinson, S., Thompson, J.C., Harris, J.M., Stopford, C.L., Richardson, A.M., Jones, M., Gerhard, A., Davidson, Y.S., Robinson, A*., et al.* (2012). Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. Brain 135, 693-708.

Sobue, G., Hashizume, Y., Yasuda, T., Mukai, E., Kumagai, T., Mitsuma, T., and Trojanowski, J.Q. (1990). Phosphorylated high molecular weight neurofilament protein in lower motor neurons in amyotrophic lateral sclerosis and other neurodegenerative diseases involving ventral horn cells. Acta Neuropathol 79, 402-408.

Song, M.S., Carracedo, A., Salmena, L., Song, S.J., Egia, A., Malumbres, M., and Pandolfi, P.P. (2011). Nuclear PTEN regulates the APC-CDH1 tumor-suppressive complex in a phosphatase-independent manner. Cell 144, 187-199.

Spreux-Varoquaux, O., Bensimon, G., Lacomblez, L., Salachas, F., Pradat, P.F., Le Forestier, N., Marouan, A., Dib, M., and Meininger, V. (2002). Glutamate levels in cerebrospinal fluid in amyotrophic lateral sclerosis: a reappraisal using a new HPLC method with coulometric detection in a large cohort of patients. J Neurol Sci 193, 73-78.

Sreedharan, J., Blair, I.P., Tripathi, V.B., Hu, X., Vance, C., Rogelj, B., Ackerley, S., Durnall, J.C., Williams, K.L., Buratti, E*., et al.* (2008). TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science 319, 1668-1672.

Stambolic, V., MacPherson, D., Sas, D., Lin, Y., Snow, B., Jang, Y., Benchimol, S., and Mak, T.W. (2001). Regulation of PTEN transcription by p53. Mol Cell 8, 317-325.

Stathopulos, P.B., Rumfeldt, J.A., Karbassi, F., Siddall, C.A., Lepock, J.R., and Meiering, E.M. (2006). Calorimetric analysis of thermodynamic stability and aggregation for apo and holo amyotrophic lateral sclerosis-associated Gly-93 mutants of superoxide dismutase. J Biol Chem 281, 6184-6193.

Stewart, H., Rutherford, N.J., Briemberg, H., Krieger, C., Cashman, N., Fabros, M., Baker, M., Fok, A., DeJesus-Hernandez, M., Eisen, A*., et al.* (2012). Clinical and pathological features of

amyotrophic lateral sclerosis caused by mutation in the C9ORF72 gene on chromosome 9p. Acta Neuropathol 123, 409-417.

Stoica, R., De Vos, K.J., Paillusson, S., Mueller, S., Sancho, R.M., Lau, K.F., Vizcay-Barrena, G., Lin, W.L., Xu, Y.F., Lewis, J.*, et al.* (2014). ER-mitochondria associations are regulated by the VAPB-PTPIP51 interaction and are disrupted by ALS/FTD-associated TDP-43. Nat Commun 5, 3996.

Su, X.W., Broach, J.R., Connor, J.R., Gerhard, G.S., and Simmons, Z. (2014). Genetic heterogeneity of amyotrophic lateral sclerosis: implications for clinical practice and research. Muscle Nerve 49, 786-803.

Subramanian, V., Crabtree, B., and Acharya, K.R. (2008). Human angiogenin is a neuroprotective factor and amyotrophic lateral sclerosis associated angiogenin variants affect neurite extension/pathfinding and survival of motor neurons. Hum Mol Genet 17, 130-149.

Sundaramoorthy, V., Walker, A.K., Yerbury, J., Soo, K.Y., Farg, M.A., Hoang, V., Zeineddine, R., Spencer, D., and Atkin, J.D. (2013). Extracellular wildtype and mutant SOD1 induces ER-Golgi pathology characteristic of amyotrophic lateral sclerosis in neuronal cells. Cell Mol Life Sci 70, 4181-4195.

Suzuki, N., Maroof, A.M., Merkle, F.T., Koszka, K., Intoh, A., Armstrong, I., Moccia, R., Davis-Dusenbery, B.N., and Eggan, K. (2013). The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. Nat Neurosci 16, 1725-1727.

Takagi, M., Absalon, M.J., McLure, K.G., and Kastan, M.B. (2005). Regulation of p53 translation and induction after DNA damage by ribosomal protein L26 and nucleolin. Cell 123, 49-63.

Takahashi, Y., Fukuda, Y., Yoshimura, J., Toyoda, A., Kurppa, K., Moritoyo, H., Belzil, V.V., Dion, P.A., Higasa, K., Doi, K.*, et al.* (2013). ERBB4 mutations that disrupt the neuregulin-ErbB4 pathway cause amyotrophic lateral sclerosis type 19. Am J Hum Genet 93, 900-905.

Tan, C.F., Eguchi, H., Tagawa, A., Onodera, O., Iwasaki, T., Tsujino, A., Nishizawa, M., Kakita, A., and Takahashi, H. (2007). TDP-43 immunoreactivity in neuronal inclusions in familial amyotrophic lateral sclerosis with or without SOD1 gene mutation. Acta Neuropathol 113, 535-542.

Tan, P.H., Beutelspacher, S.C., Xue, S.A., Wang, Y.H., Mitchell, P., McAlister, J.C., Larkin, D.F., McClure, M.O., Stauss, H.J., Ritter, M.A.*, et al.* (2005). Modulation of human dendritic-cell function following transduction with viral vectors: implications for gene therapy. Blood 105, 3824-3832.

Tan, P.H., Xue, S.A., Manunta, M., Beutelspacher, S.C., Fazekasova, H., Alam, A.K., McClure, M.O., and George, A.J. (2006). Effect of vectors on human endothelial cell signal transduction: implications for cardiovascular gene therapy. Arterioscler Thromb Vasc Biol 26, 462-467.

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.

Teyssou, E., Takeda, T., Lebon, V., Boillée, S., Doukouré, B., Bataillon, G., Sazdovitch, V., Cazeneuve, C., Meininger, V., LeGuern, E.*, et al.* (2013). Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology. Acta Neuropathol 125, 511-522.

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.

Thys, R.G., and Wang, Y.H. (2015). DNA replication dynamics of the GGGGCC repeat of C9orf72. J Biol Chem.

Tian, Y., Daoud, A., and Shang, J. (2012). Effects of bpV(pic) and bpV(phen) on H9c2 cardiomyoblasts during both hypoxia/reoxygenation and H2O2-induced injuries. Mol Med Rep 5, 852-858.

Tiloca, C., Ticozzi, N., Pensato, V., Corrado, L., Del Bo, R., Bertolin, C., Fenoglio, C., Gagliardi, S., Calini, D., Lauria, G.*, et al.* (2013). Screening of the PFN1 gene in sporadic amyotrophic lateral sclerosis and in frontotemporal dementia. Neurobiol Aging 34, 1517.e1519-1510.

Tintaru, A.M., Hautbergue, G.M., Hounslow, A.M., Hung, M.L., Lian, L.Y., Craven, C.J., and Wilson, S.A. (2007). Structural and functional analysis of RNA and TAP binding to SF2/ASF. EMBO Rep 8, 756-762.

Tran, H., Almeida, S., Moore, J., Gendron, T.F., Chalasani, U., Lu, Y., Du, X., Nickerson, J.A., Petrucelli, L., Weng, Z., and Gao, F.B. (2015). Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. Neuron 87, 1207-1214.

Tresse, E., Salomons, F.A., Vesa, J., Bott, L.C., Kimonis, V., Yao, T.P., Dantuma, N.P., and Taylor, J.P. (2010). VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD. Autophagy 6, 217-227.

