

**Modelling C9orf72-linked frontotemporal
dementia and amyotrophic lateral sclerosis**

2015

Sarah Ryan

School of Medicine

Institute of Brain, Behaviour and Mental Health

**A thesis submitted to the University of Manchester for the degree of Doctor
of Philosophy (PhD) in the Faculty of Medical and Human Sciences.**

Supervisors:

Prof Stuart Pickering-Brown and Prof David Sattelle

Table of Contents

List of Figures	6
List of Tables	7
List of abbreviations.....	8
Abstract.....	10
Declaration.....	11
Copyright statement.....	11
Acknowledgements.....	12
Chapter 1 – Introduction	13
1.1 Overview of frontotemporal lobar degeneration and amyotrophic lateral sclerosis	13
1.2 Clinical aspects of FTLN/ALS.....	16
1.2.1 bvFTD	16
1.2.2 Primary progressive aphasia	18
1.2.3 Extrapyrarnidal symptoms in FTLN	18
1.2.4 ALS.....	19
1.3 Genetics of FTLN/ALS	20
1.3.1 <i>SOD1</i>	20
1.3.2 <i>MAPT</i>	21
1.3.3 <i>GRN</i>	22
1.3.4 <i>TARDBP</i> and <i>FUS</i>	25
1.3.5 Rare mutations in FTLN/ALS.....	26
1.4 Pathological features of FTLN/ALS.....	28
1.4.1 <i>SOD1</i>	28
1.4.2 Tau	29
1.4.3 TDP-43 and <i>FUS</i>	31
1.4.4 Clinicopathological and genetic correlations in FTLN/ALS	34
1.5 The C9orf72 expansion	36
1.5.1 Discovery of the C9orf72 expansion.....	36
1.5.2 Epidemiology of C9FTLN/ALS.....	39
1.5.3 Clinical and Pathological Features of C9FTLN/ALS.....	41

1.5.4 The C9ORF72 protein.....	48
1.6 <i>C. elegans</i> as a model organism in neurodegenerative disease	50
1.6.1 Advantages of <i>C. elegans</i> as a model organism.....	50
1.6.2 Previous uses of <i>C. elegans</i> in neurodegeneration research	51
1.6.3 The <i>C. elegans</i> C9orf72 orthologue	53
1.7 Research aims.....	54
Chapter 2: Materials and methods	55
2.1 <i>C. elegans</i> strains and maintenance	55
2.2 Cell biology methods	55
2.2.1 Maintenance of cell culture	55
2.2.2 Transient transfection	56
2.2.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)	56
2.2.3 Immunoblotting	56
2.2.4 Immunofluorescence.....	57
2.3 Molecular biology methods.....	59
2.3.1 Agarose gel electrophoresis	59
2.3.2 Cloning	59
2.3.3 Transformation of plasmids into competent <i>E. coli</i>	59
2.3.4 Sequencing	60
2.3.5 Site-directed mutagenesis (SDM).....	60
2.3.6 RNA extraction and quantitative real-time PCR (qRT-PCR)	61
Chapter 3: Does loss of <i>C9ORF72</i> function cause disease?	63
3.1 Introduction	63
3.2 Materials and methods	65
3.2.1 – <i>C. elegans</i> strains and outcrossing	65
3.2.2 Sequencing of F18A1.6 mutations	65
3.2.3 RNA extraction and quantitative real-time PCR (qRT-PCR)	66
3.2.4 Quantification of swimming behaviour in <i>C. elegans</i>	66
3.2.5 Egg-laying assays.....	67
3.2.6 Life-span assays	67

3.3 Results	68
3.3.1 <i>F18A1.6</i> is knocked down in LM1006 worms	68
3.3.2 Loss of <i>F18A1.6</i> does not affect motility or life-span	68
3.4 Discussion.....	70
3.4.1 Knockout or knockdown of C9orf72 in other animal models	70
3.4.2 Genetic evidence for loss of C9orf72 function in neurodegeneration	72
3.4.3 The pathological features of C9FTLD/ALS support toxic gain of function	Error!
Bookmark not defined.	
3.4.4 Animal models expressing G ₄ C ₂ repeats	74
3.4.5 Summary.....	77
Chapter 4: Generation and characterisation of cellular models of dipeptide pathology in C9FTLD/ALS	79
4.1 Introduction	79
4.2 Methods	80
4.2.1 Generation of short DPR constructs using alternative sequences.....	80
4.2.2 Cloning strategy to increase DPR repeat-length	80
4.2.3 Generation of DPR constructs with no GFP-tag.....	84
4.2.4 Transient transfection of DPR constructs in cell culture.....	85
4.2.5 qRT-PCR.....	86
4.2.6 SDS-PAGE and Immunoblotting.....	86
4.2.7 Immunofluorescence.....	87
4.3 Results	88
4.3.1 Generation of DPR constructs using alternative codon sequences.....	88
4.3.2 Optimisation of DPR expression.....	89
4.3.3 DPR protein expression <i>in vitro</i>	92
4.3.4 Subcellular localisation of DPRs <i>in vitro</i>	93
4.4 Discussion.....	98
4.4.1 Generation of cellular models of DPR pathology	98
4.4.2 Differential distribution of DPRs <i>in vitro</i>	100
4.4.3 Summary.....	104
Chapter 5: Impact of DPRs on cellular function	105

5.1 Introduction	105
5.2 Materials and Methods	106
5.2.1 Immunofluorescence.....	106
5.2.2 Generation of “RNA-only” poly-GR constructs	108
5.3 Results	109
5.3.1 Effects of DPRs on the FTLD/ALS-related proteins, TDP-43, FUS and EWS....	109
5.3.2 Impact of alanine-rich DPRs on cytoplasmic proteins.....	114
5.3.4 Arginine-rich DPRs translocate to the nucleolus, causing nuclear stress.....	116
5.4 Discussion.....	124
5.4.1 Existing models of DPR pathology	124
5.4.2 Impact of cytoplasmic alanine-rich DPRs on cellular function.....	125
5.4.3 Arginine-rich DPRs translocate to the nucleolus, causing nuclear stress.....	127
5.4.4 Summary.....	134
Chapter 6: General discussion	135
6.1 Summary of main findings	135
6.2 Evidence for RNA vs. DPR toxicity in C9FTLD/ALS	136
6.3 Proposed mechanisms of neurodegeneration in C9FTLD/ALS.....	140
6.4 Convergent mechanisms of neurotoxicity in FTLD/ALS	144
6.4.1 Protein degradation and ER stress.....	144
6.4.2 Nucleocytoplasmic transport.....	145
6.4.3 Aberrant RNA processing.....	146
6.5 Future work	149
6.5.1 Improvement DPR expression levels in our cellular models	149
6.5.2 Quantification of DPR toxicity	150
6.5.3 Time-course of DPR pathology.....	150
6.5.4 Co-expression of DPRs	151
6.5.5 Mechanisms of DPR toxicity	152
6.5.6 Generation of <i>in vivo</i> models of DPR pathology.....	153
6.6 Concluding remarks.....	155
References.....	156

Word count: 47,055

List of Figures

Figure 1.1: Overview of the main clinical features of FTLD/ALS.....	14
Figure 1.2: Pathological lesions commonly observed in ALS tissue.....	29
Figure 1.3: Neuronal tau pathology in FTLD.....	31
Figure 1.4: FUS pathology in cortical neurons of FTLD tissue.....	31
Figure 1.5: Different types of TDP-43 pathology observed in cortical FTLD tissue.....	33
Figure 1.6: Type-B TDP-43 pathology in C9FTLD/ALS.....	46
Figure 1.7: The developmental stages of the <i>C. elegans</i> life-cycle.....	50
Figure 1.8: Comparison of the human <i>C9orf72</i> and <i>C. elegans</i> <i>F18A1.6</i> genes.....	53
Figure 3.1: Characterisation of the <i>F18A1.6</i> null mutant <i>C. elegans</i> strain, LM1006.....	68
Figure 3.2: Assessment of motility and survival in <i>F18A1.6</i> null mutant worms.....	69
Figure 4.1: Design of alternative codon sequences for DPR expression.....	81
Figure 4.2: Schematic illustrating the DPR cloning strategy.....	82
Figure 4.3: Agarose DNA gels showing repeat sequences from DPR constructs at a range of different repeat-lengths.....	88
Figure 4.4: Optimisation of DPR transfection.....	91
Figure 4.5: Confirmation of GFP-tagged DPR protein expression in HeLa cells.....	93
Figure 4.6: Differential distribution of GFP-tagged DPRs in HeLa cells at 24h post-transfection.....	95
Figure 4.7: Differential distribution of GFP-tagged DPRs in HeLa cells at 72h post-transfection.....	96
Figure 4.8: Distribution of untagged DPRs in HeLa cells.....	97
Figure 4.9: Example images of DPR pathology in patient brain tissue.....	103
Figure 5.1: Impact of DPR expression on TDP-43 distribution.....	110
Figure 5.2: Impact of DPR expression on EWS distribution.....	111
Figure 5.3: Impact of DPR expression on FUS distribution.....	112

Figure 5.4: Disruption of the nuclear membrane in HeLa cells expressing DPRs.....	113
Figure 5.5: Co-localisation of alanine-rich DPRs with components of the ubiquitin-protease system.....	115
Figure 5.6: Alanine-rich DPRs do not co-localise with ubiquitin in HeLa cells.....	116
Figure 5.7: Arginine-rich DPRs translocate to the nucleolus.....	118
Figure 5.8: Preliminary evidence suggesting arginine-rich DPRs may be exported from the nucleus over time.....	119
Figure 5.9: Poly-GR peptide, not RNA, causes a nucleolar stress phenotype.....	120
Figure 5.10: Nucleolar, but not cytoplasmic, poly-GR (green) causes a reduction in staining for the Cajal body marker, coilin.....	121
Figure 5.11: Nucleolar poly-GR causes loss of nuclear SMN-containing bodies.....	122
Figure 6.1: Proposed mechanisms of toxicity in C9FTLD/ALS.....	141

List of Tables

Table 1.1: Summary of the main genes which have been linked to FTLN/ALS.....	24
Table 1.2: Description of the major subtypes of FTLN-TDP pathology.....	32
Table 1.3: DPRs which could arise from RAN-translation of the repeat expansion.....	48
Table 1.4: Summary of the three possible mechanisms through which the C9orf72 expansion may cause neurotoxicity.....	54
Table 4.1: List of DPR constructs produced.....	89
Table 5.1: List of primary antibodies and stains used for immunofluorescence.....	107

List of abbreviations

3R/4R tau: Tau protein containing 3 or 4 microtubule-binding domains, respectively

AD: Alzheimer's disease

ALS: Amyotrophic lateral sclerosis

AP: Alanine-proline

bvFTD: Behavioural variant frontotemporal dementia

C9ALS: C9orf72-linked amyotrophic lateral sclerosis

C9FTLD: C9orf72-linked frontotemporal lobar degeneration

CB: Cajal body

CBD: Corticobasal degeneration

CpG: Cytosine-phosphate-guanine

DPR: Dipeptide repeat

ER: Endoplasmic reticulum

FET: FUS/TLS, EWSR1, and TAF15

FISH: Fluorescent *in situ* hybridisation

FTD: Frontotemporal dementia

FTDP-17: Frontotemporal dementia with Parkinsonism linked to chromosome 17

FTLD: Frontotemporal lobar degeneration

FTLD-ALS: Combined frontotemporal lobar degeneration and amyotrophic lateral sclerosis

FUS: Fused-in sarcoma

GA: Glycine-alanine

GAR: Glycine-arginine rich

GP: Glycine-Proline

GR: Glycine-arginine

GRN: Granulin

GEF: Guanine nucleotide exchange factor

hnRNP: Heterogeneous ribonucleoprotein

iPSC: Induced pluripotent stem cell

LB: Luria broth

NEB: New England BioLabs

NGM: Nematode growth media

PBS: Phosphate buffered saline

PGRN: Progranulin

PNFA: Progressive non-fluent aphasia

PR: Proline-arginine

PSP: Progressive supranuclear palsy

qRT-PCR: Quantitative real-time polymerase chain reaction

RG/RGG: arginine-glycine/arginine-glycine-glycine

SD: Semantic dementia

SDM: Site-directed mutagenesis

SDS-PAGE: Sodium dodecyl sulphate polyacrylamide gel electrophoresis

SMA: Spinal muscular atrophy

SMN: Survival motor neuron protein

SOD1: Cu/Zn superoxide dismutase 1

TBS-T: Tris buffered saline with Tween-20

TDP-43: Transactive response DNA-binding protein 43

TREX: Transcription export complex

UPS: Ubiquitin protease system

VCP: Valosin-containing protein

Abstract

Frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS) are neurodegenerative diseases with considerable clinical, genetic and pathological overlap. A GGGGCC hexanucleotide repeat expansion in a non-coding region of C9orf72 on chromosome 9 is the major cause of both FTLD and ALS. An understanding of the mechanisms through which the expansion leads to neurodegeneration will therefore be vital for development of novel therapeutics.

There are 3 possible mechanisms through which the GGGGCC expansion may cause toxicity: (i) through haploinsufficiency of C9orf72, (ii) repetitive RNA transcripts arising from the expansion may be toxic or (iii) translation of the expansion may produce toxic peptides. Whilst the expansion is located in a non-coding region of the gene, long GC-rich RNA transcripts may be translated in the absence of an ATG start codon, through a process known as repeat-associated non-ATG translation (RAN-translation). 5 distinct dipeptide repeat proteins (DPRs) have been found to arise from the expansion through RAN-translation of the sense and antisense strands in all frames: poly-GA, -GR, -PR, -AP and -GP. All 5 DPRs have been shown to aggregate in patient tissue, indicating that they may play a role in C9FTLD/ALS pathogenesis. This project aimed to generate a series of models to investigate the mechanisms of neurodegeneration in C9FTLD/ALS, using *C. elegans* and cell culture.

A transgenic worm strain which does not express the *C. elegans* orthologue of C9orf72 was first characterised. No impairments were observed in motility or life-span, demonstrating that loss of C9orf72 function does not cause an ALS-like phenotype in *C. elegans*. When considered alongside recent literature, this finding suggests that toxic gain of function mechanisms may be more important in C9FTLD/ALS pathogenesis.

The impact of DPRs on cellular function was next investigated. Constructs were generated containing alternative codon sequences for each of the 5 DPRs, to express each peptide in the absence of the repetitive GGGGCC RNA sequences found in disease. A step-wise cloning strategy was employed to progressively increase repeat-length in these constructs until physiologically-relevant lengths of >1000 repeats were obtained. DPRs were then expressed in HeLa cells, in order to assess the individual effects of each peptide on cellular function. Poly-GA formed large, star-shaped cytoplasmic inclusions which co-localised with ubiquilin-2 and p62, closely resembling the inclusions observed in patient tissue. This implies a potential role of proteasome dysfunction in C9FTLD/ALS.

Conversely, the alanine-rich DPRs, poly-GR and -PR, translocated to the nucleolus, where poly-GR in particular caused nucleolar stress. Furthermore, nucleolar poly-GR caused loss of Cajal bodies from the nucleus, and a loss or mislocalisation of survival motor neuron protein (SMN). This is of particular interest to C9ALS, since loss of SMN is selectively toxic to motor neurons. Furthermore, loss of Cajal bodies and nucleolar stress are likely to cause defects in RNA processing, which may contribute to neurotoxicity.

Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University's policy on Presentation of Theses.

Acknowledgements

This project would not have happened without the help and support of everyone in the Pickering-Brown and Sattelle labs, past and present. Firstly I would like to thank my supervisors, Prof Stuart Pickering-Brown and Prof David Sattelle, for the opportunity to work on an amazing project, and for their continued support and guidance throughout. Thanks also to my advisor, Prof Stuart Allan, for his encouraging pep talks and careers advice. I hope that the 4 of us will continue to work well together in future projects.

Special thanks go to Dr Janis Bennion Callister, who has been a teacher, a mentor and a friend. Janis took point on the DPR project, and without her enormous contribution it may never have gotten off the ground. Janis designed the DPR sequences and cloning strategy, developed the immunofluorescence protocol and performed much of the imaging. The coilin and fibrillar experiments in Chapter 5 were performed by Janis, and the figures in Section 4.2.7 were made from a combination of both our images. The initial DPR cloning was also a joint effort between us both. I cannot thank Janis enough for always looking out for me, for giving me great advice, for our long, post-6pm TC chats where we probably should have just called it a day and gone to the pub, and for basically teaching me to science.

Thanks also to Dr Sara Rollinson and “Auntie” Kate Young for regular favours, answering a million stupid questions and generally making the lab a fun place to work. Sara performed the qRT-PCR described in Chapter 4, and the data analysis for the qRT-PCR in Chapter 3. She also kindly offered to proof read my entire thesis, even though it didn't fall into her preferred category of books about dragons. Thanks also to Rosie Heartshorne, who was definitely the best Masters student ever, and a good friend. The immunoblotting of the shortest DPRs in Chapter 4 was performed by Rosie. Rosie also helped out with splitting cells and making buffers, and kept me sane in the last few months of hectic lab work. My final science-related thanks go to Dr Freddie Partridge, who patiently taught me how to inject DNA into the vulva of a microscopic worm without killing it.

I also owe huge thanks to all the amazing people in my life outside of work who made my time in Manchester brilliant. In particular, Mike Haley, Connie Baker, Rosie Griffiths, Maria Lizio, Katie Roberts, Frances Smith and Becki Mackenzie have been there through the good, the bad and the ridiculous, and my parents have been absolutely fantastic and supportive throughout, as always. Everyone on the 2nd floor has also been excellent craic and made working in AV Hill super fun.

Finally, thanks to the University of Manchester Neuroscience Research Institute (it was fun while it lasted), Faculty of Medical and Human Sciences, MRC and Wellcome Trust for funding.

Chapter 1 – Introduction

1.1 Overview of frontotemporal lobar degeneration and amyotrophic lateral sclerosis

Frontotemporal lobar degeneration (FTLD) and amyotrophic lateral sclerosis (ALS; collectively referred to as FTLD/ALS) are two neurodegenerative diseases with considerable clinical, genetic and pathological overlap. As such, FTLD and ALS are now considered to form a heterogeneous disease spectrum. Individuals present with symptoms of both diseases in ~15% of cases, and this is termed FTLD-ALS (Jokelainen 1977b; Ringholz et al. 2005; Lomen-Hoerth 2004). FTLD is the second most common cause of degenerative dementia, accounting for approximately 20% of all cases in patients under 65 years (Harvey et al. 2003). Clinical presentation is varied, and patients may display a wide range of behavioural and cognitive symptoms (Figure 1.1), including impairments in language, socially inappropriate behaviour, personality changes and emotional blunting (Neary et al. 1998; Brun et al. 1994). Large-scale progressive atrophy of the frontal or temporal lobes is observed in all cases (Neary et al. 1998; Brun et al. 1994), however the mechanisms leading to neurodegeneration in FTLD are not understood, and there are no disease-modifying treatments available. The pathogenesis of ALS is also poorly understood. ALS is a motor neuron disease characterised by progressive loss of motor neurons from the cortex, brainstem and spinal cord, resulting in muscle wasting, weakness and progressive paralysis (Figure 1.1). It is typically fatal within 2-5 years of symptomatic onset, once the paralysis reaches the respiratory muscles (Jokelainen 1977a; Chancellor et al. 1993). The only treatment available is Riluzole, which yields only a modest increase of several months in life-span (Miller et al. 2012). Therefore there are no suitable treatments available for either FTLD or ALS.

The calculated prevalence of FTLD varies, for example a study of dementia patients in Cambridge, UK estimated the prevalence of early onset FTLD (before 65 years) as approximately 15 in 100,000, with an overall prevalence of 81 in 100,000 when including late-onset cases (Ratnavalli et al. 2002). A prevalence of 2.7 in 100,000 was estimated in the Netherlands (Rosso et al. 2003), whilst an Italian study calculated an early onset prevalence of 17.6 per 100,000 in Northern Italy (Borrioni et al. 2010). In all cases, the incidence is age-dependent. The incidence of ALS has been estimated at 2.16 in 100,000 in a large study of populations in the UK, Ireland and Italy, with increased incidence in males (Logroscino et al. 2010), although there is again considerable variation between studies.

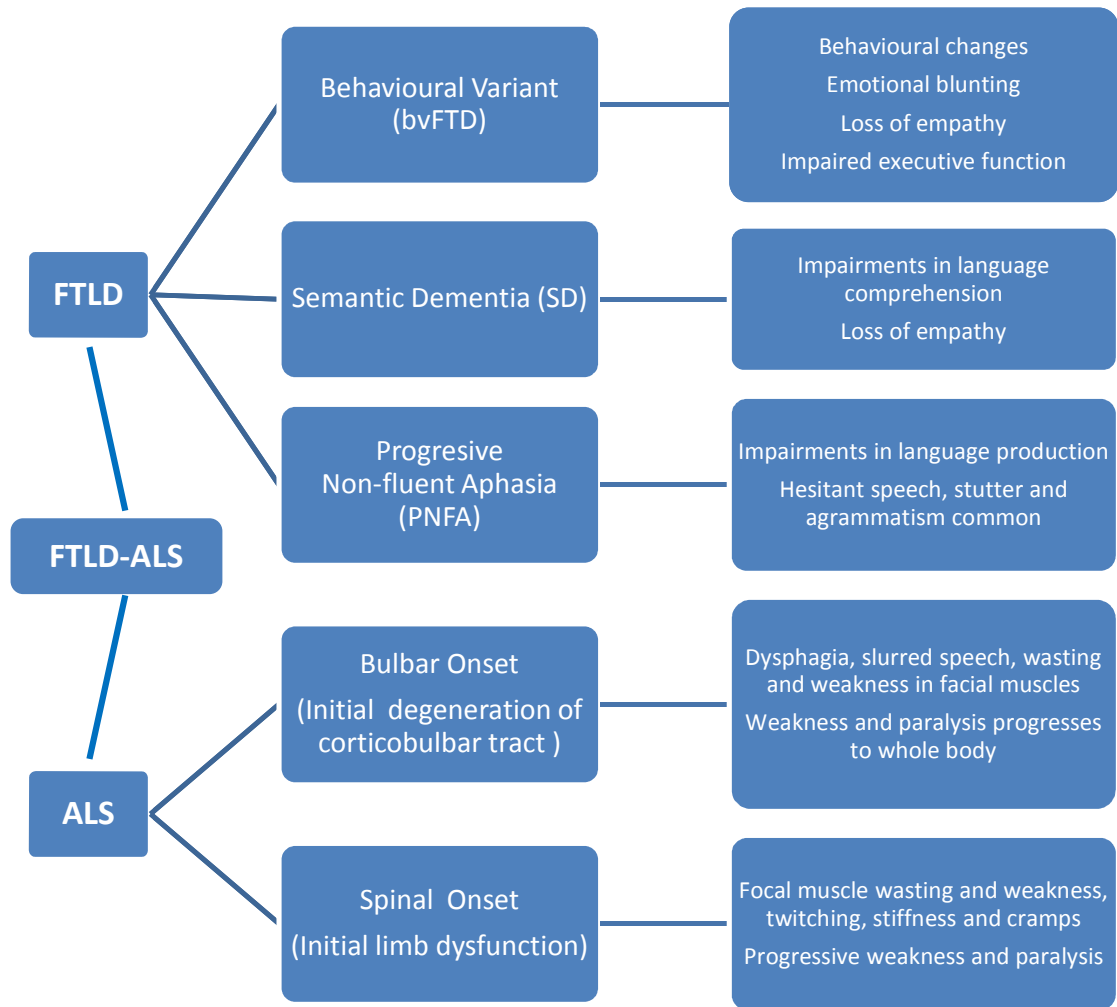


Figure 1.1: Overview of the main clinical features of FTLD/ALS. FTLD can be subcategorised into behavioural variant, semantic dementia and progressive non-fluent aphasia. In addition to the symptoms listed here, FTLD patients may also present with Parkinsonism or psychotic symptoms. Approximately 15% of patients present with a combination of FTLD and ALS symptoms. Information obtained from Neary et al. 1998; Brun et al. 1994; Brooks et al. 2000.

An average prevalence of 5.2 in 100,000 has been reported following meta-analysis of 15 studies in Europe and North American populations (Worms 2001). It is also worth noting that the prevalence of ALS is reduced by the high mortality rate. Despite the variability between epidemiological studies, it is evident that both FTLD and ALS are reasonably common diseases in Western populations. Given the continuing rise in the elderly population, the incidence of age-related neurodegenerative dementias such as FTLD is also predicted to increase dramatically in future decades. This will in turn increase the already large socio-economic burden of healthcare funding in dementia.