Turner, B.J., and Talbot, K. (2008). Transgenics, toxicity and therapeutics in rodent models of mutant SOD1-mediated familial ALS. Prog Neurobiol 85, 94-134.

Turner, M.R., Bowser, R., Bruijn, L., Dupuis, L., Ludolph, A., McGrath, M., Manfredi, G., Maragakis, N., Miller, R.G., Pullman, S.L.*, et al.* (2013). Mechanisms, models and biomarkers in amyotrophic lateral sclerosis. Amyotroph Lateral Scler Frontotemporal Degener 14 Suppl 1, 19-32.

Uemura, M., Zheng, Q., Koh, C.M., Nelson, W.G., Yegnasubramanian, S., and De Marzo, A.M. (2012). Overexpression of ribosomal RNA in prostate cancer is common but not linked to rDNA promoter hypomethylation. Oncogene 31, 1254-1263.

Uribe, D.J., Guo, K., Shin, Y.J., and Sun, D. (2011). Heterogeneous nuclear ribonucleoprotein K and nucleolin as transcriptional activators of the vascular endothelial growth factor promoter through interaction with secondary DNA structures. Biochemistry 50, 3796-3806.

van Blitterswijk, M., Baker, M.C., Bieniek, K.F., Knopman, D.S., Josephs, K.A., Boeve, B., Caselli, R., Wszolek, Z.K., Petersen, R., Graff-Radford, N.R.*, et al.* (2013a). Profilin-1 mutations are rare in patients with amyotrophic lateral sclerosis and frontotemporal dementia. Amyotroph Lateral Scler Frontotemporal Degener 14, 463-469.

van Blitterswijk, M., DeJesus-Hernandez, M., Niemantsverdriet, E., Murray, M.E., Heckman, M.G., Diehl, N.N., Brown, P.H., Baker, M.C., Finch, N.A., Bauer, P.O.*, et al.* (2013b). Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. Lancet Neurol 12, 978-988.

van Blitterswijk, M., van Es, M.A., Hennekam, E.A., Dooijes, D., van Rheenen, W., Medic, J., Bourque, P.R., Schelhaas, H.J., van der Kooi, A.J., de Visser, M.*, et al.* (2012a). Evidence for an oligogenic basis of amyotrophic lateral sclerosis. Hum Mol Genet 21, 3776-3784.

van Blitterswijk, M., van Es, M.A., Koppers, M., van Rheenen, W., Medic, J., Schelhaas, H.J., van der Kooi, A.J., de Visser, M., Veldink, J.H., and van den Berg, L.H. (2012b). VAPB and C9orf72 mutations in 1 familial amyotrophic lateral sclerosis patient. Neurobiol Aging 33, 2950.e2951-2954.

van Blitterswijk, M., van Vught, P.W., van Es, M.A., Schelhaas, H.J., van der Kooi, A.J., de Visser, M., Veldink, J.H., and van den Berg, L.H. (2012c). Novel optineurin mutations in sporadic amyotrophic lateral sclerosis patients. Neurobiol Aging 33, 1016.e1011-1017.

Van Damme, P., Bogaert, E., Dewil, M., Hersmus, N., Kiraly, D., Scheveneels, W., Bockx, I., Braeken, D., Verpoorten, N., Verhoeven, K.*, et al.* (2007). Astrocytes regulate GluR2 expression in motor neurons and their vulnerability to excitotoxicity. Proc Natl Acad Sci U S A 104, 14825-14830.

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K.J., Nishimura, A.L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P.*, et al.* (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 323, 1208-1211.

VanEtten, R.L., Waymack, P.P., and Rehkop, D.M. (1974). Letter: Transition metal ion inhibition of enzyme-catalyzed phosphate ester displacement reactions. J Am Chem Soc 96, 6782-6785.

Vanselow, B.K., and Keller, B.U. (2000). Calcium dynamics and buffering in oculomotor neurones from mouse that are particularly resistant during amyotrophic lateral sclerosis (ALS)-related motoneurone disease. J Physiol 525 Pt 2, 433-445.

Vassall, K.A., Stathopulos, P.B., Rumfeldt, J.A., Lepock, J.R., and Meiering, E.M. (2006). Equilibrium thermodynamic analysis of amyotrophic lateral sclerosis-associated mutant apo Cu,Zn superoxide dismutases. Biochemistry 45, 7366-7379.

Vazquez, F., Ramaswamy, S., Nakamura, N., and Sellers, W.R. (2000). Phosphorylation of the PTEN tail regulates protein stability and function. Mol Cell Biol 20, 5010-5018.

Venema, J., and Tollervey, D. (1999). Ribosome synthesis in Saccharomyces cerevisiae. Annu Rev Genet 33, 261-311.

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.

Walsh, M.J., Cooper-Knock, J., Dodd, J.E., Stopford, M.J., Mihaylov, S.R., Kirby, J., Shaw, P.J., and Hautbergue, G.M. (2015). Invited review: decoding the pathophysiological mechanisms that underlie RNA dysregulation in neurodegenerative disorders: a review of the current state of the art. Neuropathol Appl Neurobiol 41, 109-134.

Wang, F., Lu, Y., Qi, F., Su, Q., Wang, L., You, C., Che, F., and Yu, J. (2014). Effect of the human SOD1-G93A gene on the Nrf2/ARE signaling pathway in NSC-34 cells. Mol Med Rep 9, 2453-2458.

Wang, H., Wang, M., Reiss, K., Darbinian-Sarkissian, N., Johnson, E.M., Iliakis, G., Amini, S., Khalili, K., and Rappaport, J. (2007). Evidence for the involvement of Puralpha in response to DNA replication stress. Cancer Biol Ther 6, 596-602.

Wang, L., Gutmann, D.H., and Roos, R.P. (2011). Astrocyte loss of mutant SOD1 delays ALS disease onset and progression in G85R transgenic mice. Hum Mol Genet 20, 286-293.

Waring, S.C., Esteban-Santillan, C., Reed, D.M., Craig, U.K., Labarthe, D.R., Petersen, R.C., and Kurland, L.T. (2004). Incidence of amyotrophic lateral sclerosis and of the parkinsonism-dementia complex of Guam, 1950-1989. Neuroepidemiology 23, 192-200.

Watts, G.D., Wymer, J., Kovach, M.J., Mehta, S.G., Mumm, S., Darvish, D., Pestronk, A., Whyte, M.P., and Kimonis, V.E. (2004). Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. Nat Genet 36, 377-381.

Weidberg, H., and Elazar, Z. (2011). TBK1 mediates crosstalk between the innate immune response and autophagy. Sci Signal 4, pe39.

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.

Wheaton, M.W., Salamone, A.R., Mosnik, D.M., McDonald, R.O., Appel, S.H., Schmolck, H.I., Ringholz, G.M., and Schulz, P.E. (2007). Cognitive impairment in familial ALS. Neurology 69, 1411-1417.

White, M.K., Johnson, E.M., and Khalili, K. (2009). Multiple roles for Puralpha in cellular and viral regulation. Cell Cycle 8, 1-7.

Wild, P., Farhan, H., McEwan, D.G., Wagner, S., Rogov, V.V., Brady, N.R., Richter, B., Korac, J., Waidmann, O., Choudhary, C*., et al.* (2011). Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. Science 333, 228-233.

Williams, K.L., Warraich, S.T., Yang, S., Solski, J.A., Fernando, R., Rouleau, G.A., Nicholson, G.A., and Blair, I.P. (2012). UBQLN2/ubiquilin 2 mutation and pathology in familial amyotrophic lateral sclerosis. Neurobiol Aging 33, 2527.e2523-2510.