Understanding FTLN/ALS disease mechanisms and developing candidate therapies are therefore urgent priorities.

An important recent advance in FTLN/ALS research was the identification of a novel mutation in familial cases, which may cause pure FTLN, pure ALS or combined FTLN-ALS (DeJesus-Hernandez et al. 2011; Renton et al. 2011). The mutation is a large hexanucleotide repeat expansion in a gene named C9orf72 on chromosome 9, and is the most common known cause of FTLN/ALS (Renton et al. 2011; DeJesus-Hernandez et al. 2011). However, the pathogenic mechanisms involved in C9orf72-linked FTLN/ALS (C9FTLN/ALS) remain unknown. This project aims to investigate the mechanisms by which the C9orf72 expansion causes disease, using cell culture and *C. elegans* models.

1.2 Clinical aspects of FTLD/ALS

FTLD and ALS are heterogeneous diseases which can be subcategorised according to clinical, genetic or pathological features. Sections 1.2-1.4 provide a general overview of clinical presentation of FTLD/ALS, the genes which have been previously implicated in FTLD/ALS and the most common pathological hallmarks of both diseases. The specific features of C9FTLD/ALS will be discussed in Section 1.5.

When categorised according to clinical presentation, there are three key subtypes of FTLD: behavioural variant frontotemporal dementia (bvFTD), progressive non-fluent aphasia (PNFA) and semantic dementia (SD). PNFA and SD may also be referred to by the umbrella term of primary progressive aphasia (PPA). In addition, there is some clinical overlap between FTLD and other degenerative proteinopathies such as corticobasal degeneration (CBD), progressive supranuclear palsy (PSP) and inclusion body myopathy. FTLD patients may also present with ALS symptoms in addition to any of the above three syndromes (FTLD-ALS), or with other motor syndromes such as Parkinsonism, although this is considerably less common.

1.2.1 bvFTD

bvFTD with or without ALS is the most common form of FTLD. It is primarily characterised by changes in personality and social interaction, with symptoms such as disinhibition and socially inappropriate behaviour exhibited (Neary et al. 1998; Brun et al. 1994). Patients have been frequently reported to become rude or aggressive towards family members, carers or strangers (Brun et al. 1994). Disinhibition may lead to impulsive behaviours and decision-making, such as uncharacteristic gambling or large impulsive purchases or donations. Hypersexuality and inappropriate sexual behaviours are also common (Neary et al. 1998; Brun et al. 1994). The development of behavioural changes is usually accompanied by loss of insight, and therefore most patients are unaware of their symptoms and do not realise their behaviour may be considered inappropriate or unusual (Neary et al. 1998; Brun et al. 1994). Indeed, the responses provided by patients when questioned about their behaviour have been shown to differ considerably from responses given by caregivers (Eslinger et al. 2012).

Another common hallmark of bvFTD is apathy, which may occur very early in disease (Neary et al. 1998; Brun et al. 1994). This may manifest in a number of ways, such as lack of motivation, immobility, poor personal hygiene and abandonment of hobbies and

interests. Patients may also lose interest in family and friends, and begin to withdraw socially (Merrilees et al. 2013; Piguet et al. 2011; Eslinger et al. 2012). Social apathy is often associated with emotional blunting and loss of empathy, with patients being unaware of the emotional impact of their disinterest on family members. For example, one case study described a man who remained indifferent when his wife suffered an injury, a behaviour which was deemed out of character by his wife (Rohrer et al. 2008). Several different aspects of empathy have been shown to be affected in bvFTD, including inability to understand the intentions of others, impaired recognition of emotions and changes in moral judgement (Cerami et al. 2014; Baez et al. 2014).

Obsessive and ritualistic behaviours are frequently observed in bvFTD (Neary et al. 1998; Brun et al. 1994). These may include repetitive checking behaviours, for example, checking locks multiple times per day, excessive cleaning or obsessive re-arranging of household items (Perry et al. 2012; Hodges 2001). Many patients also develop hoarding habits, insisting on storage of old or unneeded items or waste (Perry et al. 2012). Dietary changes are also common; in particular, an increased preference for sweet foods is frequently reported in bvFTD (Neary et al. 1998; Brun et al. 1994). Dietary behaviours may also become obsessive or ritualistic, for example patients may eat the same food at the same time every day (Hodges 2001). Furthermore, over-eating and food-stuffing behaviours are common, with case studies reported of patients eating so excessively or quickly that vomiting is induced (Snowden et al. 2011; Piguet et al. 2011).

Whilst memory and language are usually well preserved in bvFTD, a range of other cognitive problems are common, including defects in attention, problem solving and executive function (Neary et al. 1998; Brun et al. 1994). Patients may therefore experience difficulties with ordinary daily tasks such as dressing themselves, and many require full-time care in the later stages of disease. The abilities to plan activities and multitask are also frequently impaired (Merrilees et al. 2010). Impairments in episodic memory and language may also occur infrequently in the early stages of bvFTD, or more frequently as the disease progresses (Pijnenburg et al. 2004; Neary et al. 1998; Brun et al. 1994).

The clinical progression of bvFTD may vary, however certain signs and symptoms have been consistently reported to occur at particular stages in disease. For example, loss of empathy, apathy and behavioural or personality changes tend to occur in the early stages of disease, and such symptoms are usually the initial reason for patient referral to specialist clinics. The majority of patients do not recognise these changes within themselves, indicating that loss of insight is also an early feature of bvFTD (Neary et al.

1998; Brun et al. 1994). Despite this decline in behavioural function, bvFTD patients may perform relatively well on neuropsychological tests early in disease (Piguet et al. 2011). As the disease progresses, however, cognitive dysfunction becomes more prominent, with memory and language impairments often occurring in the later stages (Neary et al. 1998; Brun et al. 1994). The average disease duration in bvFTD is approximately 6 years from onset (Linds et al. 2015; Hodges et al. 2003).

1.2.2 Primary progressive aphasia

The main subtypes of PPA are SD and PNFA. These disorders are characterised by impairments in language. PNFA patients experience difficulty primarily with expressive language, for example problems initiating speech and in word retrieval, slow or hesitant speech, development of a stutter, grammatical errors, and impairments in reading and writing (Neary et al. 1998). Conversely, SD is primarily a disorder of language comprehension. Symptoms typically exhibited in SD include fluent anomia, impaired understanding of words, prosopagnosia and difficulty recognising and naming objects (Neary et al. 1998). The majority of cognitive functions are preserved in PNFA, although impairments in working memory and executive function may be observed (Snowden et al. 2006). Some behavioural symptoms reminiscent of bvFTD may also present in SD and PNFA, such as personality changes and apathy (Neary et al. 1998). However, language disorders are the predominant clinical presentation in these disorders. Like bvFTD, SD and PNFA patients display loss of insight, and usually do not recognise changes in their own behaviour (Eslinger et al. 2012; Snowden et al. 2006).

1.2.3 Extrapyraxidal symptoms in FTLT

In addition to co-morbidity with ALS in ~15% of cases, FTLT may also present with extrapyramidal motor features more reminiscent of Parkinson's disease. This is most commonly observed in FTLT with Parkinsonism linked to chromosome 17 (FTDP-17; discussed in Section 1.3), but may also occur in other forms of FTLT. In these cases, patients present with a combination of dementia and motor symptoms such as bradykinesia, tremor, stooped posture and muscular rigidity (Hirschbichler et al. 2015). Hypophonia, which is also a common feature of Parkinson's disease, has also been reported in cases of FTLT (Snowden et al. 2011).

1.2.4 ALS

ALS is characterised by the selective degeneration of motor neurons in the cortex, brainstem and spinal cord. Both upper and lower motor neurons tend to be affected (Brooks et al. 2000). The disease may be clinically categorised into two groups, depending on whether symptomatic onset begins with bulbar or spinal signs (Al-Chalabi et al. 2012). Approximately two thirds of patients present with spinal onset, which manifests initially as limb dysfunction (Logroscino et al. 2010; Chancellor et al. 1993). The early symptoms include muscle wasting and weakness, twitching, stiffness and cramps, with bulbar symptoms occurring later in disease (Brooks et al. 2000). Bulbar onset ALS begins with degeneration of motor neurons in the corticobulbar tract. Therefore initial symptoms include wasting and weakness in the facial muscles, jaw and tongue, dysphagia, slow and slurred speech, sialorrhoea, inability to control facial expressions and laryngospasm (Hern et al. 1992). Interestingly, spinal onset ALS is more common in males, although the reason for this is unknown (Logroscino et al. 2010). The oculomotor nerve and motor neurons controlling the bladder/bowels are typically spared in ALS (Okamoto et al. 1993). In all cases, ALS is a severe progressive disease causing paralysis throughout the body as motor neuron loss increases. It is typically fatal in 2-5 years due to respiratory failure when paralysis reaches the intercostal muscles (Jokelainen 1977a; Chancellor et al. 1993). Furthermore, the prognosis is worse in bulbar onset ALS since dysphagia and related problems increase the incidence of pneumonia and other lung infections (Chiò et al. 2009). The majority of ALS patients do not suffer from cognitive or behavioural symptoms, however some patients additionally present with cognitive impairments or FTLD (Jokelainen 1977b; Ringholz et al. 2005). Where FTLD is present as a secondary feature of ALS, the term ALS-FTLD may be used.

1.3 Genetics of FTLD/ALS

The genetics of FTLD/ALS are complex; multiple monogenetic mutations have been identified as causative in either FTLD or ALS, and in some cases a single gene has been implicated in both disorders. The mutations known to be involved in FTLD/ALS to date are summarised in Table 1.1. FTLD is considered familial in ~25-50% of cases (Rademakers et al. 2012), whilst ALS is familial in ~5-10% of cases (Al-Chalabi et al. 2012). More commonly however, FTLD and ALS are apparently sporadic diseases, with no obvious familial link. Although there is considerable pathological heterogeneity within the FTLD/ALS spectrum, the same pathological features are observed in familial and sporadic cases. This section will provide an overview of the genetic mutations which have been previously linked to FTLD/ALS.

1.3.1 *SOD1*

The first discovery of a causative mutation in FTLD or ALS was in 1993. A series of missense mutations in Cu/Zn superoxide dismutase 1 (*SOD1*) were identified in ALS, (Rosen et al., 1993), which are found in ~20% of familial ALS cases and ~3% of sporadic ALS. *SOD1* mutations are predominantly found in pure ALS, although dementia may also present in rare cases (Al-Chalabi et al. 2012). No cases of pure FTLD with *SOD1* mutations have been reported. Over 160 *SOD1* mutations have been identified as pathogenic in ALS patients to date. Whilst the majority are single amino acid substitutions, a number of nonsense and deletion mutations have also been reported (Hosler et al. 1996; Morita et al. 1996; Sapp et al. 1995; Watanabe et al. 2000; Rosen et al. 1993). *SOD1* transgenic mice typically exhibit loss of motor neurons, paralysis or motor impairments and shortened life-span (reviewed by Joyce, Fratta, Fisher, & Acevedo-Arozena, 2011), further demonstrating the causative nature of *SOD1* mutations.

The *SOD1* protein is an abundantly expressed enzyme which catalyses the conversion of toxic superoxide free radicals into harmless O_2 and H_2O_2 (Noor et al. 2002). *SOD1* is expressed in the cytoplasm of many cell types including motor neurons and glia, although the selective vulnerability of motor neurons in *SOD1*-linked ALS has not been explained. *SOD1* is also expressed in the mitochondrial intermembrane space, although less abundantly. In disease, however, misfolded *SOD1* has been found to mislocalise to the mitochondrial space, where it accumulates excessively. Therefore a role of mitochondrial dysfunction and oxidative stress has been strongly implicated in ALS

(reviewed by Pizzuti & Petrucci, 2011; Tafuri, Ronchi, Magri, Comi, & Corti, 2015). However, despite the discovery of *SOD1* mutations in ALS being over twenty years ago, the downstream effects of *SOD1* mutations leading to disease remain poorly understood.

1.3.2 *MAPT*

Five years after the discovery of *SOD1* mutations in ALS, The first monogenetic causes of FTLD were identified as a series of missense and splice-site mutations in the *MAPT* gene on chromosome 17 (Hutton et al. 1998; Poorkaj et al. 1998; Spillantini et al. 1998). The majority of *MAPT* mutations cause a specific clinical phenotype of typical FTLD with comorbid Parkinsonism (FTLD with Parkinsonism linked to chromosome 17; FTDP-17). *MAPT* mutations are one of the most common known causes of FTLD, accounting for ~2-11% of cases, dependent on geographical location (Rohrer & Warren 2011). However, no *MAPT* mutations have been identified in ALS.

MAPT encodes the microtubule-associated protein tau, a cytoskeletal protein which plays an important role in microtubule polymerisation and stabilisation in neurons (Amos 2004). Tau is able to bind directly to microtubules via 3 or 4 microtubule-binding domains encoded by exons 9-12. In the healthy brain, alternative splicing of *MAPT* leads to the expression of 6 different tau isoforms, each containing either 3 or 4 microtubule-binding domains (termed 3R and 4R tau, respectively) produced from translation of these exons (Goedert et al. 1989). The number of microtubule-binding domains is determined by inclusion or exclusion of exon 10; alternative splicing causes exon 10 to be skipped in 3R tau isoforms. An equal ratio of 3R:4R tau is present in the healthy adult human brain. However, over 40 *MAPT* mutations have been identified in FTDP-17 to date, and many of these are clustered in or around exon and intron 10, causing splicing abnormalities and therefore altering the ratio of 3R:4R tau (Hutton et al. 1998; Grover et al. 1999; Hasegawa et al. 1999; Jiang et al. 2003). Some FTDP-17-linked mutations have been found to alter the ratio of 3R:4R tau in both directions, indicating that the balance of tau isoforms is critical for neuronal function (Hasegawa et al. 1999; van Swieten et al. 2007). Other *MAPT* mutations include deletions and missense mutations in the microtubule binding regions, suggesting that loss of microtubule-binding function is key to disease pathogenesis.

1.3.3 *GRN*

The second gene to be identified as a major cause of FTLD was *GRN*, encoding progranulin (PGRN; Baker et al. 2006; Cruts et al. 2006). This locus is only 1.7Mb away from the *MAPT* gene on chromosome 17, which caused initial difficulties in identifying *GRN* mutations in FTLD following linkage and association studies. Since the initial discovery of *GRN* mutations in 2006, over 60 different mutations have been identified in FTLD, accounting for ~5-11% of cases (Rohrer & Warren 2011). Therefore the frequency of *MAPT* and *GRN* mutations in FTLD cohorts are similar. Interestingly, *GRN* mutations may also cause the clinically related disorder, CBD, indicating a pathological link between CBD and FTLD (Masellis et al. 2006; Rademakers et al. 2007; Spina et al. 2007).

The majority of *GRN* mutations are known to cause disease through PGRN haploinsufficiency. A variety of heterozygous frame-shift mutations have been identified which introduce premature stop codons in the *GRN* gene, leading to the production of truncated pre-mRNA transcripts and nonsense-mediated decay (Gass et al. 2006; Cruts et al. 2006; Le Ber et al. 2007; van der Zee et al. 2007). PGRN expression is therefore reduced in these cases. Mutations in ATG-start codons were also identified, thus preventing normal translation of the PGRN protein (Cruts et al. 2006). Similarly, several mutations introducing partial or complete deletion of one copy of *GRN* have been reported in FTLD (Ghidoni et al. 2008; Gijssels et al. 2008). Splice-site mutations are also frequently observed in *GRN* (Gass et al. 2006; Cruts et al. 2006). These are thought to reduce PGRN levels at a post-transcriptional level, by altering splicing of regions important for protein export, thus promoting degradation of the protein. Furthermore, mutations affecting the localisation signal domain are thought to cause loss of PGRN function by preventing secretion of the protein (Mukherjee et al. 2006; Shankaran et al. 2008). The hypothesis that *GRN* mutations lead to neurodegeneration through loss of PGRN function is supported by analysis of protein expression in patients. Immunohistochemical analysis of FTLD tissue from patients with *GRN* mutations showed reduced PGRN reactivity in affected neurons (Cruts et al. 2006). Reduced PGRN levels have also been observed in the cerebrospinal fluid (CSF), plasma and serum of patients carrying a mutation (Finch et al. 2009; Ghidoni et al. 2008; Sleegers et al. 2009; Van Damme et al. 2008). Indeed, analysis of PGRN levels in the CSF by enzyme-linked immunosorbent assay (ELISA) has proved this reduction to be a reliable biomarker for disease. Therefore the common mechanism in *GRN*-linked FTLD pathogenesis is loss of PGRN function.

A common single nucleotide polymorphism (SNP) in *GRN* (rs5848) located in the 3' untranslated region has been reported to be a non-causative risk factor for FTLD, increasing the risk of non-mutation carriers developing the disease by approximately 3.2 fold (Rademakers et al. 2008). The risk locus is a binding site for a microRNA species, miR-659, which has been shown to regulate *GRN* expression *in vitro*, suggesting that this SNP may confer increased risk of disease by affecting *GRN* expression levels (Rademakers et al. 2008). However, a large study of 3 separate FTLD cohorts totalling 467 cases was unable to replicate the association between the rs5848 SNP and increased risk of disease, indicating that this finding is likely to have been a false positive result (Rollinson et al. 2011).

PGRN is a large, secreted precursor protein which is cleaved *in vivo* to produce a series of ~6kDa peptides known as granulins (GRNs; Bateman et al. 1990). The GRNs are encoded by cysteine-rich tandem repeat regions in the *GRN* gene, and therefore the structure of each peptide is partially conserved. Multiple functions have been proposed for PGRN overall, however, the individual contributions of the precursor protein and each individual GRN to these functions remain unclear.

Whilst expression is strong in brain regions important for FTLD such as the hippocampus and cortex, PGRN is also present in peripheral tissues including the gastrointestinal tract, immune system, epithelial cells and reproductive system (Daniel et al. 2000). In the brain, PGRN is primarily expressed in mature neurons, with relatively low expression observed throughout development (Petkau & Leavitt 2014). PGRN is also abundant in activated microglia and macrophages, implying a role in the regulation of neuroinflammation. Indeed, upregulation of PGRN expression has been observed in microglia and macrophages following brain and spinal cord injury in mice (Tanaka et al. 2013), as well as in animal models of various neurodegenerative diseases which are known to involve an inflammatory component, such as ALS and Alzheimer's disease (AD; Pereson et al., 2009; Philips et al., 2010). Upregulation of *GRN* expression has also been observed in post-mortem spinal cord tissue in ALS (Irwin et al. 2009). Interestingly, an increased inflammatory response has been reported in *GRN* knockout mice, suggesting the role of PGRN in inflammation may be to suppress microglial activation, rather than enhance it in mice (Tanaka et al. 2013). Gliosis and inflammation are common features of most neurodegenerative diseases, and an excessive inflammatory response has been proposed to exacerbate neuronal loss in these disorders. Therefore one mechanism through which loss of *GRN* may contribute to neurodegeneration in FTLD is through dysregulation of neuroinflammation.

Table 1.1: Summary of the main genes which have been linked to FTL/ALS to date.

Mutated Gene	Protein	Link to FTL/ALS	Normal Functions	References
<i>C9orf72</i>	C9ORF72	Hexanucleotide expansion causative in FTL, ALS and FTL-ALS	Unknown. Possible DENN-like GEF activity	(DeJesus-Hernandez et al. 2011; Renton et al. 2011)
<i>MAPT</i>	Tau	Causative in FTL	Microtubule stabilisation	(Hutton et al. 1998; Poorkaj et al. 1998; Spillantini et al. 1998)
<i>GRN</i>	Progranulin	Causative in FTL	Lysosome function Inflammation Neurotrophic factor	(Baker et al. 2006; Cruts et al. 2006)
<i>SOD1</i>	SOD1	Causative in ALS	Oxidative stress, mitochondrial function	(Rosen et al. 1993)
<i>TARDBP</i>	TDP-43	Causative in ALS/ALS-FTL Rarely reported in pure FTL	RNA processing	(Kabashi et al. 2008; Sreedharan et al. 2008; Ticozzi et al. 2011; Borroni et al. 2009)
<i>FUS</i>	FUS	Causative in ALS/ALS-FTD Rarely reported in pure FTD	RNA processing	(Kwiatkowski et al. 2009; Vance et al. 2009; Zou et al. 2012)
<i>SQSTM1</i>	p62	Causative in FTL May be causative in ALS	Protein degradation	(Miller et al. 2015; Rubino et al. 2012)
<i>VCP</i>	Valosin-containing protein (VCP)	Causative in inclusion body myopathy with FTL Rarely reported in ALS and ALS-FTL	Vesicular trafficking, protein degradation	(DeJesus-Hernandez et al. 2011; Koppers et al. 2012; Schröder et al. 2005; Johnson et al. 2010; Watts et al. 2004)
<i>UBQLN2</i>	Ubiquilin-2	Causative in FTL/ALS	Protein degradation	(Daoud et al. 2012; Deng et al. 2011; Gellera et al. 2013; Millecamps et al. 2012; Synofzik et al. 2012)
<i>MATR3</i>	Matrin 3	Rarely reported in ALS	RNA processing	(Lin et al. 2015; Johnson et al. 2014)
<i>TREM2</i>	Triggering receptor expressed on myeloid cells 2	Rarely reported in FTL	Inflammation	(Borroni et al. 2014; Guerreiro et al. 2013)
<i>CHCHD10</i>	Coiled-coil-helix-coiled-coil-helix domain containing 10	Rarely reported in FTL/ALS	Mitochondrial function	(Chaussonot et al. 2014; Bannwarth et al. 2014)
<i>CHMP2B</i>	Charged multivesicular body protein 2b (CHMP2B)	Rarely reported in FTL	Endosomal trafficking Protein degradation	(Ferrari et al. 2010; Skibinski et al. 2005; Parkinson et al. 2006)
<i>PFN1</i>	Profilin 1	Rarely reported in ALS	Actin microfilament stabilisation	(Wu et al. 2012; Tiloca et al. 2013)
<i>HNRNPA1</i>	Heterogeneous ribonucleoprotein A1	Rarely reported in ALS-like syndromes	RNA processing	(Kim et al. 2013)
<i>HNRNPA2B1</i>	Heterogeneous ribonucleoprotein A2/B1	Rarely reported in ALS-like syndromes	RNA processing	(Kim et al. 2013)

Perhaps the most well-established function of PGRN is its role in neurotrophic signalling. Overexpression of *GRN* has been shown to promote neuronal growth and survival, and to protect cultured neurons against toxic insults such as serum starvation (Ryan et al. 2009; Van Damme et al. 2008). *GRN* E has also been shown to independently exhibit neurotrophic effects *in vitro* (Van Damme et al. 2008). Conversely, depletion of *GRN* has been shown to inhibit neuronal growth and dendritic branching both *in vivo* and *in vitro*,

further demonstrating the neurotrophic role of PGRN (Gass et al. 2012; Petkau et al. 2012; Guo et al. 2010; Kleinberger et al. 2010). Therefore an alternative mechanism through which loss of PGRN function may lead to neurodegeneration is via loss of neurotrophic activity.

PGRN has also been recently found to play an important role in lysosomal function; whilst FTLD-linked *GRN* mutations are heterozygous, a rare homozygous loss of function mutation in *GRN* has been reported to cause a clinical phenotype more reminiscent of a lysosomal storage disorder (Smith et al. 2012), and *GRN* null mutant mice exhibit signs of lysosomal dysfunction (Ahmed et al. 2010; Petkau et al. 2012; Tanaka et al. 2014). Taken together, the literature shows that PGRN is a multifunctional protein with important roles in a number of cellular processes relevant to neurodegenerative disease. The varied functions of PGRN and the GRN family are not yet fully understood, however, and it is likely that additional functions of PGRN remain to be discovered. The mechanisms leading from *GRN* haploinsufficiency to neurodegeneration in FTLD therefore remain unclear.

1.3.4 *TARDBP* and *FUS*

Transactive response DNA-binding protein 43 (TDP-43) is a 43kDa DNA and RNA binding protein encoded by the *TARDBP* gene on chromosome 1. Fused-in sarcoma (*FUS*) is a functionally related DNA and RNA binding protein encoded by *FUS* on chromosome 16 (Baloh 2012; Mackenzie et al. 2010). TDP-43 and *FUS* belong to the FET (*FUS*/*TLS*, *EWSR1*, and *TAF15*) family of DNA/RNA binding proteins, and play diverse roles in RNA processing and control of gene expression. Mutations in both *TARDBP* and *FUS* have been identified in ALS (Kabashi et al. 2008; Kwiatkowski et al. 2009; Vance et al. 2009; Sreedharan et al. 2008; Yan et al. 2010), as well as in ALS-FTLD (Yan et al. 2010).