Winton, M.J., Igaz, L.M., Wong, M.M., Kwong, L.K., Trojanowski, J.Q., and Lee, V.M. (2008). Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation. J Biol Chem 283, 13302-13309.

Woerner, A.C., Frottin, F., Hornburg, D., Feng, L.R., Meissner, F., Patra, M., Tatzelt, J., Mann, M., Winklhofer, K.F., Hartl, F.U., and Hipp, M.S. (2016). Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. Science 351, 173-176.

Wood, J.D., Beaujeux, T.P., and Shaw, P.J. (2003). Protein aggregation in motor neurone disorders. Neuropathol Appl Neurobiol 29, 529-545.

Worms, P.M. (2001). The epidemiology of motor neuron diseases: a review of recent studies. J Neurol Sci 191, 3-9.

Wu, C.H., Fallini, C., Ticozzi, N., Keagle, P.J., Sapp, P.C., Piotrowska, K., Lowe, P., Koppers, M., McKenna-Yasek, D., Baron, D.M*., et al.* (2012). Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. Nature 488, 499-503.

Wu, H., Sun, S., Tu, K., Gao, Y., Xie, B., Krainer, A.R., and Zhu, J. (2010). A splicing-independent function of SF2/ASF in microRNA processing. Mol Cell 38, 67-77.

Xi, Z., Zinman, L., Moreno, D., Schymick, J., Liang, Y., Sato, C., Zheng, Y., Ghani, M., Dib, S., Keith, J*., et al.* (2013). Hypermethylation of the CpG Island Near the G(4)C(2) Repeat in ALS with a C9orf72 Expansion. Am J Hum Genet 92, 981-989.

Xiao, S., MacNair, L., McGoldrick, P., McKeever, P.M., McLean, J.R., Zhang, M., Keith, J., Zinman, L., Rogaeva, E., and Robertson, J. (2015). Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. Ann Neurol 78, 568-583.

Xu, Z., Poidevin, M., Li, X., Li, Y., Shu, L., Nelson, D.L., Li, H., Hales, C.M., Gearing, M., Wingo, T.S., and Jin, P. (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. Proc Natl Acad Sci U S A 110, 7778-7783.

Yamanaka, K., Boillee, S., Roberts, E.A., Garcia, M.L., McAlonis-Downes, M., Mikse, O.R., Cleveland, D.W., and Goldstein, L.S. (2008a). Mutant SOD1 in cell types other than motor neurons and oligodendrocytes accelerates onset of disease in ALS mice. Proc Natl Acad Sci U S A 105, 7594-7599.

Yamanaka, K., Chun, S.J., Boillee, S., Fujimori-Tonou, N., Yamashita, H., Gutmann, D.H., Takahashi, R., Misawa, H., and Cleveland, D.W. (2008b). Astrocytes as determinants of disease progression in inherited amyotrophic lateral sclerosis. Nat Neurosci 11, 251-253.

Yang, C., Kim, M.S., Chakravarty, D., Indig, F.E., and Carrier, F. (2009). Nucleolin Binds to the Proliferating Cell Nuclear Antigen and Inhibits Nucleotide Excision Repair. Mol Cell Pharmacol 1, 130-137.

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.

Yang, T.H., Tsai, W.H., Lee, Y.M., Lei, H.Y., Lai, M.Y., Chen, D.S., Yeh, N.H., and Lee, S.C. (1994). Purification and characterization of nucleolin and its identification as a transcription repressor. Mol Cell Biol 14, 6068-6074.

Yang, Y., Hentati, A., Deng, H.X., Dabbagh, O., Sasaki, T., Hirano, M., Hung, W.Y., Ouahchi, K., Yan, J., Azim, A.C*., et al.* (2001). The gene encoding alsin, a protein with three guanine-

nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. Nat Genet 29, 160-165.

Yang, Z.Z., Tschopp, O., Hemmings-Mieszczak, M., Feng, J., Brodbeck, D., Perentes, E., and Hemmings, B.A. (2003). Protein kinase B alpha/Akt1 regulates placental development and fetal growth. J Biol Chem 278, 32124-32131.

Yokoseki, A., Shiga, A., Tan, C.F., Tagawa, A., Kaneko, H., Koyama, A., Eguchi, H., Tsujino, A., Ikeuchi, T., Kakita, A.*, et al.* (2008). TDP-43 mutation in familial amyotrophic lateral sclerosis. Ann Neurol 63, 538-542.

Yoshida, S., Uebayashi, Y., Kihira, T., Kohmoto, J., Wakayama, I., Taguchi, S., and Yase, Y. (1998). Epidemiology of motor neuron disease in the Kii Peninsula of Japan, 1989-1993: active or disappearing focus? J Neurol Sci 155, 146-155.

Zamiri, B., Mirceta, M., Bomsztyk, K., Macgregor, R.B., and Pearson, C.E. (2015). Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)8•(GGCCCC)8 repeat: effect of CpG methylation. Nucleic Acids Res.

Zhan, L., Xie, Q., and Tibbetts, R.S. (2015). Opposing roles of p38 and JNK in a Drosophila model of TDP-43 proteinopathy reveal oxidative stress and innate immunity as pathogenic components of neurodegeneration. Hum Mol Genet 24, 757-772.

Zhang, D., Iyer, L.M., He, F., and Aravind, L. (2012). Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. Front Genet 3, 283.

Zhang, K., Donnelly, C.J., Haeusler, A.R., Grima, J.C., Machamer, J.B., Steinwald, P., Daley, E.L., Miller, S.J., Cunningham, K.M., Vidensky, S.*, et al.* (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature 525, 56-61.

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.

Zhang, Z., and Krainer, A.R. (2004). Involvement of SR proteins in mRNA surveillance. Mol Cell 16, 597-607.

Zhao, J., Qu, Y., Wu, J., Cao, M., Ferriero, D.M., Zhang, L., and Mu, D. (2013). PTEN inhibition prevents rat cortical neuron injury after hypoxia-ischemia. Neuroscience 238, 242-251.

Zhou, Z., Luo, M.J., Straesser, K., Katahira, J., Hurt, E., and Reed, R. (2000). The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans. Nature 407, 401-405.

Zhu, L., Xu, M., Yang, M., Yang, Y., Li, Y., Deng, J., Ruan, L., Liu, J., Du, S., Liu, X.*, et al.* (2014). An ALS-mutant TDP-43 neurotoxic peptide adopts an anti-parallel β-structure and induces TDP-43 redistribution. Hum Mol Genet 23, 6863-6877.