The significance of *TARDBP* mutations in FTLD with no associated ALS is less clear. Whilst cases of *TARDBP* mutations in patients with pure FTLD have been reported (Borroni et al. 2009; Kovacs et al. 2009), early genetic analysis of FTLD cohorts did not identify any significant mutations in *TARDBP*, and only weak association was found between *TARDBP* mutations and pure FTLD (Gijselinck et al. 2009; Rollinson et al. 2007; Schumacher et al. 2009). However, one of these studies also did not identify *TARDBP* mutations in ALS patients (Gijselinck et al. 2009), which have since been found to be causative. A larger study has since identified multiple mutations in a larger

cohort of FTLD patients, and the frequency of these mutations was low in this cohort (Borroni et al. 2010). It is possible that the early negative results were due to low frequency of FTLD-causing *TARDBP* mutations in the population, and that a larger sample size was required to capture this information. Interestingly, a single base substitution causing an arginine to serine change at amino acid 267 in TDP-43 has been reported in cases of pure FTLD and pure ALS (Borroni et al. 2009; Corrado et al. 2009). This is comparable to the C9orf72 expansion which may cause pure FTLD, pure ALS or combined FTLD-ALS, although the TDP-43 mutation is considerably less common. It remains unclear why the same mutation may cause FTLD in some cases and ALS in others. Taken together, the conflicting literature suggests that certain *TARDBP* mutations may be causative in FTLD but very rarely when compared to other known monogenetic causes. It is also possible that polymorphisms in *TARDBP* confer increased risk of FTLD without being directly causative, or form part of an as of yet unidentified haplotype containing additional disease-related mutations.

The role of *FUS* mutations in FTLD is also controversial. Rohrer *et al.* reported the absence of observable *FUS* mutations in a large-scale screen of an FTLD cohort (Rohrer et al. 2009). However, more recent studies have identified several *FUS* mutations in FTLD patients, suggesting that such mutations may be causative in rare cases (Huey et al. 2012; Van Langenhove et al. 2010). Further genetic screening of large patient cohorts and linkage studies in families are required to determine conclusively whether *TARDBP* and *FUS* mutations are causative in FTLD.

1.3.5 Rare mutations in FTLD/ALS

Numerous additional genes have been linked to familial FTLD/ALS, although with a very low mutation frequency compared to those already discussed. As such, the causative nature of these mutations remains unconfirmed in some cases. Interestingly, some common themes can be noted in the functions of genes known to be affected in FTLD/ALS, including protein degradation, RNA processing and endosomal sorting and trafficking. FTLD/ALS-linked mutations have been identified in two key components of the ubiquitin protease system (UPS), ubiquitin-2 and p62 (encoded by *UBQLN2* and *SQSTM1*, respectively), indicating that impaired protein degradation may play a role in disease pathogenesis (Rubino et al. 2012; Miller et al. 2015; Deng et al. 2011; Synofzik et al. 2012). Mutations have also been rarely reported in valosin-containing protein (VCP; encoded by *VCP*) and charged multivesicular body protein (CHMP2B; encoded by *CHMP2B*), both of which play dual roles in protein degradation and vesicular

trafficking (DeJesus-Hernandez et al. 2011; Johnson et al. 2010; Kwok et al. 2015; Ferrari et al. 2010; Parkinson et al. 2006; Skibinski et al. 2005). A further link to intracellular trafficking is provided by the discovery of mutations in the actin-binding protein, profilin 1 (encoded by *PFN1*), which is required for stabilisation of actin filaments and axonal transport (Wu et al. 2012; Tiloca et al. 2013). Recent reports of ALS-linked mutations in the RNA-binding proteins matrin 3 (encoded by *MATR3*) and heterogeneous ribonucleoproteins (hnRNP) A1 and A2/B1 (encoded by *HNRNPA1* and *HNRNPA2B1*, respectively) implicate aberrant RNA processing as a common feature of ALS pathogenesis in particular (Lin et al. 2015; Johnson et al. 2014; Kim et al. 2013). When considered alongside the more well-established roles of *TARDBP* and *FUS* mutations in ALS and perhaps FTLD, these findings present a strong argument for an important role of RNA processing defects in neurodegeneration. Other recent reports of genes rarely mutated in FTLD/ALS include coiled-coil-helix-coiled-coil-helix domain containing 10 (*CHCHD10*, encoded by *CHCHD10*), which together with *SOD1* implicates mitochondrial dysfunction in ALS (Bannwarth et al. 2014; Chausseot et al. 2014), and triggering receptor expressed on myeloid cells 1 (*TREM2*, encoded by *TREM2*) which combines with *GRN* to suggest inflammation may play a key role in FTLD (Guerreiro et al. 2013; Borroni et al. 2014). Therefore the genetics of FTLD/ALS implicate several common cellular processes and pathways in disease pathogenesis.

1.4 Pathological features of FTLD/ALS

On a macroscopic level, FTLD is characterised generally by gross atrophy of the frontal and anterior temporal lobes, regardless of the clinical subtype. Posterior brain regions may also atrophy as the disease progresses (Neary et al. 1998; Brun et al. 1994). Conversely, ALS causes selective atrophy of motor neurons in the cortex, corticospinal and corticobulbar tracts (Brooks et al. 2000). As in most neurodegenerative diseases, gliosis is typically observed in affected brain regions in FTLD/ALS (Neary et al. 1998; Brun et al. 1994). In addition, affected neurons and sometimes glia contain insoluble, ubiquitinated and hyperphosphorylated protein aggregates. The main protein components of these inclusions vary and can be used to categorise FTLD and ALS pathologically. There are three major pathological subtypes of FTLD named according to the main component of protein aggregates: FTLD-tau, FTLD-TDP (containing TDP-43) and FTLD-FUS. Similarly, ALS may be categorised according to three types of protein aggregates: SOD1, TDP-43 and FUS (summarised in Figure 1.2). In addition, a distinguishing pathological feature of ALS which does not occur in pure FTD is the presence of Bunina bodies. These are small, eosinophilic inclusions which may be found intraneuronally in motor regions of the brain and spinal cord (Okamoto et al. 2008; Shibata et al. 1996). Bunina bodies may also be present in motor neurons in cases of FTLD-ALS. However, the pathogenic implications of Bunina bodies are unclear.

In the vast majority of cases, tau, TDP-43, FUS and SOD1 pathology do not occur together in individual cases. Immunoreactivity for both TDP-43 and FUS has been observed in cytoplasmic inclusions in some cases (Deng et al. 2010), however this is extremely rare. Similarly, SOD1 pathology is not observed in ALS tissue exhibiting TDP-43 or FUS inclusions, and tau, TDP-43 and FUS pathology are not usually observed together in FTLD (Vance et al. 2009; Mackenzie et al. 2007), suggesting that several distinct mechanistic pathways lead to similar clinical presentations in both FTLD and ALS. It is particularly interesting that TDP-43 and FUS pathology is common to both FTLD and ALS. This provided relatively early evidence supporting the link between these two diseases. However it remains unclear why similar pathological features may lead to pure FTLD, pure ALS, or combined FTLD-ALS.

1.4.1 SOD1

The first protein to be implicated in ALS was SOD1, since the gene encoding this protein was found to be mutated in a small percentage of cases. SOD1 was found to

accumulate in spinal cord neurons in ALS autopsy tissue in cases where a *SOD1* mutation is present, and occasionally in sporadic cases (Shibata et al. 1996). *SOD1* does not accumulate in cases of pure FTLD. *SOD1* inclusions are large and cytoplasmic, and predominantly found in neurons in affected regions (Figure 1.2A).

Misfolding of *SOD1* is also thought to be important in ALS pathogenesis; High immunoreactivity for misfolded *SOD1* has been observed in patient tissue and transgenic rodents expressing *SOD1* mutations (Shibata et al. 1996; Mackenzie et al. 2007; Bruijn et al. 1997). Interestingly, misfolded *SOD1* appears to preferentially translocate to the mitochondria in transgenic mice and patient spinal cord tissue (Liu et al. 2004; Vijayvergiya 2005). This may cause neurotoxicity through disruption of normal mitochondrial function, leading to ATP deficit, oxidative stress or impaired calcium homeostasis. Misfolded protein may also cause toxicity by “overloading” protein degradation systems, which are unable to process the mutant *SOD1*. Therefore there are a number of ways in which *SOD1* pathology may affect cellular processes important to neuronal survival. However, the precise mechanisms of toxicity remain unclear.

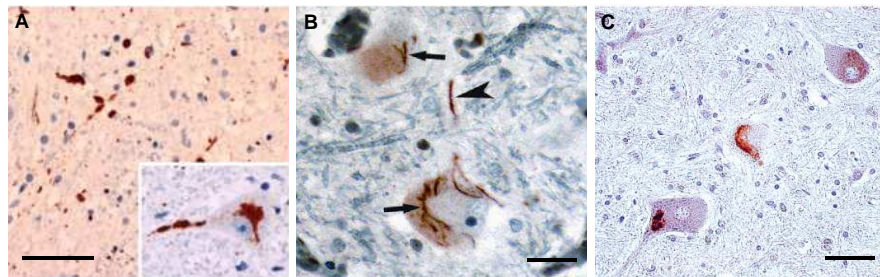


Figure 1.2: Pathological lesions commonly observed in ALS tissue. **(A)** Cytoplasmic *SOD1* inclusions are abundant in lower motor neurons. Scale bar represents 60 μ m. **(B)** *FUS* pathology in anterior horn in sporadic ALS tissue. Inclusions are predominantly neuronal and cytoplasmic (indicated by long arrows), with some dystrophic neurites (arrowheads). Scale bar represents 30 μ m **(C)** Neuronal cytoplasmic *TDP-43* inclusions in anterior horn. Scale bar represents 50 μ m Images adapted from Pokrishevsky et al. 2012; Deng et al. 2010; Sumi et al. 2009.

1.4.2 Tau

FTLD-tau is characterised by aggregation of the microtubule-associated protein tau, which forms insoluble cytoplasmic aggregates in affected neurons (Joachim et al. 1987; Love et al. 1988). Insoluble tau is hyperphosphorylated and ubiquitinated (Love et al. 1988; Murayama et al. 1990). Such aggregates are found in all cases of FTLD caused by *MAPT* mutations and many sporadic cases, overall accounting for ~40% of all FTLD. Tau pathology does not occur in pure ALS.

Tau pathology is a common feature of many neurodegenerative diseases, including AD, CBD and PSP. However, the typical characteristics of tau pathology vary between diseases, allowing for pathological classification in most cases. Unlike the paired helical filaments observed in AD, tau inclusions in bvFTD (historically referred to as Pick's disease when tau pathology is present) tend to form globose tangles known as Pick bodies (Figure 1.3), which are primarily composed of hyperphosphorylated 3R tau (de Silva et al. 2006; Taniguchi et al. 2004; Delacourte et al. 1998). However, 4R tau has also been observed in some cases, demonstrating heterogeneity in the pathological features of Pick's disease (Zhukareva et al. 2002). Conversely, straight filamentous structures are most typical of FTDP-17 (Figure 1.3), and these may contain either 3R or 4R tau, depending on which mutation is present (Mackenzie, Neumann, et al. 2011; de Silva et al. 2006; Taniguchi et al. 2004). Pick bodies may also be present in FTDP-17, albeit less frequently (de Silva et al. 2006; Pickering-Brown et al. 2000). Furthermore, aggregated tau may be cleaved, and certain truncated forms of tau have been found to be specific to particular tauopathies; a 35kDa C-terminal fragment (tau35) containing 4R tau was found to be present in brain homogenates from FTDP-17 cases, but not Pick's disease or healthy controls. Interestingly, the tau35 fragment was not present in all FTDP-17 cases, and appeared to be dependent on which *MAPT* mutation was present (Wray et al. 2008). In addition, immunoblotting of brain homogenate using antibodies against phosphorylated tau showed that a 55 and 64kDa tau doublet was present in Pick's disease, whereas additional or different sized bands were present in AD, PSP and CBD (Delacourte et al. 1996). Therefore the phosphorylation status of pathological tau may also directly relate to clinical phenotype.

The role of tau cleavage, hyperphosphorylation and aggregation in neurotoxicity is unclear. Tau phosphorylation plays an important role in regulation of tau binding to microtubules, since phosphorylated binding sites are unable to interact with microtubules. Therefore tau hyperphosphorylation may interfere with the protein's ability to stabilise microtubules, thus causing neurotoxicity through loss of tau function. However it is also possible that the presence of insoluble tau aggregates is toxic through an unknown gain of function mechanism. Whilst the mechanisms of toxicity are unclear, there is strong evidence to suggest tau pathology plays an important role in neurodegeneration, particularly since tau pathology is a common feature of many neurodegenerative diseases (Zhang et al. 2015; Hanger et al. 2014).

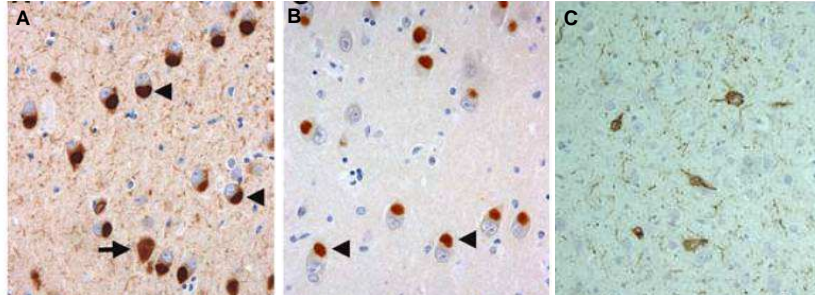


Figure 1.3: Neuronal tau pathology in FTLD. **(A)** Globose, cytoplasmic inclusions (known as Pick bodies) in FTLD brain labelled with an antibody against phosphorylated tau. Pick bodies indicated by arrowheads. A neurofibrillary tangle is also visible, indicated by long arrow. **(B)** Pick bodies detected by an antibody against 3R tau. **(C)** Filamentous tau tangles in cortical neurons in an FTDP-17 case caused by *MAPT* mutation. Images adapted from Cairns et al., 2007; de Silva et al., 2006.

1.4.3 TDP-43 and FUS

Inclusions containing the RNA-binding protein, TDP-43, are the most common type of pathological lesion in both FTLD and ALS (Arai et al. 2006; Neumann et al. 2006; Mackenzie et al. 2007), found in ~50% of FTLD cases and the majority of ALS cases (Rademakers et al. 2012). This includes all cases caused by mutations in *TARDBP*, *GRN* and *C9ORF72*, as well as many sporadic cases (Dobson-Stone et al. 2012; Tsai et al. 2012; Neumann et al. 2006; Davidson et al. 2007). FUS aggregations are also present in a smaller number of cases, including all cases where a *FUS* mutation is present (Neumann et al. 2009). Examples of FUS-positive inclusions in FTLD and ALS are shown in Figures 1.4 and 1.2B, respectively.

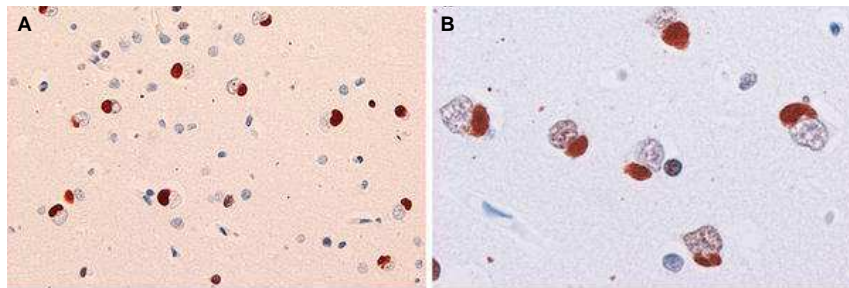


Figure 1.4: FUS pathology in cortical neurons of FTLD tissue **(A, B)**. Inclusions are rounded in structure and predominantly found in the cytoplasm, indicating a mislocalisation of FUS from the nucleus. Images adapted from Mackenzie et al., 2011.

Aggregated TDP-43 and FUS are hyperphosphorylated and ubiquitinated (Neumann et al. 2006; Neumann et al. 2009). In brain but not usually in spinal cord, TDP-43 is frequently truncated in the N-terminal region (Neumann et al. 2006; Davidson et al. 2007). Inclusions are predominantly intraneuronal and cytoplasmic in most cases, with some nuclear intraneuronal inclusions, glial inclusions and dystrophic neurites. The predominant type of inclusion varies between cases however, allowing for further categorisation of pathological phenotype, particularly within TDP-positive cases. In FTLD in particular, TDP-43 pathology has been subcategorised into four separate groups, based on the predominant type and location of each type of inclusion (summarised in Table 1.2 and Figure 1.5).

Table 1.2: Description of the major subtypes of FTLD-TDP pathology. Information obtained from Mackenzie et al. 2011.

Subtype	Description	Common Clinical Presentations	Associated Genes
A	Predominant neuronal cytoplasmic inclusions and short dystrophic neurites. Abundant pathology in cortical layer 2.	bvFTD, PNFA	<i>GRN</i> <i>C9orf72</i> (rarely)
B	Moderately abundant neuronal cytoplasmic inclusions, with few dystrophic neurites. Pathology present throughout all cortical layers.	bvFTD, ALS-FTLD	<i>C9orf72</i> <i>GRN</i> (rarely)
C	Predominant long dystrophic neurites, with few neuronal cytoplasmic inclusions. Abundant pathology in cortical layer 2.	bvFTD, SD	No known genetic associations
D	Predominant short dystrophic neurites and neuronal nuclear inclusions, with few neuronal cytoplasmic inclusions. Pathology present throughout all cortical layers.	Inclusion body myopathy with FTD	<i>VCP</i>

TDP-43 and FUS both belong to the FET family of DNA/RNA-binding proteins, and are able to interact with many proteins within this family as well as each other, suggesting a

degree of functional overlap (Lanson et al. 2011). However, the literature suggests that TDP-43 and FUS do not perform the majority of their functions as part of a protein complex, and that the two proteins regulate distinct sets of RNA-binding partners (Baloh 2012). They play an important role in various aspects of RNA processing, including regulation of pre-mRNA splicing, transcriptional control and mRNA stabilisation and transport (Mackenzie et al. 2010).

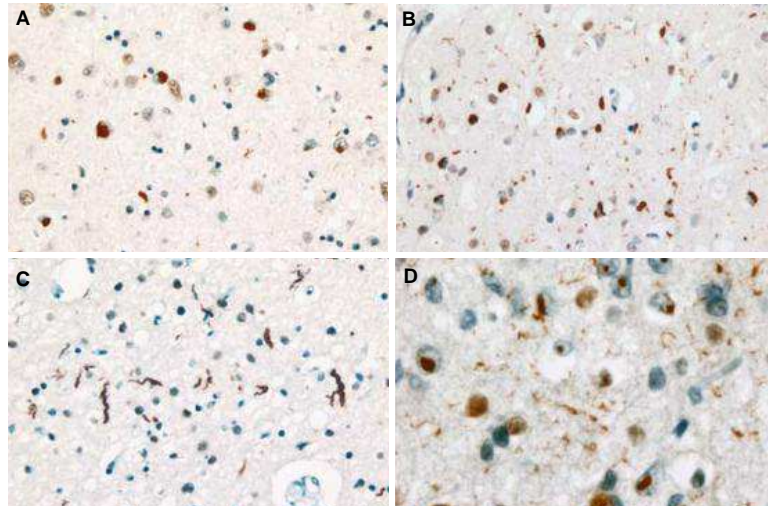


Figure 1.5: Different types of TDP-43 pathology observed in cortical FTLN tissue. **(A)** Type-A pathology predominantly consists of neuronal cytoplasmic inclusions and short dystrophic neurites. **(B)** Type-B pathology consists of moderately abundant neuronal cytoplasmic inclusions and few short dystrophic neurites. **(C)** Type-C pathology predominantly consists of long dystrophic neurites with few neuronal cytoplasmic inclusions. **(D)** Type-D pathology is predominantly nuclear, with abundant neuronal nuclear inclusions and short dystrophic neurons and few cytoplasmic inclusions. Figure adapted from Cairns et al. 2007.

The varied functions of TDP-43 and FUS require both proteins to be active in the cytoplasm and the nucleus, although they are predominantly found in the nucleus in healthy cells (Aoki et al. 2012; Ayala et al. 2008). However, cytoplasmic TDP-43 aggregations are a common feature of FTLN/ALS, and cytoplasmic FUS aggregations are also present in a smaller number of cases. Expression of ALS-linked mutant *TARDBP* and *FUS* in transgenic mice further demonstrates that such mutations promote translocation of the TDP-43 and FUS to the cytoplasm (Lanson et al. 2011; Murakami et al. 2012; Chen et al. 2011; Huang et al. 2011; Zhou et al. 2010). This mislocalisation of TDP-43/FUS is thought to be important in neurodegenerative disease. Indeed, the majority of *TARDBP* and *FUS* mutations identified in FTLN/ALS are observed in the nuclear localisation signal (NLS) domain, which is located in the C-terminal of both proteins (Kabashi et al. 2008; Sreedharan et al. 2008; Van Deerlin et al. 2008). The

mislocalisation of these proteins as a primary event in disease suggests a loss of RNA processing function in the nucleus may be important for neurodegeneration. Knockdown of TDP-43 or FUS homologues in a variety of species lends support to this theory. Loss of TDP-43 or FUS homologue function has been found to be toxic in mice (Hicks et al. 2000; Kraemer et al. 2010; Wu et al. 2012) and zebrafish (Kabashi et al. 2010). Of particular note, Wu *et al.* demonstrated that knockdown of *TARDBP* in mouse spinal cord caused an ALS-like phenotype, with mice exhibiting motor impairments, muscle weakness and atrophy, motor neuron degeneration and astrogliosis in the spinal cord. This provides strong evidence for a loss of function mechanism in neurodegeneration.

In addition, studies in transgenic animals and cell culture suggest that TDP-43 and FUS aggregates are not directly toxic to cells, which argues against a toxic gain of function mechanism (Baloh 2012). However, the literature is conflicting, and toxicity may precede formation of protein aggregates. Aberrant cytoplasmic activity of predominantly nuclear RNA-binding proteins could potentially disrupt or enhance signalling pathways key to cell survival or death, indirectly leading to neurotoxicity. Indeed, the RNA-binding partners of ALS-related mutant FUS have been reported to be significantly different to those interacting with wild-type FUS, and these changes correlated with relocation of mutant FUS to the cytoplasm (Hoell et al. 2011). Overexpression of human wild-type *TARDBP* and *FUS* have also been shown to cause neurotoxicity in rodent and invertebrate transgenic models (Xu et al. 2010; Tatom et al. 2009; Ash et al. 2010; Mitchell et al. 2013; Xia et al. 2012), further suggesting that toxic gain of TDP-43/FUS function may be relevant to disease. Taken together, the literature suggests that the molecular pathways involved in FTL/ALS pathogenesis are complex and may involve a combination of loss and gain of function mechanisms.

1.4.4 Clinicopathological and genetic correlations in FTL/ALS

Interestingly, the combination of clinical symptoms present in an individual tends to correlate well with specific patterns of neuronal atrophy observed in imaging studies. For example, shrinkage of the orbitofrontal cortex correlates well with disinhibition (Peters et al. 2006; Rosen et al. 2002), which can be explained by loss of the inhibitory function of the orbitofrontal cortex within the limbic system. Impaired recognition of emotions in others has been found to correlate with neuronal loss in the amygdala and insular cortex (Cerami et al. 2014). In addition, atrophy of the superior longitudinal fasciculus is a particular hallmark of bvFTD, and correlates with the severity of behavioural symptoms (Borroni et al. 2007). Atrophy of perisylvian cortical regions has been observed in PNFA

but not bvFTD or SD, and severe temporal atrophy is a feature of SD (Snowden et al. 2007). Furthermore, atrophy tends to be bilateral in SD, whereas PNFA predominantly affects the left hemisphere (Neary et al. 1998).