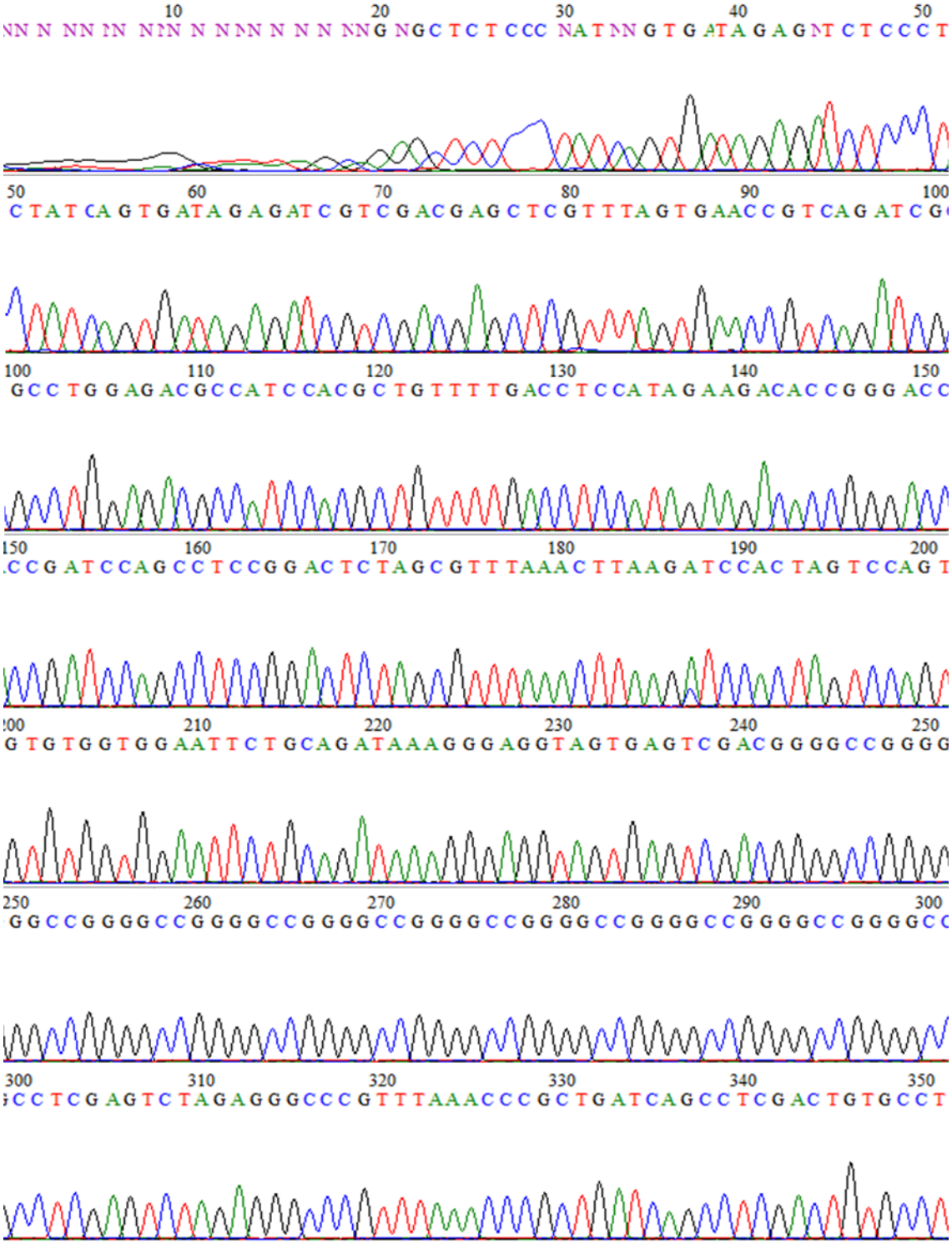
Zinszner, H., Sok, J., Immanuel, D., Yin, Y., and Ron, D. (1997). TLS (FUS) binds RNA in vivo and engages in nucleo-cytoplasmic shuttling. J Cell Sci 110 (Pt 15), 1741-1750.

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 U S A 110, E4968-4977.

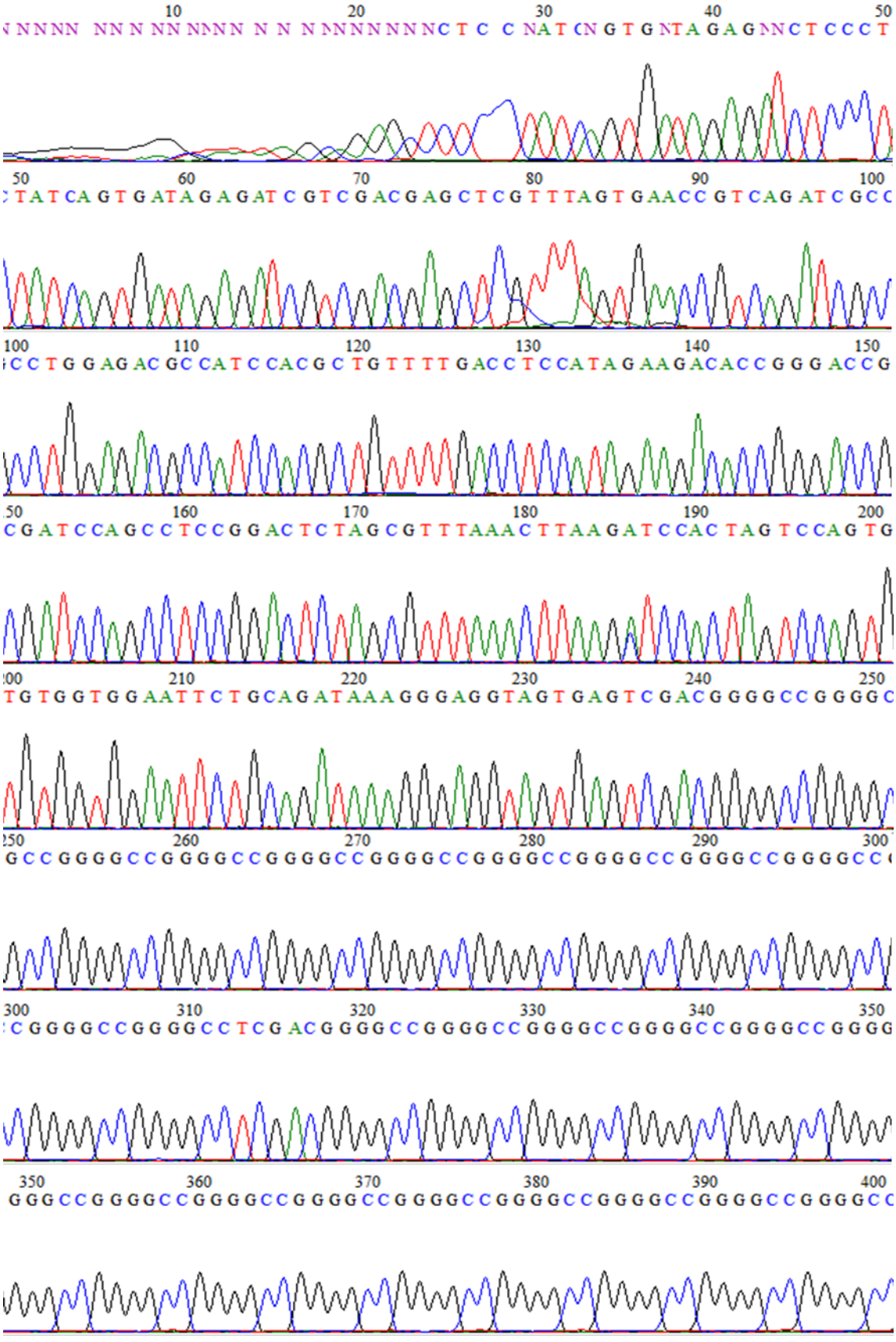
Züchner, S., Mersiyanova, I.V., Muglia, M., Bissar-Tadmouri, N., Rochelle, J., Dadali, E.L., Zappia, M., Nelis, E., Patitucci, A., Senderek, J.*, et al.* (2004). Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. Nat Genet 36, 449-451.

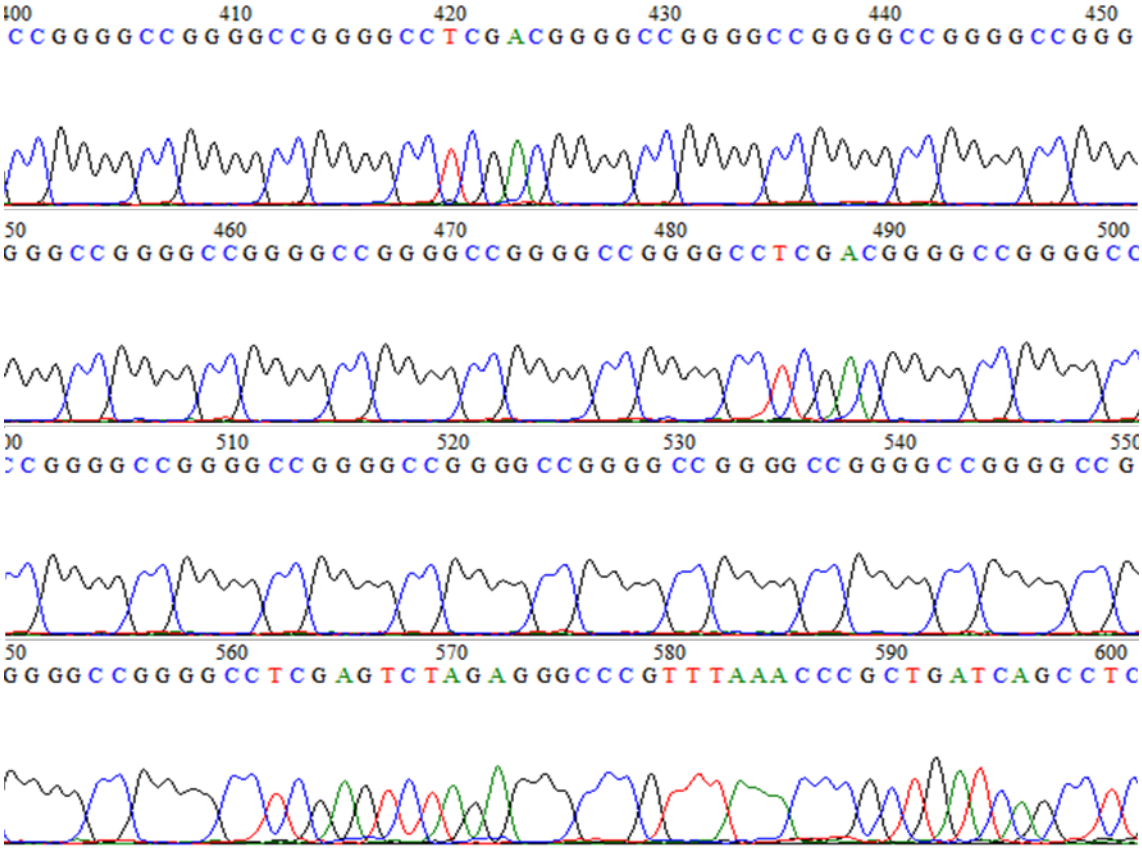
Appendices

Appendix 1 Sequence chromatogram of pcDNA5/FRT/TO-(G4C2)10 sequenced using CMV-Forward primer.