Correlations also exist between clinical presentation and genotype. *GRN* mutations have been linked to bvFTD and PNFA, rather than SD, whereas *MAPT* mutations are predominantly linked to FTDP-17 and may also cause impairments in semantic language (Snowden et al. 2006; Pickering-Brown et al. 2008). Co-morbid ALS is less frequent in FTLD-TDP cases with *GRN* mutations compared to those without. Conversely, Parkinsonism has been more frequently noted in *GRN*-linked FTLD-TDP cases than in other forms of FTLD-TDP (Chen-Plotkin et al. 2011). Rohrer and colleagues also described a case of a *GRN* mutation-carrying patient exhibiting bradykinesia, tremor and rigidity occurring 2 years after onset of behavioural symptoms typical of bvFTD (Rohrer et al. 2008). Interestingly, Parkinsonism is also a common feature of CBD, which may also be caused by *GRN* mutations (Spina et al. 2007; Kelley et al. 2009). This implies a particular link between extrapyramidal motor symptoms and *GRN*. Parkinsonism has also been reported in cases of FTLD-FUS (Snowden et al. 2011). Hypophonia, which is also a common feature of Parkinson's disease, was also observed in these cases (Snowden et al. 2011). Therefore Parkinsonism motor symptoms may occur in FTLD patients with a variety of genetic backgrounds.

Atypical clinical presentation reflecting parietal defects has also been reported in cases of *GRN*-linked FTLD; several patients within a large kindred presented with symptoms such as dyscalculia, limb apraxia and problems with visual processing in addition to a primary syndrome of bvFTD or PNFA (Rohrer et al. 2008; Pickering-Brown et al. 2008). These patients exhibited parietal atrophy in addition to the frontal and temporal neurodegeneration typical of FTLD (Rohrer et al. 2008). Finally, bvFTD with a strongly prominent presentation of obsessive compulsive symptoms has been reported in FTLD-FUS cases, and this was associated with neuronal loss in subcortical regions, particularly the caudate nucleus (Snowden et al. 2011). Therefore several links have been made between genetic mutations and clinical presentation. The clinical presentation of C9FTLD/ALS will be discussed in detail in Section 1.5.

1.5 The C9orf72 expansion

The recently discovered C9orf72 expansion is the most common known monogenetic cause of either FTLN or ALS (DeJesus-Hernandez et al. 2011; Renton et al. 2011) . Until its discovery in 2011, known mutations in genes linking both FTLN and ALS were rare. Despite the lack of knowledge regarding common genetic causes, links between these two apparently clinically distinct disorders were already evident. It is well documented that both diseases often present within one family, and in some cases even in one individual. This suggested there may be more common genetic factors contributing to both FTLN and ALS. Furthermore, the finding that neuronal and glial TDP-43 inclusions were a common hallmark of both FTLN and ALS implied some similarities in disease pathogenesis, and *TARDBP* mutations had also been linked to both FTLN and ALS. The identification of the C9orf72 expansion in large patient cohorts but not controls was a major breakthrough, definitively linking FTLN with ALS as part of a heterogeneous disease spectrum. This section will discuss the C9orf72 expansion further, including the background to its discovery.

1.5.1 Discovery of the C9orf72 expansion

Prior to the discovery of the C9orf72 expansion, multiple studies of individual families affected by both FTLN and ALS linked the diseases to a common region including the C9orf72 gene on chromosome 9p, although none of these studies identified a common mutation in this region (Boxer et al. 2011; Le Ber et al. 2009; Morita et al. 2006; Pearson et al. 2011; Vance et al. 2006; Valdmanis et al. 2007). Linkage analysis of a large Canadian FTLN/ALS family known as VSM-20 initially narrowed the region of interest to an 11Mb region containing 103 genes (Vance et al. 2006). This area overlapped with a 14Mb region identified through similar analysis of a large Scandinavian family in a separate study (Morita et al. 2006). Further analysis of the VSM-20 family further narrowed the linkage region to a 3.7Mb region of chromosome 9p, containing 10 candidate genes (Boxer et al. 2011). Several genome-wide association studies (GWAS) have also confirmed the significance of this locus in disease (Shatunov et al. 2010; Van Deerlin et al. 2010; van Es et al. 2009). Laaksovirta *et al.* reported that a locus on 9p21 was strongly associated with ALS in a Finnish cohort of 442 patients, including 93 known familial cases. Approximately 4% of the 442 ALS patients also presented with FTLN symptoms (Laaksovirta et al. 2010). The association signal was stronger when the familial cases were considered alone, with the 9p21 locus accounting for 47.3% of

cases. In apparently sporadic ALS cases where no known family members are affected, some association between the 9p21 locus and disease was observed, however this did not exceed the Bonferroni threshold for genome-wide significance. In a larger GWAS of 2323 sporadic ALS patients, locus 9p21.2 did exhibit genome-wide significance (van Es et al. 2009), and a meta-analysis study of 4312 ALS patients from 8 countries confirmed this (Shatunov et al. 2010). These data pointed to a monogenetic cause which is most commonly found in familial ALS, but may also occur in apparently sporadic cases. The finding also highlights the importance of obtaining large patient cohorts for genetic studies, since the smaller-scale studies were unable to detect a significant association between sporadic FTLD/ALS and the 9p21 locus, which became statistically significant in a larger study. A 232kb block of linkage disequilibrium was identified at 9p21 containing a 42-SNP haplotype, which was significantly linked to ALS (Laaksovirta et al. 2010; Mok et al. 2012). This haplotype was found to dramatically increase risk of disease, particularly in Northern European populations (Mok et al. 2012). More recently, a study of FTLD-ALS patients in the UK confirmed a common genetic risk factor at locus 9p21 for both diseases (Rollinson, et al. 2011).

These linkage and association studies directed further analysis of the 9p21 locus; however, despite sequencing all genes in this region, no mutations were identified. High-throughput next-generation sequencing was therefore employed, which led to the discovery of the C9orf72 mutation in FTLD/ALS kindreds by two independent groups (DeJesus-Hernandez et al. 2011; Renton et al. 2011). The mutation is a large GGGGCC repeat expansion in the first intron or promoter region of the gene, depending on alternative splicing of exon 1. DeJesus-Hernandez *et al.* initially discovered the expansion through analysis of the previously described VSM-20 family, which included multiple cases of bvFTD with and without ALS, as well as several cases of pure ALS with no associated dementia. Renton and colleagues simultaneously discovered the same expansion in a Welsh family known as GWENT#1, for which 9 probands were available. Like VSM-20, the GWENT#1 family included a mixture of bvFTD and ALS cases, including some cases of combined disease (Pearson et al. 2011). Following identification of the expansion in these families, both groups then screened larger FTLD/ALS cohorts, confirming the presence of the expansion in other families and sporadic cases (DeJesus-Hernandez et al. 2011; Renton et al. 2011). As suggested by previous linkage and association studies, patients carrying the expansion may present with FTLD, ALS or FTLD-ALS (DeJesus-Hernandez et al. 2011; Renton et al. 2011).

The number of repeats present in affected patients has been shown to vary considerably. In the original reports of expansion-carrying families, DeJesus-Hernandez *et al.* estimated repeat-length as between ~700-1600, whereas Renton and colleagues observed a minimum of ~250 repeats (DeJesus-Hernandez *et al.* 2011; Renton *et al.* 2011). Since these early studies, variability in expansion size has been reported in various C9FTLD/ALS cohorts internationally, with some studies reporting repeat-lengths as high as ~4500-5000 (Buchman *et al.* 2013; Dols-Icardo *et al.* 2014; van Blitterswijk *et al.* 2013; Nordin *et al.* 2015). Repeat-length remains highly variable even within expansion-carrying families; indeed, Dols-Icardo and colleagues reported dramatic differences in repeat number between monozygotic twins (Dols-Icardo *et al.* 2014). Interestingly, the expansion size has been shown to vary considerably not just between individuals, but also between different tissues within individuals. Notable differences in repeat-length have been observed between cell lines such as fibroblasts and lymphoblasts obtained from patients and *in vivo* samples from the same individuals (Beck *et al.* 2013; Hübers *et al.* 2014). Furthermore, repeat-length variation has been demonstrated between DNA samples extracted from blood, other peripheral tissues and brain or spinal cord (Fratta *et al.* 2015; Nordin *et al.* 2015). Therefore a high degree of instability is inherent within the C9orf72 repeat expansion. This raises important issues for genetic counselling of affected families, and also limits the use of blood samples for analysis of repeat-length in research studies.

Healthy controls may also carry the expansion, although with a much lower repeat-length. The threshold of repeat-length required for disease remains unclear. DeJesus-Hernandez *et al.* reported that up to 23 repeats may be present in healthy controls, whilst Renton *et al.* noted a maximum of 15 repeats in control samples (DeJesus-Hernandez *et al.* 2011; Renton *et al.* 2011). One study reported that short repeat-lengths of 20-22 segregated with disease in an FTLD/ALS family (Gómez-Tortosa *et al.* 2013), however, the majority of the literature suggests that much longer lengths are required to cause disease, with most cases exhibiting hundreds or thousands of repeats. Given the variability in expansion size between tissue types, it is possible that longer repeats were present in other tissues within the described family. It is also possible that the family may carry another, unidentified, mutation leading to FTLD/ALS. Finally, the presence of an additional mutation in the region flanking the expansion has been recently found to prevent detection of large expansions by certain assays (Rollinson *et al.* 2015; discussed in Section 1.5.2). Therefore this family may harbour a second, unknown mutation which does not cause disease itself, but which prevents accurate estimation of the expansion size in these cases. In a study of 7579 control cases, 11 were found to

possess the repeat expansion (Beck et al. 2013; 0.15%). However, it is possible that these cases may go on to develop FTL/ALS later in life. In addition, the expansion was defined as containing >32 repeats in this study. If the threshold required for disease is in the order of hundreds of repeats, this approach may capture non-pathogenic control expansions of so-called intermediate repeat-length. Since the expansion has not been found in control cases in other large cohorts (DeJesus-Hernandez et al. 2011; Renton et al. 2011), this report of expansions in only 0.15% of controls does not offer persuasive evidence to counter the pathogenic nature of the expansion.

The C9orf72 mutation is by far the most common genetic cause of FTL/ALS known to date (DeJesus-Hernandez et al. 2011; Renton et al. 2011). Therefore a detailed understanding of the gene, the protein it encodes and the downstream pathways leading to disease will be vital for development of novel therapeutics.

1.5.2 Epidemiology of C9FTL/ALS

Since its discovery, the presence of the C9orf72 expansion and its high frequency in FTL/ALS has been confirmed in a number of populations across the world, including European countries (García-Redondo et al. 2013; Ratti et al. 2012), Australia (Dobson-Stone et al. 2012), and the USA (Majounie et al. 2012). In a study of 2668 FTL patients from 15 Western European countries, the overall frequency of the expansion was 9.98%. This was increased to 18.52% when limited to familial cases only (van der Zee et al. 2013). Variability between countries was observed, with overall frequencies as high as 29.3% in Finland, 20.73% in Sweden and 25.5% in Spain, and as low as 4.82% in Germany (van der Zee et al. 2013). In a separate study, the expansion was present in 39.3% of white Europeans and Americans with familial ALS and 24.8% of white Europeans with familial FTL, whilst much lower frequencies were observed in black and Hispanic cohorts (Majounie et al. 2012).

Recently, a 10-base deletion adjacent to the C9orf72 expansion was discovered in two brothers in a Manchester FTL/ALS cohort (Rollinson et al. 2015). This deletion prevented detection of the expansion by repeat-primed PCR using a previously established method (Renton et al. 2011), presumably by creating secondary structures which prevented polymerase read-through into the expansion. An alternative PCR assay (DeJesus-Hernandez et al. 2011) was able to detect the expansion in these cases despite the presence of the deletion, highlighting the importance of genetic testing via multiple methods. Following re-analysis of patient samples in the Manchester cohort

using the second method, the total number of expansion-carriers was increased by ~25%, indicating that the 10-base deletion had previously hidden a significant proportion of carriers (Rollinson et al. 2015). Re-analysis of a London FTL/ALS cohort did not reveal an increase in the number of expansion-carriers, suggesting that the 10-base deletion may be confined to a local geographic area (Rollinson et al. 2015); however, other mutations may exist in cohorts worldwide which mask a proportion of expansion-carriers in a similar manner. Indeed, analysis of a large Flanders-Belgian FTL/ALS cohort identified several more small deletion and insertion mutations in the regions flanking the C9orf72 expansion, which were more frequent in expansion-carriers than controls (van der Zee et al. 2013). Therefore the discussed figures regarding expansion frequency in different populations should be considered minimal estimations, since the true frequency could be higher.

The low incidence of the C9orf72 expansion in non-Caucasian populations has led to suggestions that the mutation originally arose through genetic instability of a founder haplotype, leading to a spontaneous expansion in a single common ancestor of European descent (Smith et al. 2013; Mok et al. 2012). However, the variability of repeat-length between both individuals and tissues argues against this; a recent report of repeat-lengths in tissues of different developmental origins found that ~90 repeats were consistently found in tissues of non-neuronal origin, whereas longer repeats were present in the CNS (Fratta et al. 2015). This suggests that the longer expansion occurred after developmental differentiation of the central and peripheral nervous systems. The authors propose that inheritance of the expansion from a single founder would require repeat shrinkage events to have occurred independently in all non-CNS tissues in this case, each coincidentally resulting in a repeat-length of ~90. Whilst the origin of the repeat expansion remains unclear, the presence of a 42 SNP haplotype in the majority of cases suggests that this is a risk haplotype (Laaksovirta et al. 2010; Mok et al. 2012), proposed to cause inherent instability leading to expansion of an endogenous microsatellite region.

It is also worth noting that the presence of the C9orf72 expansion has been studied in other neurological diseases, and the only strong evidence for a causative role in disease is in FTL/ALS. The absence of a causative link between the expansion and disease has been noted in Parkinson's disease (Harms, Neumann, et al. 2013; Theuns et al. 2014), AD (Xi et al. 2012) and spinocerebellar ataxia (Fogel et al. 2012). One study identified expansions of >32 repeats in 11/905 AD cases (1.2%), however, many of these patients displayed behavioural symptoms reflective of bvFTD, and no pathological diagnosis was available in these cases (Beck et al. 2013). It is therefore possible that

some of these cases were misdiagnosed with AD. A screen of 102 patients with a clinical diagnosis of probable dementia with Lewy bodies (DLB) identified the expansion in 2 cases, however, this is too rare to imply a causative role of the expansion, particularly given the lack of confirmed DLB diagnosis at autopsy (Snowden, et al. 2012). A small number of studies have also identified C9orf72 expansions in very rare cases of bipolar disorder and schizophrenia (Galimberti et al. 2014). Again, the frequency is too low to imply causation, considering the common nature of both diseases which would be likely to occur at a low frequency in any given patient or control cohort. Furthermore, the expansion does not occur in all bipolar/schizophrenic patient cohorts, and does not segregate with disease (Fahey et al. 2014; Huey et al. 2013; Yoshino et al. 2015). However, C9FTLD/ALS is commonly associated with psychotic symptoms, therefore it is possible that in some cases expansion-carrying psychiatric patients observed in middle age are experiencing these symptoms prior to development of FTLD later in life. This may result in either initial misdiagnosis or co-morbidity of FTLD with other psychiatric conditions.

1.5.3 Clinical and Pathological Features of C9FTLD/ALS

Since the discovery of the C9orf72 expansion a great deal of work has been undertaken to characterise the clinical presentation and pathological hallmarks of C9FTLD/ALS, and several key differences have been found between expansion-carrying and non-carrying patients. Firstly, combined FTLD-ALS is more common in expansion-carriers than in non C9orf72-linked FTLD or ALS (non-C9FTLD/ALS) cohorts. It is worth noting however, that the expansion is not the only cause of combined disease, and FTLD-ALS has been noted in patients with a family history but no C9orf72 involvement. This suggests that one or more mutations linking FTLD and ALS may remain to be discovered.

A shortened life-span has been observed in ALS cases with the expansion compared to cases without (DeJesus-Hernandez et al. 2011; Cooper-Knock et al. 2012; Irwin et al. 2013). This appears specific to ALS cases, since there is no difference in disease duration in C9FTLD compared to non C9orf72-linked cases (Devenney et al. 2014). However, some studies have suggested that more rapidly progressing atrophy of the frontotemporal lobes and cognitive decline may be observed in C9FTLD (Irwin et al. 2013). C9ALS may present with either bulbar or limbic onset (Cooper-Knock et al. 2012), although there is some suggestion that bulbar onset is more common in expansion-carriers than in non-carriers. Brettschneider and colleagues reported bulbar onset in 57% of C9ALS cases in a Pennsylvania cohort (Brettschneider et al. 2012),

whereas only a third of ALS cases overall tend to present with bulbar onset. Bulbar onset and symptoms throughout disease progression was also common in a Manchester C9FTLD/ALS cohort. Increased incidence of bulbar symptoms could explain the shortened life-span in C9ALS, since bulbar onset disease tends to have a worse prognosis due to increased risk of lung infections. The majority of cases in the Manchester cohort later developed limbic symptoms with both upper and lower motor neuron involvement and widespread denervation measured by electromyography, in keeping with a typical ALS diagnosis and progression. However, 2 out of 9 C9FTLD/ALS patients did not develop limbic involvement, with clinical presentation more closely resembling progressive bulbar palsy. Interestingly, extrapyramidal pathology is more frequently noted in C9orf72-linked cases of FTLN (Cooper-Knock et al. 2012), and several studies have shown that Parkinsonism is also a relatively common symptom of C9FTLD/ALS (Boeve et al. 2012; Luigetti et al. 2013; O'Dowd et al. 2012; Snowden et al. 2012).

The most common form of C9orf72-linked dementia is bvFTD, although PNFA and SD have also been reported infrequently in expansion-carrying patients (Galimberti et al. 2013; Mahoney et al. 2012; Snowden et al. 2012). In a Manchester-based cohort of 398 FTLN patients with or without co-morbid ALS, the clinical presentation of 38 expansion-carriers was described in detail and compared to the remaining non-C9FTLD cases (Snowden et al. 2012). bvFTD with or without associated ALS was the predominant clinical diagnosis, with 3/38 cases categorised as PNFA and just 1 case of SD. The proportion of bvFTD cases was high compared to the non-C9FTLD cases. No significant differences were observed in gender, age-of-onset or disease duration between C9FTLD and non-C9FTLD cases. Imaging showed typical signs of bvFTD in the C9orf72-linked group, with variability in the relative degree of atrophy in frontal versus temporal lobes, and left versus right hemispheres. Behavioural symptoms typical of bvFTD were observed in most cases; an increased incidence of apathy and reduced motivation or initiative was particularly noted in C9FTLD compared to non-C9FTLD, with ~88% of cases affected. Disinhibition was also common in C9FTLD (~66%), presenting as loss of manners and social appropriateness and/or impulsive behaviour such as excessive spending or unusually large donations to charity. Emotional processing deficits including loss of empathy were observed, but no more frequently than in non-C9FTLD. Dietary changes were mostly present at the same frequency in both groups, with one specific exception; a reduced incidence of acquired sweet food preference was noted in C9FTLD compared to non-C9FTLD in the Manchester cohort (Snowden et al. 2012), however there is some conflicting literature on the subject. Kertesz et al. reported

both gluttony and acquired preference for sweet foods as common behavioural symptoms in C9FTLD (Kertesz et al. 2013).

Many behavioural characteristics of the Manchester cohort were replicated in a London-based cohort of 223 FTLN patients including 18 C9FTLD cases (Mahoney et al. 2012). This study confirmed that there were no differences in age-of-onset or disease duration in C9FTLD compared to non-C9FTLD. bvFTD with or without co-morbid ALS was again the most common clinical presentation in C9FTLD, with symptoms such as disinhibition and apathy frequently occurring (76% and 63% of cases respectively). Furthermore, Mahoney *et al.* also found an increased incidence of anxiety, with 52% of C9FTLD patients affected.

In addition to the range of behavioural symptoms observed in C9FTLD, cognitive impairments are also exhibited in most cases. Executive function was impaired in all C9FTLD cases in the London cohort (Mahoney et al. 2012), and almost all cases in the Manchester group (Snowden et al. 2012). Whilst memory and learning tend to be relatively well-preserved in FTLN compared to other forms of dementia, impairments in episodic memory were also reported in the majority of C9FTLD cases in both the Manchester and London cohorts (Mahoney et al. 2012; Snowden et al. 2012). Impaired object recognition was observed infrequently, and several patients in the London study presented with prosopagnosia. Finally, language impairments were common in C9FTLD. Whilst the language disorders PNFA and SD were only rarely diagnosed in expansion-carriers, many C9orf72-linked bvFTD cases also exhibited some level of language impairment, with aphasia and word-finding difficulties common (Mahoney et al. 2012; Snowden et al. 2012). Therefore a range of behavioural and cognitive symptoms may be observed in C9FTLD.

Perhaps the most noticeable difference between FTLN patients with and without the expansion is an increased incidence of psychotic symptoms such as hallucinations and delusions in expansion-carriers (Dobson-Stone et al. 2012; Galimberti et al. 2013; Snowden et al. 2012; Devenney et al. 2014). This was first highlighted in the Manchester cohort, where psychotic symptoms were significantly more common in C9FTLD and the presence of psychosis notably increased the likelihood of the expansion being present (Snowden et al. 2012). 38% of C9FTLD patients exhibited florid psychosis, with an additional 28% exhibiting less severe psychological symptoms such as paranoia and irrational thoughts. In comparison, less than 4% of non C9orf72-linked cases showed similar symptoms. This finding has been replicated internationally, with increased incidence of psychosis reported in Italian (Galimberti et al. 2013),

Australian (Devenney et al. 2014; Dobson-Stone et al. 2012) and Canadian cohorts (Kertesz et al. 2013). One report of a Finnish cohort of 20 C9FTLD cases and 50 cases where no expansion was present showed that whilst the incidence of psychosis was approximately doubled in expansion-carriers, the difference was not statistically significant (Kaivorinne et al. 2013). It is possible that this would become significant in a larger study resembling the Manchester cohort in size.

One of the most common forms of psychotic symptom exhibited by C9FTLD patients is delusional thinking. Delusions are often persecutory or paranoid in nature, with patients believing friends, family members, TV personalities or strangers wish to harm them. Some patients have been reported to call the police, carry weapons or barricade themselves in their homes for protection against perceived threats (Kertesz et al. 2013; Snowden et al. 2012). This type of behaviour can be particularly distressing for family members and carers; one patient was reported to have threatened to shoot his wife for example, since he believed she wished to harm him (Snowden et al. 2012). Alternatively, many patients present with somatoform delusions. Some examples of cases reported in the Manchester cohort include a patient who believed shards of plastic were embedded in his skull, another who reported pain where no medical cause could be found and without apparent distress, and one who believed he was infested by parasitic mites (Snowden et al. 2012). Another case report from Liverpool described a C9FTLD patient with delusional pregnancy (Larner 2013). The incidence of visual and/or auditory hallucinations is also significantly greater in C9FTLD (Kertesz et al. 2013), with patients frequently maintaining conversations or laughing in response to non-existent figures believed to be in the room. The combination of hallucinations and delusional thinking may also result in reports of paranormal, supernatural or religious phenomena from patients such as visions of the devil or disembodied spirits, belief that deceased friends or relatives are attempting to contact the patient or reportedly hearing the voice of God (Snowden et al. 2012). A common feature of C9FTLD was that patients would experience one single form of delusion, which may be persistent and lead to irrational or unusual behaviours. No mono-delusional psychosis was reported in non-C9FTLD in the Manchester cohort (Snowden et al. 2012).

In addition to the large number of C9FTLD patients exhibiting florid psychosis, many other patients in the Manchester cohort exhibited less severe psychological symptoms such as irrational or bizarre behaviours and thought patterns (Snowden et al. 2012). Several patients were observed to misuse objects, for example using the toilet brush to brush hair or teeth. One patient became obsessed with searching for important documents in the garden at night, despite lack of a rational reason to believe they would

be outdoors, and another was reported to regularly wrap faeces in kitchen foil and bake it in the oven. No cognitive impairments were observed which could account for these types of irrational behaviour, for example there were no object recognition deficits in patients misusing everyday objects (Snowden et al. 2012).

An increased incidence of complex repetitive behaviours was also observed in the Manchester cohort (Snowden et al. 2012). These may be closely linked to delusions if present. For example, the patient experiencing visions of the devil developed repetitive behaviours which she believed were protective, and the patient who reported shards of plastic lodged in his skull would repeatedly pluck at his scalp in an attempt to remove them. One patient experiencing persecutory delusions would call the police every evening at the same time demanding protection from perceived threats, demonstrating routine and regularity in his repeated behaviour (Snowden et al. 2012). Other obsessive or repetitive behaviours observed in C9FTLD cohorts were not necessarily related to delusions, for example one patient was reported to use a saline nose-spray so frequently that it caused nose bleeds (Kertesz et al. 2013), whilst another combed his hair so frequently and aggressively that his scalp would bleed (Snowden et al. 2012). Obsessive cleaning routines or collecting and hoarding of unnecessary items have also been frequently reported, with one patient even insisting on storage of spoiled food waste (Kertesz et al. 2013; Snowden et al. 2012).

Interestingly, the 10 base deletion mutation identified in a proportion of C9FTLD patients in the Manchester cohort (Rollinson et al. 2015) was found to be protective against psychosis; in an assessment of the clinical records of 50 C9FTLD/ALS patients, only 1/14 patients carrying the deletion exhibited psychotic symptoms compared to 30/36 patients without the deletion (Snowden et al. 2015). Pathological features were also studied in cases where autopsy tissue was available (5 deletion-carriers and 6 without the deletion). All 5 deletion-carriers exhibited type-A TDP-43 pathology, whereas 5/6 non deletion-carriers displayed type-B pathology, with only 1 patient exhibiting type-A lesions. The reasons for this difference in pathological phenotype remain unclear, however, this study suggests a causative link between TDP-43 pathology and clinical presentation in FTLN (Snowden et al. 2015).

Combining the literature discussed in this section, it is clear that there are several key clinical features which may distinguish C9FTLD/ALS from cases with no C9orf72 involvement, including a number of non-cognitive aspects. A number of pathological differences have also been observed between C9- and non-C9FTLD/ALS cases. MRI data has also shown substantial atrophy in the thalamus and cerebellum in expansion-

carrying FTLN patients compared to non-carriers (Irwin et al. 2013; Mahoney et al. 2012). This is particularly interesting since these areas are known to exhibit relatively high C9orf72 expression (Renton et al. 2011). Pathological protein aggregates are also frequently found in the cerebellum in C9FTLD/ALS. The C9ORF72 protein does not appear to aggregate in patient tissue (Brettschneider et al. 2012). However, numerous immunohistochemical studies have shown that type-B TDP-43 pathology (Table 1.2) is a typical feature of C9FTLD/ALS brain tissue. TDP-43 is therefore usually mislocalised to the cytoplasm, where it forms insoluble inclusions in neurons and sometimes glia (Figure 1.6; Bigio et al. 2013; Dobson-Stone et al. 2012; Cooper-Knock et al. 2012).

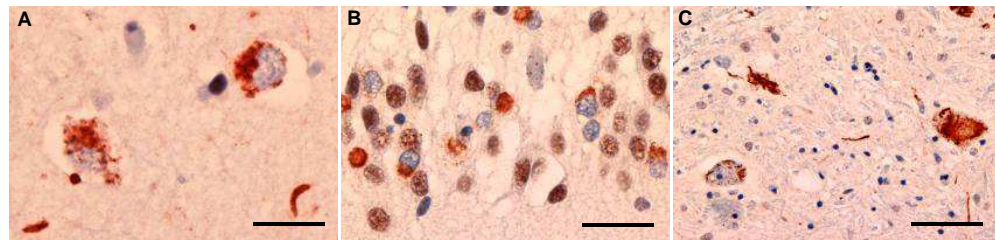


Figure 1.6: Type-B TDP-43 pathology in the VSM-20 C9FTLD/ALS family. Abundant neuronal cytoplasmic inclusions with few short dystrophic neurites in (A) cortex, (B) hippocampus and (C) spinal cord. Scale bars represent: (A) 15 μ m, (B) 30 μ m and (C) 100 μ m. Figure adapted from DeJesus-Hernandez et al. 2011.

Interestingly, a study of four C9ALS cases with no associated dementia found that unidentified star-shaped cytoplasmic protein aggregates were present in the cortex, basal ganglia and cerebellum (Troakes et al. 2012). These aggregates were immunopositive for ubiquitin and p62, which is a common protein marker of aggregates found in neurodegenerative disease. The majority of inclusions were negative for TDP-43, tau and FUS, although typical TDP-43 pathology was also observed in the spinal cord (Troakes et al. 2012). Cooper-Knock *et al.* also observed TDP-43-negative cytoplasmic inclusions in the frontal cortex and hippocampus of expansion-carrying ALS patients (Cooper-Knock et al. 2012). A number of additional studies confirmed the presence of unidentified TDP-43-negative inclusions in the cortex, cerebellum and hippocampal region, which often co-localised with ubiquitin and related proteins such as p62 and ubiquilin-2 (Al-Sarraj et al. 2011; Brettschneider et al. 2012). Ubiquilin-2 pathology in particular was distinctive in C9FTLD/ALS cases, with cytoplasmic ubiquilin-2-positive protein aggregates found in neurons of the cerebellar granular layer, hippocampal molecular layer and CA regions (Brettschneider et al. 2012).

An interesting phenomenon which has recently been described in other repeat expansion disorders is the ability of long, repetitive GC-rich RNA transcripts to form stable G-quadruplex structures, which initiate translation in non-coding regions of the gene in the absence of an ATG start-codon (Zu et al. 2011). This is described as repeat associated non-ATG (RAN) translation. RAN-translation was first discovered when Zu and colleagues removed the ATG start-codon from the 5' end of a minigene expressing a poly-glutamine expansion contained within the ataxin 8 gene (*ATXN8*), which is known to cause spinocerebellar ataxia 8 (SCA8). The start-codon was removed to allow expression of the expanded RNA sequence in the absence of poly-glutamine protein. Surprisingly, expression of the repeat expansion continued to produce the protein, despite removal of the only 5' start codon. Furthermore, Zu *et al.* (2011) noted that translation of the repeat sequence occurred bi-directionally and in all three frames in this *in vitro* model. Antibodies raised against the repetitive peptides arising from the *ATXN8* expansion were then generated, and immunohistochemical analysis of SCA8 patient tissue demonstrated that RAN-translation also occurred *in vivo* (Zu et al. 2011). Since this initial discovery, RAN-translation has been found to occur in several other repeat expansion disorders, including myotonic dystrophy, fragile X-associated tremor ataxia syndrome and Huntington's disease (Zu et al. 2011; Todd et al. 2013; Bañez-Coronel et al. 2015).

The discovery of RAN-translation suggests that despite its location in a non-coding region of the gene, the C9orf72 expansion could be translated to produce novel peptides, which were hypothesised to aggregate in these TDP-43-negative inclusions. In the absence of a start-codon, translation could begin in any frame and on either the forwards or reverse strand. Therefore there are 5 possible dipeptide repeats (DPRs) which could potentially arise from the GGGGCC expansion (Table 1.3).

Since the discovery of RAN-translation, RNA transcripts arising from the C9orf72 expansion have been shown to form stable G-quadruplex structures as predicted (Fratta et al. 2012). In 2013, two key immunohistochemical studies demonstrated that not only are the putative DPRs listed in Table 1.3 indeed present in patient tissue, but that they also form intraneuronal inclusions in affected brain regions (Ash et al. 2013; Mori et al. 2013). This has been replicated in multiple studies involving different patient cohorts (Mann et al. 2013; Mori et al. 2013; Mackenzie et al. 2013). This finding is particularly interesting since the DPRs are not mutated forms of normal proteins, but novel peptides which are not produced in any form in the healthy brain. The pathogenic implications of DPR generation and aggregation in disease are unclear.

Table 1.3: DPRs which could arise from RAN-translation of the repeat expansion in different frames and on the forwards and reverse strands.

FORWARDS			REVERSE		
DNA Sequence	Amino Acids	Code	DNA Sequence	Amino Acids	Code
GGG GCC	Glycine, Alanine	GA	GGC CCC	Glycine, Proline	GP
GGG CCG	Glycine, Proline	GP	GCC CCG	Alanine, Proline	AP
GGC CGG	Glycine, Arginine	GR	CCC CGG	Proline, Arginine	PR
GCC GGG	Glycine, Alanine	GA	CCC GGC	Proline, Glycine	GP
CCG GGG	Proline, Glycine	GP	CCG GCC	Proline, Alanine	AP
CGG GGC	Arginine, Glycine	GR	CGG CCC	Arginine, Proline	PR

In addition to TDP-43 pathology and DPR inclusions, mRNA transcripts arising from the G₄C₂ expansion have been found to form foci in patient tissue. RNA fluorescence *in situ* hybridisation (FISH) analysis has consistently shown that foci are primarily nuclear, and found in neurons of affected areas such as the frontal cortex, cerebellum, hippocampus and spinal cord (DeJesus-Hernandez et al. 2011; Mizielinska et al. 2013; Lagier-Tourenne et al. 2013). Both sense and antisense direction RNA foci are abundant in these areas (Mizielinska et al. 2013; Lagier-Tourenne et al. 2013; Lee et al. 2013). Interestingly, whilst DPR pathology appears to be limited to neurons, RNA foci were also present in microglia, astrocytes and oligodendrocytes in patient tissue, albeit less frequently than in neurons (Mizielinska et al. 2013). The respective roles of RNA foci, DPR inclusions and TDP-43 pathology in C9FTLD/ALS pathogenesis remain unclear.

1.5.4 The C9ORF72 protein

Very little is known about the C9orf72 gene and the protein it encodes (also named C9ORF72). There are 3 transcripts of the C9orf72 gene coding for 2 predicted isoforms of C9ORF72, both of which are produced in the human brain through alternative splicing. Isoform A is thought to be 481 amino acids in length, arising from transcripts 1 and 3 and encoded by exons 2-11. Isoform B is truncated at 222 amino acids and encoded by only exons 2-5 (DeJesus-Hernandez et al. 2011). Renton *et al.* found that C9orf72 mRNA was expressed in a variety of neuronal regions including the frontal cortex, hippocampus, thalamus and basal ganglia (Renton et al. 2011). A particularly high expression level was observed in the cerebellum, which is interesting when

combined with data from imaging studies showing atrophy of the cerebellum in C9FTLD/ALS (Irwin et al. 2013; Mahoney et al. 2012). An early immunofluorescence study suggested that C9ORF72 is a predominantly cytoplasmic protein in neurons (DeJesus-Hernandez et al. 2011) although the literature is somewhat conflicting, with a second study observing more prominent nuclear expression (Farg et al. 2014). This could be due to variation between different cell lines versus *ex vivo* tissue, or to non-specific binding of the relatively new C9ORF72 antibodies. More recently, antibodies individually targeting the two isoforms of C9ORF72 were generated; isoform A was found to be primarily cytoplasmic in cerebellar Purkinje neurons, whilst isoform B was localised in the nuclear membrane (Xiao et al. 2015). In addition to neurons, C9ORF72 is also expressed in many peripheral tissues, with particular abundance in lymphoblasts and testes (Suzuki et al. 2013; DeJesus-Hernandez et al. 2011).

The function of C9ORF72 is poorly understood. Computational studies have demonstrated homology in the secondary structure of C9ORF72 and a protein family known as differentially expressed in normal and neoplasia (DENN) related proteins (Levine et al. 2013; Zhang et al. 2012). DENN is a Rab GDP/GTP exchange factor (GEF) with an important role in endosomal trafficking and autophagy (Marat et al. 2011). It has been proposed that C9ORF72 may regulate membrane trafficking, either through regulation of DENN or DENN-related proteins, or by directly exhibiting RabGEF activity (Levine et al. 2013; Zhang et al. 2012). This is particularly interesting since other rare mutations in genes important for vesicular trafficking have been previously linked to FTLD/ALS (Table 1.1). A recent study demonstrated vesicular co-localisation C9ORF72 with several members of the Rab protein family in neuronal cell lines and ALS spinal cord tissue, supporting a role of C9ORF72 in endosomal trafficking and autophagy (Farg et al. 2014). A co-immunoprecipitation study has also identified proteins involved in nuclear import as binding partners of C9ORF72 *in vivo*, implying a role of C9ORF72 in nucleocytoplasmic transport (Xiao et al. 2015). However, further work is required to replicate these findings on a larger scale.

1.6 *C. elegans* as a model organism in neurodegenerative disease

1.6.1 Advantages of *C. elegans* as a model organism

Caenorhabditis elegans (*C. elegans*) is a free-living nematode worm, approximately 1.0-1.5mm long, which is an extremely powerful tool in the study of genetic disease. Practically, *C. elegans* is an ideal model organism. It is highly economical, convenient to store and maintain and has a high reproductive capability (250-300 offspring). In addition the *C. elegans* life-span is short, totalling ~25 days when stored at 20°C. Worms hatch and develop into young adults via 4 larval stages (L1-4) within just 3 days of egg-laying. This rapid life-cycle is particularly useful for study of neurodegeneration and other diseases of aging, since it allows monitoring of aging worms through to end of life in approximately 1 month. The different developmental stages of the *C. elegans* life-cycle are shown in Figure 1.7.

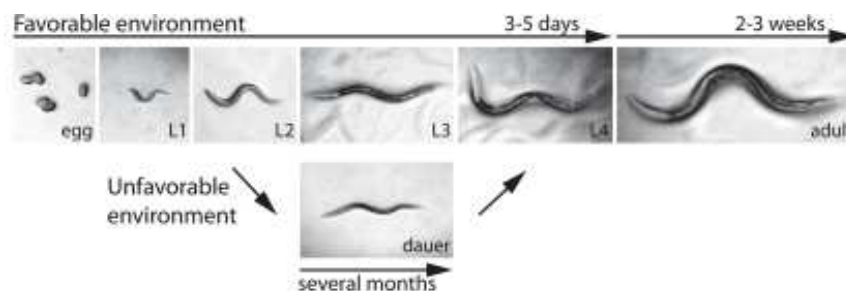


Figure 1.7: The developmental stages of the *C. elegans* life-cycle. Worms develop from egg to early adulthood in ~72h via 4 larval stages (L1-4). At the end of each larval stage the outer cuticle is moulted and replaced. Starved worms may also interrupt the normal life-cycle and either arrest at L1, or develop from L1 to dauer stage, where they are able to survive for long periods of time without food, and continue the normal life-cycle from stage L4 once access to nutrients is re-established. Worms begin to produce eggs from stage L4, and to lay eggs from early adulthood at L4+1 day. Images obtained from Fielenbach & Antebi 2008.

Worms are naturally found in the environment in many countries of moderate temperature, including the UK. They feed on microbes, particularly bacteria, and as such may be found in soil or on rotting fruit. However, *C. elegans* are unable to survive at temperatures much higher than 25°C, and have never been found to act as parasites in living plants or animals. In laboratory conditions, *C. elegans* can be conveniently stored on agar plates seeded with a bacterial lawn (typically *E. coli*) as a source of nutrients.

The majority of worms are self-fertilising hermaphrodites, and transfer of a single adult worm onto a seeded agar plate will result in a fully populated plate in a matter of days. Therefore worm stocks are inexpensive and easy to maintain. Interestingly, male worms may also occur at a very low frequency in the population (~0.1%), and this is due to a genetic error; worms with two copies of the X-chromosome are hermaphrodites, whereas loss of one copy of the chromosome results in a male worm. There is no Y-chromosome in the *C. elegans* genome. Male worms may be generated for genetic crosses by introducing a toxic insult such as heat-shock or ethanol treatment to L4 stage worms during egg production, and interrupting meiosis.

The *C. elegans* worm is a very simple invertebrate organism with approximately 1000 somatic cells and only 302 neurons. This simplicity is ideal for use as a model organism. *C. elegans* remains the only animal for which a complete nervous system wiring diagram is available (White et al. 1986). Despite their simplicity however, it is important to note that a considerable amount of similarity remains between *C. elegans* and humans, both genetically and physiologically. More than 60% of known human disease genes have a *C. elegans* homologue (Sleigh & Sattelle 2010). Furthermore, the major neurotransmitter systems and basic cell biology of neurons are well conserved. Indeed, many mutant phenotypes in *C. elegans* can be rescued by expression of the homologous human wild-type gene, demonstrating functional as well as structural and genetic conservation (Sleigh & Sattelle 2010). Therefore *C. elegans* may be considered a relevant model for human disease.

Finally, the worm is highly amenable to genetic manipulation, and, since it was the first organism to have its genome fully sequenced, a wealth of information is available to aid genetic studies in worm. Compared to many other *in vivo* transgenic models, the technology to create mutant strains is relatively simple and rapid (Ségalat 2007). As such, *C. elegans* is an incredibly useful tool in the study of genetic disease.

1.6.2 Previous uses of *C. elegans* in neurodegeneration research

C. elegans has previously been used in numerous studies to successfully model a variety of neurodegenerative and neuromuscular diseases (Grice et al. 2011; Joyce et al. 2011; Lanson & Pandey 2012; Ségalat 2007; Sleigh & Sattelle 2010). For example, treatment with 1-methyl-4-phenylpyridinium (MPP), a chemical which inhibits mitochondrial function and which is known to induce Parkinsonism in mammals including humans, resulted in loss of dopaminergic neurons, motor impairment and

reduced viability in *C. elegans* (Braungart et al. 2004). Exposure of MPP-treated worms to drugs approved for use in Parkinson's disease such as lisuride reduced the motor impairment, further highlighting the similarities between the *C. elegans* model and human disease (Braungart et al. 2004). Comparability to human disease is also demonstrated in *C. elegans* models of AD. Expression of human amyloid- β (A β) in *C. elegans* caused formation of A β -immunoreactive inclusions comparable to those present in human autopsy tissue (Link 1995). In addition, expression of human mutant or wild type (WT) tau protein produced an uncoordinated phenotype in *C. elegans*, as well as glycogen synthase-3 β (GSK-3 β) linked hyperphosphorylation of tau (Kraemer et al. 2003). This may be relevant to both AD, in which aberrant GSK-3 activity has been strongly implicated in tau hyperphosphorylation, and FTDP-17, which is caused by tau mutations. Indeed, the uncoordinated phenotype exhibited in transgenic *C. elegans* could be compared to the Parkinsonism symptoms present in FTDP-17 patients.

Numerous transgenic *C. elegans* models of ALS have also been created, including expression of various mutant forms of human *SOD1*. *SOD1* mutations produce a motor phenotype when expressed in *C. elegans* (Wang et al. 2009). Whilst these mutations did not cause neurodegeneration or reduced life span, the motor phenotype may still be sufficiently useful as a biomarker in pharmacological screening studies. *C. elegans* strains expressing WT or ALS-linked mutant TDP-43 have also been created (Ash et al. 2010; Liachko et al. 2010). Both the WT and mutant forms of human TDP-43 were toxic in *C. elegans*, although the mutant forms produced a more severe phenotype. Motor impairments, reduced viability and loss of motor neuron synapses have been reported in these models (Ash et al. 2010; Liachko et al. 2010). Ash *et al.* also demonstrated the functional orthology of TDP-43 and the *C. elegans* homologue, TDP-1, by showing that TDP-1 can support alternate splicing of human *CFTR* in a cell-based assay. A final example of how *C. elegans* has been utilised to model ALS is with *FUS* mutant worms. Murakami *et al.* showed that neuronal expression of mutant *FUS* caused age-dependent degeneration and paralysis in *C. elegans*, as well as mislocalisation of *FUS* protein to the cytoplasm where it formed insoluble aggregates (Murakami et al. 2012). These observations exhibit remarkable similarity to the pathological and clinical features of *FUS*-linked ALS, further demonstrating the validity of *C. elegans* as a model of human disease.

1.6.3 The *C. elegans* C9orf72 orthologue

There is an orthologue of C9orf72 in *C. elegans*, named *F18A1.6* or *alfa-1* (ALS/FTLD associated gene homologue). The gene is approximately 4Kb in length and located on the reverse strand of chromosome 2 (<http://www.wormbase.org>). It encodes ALFA-1. Like human C9orf72, there are 3 transcripts of *F18A1.6*, which produce 2 isoforms of the protein named a and b. These are 731 and 734 amino acids in length, respectively. No information is currently available regarding the distribution or function of the protein. Aside from an additional worm-specific region in *F18A1.6*, there is a high level of structural similarity between *C. elegans* *F18A1.6* and human C9orf72 (Figure 1.8). BLAST analysis using the online database, Ensembl, shows that the worm protein has 88% similarity to human C9ORF72 based on tertiary structure predictions, with 23% identical amino acid sequence. This implies that there may also be functional similarities between the proteins encoded by these genes. Therefore *C. elegans* may be an extremely useful genetic tool for investigation of C9orf72 function in health and disease. This project aimed to utilise *C. elegans* models, as well as human cell culture, in order to investigate the pathogenesis of C9FTLD/ALS.

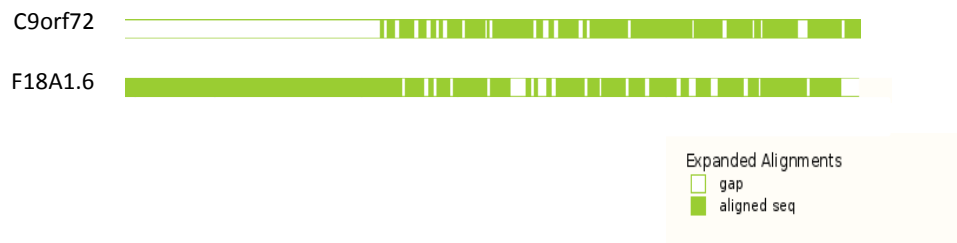


Figure 1.8: Comparison of the human C9orf72 and *C. elegans* F18A1.6 genes. Image adapted from the Ensembl C9orf72 GeneTree. Green areas depict regions of alignment between the 2 orthologues. The major difference between species is the presence of a large worm-specific region in F18A1.6, which is not present in human C9orf72.


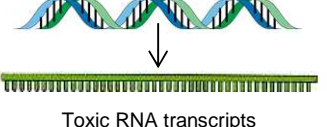
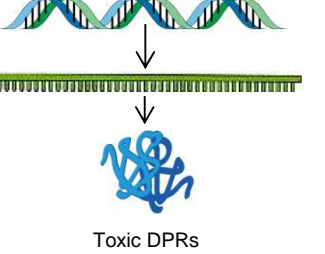
1.7 Research aims

There are three possible mechanisms through which the C9orf72 expansion could cause neurodegeneration and disease (Table 1.4). Firstly, the expansion may interfere with normal expression of the C9ORF72 protein, reducing the amount of protein produced. Since the normal function of C9ORF72 is poorly understood, it is unclear how likely loss of C9ORF72 function would be to cause neurotoxicity. The second mechanism through which the expansion may cause toxicity is through production of toxic RNA species. Long, repetitive mRNA transcripts arising from the expansion could theoretically be toxic in a number of ways, such as by sequestering other RNA species and proteins, or by physically interfering with cellular processes such as axonal transport. Finally, the expansion may be translated to produce toxic peptides, which may impact on neuronal survival via unknown mechanisms. This project aimed to investigate which of these hypotheses is most likely to give rise to the disorder. This was achieved through generation and analysis of a series of models, each designed to address one of these hypotheses individually.

The main aims of this project were therefore to:

1. Generate and validate a range of *in vitro* and *in vivo* models of C9FTLD/ALS, using *C. elegans* and human cell culture.
2. Utilise these models to investigate pathogenic mechanisms in C9FTLD/ALS

Table 1.4: Summary of the three possible mechanisms through which the C9orf72 expansion may cause neurotoxicity, and related research questions.

Haploinsufficiency	Toxic RNA	Toxic Peptides
<p>Does loss of <i>C9ORF72</i> function cause disease?</p>  <p>Reduced mRNA expression</p>	<p>Are long, repetitive RNA transcripts toxic?</p>  <p>Toxic RNA transcripts</p>	<p>Is the expansion translated to produce toxic peptides?</p>  <p>Toxic DPRs</p>

Chapter 2: Materials and methods

Standard protocols and procedures used throughout this project are detailed in this chapter. Where a protocol differed from the standard method described here, or more detail is required, this will be included in the methods section of the relevant experimental chapter (Chapters 3-5).

2.1 *C. elegans* strains and maintenance

C. elegans strains were obtained from the Caenorhabditis Genetics Centre (CGC) at the University of Minnesota. All worms were maintained at 20°C on 35mm plastic petri-dishes containing nematode growth media (NGM; agar 1.7% w/v; NaCl 50mM; peptone 0.25% w/v; CaCl₂ 1mM; cholesterol 5µg/ml; KH₂PO₄ 25mM; MgSO₄ 1mM). NGM plates were seeded with the *E. coli* strain, OP50 (obtained from the CGC), to generate a bacterial lawn and provide a food source for worms. Individual hermaphrodites were transferred onto fresh OP50-seeded NGM plates as required and allowed to reproduce by self-fertilisation.