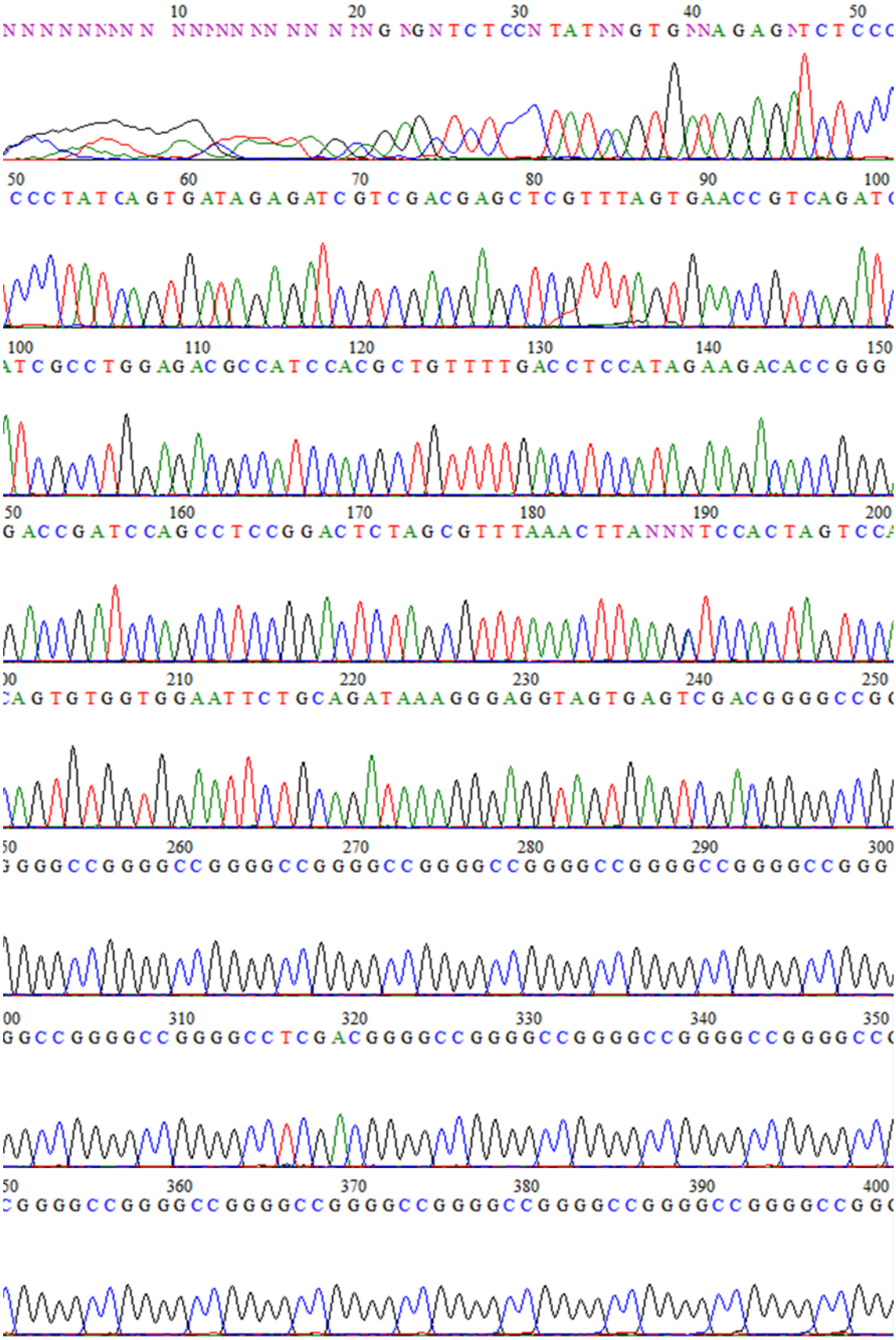


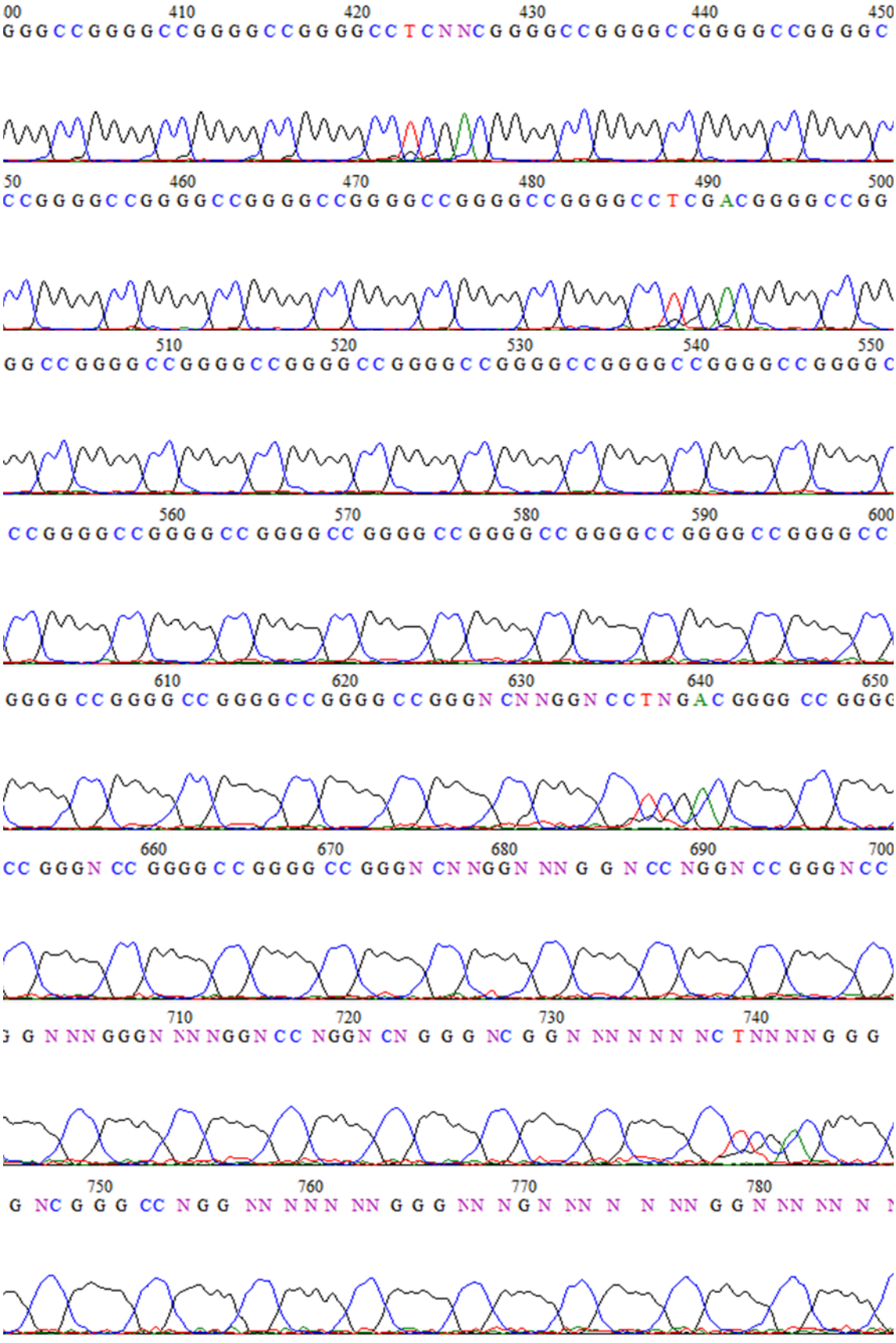
Appendix 2 Sequence chromatogram of pcDNA5/FRT/TO-(G4C2)51 sequenced using CMV-Forward primer.

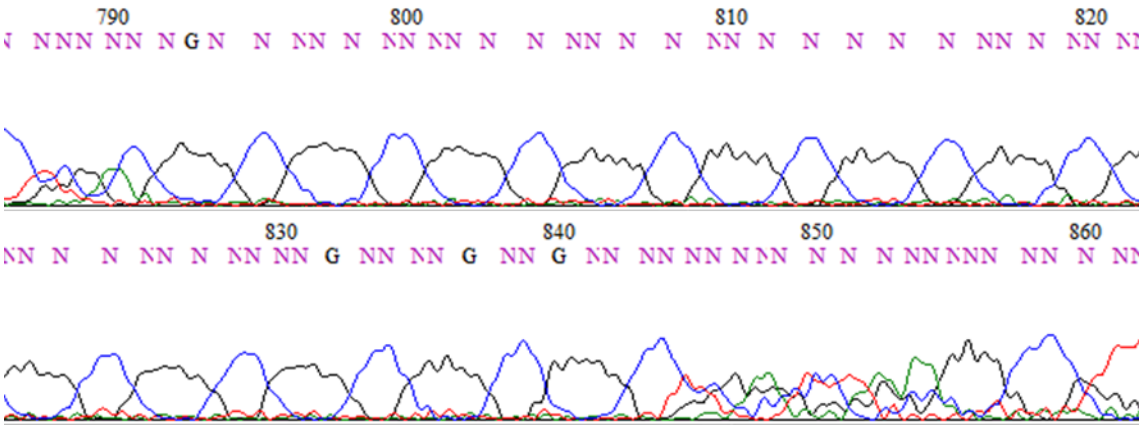



400 410 420 430 440 450
CCGGGGCCGGGGCCGGGGCCTCGACGGGGCCGGGGCCGGGGCCGGGGCCGGG
50 460 470 480 490 500
GGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCTCGACGGGGCCGGGGCC
00 510 520 530 540 550
CCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCGGGGCCG
50 560 570 580 590 600
GGGGCCGGGGCCTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTC

Appendix 3 Sequence chromatogram of pcDNA5/FRT/TO-(G4C2)51 sequenced using CMV-Forward primer.




00 410 420 430 440 450
G G G C C G G G G C C G G G G C C G G G G C C T C N N C G G G G C C G G G G C C G G G G C C G G G G C
50 460 470 480 490 500
C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C T C G A C G G G G C C G G
510 520 530 540 550
G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C
560 570 580 590 600
C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C G G G G C C
610 620 630 640 650
G G G G C C G G G G C C G G G G C C G G G G C C G G G N C N N G G N C C T N G A C G G G G C C G G G G
660 670 680 690 700
C C G G G N C C G G G G C C G G G G C C G G G N C N N G G N N N G G N C C N G G N C C G G G N C C
710 720 730 740
G G N N N G G G N N N N G G N C C N G G N C N G G G N C G G N N N N N N C T N N N N G G G
750 760 770 780
G N C G G G C C N G G N N N N N N N G G G N N N G N N N N N N N G G N N N N N


790
800
810
820
830
840
850
860

Appendix 4 Arrays were hybridised, washed and stained in three separate batches on different days.

Batch 1	Batch 2	Batch 3
0_0_A	0_0_C	0_0_B
0_5_A	0_1_A	0_1_C
0_5_B	0_1_B	0_5_C
10_0_A	10_0_C	10_5_A
10_0_B	51_0_A	51_0_B
10_5_B	51_5_B	51_0_C
10_5_C	51_5_C	102_0_B
51_5_A	102_0_C	102_1_A
102_0_A	102_5_B	102_1_C
102_1_B	102_5_C	102_5_A

Appendix 5 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, Gene symbol, Fold change, and P-value are included for a comparison between C9ORF72-ALS patients and control LCM MNs.