2.2 Cell biology methods

2.2.1 Maintenance of cell culture

HeLa cells (ECACC #93021013) were maintained at 37°C with 5% CO₂ in Dulbecco's Modified Medium (DMEM; Sigma). DMEM was supplemented with 10% v/v fetal calf serum (Gibco), 2mM L-glutamine (Sigma), 100U/ml penicillin and 100µg/ml streptomycin (Sigma) prior to use. Cells were passaged as required, and fresh cells obtained once the passage number reached ~20-25. When passaging cells, media was first removed by aspiration and cells washed once in warm phosphate buffered saline (PBS; Sigma). Cells were detached from the base of the flask by incubation with 2ml trypsin solution (Sigma) at 37°C for ~5min, before dilution in 8ml DMEM to prevent excessive trypsin digestion. Cells were pelleted by centrifugation at 20,000RPM for 5min and re-suspended in 1ml of fresh DMEM. ~1ml of cell suspension was transferred to a T75 culture flask (Corning) containing ~20ml fresh DMEM. Culture flasks were then maintained at 37°C as standard.

2.2.2 Transient transfection

HeLa cells were seeded on plastic plates in DMEM at appropriate densities to achieve ~70% confluency following overnight incubation at 37°C. Transient transfections were performed using either JetPrime (PolyPlus), Lipofectamine 2000 (Life Technologies) or FuGene HD (FuGene; Promega). For JetPrime transfections, plasmid and reagent were diluted in JetPrime Transfection Buffer (PolyPlus) to achieve a final volume of 200µl per well of a 6-well plate. For transfection with FuGene or Lipofectamine, plasmids and transfection reagents were diluted in Opti-MEM serum-free media (Gibco), again to a total volume of 200µl per well. Plasmid concentration and ratio of DNA to reagent were varied for optimisation purposes until a final protocol was established (optimisation described in Chapter 4). Transfection reaction mixtures were mixed thoroughly, briefly centrifuged and incubated at room temperature for 15min before adding to cells by gently pipetting onto the media. Wells were mixed by swirling and plates returned to the 37°C incubator. The media was replaced ~4h after transfection to avoid excessive toxicity. Transfected cells were maintained at 37°C with 5% CO₂ as standard until harvest at 24, 48 or 72h post-transfection.

2.2.2 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Cells were washed in chilled PBS and harvested in RIPA buffer (Tris-HCl pH8 50mM; NaCl 150mM; NP-40 1% v/v; NaF 1mM; sodium deoxycholate 0.5% w/v; SDS 0.1% w/v) containing protease and phosphatase inhibitors (Cell Signalling). Sample preparation was performed on ice. Lysate was cleared of debris by centrifugation at 14,000rpm at 4°C for 15min and collection of the supernatant. Cleared lysate was mixed with 5x SDS sample buffer (tris-HCl pH6.8 125mM; SDS 2.5% w/v; bromophenol blue 0.025%; glycerol 12.5%, dithiothreitol 12.5%) and boiled for 5min to denature proteins.

Proteins were separated by electrophoresis on pre-cast 4-15% polyacrylamide gels (BioRad) in Tris running buffer (Tris 25mM; glycine 192mM; SDS 0.1% w/v). Precision Plus Protein Kaleidoscope protein ladder (BioRad) was used to identify band sizes. Gels were run at ~150V until the protein ladder reached the bottom of the gel, unless otherwise specified.

2.2.3 Immunoblotting

Proteins separated by SDS-PAGE were transferred onto 0.02% nitrocellulose membrane (Whatman) by semi-dry transfer in methanol transfer buffer (methanol 20%

v/v; tris 48mM; glycine 39mM; SDS 0.0373% w/v) at 15V for 1h. Membranes were blocked for a minimum of 30min in 4% bovine serum albumin (BSA) w/v in tris-buffered saline (Tris-Cl 50mM; NaCl150mM) with 1% tween-20 (TBS-T) v/v. Primary antibodies were diluted as appropriate in blocking solution, and membranes were incubated in antibody solution overnight at 4°C with agitation. Details of the antibodies used are provided in the relevant chapters.

Membranes were washed in TBS-T for 3 x 10-15min before incubation with secondary antibodies for ~1hr at room temperature with agitation. Horseradish-peroxidase conjugated secondary antibodies were diluted as appropriate in 5% milk w/v in TBS-T. Membranes were washed in TBS-T for 3 x 10min prior to imaging by chemiluminescence. Images were captured using the SynGene G:Box ChemiXX6 and associated software, GenSys.

2.2.4 Immunofluorescence

HeLa cells were seeded on 22mm glass coverslips pre-treated with 1M HCl, washed in distilled H₂O and stored in 70% ethanol. Coverslips were further washed in distilled water immediately prior to use. Cells were transfected as described above. At the appropriate time-point, cells were washed with warm PBS and fixed with 4% paraformaldehyde w/v in PBS at room temperature for 20min. Cells were washed 3 times in PBS and solubilised with 0.1% TritonX-100 v/v in PBS for 10min at room temperature. 3 further PBS washes were performed, followed by blocking in 1% fish skin gelatin v/v in PBS for ~1h. Primary and secondary antibodies were also diluted in PBS as appropriate. All antibody incubations were performed at room temperature for ~30min unless otherwise stated. Coverslips were placed face-down on 50µl drops of antibody solution and protected from light during the incubations. The primary and secondary antibodies used are detailed in the relevant experimental chapters.

NucBlue DAPI nuclear stain for fixed cells (Life Technologies) was also diluted in PBS (1 drop in 500µl), and incubations performed using the same method for 10min at room temperature. In some cases, DAPI was added to the secondary antibody to avoid excessive coverslip manipulations.

Each antibody or DAPI incubation was followed by 3 washes in PBS. Coverslips were air dried overnight and mounted on glass slides using ProLong Gold mounting reagent (Life Technologies). Coverslips were protected from light wherever practical throughout the whole procedure.

Images were captured at x60 magnification using an Olympus BX51 upright microscope with a 10xl 0.30 Plan Fln objective and a Coolsnap ES camera (Photometrics). Specific band pass filter sets for DAPI, FITC and Texas Red were used to prevent bleed through between channels. Image capture software used was MetaVue (Molecular Devices), and images were processed and analysed using ImageJ (<http://rsb.info.nih.gov/ij>).

2.3 Molecular biology methods

2.3.1 Agarose gel electrophoresis

DNA fragments were separated by electrophoresis on 0.8-2% agarose gels containing 0.5µg/ml ethidium bromide in tris-acetate-EDTA (TAE) buffer (tris 40mM; acetic acid 20mM; EDTA 1mM). DNA samples were mixed with 5x DNA loading buffer (Bioline) containing bromophenol blue prior to loading. Gels were run at ~90V for ~45-90min, or until adequate separation of the required bands was achieved. Bioline hyperladders 1kb+ and 100bp were used to identify band sizes. Bands were visualised using UV light and excised if required. Images were captured using a SynGene G:Box ChemiXX6 and associated software. Where required, DNA was extracted from excised bands using the QIAquick Gel Extraction kit (QIAGEN).

2.3.2 Cloning

All cloning was performed using Quick Ligase (New England Biolabs; NEB). Vectors were digested as appropriate (detailed in relevant experimental chapters) and digested fragments were purified by gel electrophoresis and extraction. DNA concentration of purified fragments was quantified by NanoDrop (ThermoFisher). Where sufficient quantities of purified DNA were obtained to do so, molar ratios of insert to vector were calculated, and appropriate volumes of each were added to maintain a ratio of ~3:1 in the ligation mix. ~1µl ligase was added per 10µl DNA, followed by addition of 2x Ligase Buffer (NEB). Ligations were performed for 5min at room temperature, followed by immediate transformation into competent *E. coli* to maintain transformation efficiency.

2.3.3 Transformation of plasmids into competent *E. coli*

Competent *E. coli* were thawed on ice before addition of 2µl plasmid or ligation product. Several different strains of *E. coli* were used throughout this project, including BL21-Gold DE3 (Agilent Technologies), XL10-Gold (Agilent Technologies) and 2 different strains which are optimised for transformation of repetitive or toxic plasmids (NEB #c2992 and #c3040). If required by the competent cells, β-mercaptoethanol was added prior to transformation in the quantities specified by the manufacturer's protocol, and incubated on ice for 10min. 2µl plasmid or ligation product was added to each tube of competent cells. Cells were mixed gently with a pipette tip and incubated on ice for ~20-40min before heat-shocking at 42°C for 30s and incubating on ice for a further 2min.

300µl SOC (super optimal broth with catabolite repression; NEB) media was added to each transformation reaction. Cells were grown at 37°C for ~1h with agitation, seeded on luria broth (LB) agar plates containing kanamycin at 30mg/ml and grown overnight at 37°C.

2.3.4 Sequencing

Sequencing of DNA extracted from *C. elegans* was performed by GATC Biotech (detailed in Chapter 3). All other sequencing was performed by the University of Manchester Sequencing Service.

Sequencing reactions were prepared with between 500-2000ng DNA, depending on plasmid size. 1µl BigDye Terminator (BDT; ThermoFisher), 4µl BDT buffer (ThermoFisher) and 1µl 3.2pMol primer were added per reaction, and diluted in sterile water for injection (Braun) to a total volume of 20µl. Reactions were run using the following thermocycler protocol: 96°C 1min, [96°C 10s, 50°C 5s, 60°C 4min]x25 cycles. DNA was precipitated by addition of 2µl 3M sodium acetate and 100µl 100% ethanol followed by incubation at room temperature for 15min protected from light, and pelleted at 14,000rpm for 15min. Supernatants were discarded immediately after centrifugation. Pellets were washed with 140µl 70% ethanol and centrifuged for a further 10min at 14,000rpm. Supernatants were again discarded and pellets thoroughly air dried. Precipitated DNA samples were then sequenced by the University of Manchester Sequencing Service.

2.3.5 Site-directed mutagenesis (SDM)

SDM was performed using the Agilent Technologies QuikChange Lightning kit. Primers were designed using the associated online tool provided by Agilent, and are detailed in relevant experimental chapters. 30ng plasmid was mixed with 1µl dNTPs, 1µl Quick Lightning Enzyme (Agilent), 1.5µl Quick SDM Reagent (Agilent) and 1µl each of 8.06pMol forward and reverse primer stocks. Reaction mixes were then diluted in 10x Quick Lightning buffer (Agilent) and water for injection (Braun) to a total volume of 50µl. Reactions were run using the following thermocycler protocol: 95°C 2m, [95°C 20s, 60°C 10s, 68°C 30s]x18, 68°C 5m. The extension time was altered according to plasmid length; an extension time of ~30s per 1kb was used. Following the thermocycler reaction, 2µl Dpn1 (Agilent) was added and incubated at 37°C for 5min before immediate transformation into competent *E. coli* according to the standard protocol

described above and using the competent cells provided in the QuikChange Lightning kit (Agilent).

2.3.6 RNA extraction and quantitative real-time PCR (qRT-PCR)

qRT-PCR was performed on *C. elegans* (Chapter 3) and cultured HeLa cells (Chapter 4), in order to compare mRNA expression levels of genes or plasmids, relative to a reference. The qRT-PCR data presented in Chapter 4 was obtained by Dr Sara Rollinson in the Pickering-Brown lab.

Worms or cells were lysed in Trizol (Life Technologies). RNA was extracted by addition of chloroform, precipitated in isopropanol and pelleted at 12,000rpm for 10mins at 4°C. Pellets were washed in 70% ethanol and dried before dissolving in 20µl RNase-free sterile water. RNA samples were stored at -80°C until required.

50ng random hexamers was added to 1µg RNA and made up to 13µl total volume in RNase-free water. Samples were denatured at 65°C for 5mins and immediately transferred to ice for 1min. cDNA synthesis reactions were prepared using the Roche Transcriptor (cDNA reverse transcriptase) kit (containing reverse transcriptase buffer, dNTPs and RNase inhibitors) at the concentrations recommended by the manufacturer. Samples were incubated at 25°C for 10mins, 55°C for 60mins and 85°C for 5mins to synthesise cDNA.

qRT-PCR reaction mixes were prepared with 100ng cDNA template, 10pM forward and 10pM reverse primer and 1x Power Sybr Green MasterMix (Applied Biosystems). 3-4 technical replicates were prepared for each reaction, as well as 3-4 control reactions with no cDNA template. qRT-PCR reactions were performed according to the manufacturer's protocol for Sybr Green. Primers were designed to overlap exon junctions in spliced mRNA where possible, preventing binding to any residual DNA. Relative amounts of mRNA were calculated compared to reference genes, to correct for differences in total RNA content between samples. Primers and reference genes are detailed in the relevant experimental chapters. All qRT-PCR data was analysed using the delta delta Ct method as described previously (Livak & Schmittgen 2001).

2.4 Data analysis and statistics

All statistical analysis was performed using GraphPad Prism. Swimming behaviour and egg-laying were quantified in transgenic and wild-type *C. elegans* strains in Chapter 3, and data sets were compared using unpaired t-tests. *C. elegans* life-span data was compared between strains using a non-parametric Mantel-Cox test. All behavioural tests were performed in triplicate as minimum. A minimum of 20 technical replicates were performed for thrashing and life-span assays, and 12 were performed for egg-laying assays. One independent replicate was considered as the mean value obtained from technical replicates from a single plate of worms. Independent replicates were performed on different days and using separate plates of worms. Full methods for all behavioural assays including *n* numbers for each test group are detailed in Section 3.2.

For immunofluorescence, a minimum of 10 technical replicates from a single flask of cells were considered before choosing representative images. Independent replicates were considered as images obtained from separate flasks of cells, prepared on different days. All image processing was performed using ImageJ (<http://rsb.info.nih.gov/ij>). An unpaired t-test was performed to compare the number of SMN-containing nuclear bodies in cells expressing DPRs or control, and this is detailed fully in Chapter 5. Full methods for immunofluorescence experiments including *n* numbers for each test group are detailed throughout Chapters 4 and 5.

Chapter 3: Does loss of *C9ORF72* function cause disease?

3.1 Introduction

As discussed in Section 1.7, there are three possible mechanisms through which the *C9orf72* expansion could cause FTL/ALS (Table 1.4):

1. Haploinsufficiency
2. Toxic gain of function due to repetitive RNA transcripts
3. Toxic gain of function from DPRs arising from translation of the expansion.

Evidence from analysis of patient samples exists to support all 3 mechanisms as being relevant to C9FTLD/ALS. Firstly, *C9orf72* mRNA expression is known to be reduced in expansion-carrying patients compared to non-C9FTLD/ALS cases, implying that loss of function may be important to disease pathogenesis (DeJesus-Hernandez et al. 2011). Since the function of *C9ORF72* protein is poorly understood, the consequences of reduced expression cannot be predicted. Early work suggests that *C9ORF72* may play a role in both endosomal trafficking and autophagy (Levine et al. 2013; Zhang et al. 2012). These are both important cellular processes which have been previously been linked to neurodegeneration, therefore it is feasible that disruption to *C9ORF72* function may be neurotoxic. However, further work is required to confirm that these functions may be attributed to *C9ORF72*.

Evidence supporting gain of function mechanisms of toxicity in C9FTLD/ALS can also be found from analysis of patient tissue. Nuclear RNA foci containing both sense and antisense repeat RNA are abundant in affected areas of patient brain, where they co-localise with a number of RNA-binding proteins (DeJesus-Hernandez et al. 2011; Mizielinska et al. 2013; Lagier-Tourenne et al. 2013). This may cause neurotoxicity through loss of function of the sequestered proteins, with the production of RNA as the primary toxic event. Repetitive RNA transcripts have also been implicated as toxic in a number of repeat expansion disorders such as Huntington's disease (reviewed by Fiszer & Krzyzosiak, 2013). Therefore the presence of RNA foci is likely to affect cellular function in some way. However, it is not clear whether this is detrimental to neuronal health, and functional studies are required to investigate this. The final possibility is that the DPRs produced through RAN-translation of the repeat are toxic. RNA transcripts arising from the expanded region have been shown to form G-quadruplex structures, triggering RAN-translation of 5 distinct dipeptides (Reddy et al. 2013). All 5 DPRs which

could potentially arise from the G₄C₂ repeat have been found to aggregate in patient tissue, and their impact on neuronal health has only recently begun to be studied (Mori et al. 2013; Ash et al. 2013). Repetitive expanded proteins have been well studied in polyglutamine repeat disorders, where they are proposed to interfere with a number of cellular processes. It has been suggested for example that large expanded proteins may overwhelm the ubiquitin-protease system (UPS), blocking the proteasome pore and preventing degradation of other, normal proteins (Li & Li 2011; Ortega & Lucas 2014). DPRs may be toxic in a number of ways; however, like RNA foci, their presence in the cell does not necessarily indicate initiation of neurotoxic events. Therefore functional studies are required to assess the impact of DPR on cellular processes and survival.

In order to distinguish between toxic effects which may be triggered through loss of C9orf72 function, production of expanded RNA and translation of DPRs, animal or cell-based models are required to address each of these hypotheses individually. This chapter aimed to investigate the hypothesis that loss of C9orf72 function causes neurotoxicity in C9FTLD/ALS. To address this question, a transgenic worm which is null for the *C. elegans* orthologue of C9orf72 was developed and assessed for behavioural phenotypes relevant to FTLD/ALS.

3.2 Materials and methods

3.2.1 – *C. elegans* strains and outcrossing

C. elegans strains N2 (wild-type) and VC40174 were obtained from the CGC. VC40174 is an *F18A1.6* null mutant with a single-point base substitution which introduces a stop codon at amino acid 54 or 57, depending on alternative splicing. This mutant strain was generated by random mutagenesis, and therefore it is likely that other mutations are present throughout the genome. VC40174 was therefore outcrossed through three generations, in order to remove mutations in genes besides *F18A1.6*. All worms were maintained on OP50-seeded NGM plates at 20°C as standard, and transferred to fresh plates as appropriate.

VC40174 males were generated by heat-shocking hermaphrodites at 30°C for 6-8hrs at developmental phase L4. This disrupts meiosis and causes a small proportion of progeny to lose one copy of the sex chromosome, resulting in male worms. Heat-shocked worms were allowed to grow and lay eggs on OP50-seeded NGM plates at 20°C. Progeny were screened and males obtained for crosses. Male VC40174 worms were crossed with hermaphrodites of a double-mutant strain, GE1708, also obtained from the CGC. GE1708 has two appropriately located mutations in *unc-4* and *dpy-2* on the opposite strand to the *F18A1.6* mutation in VC40174, and exhibits an easily identifiable motor and morphological phenotype which was used as a marker during outcrossing. The GE1708 mutations are recessive, and therefore worms exhibiting a phenotype were excluded as homozygous self-progeny of the GE1708 hermaphrodites. From each cross, males which were heterozygous for the *F18A1.8* mutation were obtained for use in the next cross with GE1708 hermaphrodites. After three crosses, individual hermaphrodites which were heterozygous for both the VC40174 and GE1708 mutations were transferred to separate OP50 NGM plates and allowed to self-fertilise. Progeny were sequenced and a plate containing homozygous *F18A1.6* mutants was selected. This new strain was named LM1006.

3.2.2 Sequencing of *F18A1.6* mutations

Individual worms were lysed with proteinase K (Roche) in ThermoPol Reaction Buffer (New England Biolabs). Worms in lysis reaction mix were incubated on ice for 30mins, then at 60°C for 1hr and 95°C for 15mins. PCR was performed with Quick Load Taq

MasterMix (New England Biolabs) using the thermocycler protocol recommended by the manufacturer and the primers listed below. Samples were separated by gel electrophoresis on 0.8% agarose in tris-acetate-EDTA (TAE) buffer, and the DNA was extracted using the QIAgen QIAquick Gel Extraction Kit. Sequencing was performed by GATC Biotech.

F18A1.6 PCR and sequencing primers:

Forward: TACCACATACGATGCGCTTAAA

Reverse: CTTTCATCTCTTAATGATTTTCCACTGC

3.2.3 RNA extraction and qRT-PCR

qRT-PCR was performed to compare the expression levels of *F18A1.6* mRNA in LM1006 worms and wild-type. Developmentally mixed populations of N2 or LM1006 worms were grown on OP50 NGM plates, collected in M9 buffer (22mM KH₂PO₄, 42mM Na₂HPO₄, 85.5mM NaCl, 1mM MgSO₄) and pelleted at 13,000rpm. Worms were lysed, RNA was extracted, cDNA generated and qRT-PCR performed according to the standard protocol in Section 2.3.6. *cdc-42* was used as a reference for qRT-PCR. Relative amounts of *F18A1.6* mRNA were calculated using *cdc-42* data to correct for differences in total RNA content between N2 and LM1006 samples, and compared by unpaired t-test using Microsoft Excel.

Primers used are as follows:

cdc-42 forwards ATGTCCGAGAAAAATGGGTG

cdc-42 reverse ATCCGTTGACACTGGTTTCTG

F18A1.6 forwards AGAACACAAAAATTCCACAGATAGAACT

F18A1.6 reverse GGCAAGATTATCGAAATAGAGCATT

3.2.4 Quantification of swimming behaviour in *C. elegans*

Developmentally synchronised plates of N2 and LM1006 hermaphrodites were obtained by allowing several adults to lay eggs over a short period of time (~1-2h), then removing the adults from the plate. Progeny were grown at 20°C on OP50 NGM plates until ready

for testing. Swimming speed was quantified by manual thrashing assay using young adults at developmental stage L4+1 day (n=5), and on aged worms at L4+4 days (n=4). Aging worms were removed from their progeny onto fresh plates as necessary. Individual worms were transferred to 50 μ L drops of M9 buffer on glass coverslips and allowed 1min recovery time. A 1min video was then captured of each worm using Microsoft Windows Movie Maker. The number of thrashes per 1min video clip was counted manually. ~20 worms per plate were assessed and the mean thrashing rate per minute recorded as one independent replicate. Replicate plates were assessed on different days. Thrashing rates were compared between N2 and LM1006 worms with an unpaired T-test using GraphPad Prism.

3.2.5 Egg-laying assays

Developmentally synchronised plates of N2 and LM1006 hermaphrodites were obtained as described above. Egg-laying behaviour was assessed at stage L4+1. Individual worms were transferred into wells containing 100 μ l 0.75mg/ml imipramine (Sigma-Aldrich) in M9 to stimulate egg-release and incubated at 20°C for 90mins. The number of eggs in each well was manually counted. ~12 worms per plate were assessed and the mean number of eggs recorded as an independent replicate. Replicate plates were assessed on different days. The number of eggs laid was compared between N2 (n=3) and LM1006 (n=4) worms with an unpaired T-test using GraphPad Prism.

3.2.6 Life-span assays

Developmentally synchronised plates of N2 and LM1006 worms were obtained as described above. At L4+1, 20 worms were transferred to a fresh OP50 NGM plate. The number of surviving worms on each plate was recorded daily. Worms were transferred to fresh plates as necessary, and excluded from analysis if killed during transfer. Plates of the same strain were set up on different days (n=5). The mean percentage survival per day was plotted, and the survival curves compared by Mantel-Cox test using GraphPad Prism.

3.3 Results

3.3.1 *F18A1.6* is knocked down in LM1006 worms

The relevant region of *F18A1.6* was sequenced in LM1006 worms to confirm the single-point mutation introducing a stop codon remained following three generations of outcrossing (Figure 3.1A). qRT-PCR was also used to demonstrate reduction in *F18A1.6* mRNA content, presumably due to nonsense-mediated decay in LM1006. A 74% reduction in *F18A1.6* mRNA was observed in LM1006 compared to N2 ($P < 0.001$; Figure 3.1B). Therefore the LM1006 strain can be considered an *F18A1.6* null mutant, and is a suitable model to test the haploinsufficiency hypothesis of C9FTLD/ALS in *C. elegans*.

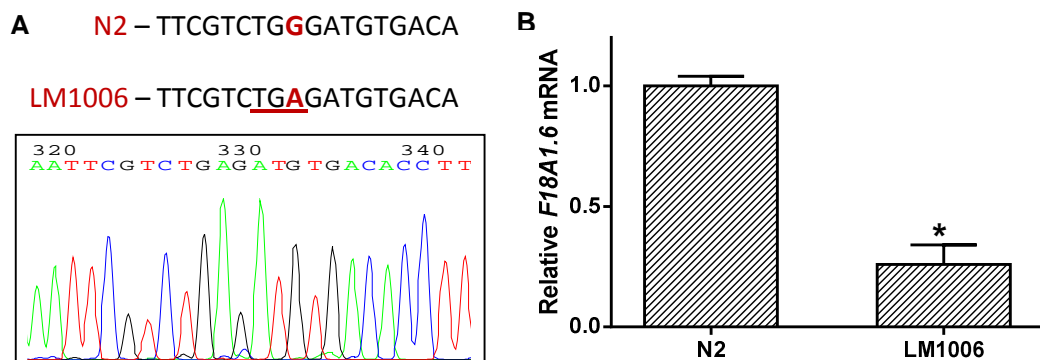


Figure 3.1: Characterisation of the *F18A1.6* null mutant *C. elegans* strain, LM1006. **(A)** Section of the DNA sequence demonstrating the presence of a stop codon at the expected locus in LM1006 compared to the wild-type gene (N2). The single-point base substitution is highlighted in red. The stop codon is underlined. Raw sequencing data is also displayed. **(B)** Relative expression of *F18A1.6* mRNA in LM1006 compared to N2, as determined by qRT-PCR ($P < 0.001$). Expression levels of *cdc-42* mRNA was used to control for variation in total RNA content. $n=1$. Error bars indicate standard deviation.

3.3.2 Loss of *F18A1.6* does not affect motility or life-span

Manual thrashing assays were performed to compare the swimming speed of LM1006 *F18A1.6* null worms and wild-type N2. No difference was observed in thrashing rate of young adults between strains (Figure 3.2A). Since FTLN/ALS is an age-related neurodegenerative disease, thrashing rate was also assessed in aged worms at developmental stage L4+4 days. Again, no motor phenotype was observed in *F18A1.6*

null worms (Figure 3.2B). In addition, the egg-laying behaviour of young adult LM1006 worms was assessed. Reductions in egg-laying are generally indicative of a motor deficit of the vulval muscles (Schafer 2005). However, no difference was observed in number of eggs laid by LM1006 and N2 adult hermaphrodites (Figure 3.2C). Since ALS is known to significantly impact survival in patients, the life-span of LM1006 worms was assessed. No difference was observed in percentage survival between *F18A1.6* null mutants and wild-type (Figure 3.2D). Therefore *F18A1.6* function is not required for survival or motility in adult worms.

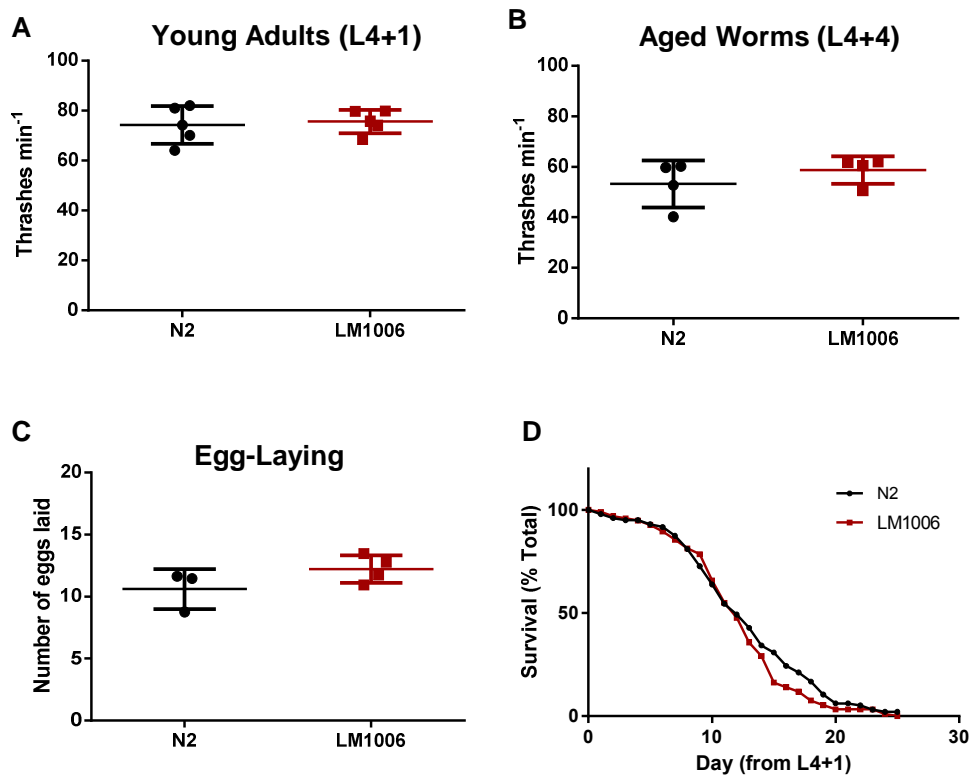


Figure 3.2: Assessment of motility and survival in *F18A1.6* null mutant worms. Manual thrashing assays were performed using young adults ($n=5$) at stage L4+1 (A) and aged worms ($n=4$) at L4+4 (B). Each data point represents the mean thrashing rate of a plate of worms. $P=0.74$ and 0.36 for young adults and aged worms, respectively. (C) Quantification of egg-laying behaviour in L4+1 *F18A1.6* null mutants ($n=4$) compared to wild-type ($n=3$). $P=0.18$. Bars indicate mean \pm standard deviation on all graphs. (D) Comparison of life-span in LM1006 and N2. Percentage survival was recorded from developmental stage L4+1 onwards. $n=5$. Each data point represents the mean percentage survival of 5 plates of worms. $P=0.81$, as determined by Mantel-Cox test.

3.4 Discussion

This chapter aims to address the hypothesis that C9orf72 haploinsufficiency causes neurodegeneration, using a *C. elegans* model. Our *F18A1.6* null worm was the first animal model of its kind, exhibiting knockout of a C9orf72 orthologue. The data presented in this section demonstrates that loss of *F18A1.6*, the *C. elegans* orthologue of C9orf72, does not affect motility, egg-laying behaviour or survival in *C. elegans*, suggesting that the C9orf72 expansion may be more likely to cause disease through toxic gain, rather than loss of function. This section will discuss these findings in the context of the published evidence for loss vs. gain of function mechanisms of neurotoxicity in C9FTLD/ALS.

3.4.1 Knockout or knockdown of C9orf72 in other animal models

Since the completion of this work, an independent group has reported a conflicting observation of a motor impairment and neurodegenerative phenotype in a second transgenic *C. elegans* strain which does not express *F18A1.6* (Therrien et al. 2013). This strain is homozygous for the ok3062 allele of *F18A1.6*, which contains a large deletion in a worm-specific region of the gene and an 8 amino acid insertion at the same locus. Therrien *et al.* used non-quantitative RT-PCR to show reduced expression of *F18A1.6* mRNA, suggesting that the deletion mutation was sufficient to cause knockout of the gene. As such, ok3062 and LM1006 can both be considered *F18A1.6* null mutants, and therefore would be expected to yield similar results in behavioural assays. However, Therrien and colleagues observed an age-dependent motility defect in ok3062 worms (Therrien et al. 2013), whereas we observed no such defects in the LM1006 null mutants.

The conflicting data appears to be at least partly due to analysis being performed using worms at different development stages. Our thrashing assays were performed in young adult worms at stage L4+1 day, and in aged worms at L4+4 days. The motor defects observed by Therrien *et al.* did not develop until approximately L4+8 days, and progressed with age. Our preliminary assays (data not shown) showed that the thrashing behaviour of wild-type worms became erratic from L4+5 days onwards, preventing the generation of reliable motility data at the stages considered key by Therrien *et al.* However, our life-span assay (Figure 3.2D) can be considered comparable to their motility assay; Therrien *et al.* subjectively considered worms paralysed if they did not respond to a tactile stimulus, thus generating discrete data with

only two possible outcomes: paralysed or not paralysed. Our life-span assay used the same method of tactile-stimulation to help determine whether worms were alive or dead, however, our results showed no reduction in motor response to tactile stimulation in null mutant worms compared to wild-type, from early adulthood until end-of-life. Surprisingly, Therrien *et al.* also performed a life-span assay on ok3062 worms using the same method of monitoring response to tactile stimulation, and observed no difference between null mutants and wild-type (Therrien et al. 2013). The contradiction between the paralysis and life-span assays within this report was not explained. It would therefore be useful to assess thrashing rate in LM1006 and ok3062 worms at the time-points investigated by Therrien *et al.*, in order to determine whether the motility phenotype exists but does not develop until L4+8 days, or whether the conflicting results are simply due to methodological differences.

Also in contrast to our findings from *C. elegans*, knockdown of the zebrafish C9orf72 orthologue (zC9orf72) has been reported to cause a motor phenotype, as well as abnormalities in axonal morphology (Ciura et al. 2013). These deficits were rescued by expression of human C9orf72. This suggests that loss of function mechanisms may contribute to disease pathogenesis in C9FTLD/ALS. However, the morphological and behavioural phenotypes were observed in zebrafish larvae at 48-96h post-fertilisation, whereas zebrafish do not reach adulthood until 90 days post-fertilisation. zC9orf72 expression was enriched in the central nervous system during early development, suggesting that the protein may play a developmental role in fish. It is therefore possible that this phenotype is a developmental, rather than degenerative, abnormality. Since FTLN/ALS is an age-related neurodegenerative disease, further studies in adult and aging fish could provide a more persuasive argument for a loss of function mechanism of C9orf72-linked neurodegeneration.

Knockdown or knockout of C9orf72 has now also been achieved in mice by at least 2 independent groups. Panda *et al.*, first reported generation of a knockout mouse using transcription activator-like effector nucleases (TALENs). However, no behavioural or pathological testing was reported, with the authors focusing on describing the TALEN technology used to generate the transgenic line (Panda et al. 2013). A more detailed report on these mice at a suitable age would be useful. Secondly, Lagier-Tourenne *et al.*, used antisense oligonucleotides to reduce C9orf72 mRNA expression in mice by ~70% in spinal cord and ~60% in brain (Lagier-Tourenne et al. 2013). Following 18 weeks of treatment, mice were evaluated for behavioural and pathological phenotypes. No inclusions were identified in brain tissue using ubiquitin, p62 and TDP-43 antibodies, and TDP-43 was not mislocalised to the cytoplasm. Rota-Rod tests demonstrated that

knockdown mice exhibited normal motor function, strength and coordination, whilst an open-field test did not reveal any abnormalities or anxiety-like behaviours. This study of a mammalian model strongly supports our findings that loss of C9orf72 function does not appear to cause an FTL/ALS-like phenotype. Due to the methodological concerns described above, the only two studies reporting abnormalities in knockdown animals do not provide persuasive evidence that loss of C9orf72 function causes a neurodegenerative phenotype. When taken together therefore, the evidence from knockdown animals including our own suggests that toxic gain of function mechanisms may be more important in disease pathogenesis.

3.4.2 Genetic evidence for loss of C9orf72 function in neurodegeneration

Perhaps the most persuasive evidence supporting a loss of function mechanism of toxicity in C9FTLD/ALS is the observation that expression of all 3 C9orf72 mRNA transcripts is reduced in expansion-carriers compared to controls (Belzil et al. 2013; Ciura et al. 2013; DeJesus-Hernandez et al. 2011; Gijssels et al. 2012; Waite et al. 2014). Particularly noticeable differences were observed in the frontal cortex and cerebellum, where C9orf72 is most abundantly expressed in normal brain (Waite et al. 2014). Expression levels were reduced by as much as 50%, which could convincingly contribute to a loss of function phenotype. Reduced C9orf72 mRNA levels have also been reported in iPSC-derived neuronal lines developed from C9FTLD/ALS patients (Almeida et al. 2013; Donnelly et al. 2013). However the literature is conflicting, with one study finding that transcript 2 mRNA was slightly increased in iPSC-derived neurons from expansion-carriers (Almeida et al. 2013), and another reporting no differences in expression of any transcripts (Sareen et al. 2013). These inconsistencies may be due to instability of the repeat during the reprogramming procedure introducing variability between iPSC-derived cell lines. Alternatively, the variability may simply be due to inherent variation in expression levels between individuals. A large degree of variation is generally observed in qRT-PCR analysis of C9orf72 expression levels in patients, however the sample size tends to be sufficiently large that an overall reduction in expression may be observed. The studies investigating expression levels in iPSC-derived neurons tend to be restricted to cell lines generated from much smaller groups of patients.

The literature is similarly conflicting with regards to levels of C9ORF72 protein expression. Waite *et al.* reported reduced C9ORF72 protein expression in the frontal cortex, but not cerebellum, of C9ALS patients compared to controls (Waite et al. 2014).

However, this reduction was not demonstrated by immunoblotting in all 14 C9ALS cases, introducing considerable variation into the data. Levels of mRNA expression also varied considerably between individuals within control and case groups, however, since the reduction in mRNA expression has been well replicated this finding is more persuasive. Gijselinck *et al.* immunoblotted tissue lysates with 3 separate antibodies to compare C9ORF72 protein levels between expansion-carriers and non-carriers and found no differences, commenting that antibody specificity was a problem in this study (Gijselinck *et al.* 2012). Whilst changes in mRNA expression are often a good indicator of changes in protein expression, this is not always the case, since compensatory alterations to translation efficiency or protein degradation may occur to maintain regulation of protein levels (Schwanhäusser *et al.* 2011; Taylor *et al.* 2013; Tian *et al.* 2004; Vogel *et al.* 2010). Therefore larger studies of C9FTLD/ALS cases compared to non-expansion carrying controls using well-characterised antibodies would be useful to confirm that C9ORF72 expression is also reduced at the protein level in C9FTLD/ALS.

If the expression of C9ORF72 is reduced at the protein level, this could potentially contribute to neurotoxicity through a loss of function mechanism. However, one case-study found that the clinical and pathological features of disease in a homozygous expansion-carrying FTD patient were no more severe than those of previously described heterozygous expansion-carriers (Fratta *et al.* 2013). If the expansion caused disease through C9orf72 haploinsufficiency, a more severe or significantly distinct disease phenotype would be expected in homozygous carriers compared to heterozygotes. Therefore this case-study supports our findings, suggesting a toxic gain of function mechanism is the most likely cause of C9orf72-linked neurodegeneration. It is important to note, however, that this was the first published report of a patient homozygous for the C9orf72 expansion. Analysis of large numbers of homozygous tissue samples would be required to draw reliable conclusions from this data, particularly given the high level of clinical and pathological variation observed in FTD/ALS.

The reduction in C9orf72 expression appears to be at least partially due to methylation of cytosine-phosphate-guanine (CpG) dinucleotide sites upstream of the expansion. A CpG island flanking the repeat region at the 5' end has been found to be hypermethylated in many cases of C9FTLD/ALS (Xi *et al.* 2013). The CpG island is located in the promoter region of both long transcripts, and is upstream of the promoter for the short transcript. The increase in methylation level correlated with shorter disease duration, suggesting the consequential reduction in mRNA expression may be toxic (Xi *et al.* 2013; Xi *et al.* 2014). However, this may simply be due to the previously reported observation that survival tends to be shorter in C9ALS compared to non-C9ALS. In

contrary to these findings, Russ *et al.* found that hypermethylation upstream of the C9orf72 microsatellite region was associated with longer disease duration, implying a protective role of reducing C9orf72 expression (Russ et al. 2015). An *in vitro* study also reported that increased CpG methylation in patient-derived lymphoblasts correlated with a reduced burden of RNA foci and DPR inclusions, suggesting that methylation may in fact be protective. Furthermore, demethylation of these cells increased their sensitivity to oxidative and autophagic stress. This demonstrates a protective function of promoter CpG methylation and subsequent reduction in C9orf72 mRNA in expansion-carrying cells. This observation supports our findings that loss of C9orf72 does not cause a neurodegenerative phenotype, and suggests that expression of expanded C9orf72 is in fact toxic, perhaps due to increased generation of toxic RNA and DPRs. More recently, the expansion region itself has also been found to be methylated in C9FTLD/ALS, although the function implications of this are unclear (Xi et al. 2015). It is possible that methylation of the expansion is a protective mechanism to prevent both transcription and formation of G-quadruplex structures, thus reducing production of DPRs.

It is also worth noting that if reduced C9orf72 expression was causative in disease, missense mutations or premature stop codons in C9orf72 would be likely to also segregate with FTL/ALS, whereas no pathogenic mutations besides the expansion have been identified to date (Harms et al. 2013). Overall therefore, the genetic evidence argues against the hypothesis that loss of function is the primary toxic event in C9FTLD/ALS pathogenesis, in agreement with the data presented in this chapter.

3.4.3 Animal models expressing G₄C₂ repeats

Several studies using animal models expressing G₄C₂ repeats support our findings. An early study published whilst the work presented in this chapter was undertaken used *Drosophila* and cell culture to provide further evidence for a toxic gain of function mechanism. Xu *et al.* expressed constructs containing either 30 or 3 G₄C₂ repeats in *Drosophila* and a mouse cell line (Xu et al. 2013). Expression of 30 repeats was sufficient to reduce viability in mammalian cells and induce neurodegeneration in the eye and motor neurons of *Drosophila*, whereas expression of 3 repeats had no effect. An age-related motor impairment was also observed in flies expressing 30, but not 3, repeats. In addition, Xu and colleagues demonstrated binding of a transcriptional activator, *purα*, to G₄C₂ repeats *in vivo* and *in vitro*. This, along with binding of several other RNA-binding proteins such as hnRNP-A1 and hnRNP-H, has since been demonstrated in iPSC-derived neurons obtained from C9FTLD/ALS patients or other

G₄C₂-expressing cell lines (Lee et al. 2013; Sareen et al. 2013). Interestingly, overexpression of pur- α rescued repeat-induced neurodegeneration in drosophila and improved viability in mammalian cells (Xu et al. 2013). This study not only supports a toxic gain of function mechanism of C9orf72-linked neurodegeneration, but also proposes a mechanism in which one or more DNA/RNA processing proteins are sequestered by expanded C9orf72 transcripts, resulting in aberrant gene expression. Sequestering of RNA species and proteins has been previously implicated in other repeat expansion disorders such as myotonic dystrophy (Osborne & Thornton 2006). Alterations in RNA processing have also been previously implicated in FTL/ALS, since the RNA processing proteins, TDP-43 or FUS, exhibit pathological changes in the majority of cases. Mutations in both *TARDBP* and *FUS* have been identified in FTL/ALS (Table 1.1), and the proteins typically translocate from the nucleus to the cytoplasm in affected neurons, where they form insoluble aggregates (Davidson et al. 2007; Neumann et al. 2006; Neumann et al. 2009). This has been proposed to cause neurotoxicity through loss of RNA processing function in the nucleus (Baloh 2012). The combination of *in vivo* invertebrate and mammalian *in vitro* models generated by Xu *et al.* (Xu et al. 2013) provides persuasive evidence for a role of toxic RNA and aberrant gene expression in C9orf72-linked neurodegeneration. However, one limitation of this study was the construct design, which contained two sequences of 15 repeats connected by a hexanucleotide motif that does not appear in the human disease gene. In addition, the expansion is known to be hundreds of repeats long in disease, whereas up to 30 repeats may be observed in healthy individuals (DeJesus-Hernandez et al. 2011; Renton et al. 2011).

More recently, constructs containing pure G₄C₂ sequences of 80 repeats have been expressed in mouse using an inducible system (Hukema et al. 2014). Expression was initiated at ~3-4 weeks of age and maintained for 12 weeks prior to analysis. Behavioural testing did not reveal any abnormalities at this time-point, although the authors did not include details of which types of behavioural assays were performed to draw this conclusion. No neuronal loss was observed, however immunohistochemical staining identified ubiquitin-positive, primarily nuclear inclusions in brain regions showing the highest expression levels (striatum and cuneate nucleus). These inclusions were negative for TDP-43 and poly-GA, and the main protein component(s) were not identified. The authors describe this transgenic mouse as an “RNA-only” model, since no DPR pathology was detected, however, no attempts to detect the remaining 4 DPRs were reported, and the unidentified nuclear inclusions appear to resemble the poly-PR inclusions found in patient tissue (Mann et al. 2013). Therefore further analysis of this

tissue using other DPR antibodies would be useful. It is possible that this model does not produce DPRs or cause a neurodegenerative phenotype, however the time-point is too early to assess that, since C9FTLD/ALS symptoms do not tend to develop until middle age in the human condition. Should re-labelling of the tissue identify the nuclear inclusions as poly-PR positive, this would be particularly interesting since it may indicate a time-course of disease pathogenesis, where some DPRs begin to aggregate before others. Further study of this mouse at later time-points will therefore be of value. Although the evidence is tentative at this stage, this study demonstrates that expression of repeats without disruption of C9orf72 function can at the very least cause a pathological phenotype, which may or may not develop into a neurodegenerative phenotype in aged mice.

A second transgenic mouse was recently created using an adeno-associated viral vector delivery system to express constructs containing G_4C_2 sequences at 2 or 66 in length with no ATG-start codon throughout the central nervous system of post-natal day 0 mice (Chew et al. 2015). By 6 months of age, $(G_4C_2)_{66}$ mice began to show subtle abnormalities in several behavioural tests. Open-field tests showed a reduced preference for exploration in open areas compared to $(G_4C_2)_2$ mice, implying an anxiety-like phenotype. $(G_4C_2)_{66}$ mice also exhibited reduced social behaviour when introduced to a novel animal compared to $(G_4C_2)_2$ mice, although this may be an extension of the anxiety-like phenotype as opposed to indication of abnormalities in social interaction. Reduced exploration of open spaces and areas containing novel animals was not due to motor dysfunction, since $(G_4C_2)_{66}$ mice displayed a hyperactivity phenotype in the open-field test, travelling further and at higher speed than $(G_4C_2)_2$ mice. Again, this may be due to increased anxiety, although the authors suggest hyperactivity may be linked to behavioural disinhibition in mice. Finally, $(G_4C_2)_{66}$ mice performed poorly in a Rota-Rod test assessing balance and coordination, but only from the second day of testing onwards. This suggests the deficit may be due to impairment in motor learning. It is particularly interesting to note that the most obvious behavioural defects at the early developmental time-point of 6 months are indicative of an anxiety-like phenotype, since anxiety is commonly reported as one of the earliest symptoms exhibited by C9FTLD patients, often beginning years before diagnosis (Mahoney et al. 2012).

Tissue was also harvested after 6 months and analysed for features of C9FTLD/ALS. Significant neuronal loss was reported in the cortex overall, motor cortex specifically and cerebellar purkinje layer of $(G_4C_2)_{66}$, but not $(G_4C_2)_2$, mice. *In situ* hybridisation detected nuclear RNA foci in ~40-54% of cells in the cortex, cerebellar purkinje layer and hippocampus, partially correlating with the pattern of neurodegeneration. Foci were less

abundantly present in the amygdala, thalamus, cerebellar granular and molecular layers and ventral horn of the spinal cord. All 3 sense direction DPRs were also present, demonstrating that 66 repeats is sufficient to initiate RAN-translation. Poly-GP was detected by ELISA in brain homogenates, and poly-GA, -GR and-GP were found to aggregate or diffusely-express in the cortex, hippocampus, cerebellum and spinal cord. DPR inclusions were frequently ubiquitin-positive, and most commonly found in neurons with some glial expression. Therefore the neuropathology of $(G_4C_2)_{66}$ mice closely resembled disease pathology found in patient tissue. The presence of antisense DPRs was not assessed. No RNA foci or DPR immunoreactivity was observed in $(G_4C_2)_2$ mice, as expected. Interestingly, phosphorylated TDP-43 was also found to aggregate in a small percentage of cortical and hippocampal cells (~7-8%). Of 250 cortical cells containing TDP-43 pathology from 5 mice in total, all cells also contained RNA foci, and ~75% of cells also contained a poly-GA inclusion. This is contrary to findings from patient tissue which have not found correlation between poly-GA and TDP-43 pathology (Gomez-Deza et al. 2015), and supports a link between the different types of C9FTLD/ALS pathology and neuronal loss. This study describes the first mammalian model to closely resemble the pathological features of disease, and presents strong evidence for toxic gain of function mechanisms of toxicity in C9FTLD/ALS.