Transcript ID	Gene Symbol	Fold Change	P-value
228006_at	PTEN	11.2546	1.35E-05
201125_s_at	ITGB5	5.5843	0.00119
212777_at	SOS1	3.35696	0.01979
209341_s_at	IKBKB	3.16884	0.11631
202686_s_at	AXL	3.10927	0.01269
201124_at	ITGB5	2.86889	0.02358
213093_at	PRKCA	2.80256	0.00332
226731_at	ITGA1	2.73413	0.0494
203685_at	BCL2	2.70459	0.06485
202723_s_at	FOXO1	2.60666	0.0276
203809_s_at	AKT2	2.55388	0.07866
210482_x_at	MAP2K5	2.43459	0.14357
1555612_s_at	G6PC	2.42838	0.06008
1552610_a_at	JAK1	2.33472	0.18949
217399_s_at	FOXO3	2.2105	0.16448
202887_s_at	DDIT4	2.16861	0.00045
223196_s_at	SESN2	2.10237	0.0907
208536_s_at	BCL2L11	2.0825	0.3364
203379_at	RPS6KA1	2.08204	0.14381
226068_at	SYK	2.03705	0.07781
202284_s_at	CDKN1A	2.01865	0.16422
225691_at	CDK12	2.01056	0.079
236664_at	AKT2	2.00531	0.278
201834_at	PRKAB1	2.00507	0.32642
202830_s_at	SLC37A4	2.00107	0.14823
212590_at	RRAS2	1.87414	0.15224
232068_s_at	TLR4	1.86715	0.20452
222999_s_at	CCNL2	1.82949	0.04503
201739_at	SGK1	1.82899	0.01689
202340_x_at	NR4A1	1.82347	0.34816
202530_at	MAPK14	1.76965	0.42941
209666_s_at	CHUK	1.73936	0.28304
226441_at	MAP3K2	1.71589	0.09514
204054_at	PTEN	1.69351	0.3928
202431_s_at	MYC	1.67664	0.40172
202847_at	PCK2	1.6713	0.37293
221060_s_at	TLR4	1.6664	0.50849

236459_at	PRKCE	1.62936	0.24505
224889_at	FOXO3	1.62723	0.51267
209342_s_at	IKBKB	1.61255	0.38191
239201_at	CDK15	1.5972	0.22748
1558143_a_at	BCL2L11	1.57667	0.05218
204297_at	PIK3C3	1.57353	0.24559
202853_s_at	RYK	1.55959	0.0062
210449_x_at	MAPK14	1.55875	0.50136
212589_at	RRAS2	1.5542	0.33377
212628_at	PKN2	1.53559	0.24039
208712_at	CCND1	1.51291	0.32986
209364_at	BAD	1.49379	0.43723
206398_s_at	CD19	1.48899	0.4906
202449_s_at	RXRA	1.44284	0.36684
202426_s_at	RXRA	1.44084	0.55632
241453_at	PTK2	1.43103	0.41771
207540_s_at	SYK	1.42509	0.46612
201234_at	ILK	1.42441	0.50657
241387_at	PTK2	1.41725	0.44044
225066_at	PPP2R2D	1.4018	0.26905
241722_x_at	MCL1	1.38925	0.55755
212332_at	RBL2	1.35227	0.25134
208824_x_at	CDK16	1.346	0.05766
208823_s_at	CDK16	1.33674	0.33747
200797_s_at	MCL1	1.33354	0.16695
227073_at	MAP3K2	1.31735	0.50772
225690_at	CDK12	1.30615	0.51302
206854_s_at	MAP3K7	1.28559	0.56793
226310_at	RICTOR	1.28358	0.42508
211333_s_at	FASLG	1.28012	0.61041
206952_at	G6PC	1.27975	0.52706
206853_s_at	MAP3K7	1.26817	0.33893
208820_at	PTK2	1.25554	0.33508
202161_at	PKN1	1.23813	0.67904
209184_s_at	IRS2	1.23611	0.24619
202210_x_at	GSK3A	1.23529	0.59625
219226_at	CDK12	1.2304	0.71416
211087_x_at	MAPK14	1.2267	0.61045
211561_x_at	MAPK14	1.21891	0.60994
202724_s_at	FOXO1	1.21878	0.7646
212312_at	BCL2L1	1.20808	0.70595
215195_at	PRKCA	1.20301	0.71054
201984_s_at	EGFR	1.20112	0.42889

225606_at	BCL2L11	1.19905	0.70546
204906_at	RPS6KA2	1.19762	0.33186
229705_at	PIK3C3	1.18909	0.0324
205271_s_at	CDK20	1.187	0.8061
207163_s_at	AKT1	1.18606	0.42615
1569272_at	PIK3C3	1.17994	0.67476
235254_at	MAP3K2	1.16993	0.75392
223195_s_at	SESN2	1.15593	0.71007
214265_at	ITGA8	1.13662	0.55785
204924_at	TLR2	1.1327	0.77274
228177_at	CREBBP	1.12985	0.82862
204131_s_at	FOXO3	1.12735	0.51347
1556655_s_at	CDK12	1.12495	0.86831
231017_at	STK11	1.12286	0.60075
228248_at	RICTOR	1.11708	0.75622
202160_at	CREBBP	1.11546	0.39302
238733_at	MDM2	1.10799	0.78926
226979_at	MAP3K2	1.10251	0.57574
203984_s_at	CASP9	1.09459	0.78315
205386_s_at	MDM2	1.09264	0.87372
208641_s_at	RAC1	1.08706	0.75898
240964_at	PTEN	1.08698	0.87875
226156_at	AKT2	1.08531	0.86743
225471_s_at	AKT2	1.08322	0.86043
203836_s_at	MAP3K5	1.07841	0.84902
237451_x_at	CASP9	1.07679	0.71896
220587_s_at	MLST8	1.07642	0.82245
202685_s_at	AXL	1.0668	0.90898
211372_s_at	IL1R2	1.06204	0.88368
227627_at	SGK3	1.04972	0.87686
225160_x_at	MDM2	1.04866	0.80214
244616_x_at	MDM2	1.04365	0.90974
227426_at	SOS1	1.04007	0.87069
211832_s_at	MDM2	1.03989	0.91227
203837_at	MAP3K5	1.03634	0.79284
217492_s_at	PTEN	1.03561	0.9246
1557970_s_at	RPS6KA2	1.03377	0.9432
212629_s_at	PKN2	1.03248	0.93566
212719_at	PHLPP1	1.02015	0.88671
209185_s_at	IRS2	1.00973	0.97333
226299_at	PKN3	1.00887	0.98546
202670_at	MAP2K1	1.00733	0.98816
229711_s_at	MDM2	1.00646	0.96837

225694_at	CDK12	1.00449	0.99428
632_at	GSK3A	1.004	0.9922
224891_at	FOXO3	1.00182	0.99447
208640_at	RAC1	1.00142	0.996
207239_s_at	CDK16	1.00006	0.99989
213012_at	NEDD4	-1.0189	0.95973
223049_at	GRB2	-1.0198	0.90722
204756_at	MAP2K5	-1.0261	0.93618
215394_at	PIK3C3	-1.0303	0.95529
224999_at	EGFR	-1.0322	0.93724
205798_at	IL7R	-1.0369	0.93805
212331_at	RBL2	-1.0382	0.89007
232876_at	MAPK14	-1.0426	0.93159
201983_s_at	EGFR	-1.043	0.9245
1555804_a_at	MAP3K19	-1.052	0.89459
205498_at	GHR	-1.053	0.89839
200980_s_at	PDHA1	-1.0565	0.80341
226218_at	IL7R	-1.0576	0.90887
211808_s_at	CREBBP	-1.0638	0.89116
202288_at	MTOR	-1.0738	0.89435
200798_x_at	MCL1	-1.0766	0.77284
209112_at	CDKN1B	-1.0849	0.59406
214328_s_at	HSP90AA1	-1.0922	0.75804
208743_s_at	YWHAB	-1.0965	0.85556
201389_at	ITGA5	-1.098	0.85649
208456_s_at	RRAS2	-1.1005	0.91845
212912_at	RPS6KA2	-1.1034	0.70576
204247_s_at	CDK5	-1.1084	0.78323
217718_s_at	YWHAB	-1.1126	0.69372
210655_s_at	FOXO3	-1.1146	0.78924
204369_at	PIK3CA	-1.117	0.84573
201452_at	RHEB	-1.1221	0.7603
226046_at	MAPK8	-1.1322	0.83582
239300_at	PIK3C3	-1.1323	0.36348
210775_x_at	CASP9	-1.134	0.84706
225363_at	PTEN	-1.1397	0.55908
218852_at	PPP2R3C	-1.1406	0.81561
44654_at	G6PC3	-1.1408	0.50127
210969_at	PKN2	-1.1422	0.69294
209390_at	TSC1	-1.145	0.55561
210211_s_at	HSP90AA1	-1.1507	0.74048
1567458_s_at	RAC1	-1.1544	0.43336
202647_s_at	NRAS	-1.1576	0.6594

1560074_at	PRKCA	-1.1632	0.59213
243492_at	THEM4	-1.1728	0.76346
214172_x_at	RYK	-1.1758	0.31314
201437_s_at	EIF4E	-1.1808	0.43703
1552611_a_at	JAK1	-1.1818	0.4436
224985_at	NRAS	-1.182	0.50934
229664_at	MAPK8	-1.1907	0.38632
215037_s_at	BCL2L1	-1.1949	0.60974
217542_at	MDM2	-1.2027	0.61732
204798_at	MYB	-1.2095	0.72784
211814_s_at	CCNE2	-1.2099	0.77485
211550_at	EGFR	-1.2172	0.62589
1555864_s_at	PDHA1	-1.2181	0.59782
1552559_a_at	CDK15	-1.2221	0.64683
1553118_at	THEM4	-1.2253	0.41306
226312_at	RICTOR	-1.2291	0.45219
211938_at	EIF4B	-1.2444	0.29075
211937_at	EIF4B	-1.2454	0.35644
204053_x_at	PTEN	-1.2478	0.56041
201835_s_at	PRKAB1	-1.2561	0.35571
217557_s_at	MDM2	-1.2589	0.63174
200979_at	PDHA1	-1.2615	0.28516
235980_at	PIK3CA	-1.2619	0.26671
221759_at	G6PC3	-1.2629	0.51812
214660_at	ITGA1	-1.2897	0.67717
211536_x_at	MAP3K7	-1.3006	0.63566
212688_at	PIK3CB	-1.3077	0.21645
209269_s_at	SYK	-1.3122	0.67354
215152_at	MYB	-1.3244	0.58063
217717_s_at	YWHAB	-1.3264	0.37399
217373_x_at	MDM2	-1.3282	0.4774
210865_at	FASLG	-1.3454	0.37793
1861_at	BAD	-1.3604	0.55135
201453_x_at	RHEB	-1.3606	0.33799
232274_at	CCNL2	-1.3634	0.32902
201648_at	JAK1	-1.3726	0.24825
211027_s_at	IKBKB	-1.3764	0.51877
217289_s_at	SLC37A4	-1.3774	0.48497
200796_s_at	MCL1	-1.3778	0.59702
211711_s_at	PTEN	-1.3861	0.07704
207005_s_at	BCL2	-1.3874	0.48661
221772_s_at	PPP2R2D	-1.3948	0.37502
205403_at	IL1R2	-1.3959	0.49143

220357_s_at	SGK2	-1.4124	0.60431
204531_s_at	BRCA1	-1.424	0.43404
1552734_at	RICTOR	-1.4243	0.42119
216976_s_at	RYK	-1.426	0.53413
230573_at	SGK2	-1.4367	0.34697
209953_s_at	CDC37	-1.4471	0.28938
206248_at	PRKCE	-1.4513	0.26208
237891_at	MDM2	-1.477	0.46615
221695_s_at	MAP3K2	-1.4853	0.60004
212780_at	SOS1	-1.4936	0.02532
1565483_at	EGFR	-1.5092	0.42247
227633_at	RHEB	-1.5105	0.40547
242674_at	EIF4E	-1.5226	0.17335
231228_at	BCL2L1	-1.5388	0.33907
206665_s_at	BCL2L1	-1.5422	0.43553
213404_s_at	RHEB	-1.5462	0.14498
217620_s_at	PIK3CB	-1.5471	0.33664
231854_at	PIK3CA	-1.5474	0.46198
226101_at	PRKCE	-1.5479	0.16316
214621_at	GYS2	-1.5577	0.37259
215075_s_at	GRB2	-1.5756	0.1419
201436_at	EIF4E	-1.5766	0.13533
208711_s_at	CCND1	-1.5776	0.45561
211370_s_at	MAP2K5	-1.5799	0.48471
41657_at	STK11	-1.5864	0.34306
225715_at	RPTOR	-1.5905	0.2525
222343_at	BCL2L11	-1.5914	0.36242
232086_at	PIK3C3	-1.5963	0.38617
221427_s_at	CCNL2	-1.6199	0.04162
224341_x_at	TLR4	-1.6213	0.36125
220038_at	SGK3	-1.6288	0.39196
211851_x_at	BRCA1	-1.6346	0.4612
1560359_at	ITGA1	-1.6459	0.49152
1553096_s_at	BCL2L11	-1.6518	0.05439
221180_at	MAP3K19	-1.689	0.28863
1553088_a_at	BCL2L11	-1.6899	0.10844
211968_s_at	HSP90AA1	-1.6902	0.25214
239188_at	PPP2R3C	-1.6926	0.34661
229253_at	THEM4	-1.702	0.07348
203684_s_at	BCL2	-1.7119	0.40501
242071_x_at	ITGA8	-1.7283	0.2544
206341_at	IL2RA	-1.7288	0.33477
211537_x_at	MAP3K7	-1.7619	0.39267

215735_s_at	TSC2	-1.7706	0.11829
211969_at	HSP90AA1	-1.7979	0.28924
226048_at	MAPK8	-1.8355	0.09455
225697_at	CDK12	-1.8525	0.06182
1565484_x_at	EGFR	-1.8622	0.36031
1556654_at	CDK12	-1.8706	0.16327
207821_s_at	PTK2	-1.8783	0.33951
210226_at	NR4A1	-1.8816	0.25831
201020_at	YWHAH	-1.9019	0.08157
209239_at	NFKB1	-1.996	0.11918
235011_at	MAP3K2	-2.0338	0.26344
211607_x_at	EGFR	-2.0762	0.11823
237718_at	EIF4E	-2.0795	0.16652
206923_at	PRKCA	-2.0808	0.0805
1552798_a_at	TLR4	-2.1097	0.17767
239092_at	ITGA8	-2.1243	0.1596
240437_at	CASP9	-2.1861	0.10184
1560689_s_at	AKT2	-2.2126	0.16275
210984_x_at	EGFR	-2.2755	0.19676
205034_at	CCNE2	-2.367	0.0156
1554826_at	CDK15	-2.3986	0.23223
211453_s_at	AKT2	-2.411	0.03222
201435_s_at	EIF4E	-2.4361	0.08987
211269_s_at	IL2RA	-2.4806	0.17388
1555780_a_at	RHEB	-2.5578	0.23684
211297_s_at	CDK7	-2.6093	0.03151
204132_s_at	FOXO3	-2.7727	0.02473
210671_x_at	MAPK8	-2.7942	0.043
235666_at	ITGA8	-3.2363	0.0632
210477_x_at	MAPK8	-3.4604	0.02917
204292_x_at	STK11	-4.5009	0.00704