3.4.5 Summary

This chapter has provided evidence that loss of C9orf72 function is unlikely to be the primary cause of neurodegeneration in C9FTLD/ALS. A major limitation of our findings from *C. elegans* is that the expression profile of *F18A1.6* is unknown, and therefore it is possible that *F18A1.6* is not expressed in neurons. If this is indeed the case, the absence of a neuronal phenotype in our null mutant strain is unsurprising, and conclusions relating to human disease mechanisms may more reliably be drawn from experiments using human cell culture or mammalian models. However, the majority of the literature supports our findings, suggesting that toxic gain of function mechanisms may be more important in disease pathogenesis. It is possible that loss of C9orf72 function may act as a disease modifier, despite the evidence that it is unlikely to be the primary toxic event in FTLD/ALS. Haploinsufficiency of a gene containing a pathogenic repeat expansion has been previously reported to act as a disease modifier rather than the primary cause of toxicity in myotonic dystrophy; myotonic dystrophy type 1 is caused by a CTG repeat expansion in the 3' untranslated region of *DMPK*, and *DMPK* knockout mice have been shown to exhibit only mild clinical and pathological features of myotonic

dystrophy, whilst RNA toxicity is thought to play a more important role in disease pathogenesis (Reddy et al. 1996; Jansen et al. 1996).

Since it is now well-established that the expansion is transcribed and translated in disease, a role of toxic RNA or DPRs in C9FTLD/ALS pathogenesis is likely and could lead to dysfunction of many cellular processes required for neuronal function and survival. The studies of animal and cellular models expressing G₄C₂ repeats described in this section support this, but are limited by 2 key weaknesses. Firstly, the number of repeats expressed is relatively short in each of these studies; the longest constructs described were 88 repeats in length, whereas patient samples tend to contain more than 1000 repeats. Development of *in vivo* and *in vitro* models expressing a greater number of repeats would therefore be useful and more closely reflect the human disease condition. Secondly, expression of pure G₄C₂ repeats prevents separation of any toxic phenotypes which may be caused by RNA and DPRs. Therefore models expressing repeat RNA or repeat peptides alone are required in order to fully understand their respective roles in disease pathogenesis. This project aimed to address these issues using models that express dipeptides, but not G₄C₂ RNA, at longer repeat lengths which are more relevant to human disease. The remainder of the project therefore focused on investigation of the dipeptides translated from the C9orf72 expansion, and the downstream consequences of this which may lead to neurodegeneration in C9FTLD/ALS.

Chapter 4: Generation and characterisation of cellular models of dipeptide pathology in C9FTLD/ALS

4.1 Introduction

The data provided in Chapter 3 suggests that loss of C9orf72 function is unlikely to be the primary toxic event in FTLD/ALS pathogenesis. This conclusion is strongly supported by a wealth of emerging literature suggesting that toxic gain of function mechanisms may be more relevant to C9FTLD/ALS, either due to toxic RNA transcripts or toxic dipeptides arising from the expansion region. However, the respective roles of RNA transcripts and DPRs in neurotoxicity remain unclear. The remainder of this project therefore aimed to develop and characterise a series of models of dipeptide pathology.

As mentioned in Section 1.5, 5 distinct DPRs have been shown to arise through RAN-translation of the C9orf72 expansion in patient tissue and form aggregates within neurons of affected areas: poly-GA, -AP, -GR, -PR and -GP (Ash et al. 2013; Mori et al. 2013). Therefore models expressing all 5 DPRs are required, to study their respective roles in disease pathogenesis. A major drawback of modelling dipeptide pathology by expressing G₄C₂ repeats in cells or by developing patient-derived iPSCs is the difficulty of separating out effects of DPRs and effects of repetitive RNA, which may also be toxic. In order to investigate their individual contribution to neurotoxicity, models separately encapsulating the features of either repetitive RNA expression or DPR pathology are required. To address this, we designed constructs to code for each DPR using alternative codons to those found in the expansion, in a less repetitive, randomised manner. Therefore we were able to express each DPR *in vitro*, in the absence of the repetitive GGGGCC mRNA transcripts arising from the expansion in disease. An additional advantage of this method is the ability to study the individual contribution of each DPR to toxicity, since RAN-translation does not occur in the absence of long, repetitive GC-rich transcripts. This chapter will describe the generation of a library of DPR constructs of different lengths, which were used to create and characterise a series of cellular models of dipeptide pathology.

4.2 Methods

4.2.1 Generation of short DPR constructs using alternative sequences

DNA sequences coding for each DPR in a less repetitive manner were designed using alternative codons to produce the same amino acids found in disease (Figure 4.1). Restriction sites for the type-2S restriction endonuclease enzymes, FokI and BbsI, were placed at the beginning and end of the repeat sequence, respectively. The advantage of type-2S enzymes is that the restriction site is several bases away from the cleavage site, which allowed precise cleavage at the beginning and end of the repetitive sequence, without additional cleavage within the repeat section. Restriction sites for the standard endonucleases, EcoRI and BamHI were also placed several bases before and after the repeat sequence respectively, in order to facilitate cloning and to incorporate a Kozac sequence to promote mammalian expression and a start codon.

Short constructs were initially manufactured by Eurofins MWG Operon for each of the five DPRs. Due to technical difficulties generating the constructs, the longest repeats produced by Eurofins MWG Operon were 36 repeats for poly-GA, -GR and -PR, and 22 repeats for poly-AP and -GP. 1 repeat was considered to be 6 bases, or 1 dipeptide unit. Once manufactured by Eurofins MWG Operon, the whole insert (including repeats, type-2S restriction sites, start codon and Kozac sequence) was isolated from the vector by double-digest with EcoRI and BamHI (New England BioLabs; NEB) according to the manufacturer's protocol. The insert and vector were then separated by agarose gel electrophoresis (protocol in Section 2.3.1) and extracted using the QIAquick Gel Extraction kit (QIAGEN). Simultaneously, the mammalian expression vector, pEGFP-N1, was also digested using the same restriction sites and the vector backbone isolated using the same gel purification method. The repeat inserts were then cloned into the pEGFP-N1 backbone by ligation with Quick Ligase (NEB) for 5min at room temperature, to create five different DPR constructs of 22 or 36 repeats in length. pEGFP-N1 is a convenient plasmid for efficient expression in mammalian cell culture, driven by the CMV promoter and including a GFP-tag downstream of the insert.

4.2.2 Cloning strategy to increase DPR repeat-length

The repeat length of each DPR construct was progressively increased in a step-wise manner by performing a series of identical cloning steps, each starting with the product

of the previous step and aiming to approximately double the current repeat-length. This process was repeated until a physiologically-relevant length of at least 1000 repeats was obtained (strategy outlined in Figure 4.2).

Figure 4.1: Design of alternative codon sequences to express DPRs in the absence of GGGGCC RNA. **(A)** Schematic illustrating how each DPR arises through RAN-translation of the G_4C_2 expansion in different frames and on both the forward and reverse strands. DNA sequence shown in black, amino acid codes forming dipeptides in colour below. **(B)** Alternative DNA sequences used to generate DPR codons in a less repetitive, randomised manner. DNA sequences shown in black, amino acid codes in colour below. Each sequence is preceded by an ATG start codon since expression is not driven by RAN-translation in the absence of the 100% GC-rich RNA found in disease. **(C)** Design of the transgene within pEGFP-N1 containing alternative repeat-sequences, for expression of DPRs in mammalian cell culture. The repeat region (red) contains 1 of the 5 alternative sequences shown in **B**. Cleavage sites of restriction enzymes used for cloning indicated by arrows.

4.2.2.1 Isolation of repeat fragments and vector backbones

pEGFP-N1::DPR constructs were digested with BamHI and EcoRI, and repeat-containing fragments obtained by gel purification as described above. This generated DNA fragments containing the entire insert, including the repeat. Repeat fragments were then isolated by digestion with Bbs1 and Fok1. The initial BamHI and EcoRI digest was required since Fok1 frequently exhibits non-specific binding and cleavage, and therefore digestion of the whole vector with Fok1 prevented clear identification of a band corresponding to the repeat fragment on an agarose gel. Over-digestion of the insert was also prevented by diluting Fok1 5-fold compared to the manufacturer's protocol and adding to the reaction only for the last 5min of an ~30min Bbs1 digestion, followed by immediate loading on an agarose gel. Constructs were also separately linearised with Bbs1 alone, and treated with 1µl calf intestinal phosphatase (NEB) to prevent re-ligation. Both the repeat fragments and linearised vectors were gel purified prior to cloning.

4.2.2.2 Ligations

Repeat fragments were cloned into the Bbs1-cleaved vectors using Quick Ligase (NEB) to generate constructs with doubled repeat numbers. In the early stages of the DPR cloning, an initial ligation step was undertaken where Quick Ligase and ligase buffer were added directly to the isolated repeats and incubated at room temperature for 5min, prior to addition of ~20ng vector and additional ligase and buffer for a further 5min reaction. This initial reaction was performed to encourage ligation of 2 or more repeat fragments before cloning back into the vector, in order to increase repeat-length as quickly as possible. However, while this step was occasionally successful in increasing repeat number by more than 2-fold, this occurred very infrequently. The initial ligation step was eventually considered to reduce the efficiency of either the cloning or subsequent transformation, and removed from the protocol.

4.2.2.3 Transformation, growth and screening of cloning products

Cloning products were transformed into competent *E. coli* immediately following the ligation reaction, according to the standard method described in Section 2.3.3. Several strains of *E. coli* were trialled until a 5-alpha F' β strain (NEB #c2992) was identified to best tolerate the repeats and used as standard for the remainder of the cloning. Transformed cells were plated on LB-agar containing 30mg/ml kanamycin and grown at

37°C overnight. If colonies were present, these were used to inoculate ~3-4ml overnight cultures in LB containing 30mg/ml kanamycin, which were grown at 37°C overnight in a shaking incubator. Plasmids were then isolated from 1.5ml of culture using the QIAprep Spin Miniprep kit (QIAGEN), and the remaining culture stored at 4°C. Plasmids were screened for successful increase in repeat-length by digestion with BamHI and EcoRI, and gel electrophoresis in order to determine the size of the insert. Any plasmids containing longer repeats were identified and the relevant bacterial cultures used to inoculate ~50ml overnight cultures in LB containing 30mg/ml kanamycin, in order to isolate greater yields of plasmid using the QIAGEN HiSpeed Plasmid Midi Kit. All midi preps were also screened for the correct repeat size by digest with EcoRI and BamHI, since longer repeats became unstable within bacterial cultures and frequently reduced in size.

4.2.2.4 Sequencing

Sequencing was performed on both the forwards and reverse strands by the University of Manchester Sequencing Service (detailed protocol in Section 2.3.4), using commercially available primers flanking the multiple cloning site in the pEGFP-N1 vector:

Forward: TAACAACCTCCGCCCCATT

Reverse: GTCCAGCTCGACCAGGATGGG

For the majority of constructs above ~100 repeats, sequencing of the entire repeat sequence was not technically possible. All constructs were therefore validated in 3 ways: (i) by sequencing of both ends of the repeat, (ii) size-screening by restriction digest and gel electrophoresis as described above and (iii) expression *in vitro* to ensure GFP expression was observed, ensuring that no frame-shift mutations had been introduced within the repeat.

4.2.3 Generation of DPR constructs with no GFP-tag

“Untagged” versions of the DPR vectors were generated using SDM to introduce a stop codon between the repeat sequence and GFP gene, preventing translation of the GFP-tag. SDM was performed using the Agilent Technologies QuikChange Lightning kit and the standard protocol described in Section 2.3.5. A stop codon was introduced by

changing a CAA, coding for glutamine, to TAA. The single-point substitution was identical for each DPR construct, however, the primers incorporated part of the repeat sequence, and therefore separate primers were required for each DPR (listed below).

Primers:

Poly-GA forward: CCG GTG CTG TCT TCC AAC GGG ATC CAC CG

Poly-GA reverse: CGG TGG ATC CCG TTG CAA GAC AGC ACC GG

Poly-AP forward: GGC CCC TGT CTT CCA ACG GGA TCC ACC

Poly-AP reverse: GGT CGA TCC CGT TGG AAG ACA GGG GCC

Poly-PR forward: CCC GCG AGT CTT CCA ACG GCA TCC ACC

Poly-PR reverse: GGT GGA TCC CGT TGG AAG ACT CGC GGG

Poly-GR forward: GCG GCA GAG TCT TCC AAC GGG ATC CAC CG

Poly-GR reverse: CGG TGG ATC CCG TTG GAA GAC TCT GCC GC

Poly-GP forward: GGG GCC TGT CTT CCA ACG GGA TCC ACC

Poly-GP reverse: GGT GGA TCC CGT TAG AAG ACA GGC CCC

The initial cloning steps to increase repeat-length were performed in parallel in both the tagged and untagged pEGFP-N1 vectors, however, due to the technical difficulty of the cloning this was considered to be an inefficient method and elongation of the untagged vectors were discontinued early in the cloning process.

4.2.4 Transient transfection of DPR constructs in cell culture

HeLa cells were transfected with DPR constructs or an empty pEGFP-N1 vector as a GFP-only control. Transfections were performed using JetPrime (PolyPlus), Lipofectamine 2000 (Life Technologies) or FuGene HD (Promega) for initial optimisation of the protocol. The amounts of DNA and transfection reagent were also varied for optimisation purposes.

Following optimisation, the remainder of transfections were performed with FuGene HD. For a single well of a 6-well plate, 800ng of plasmid and 7.2µl FuGene HD were diluted

in Opti-MEM (Life Technologies) to produce a total volume of 200µl, and incubated at room temperature for 15min before addition to a well containing 2ml DMEM. Amounts of plasmid, reagent and Opti-MEM were adjusted as appropriate for transfection in different sized plates. Plates were mixed by gentle swirling and incubated at 37°C. To prevent excessive toxicity of DPR constructs, the media was replaced after ~4h.

4.2.5 qRT-PCR

qRT-PCR was performed in order to quantify mRNA expression of DPRs for optimisation of the transfection protocol. Lower expression was generally observed following transfection with arginine-rich, compared to alanine-rich, DPRs. Optimisation was therefore performed using one of each (poly-GA and poly-GR), as well as pEGFP-N1 as a GFP-only control. HeLa cells were seeded in 12-well plates and transfected as described above.

Lysate was harvested 24h post-transfection. RNA was extracted, cDNA produced and qRT-PCR performed as described in Section 2.3.6. Primers (listed below) were designed to detect GFP mRNA expression, since primers within the repeat sequences would bind to multiple sites. There are no introns in the GFP sequence in the pEGFP-N1 vector, and therefore primers could not be designed to overlap exon junctions. To avoid amplification of DNA as well as RNA, lysates were therefore DNase-treated (Promega) prior to cDNA production. Primers against the *GAPDH* gene were used as a reference.

Primers:

GFP forward: GTCCGCCCTGAGCAAAGA

GFP reverse: TCCAGCAGGACCATGTGATC

GAPDH forward: CCTGTTCGACAGTCAGCCG

GAPDH reverse: CGACCAAATCCGTTGACTCC

4.2.6 SDS-PAGE and Immunoblotting

Lysate was collected in RIPA buffer and proteins separated by SDS-PAGE and blotted as described in Section 2.2.2-3. DPRs were detected using a rabbit anti-GFP antibody (Santa Cruz) in 5% milk in TBS-T w/v. For longer dipeptides (GA₁₀₂₀ and GR₁₁₃₆), protein

expression was too low to be detected using the standard protocol. Lysate from a minimum of 1 whole 6-well plate was therefore collected and protein concentrated using 100kDa filter columns (Amicon). The entire sample was then mixed with 5x SDS buffer, boiled for 5min and loaded into a single well of an acrylamide gel for SDS-PAGE. Pre-cast 4-15% acrylamide gradient gels (BioRad) were used to facilitate movement of large proteins from the well and into the gel. Gels were run at 120V for ~3h, to allow clear separation of high molecular-weight bands. Proteins were transferred onto 0.2% nitrocellulose by wet transfer at 20V overnight at 4°C. The wells were not removed prior to transfer, to avoid removal of large proteins which may not have moved into the gel during SDS-PAGE. Membranes were immunoblotted and imaged according to the standard protocol (Section 2.2.3) using the rabbit anti-GFP antibody (Santa Cruz).

Dot blots were also performed for GR₁₁₃₆. Lysates were collected from whole 6-well plates and concentrated as described above, mixed with 5x SDS buffer and boiled for 5min. Lysate expressing untransfected control lysates was also collected and prepared using the same method. 5µl of each sample was pipetted directly onto 0.2% nitrocellulose, allowed to dry at room temperature and immunoblotted as above using the anti-GFP antibody.

4.2.7 Immunofluorescence

HeLa cells were seeded on HCl-treated glass coverslips in 6-well plates and allowed to adhere overnight. Transfections were performed the following day according to the standard protocol described above. Cells were fixed, solubilised and stained at various time-points between 24-72h, according to the protocol in Section 2.2.4. Cells expressing GFP-tagged DPRs were stained with NucBlue nuclear DAPI stain for fixed cells (Life Technologies; 1 drop in 500µl PBS) for 5min at room temperature. Cells expressing untagged DPRs were also stained with the relevant rabbit anti-DPR primary antibody (Protein Tech; anti-GA, -GR, -AP or -GP and Genscript anti-PR) at a concentration of 1/100 in PBS, and goat anti-rabbit AlexaFluor-488 secondary antibody (Life technologies; 1 drop in 500µl PBS). All antibody incubations were performed at room temperature for ~30min. Images were captured and processed as described in Section 2.2.4.

4.3 Results

4.3.1 Generation of DPR constructs using alternative codon sequences

Constructs were generated using a stepwise cloning method to progressively increase repeat number to a pathological level. A library of constructs containing different lengths of each repeat has therefore been developed, with lengths between ~22-1000 repeats for all except poly-GP, which remains at 43 repeats in the untagged vector and at 22 repeats with GFP-tag, due to technical difficulties with cloning of this sequence. This is demonstrated by the example ethidium bromide-labelled agarose gel images in Figure 4.3, showing the varying sizes of repeat-fragments obtained from different length constructs digested by BamHI and EcoRI.

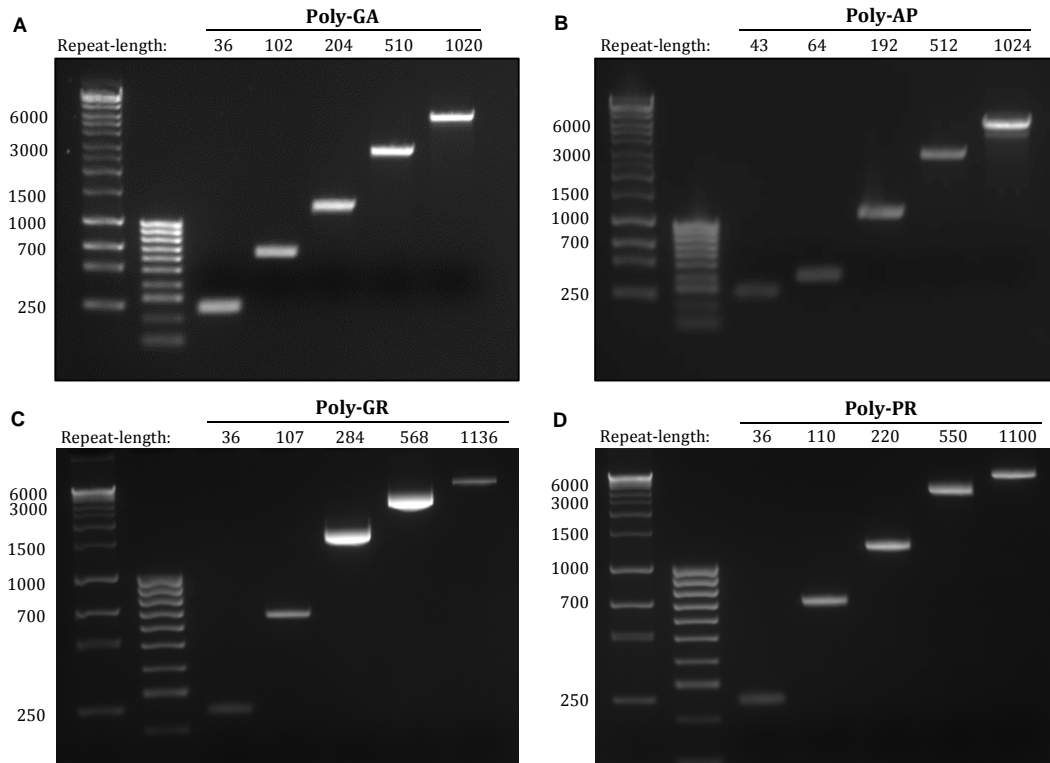


Figure 4.3: Ethidium bromide-labelled agarose gels showing repeat sequences from GFP-tagged (A) poly-GA, (B) poly-AP, (C) poly-GR and (D) poly-PR constructs at a range of different repeat lengths. Repeat sequences were isolated from DPR constructs by double-digest with BamHI and EcoRI followed by gel purification. Band sizes shown as number of bases. DNA ladders shown in left hand lanes are Bioline Hyperladders 1kb+ and 100bp.

1 repeat is defined as 6 bases or 2 amino acids, therefore the predicted band sizes are repeat-length multiplied by 6. The majority of DPR constructs are GFP-tagged on the C-terminal, however, a smaller library of “untagged” constructs which do not include the GFP-tag were also generated, by introducing a stop codon immediately after the repeat sequence. The full library of constructs produced is detailed in Table 4.1.

Table 4.1: List of DPR constructs produced according to repeat-length, where one repeat is defined as 6 bases or 2 amino acids.

GFP-tagged					Untagged				
GA	AP	GR	PR	GP	GA	AP	GR	PR	GP
36	22	36	36	22	36	22	36	36	22
102	43	63	55		58	106	71	50	43
204	64	71	110		71			71	
510	192	107	220		101			106	
816	256	142	440						
1020	512	284	550						
	1024	568	880						
		1136	1100						

4.3.2 Optimisation of DPR expression

Once the DPR construct library was complete, a protocol was required to efficiently express the peptides in mammalian cell culture. To optimise this protocol, 3 different transfection reagents were compared in HeLa cells: FuGene HD (Promega), JetPrime (PolyPlus) and Lipofectamine 2000 (Life Technologies). Each reagent was tested for transfection of pEGFP-N1, GA₁₀₂ and GR₇₁, using a range of plasmid concentrations. mRNA expression levels were compared by qRT-PCR using *GAPDH* as a reference gene. DPR expression levels relative to *GAPDH* were then compared for each plasmid. Comparisons are presented in Figures 4.4A-C as fold-changes in relative expression, compared to the FuGene treatment containing 50ng/well DNA for each plasmid. For all 3 plasmids, 400ng/well transfected with FuGene appeared to produce the largest increase in efficiency, with fold changes of up to 281. However, it was later noted that due to human error, the ratio of DNA-to-reagent used for all 3 400ng treatments was in fact 1:9, whereas all other treatments used a 1:3 ratio. This most likely explained the dramatic increase in transfection efficiency at 400ng. When the 400ng data was removed from analysis, Lipofectamine performed more efficiently for expression of pEGFP-N1 and

GA₁₀₂, whereas FuGene was more efficient for GR₇₁. Since poly-GR expression is generally lower than that of poly-GA or pEGFP-N1 (demonstrated in Figure 4.4D), FuGene was chosen to achieve the best overall transfection efficiency across all plasmids. Furthermore, FuGene remained the most efficient reagent for all 3 plasmids when used at a ratio of 1:9.

It was unclear from this data whether the dramatic increase in transfection efficiency using FuGene at a 1:9 DNA-to-reagent ratio with 400ng plasmid was due to the relatively high plasmid concentration, high reagent concentration or a combination of both. A further optimisation step was therefore performed to test this. Cells were transfected with each of the 3 plasmids (pEGFP-N1, GA₁₀₂ or GR₇₁) at either 400ng or 800ng per well, in order to determine whether increasing DNA concentration was responsible for the improved efficiency. In addition, each transfection was performed twice, using either the standard 1:3 ratio of DNA-to-reagent or the 1:9 ratio used in error in the previous experiment. All transfections were performed with FuGene. Data is presented in Figure 4.4E as fold-changes in relative mRNA expression using the 400ng pEGFP-N1 1:9 transfection as a reference. For all 3 plasmids, increasing the amount of DNA per well from 400ng to 800ng dramatically decreased efficiency. Furthermore, a DNA-to-reagent ratio of 1:9 was consistently more effective than a 1:3 ratio. Since FuGene does not appear to cause excessive toxicity when used at this concentration, transfection of 400ng/well of plasmid in a 12-well plate using a 1:9 ratio of DNA-to-reagent was adopted as the standard protocol for all future transfections. For transfection in 6-well plates, the amounts were doubled.

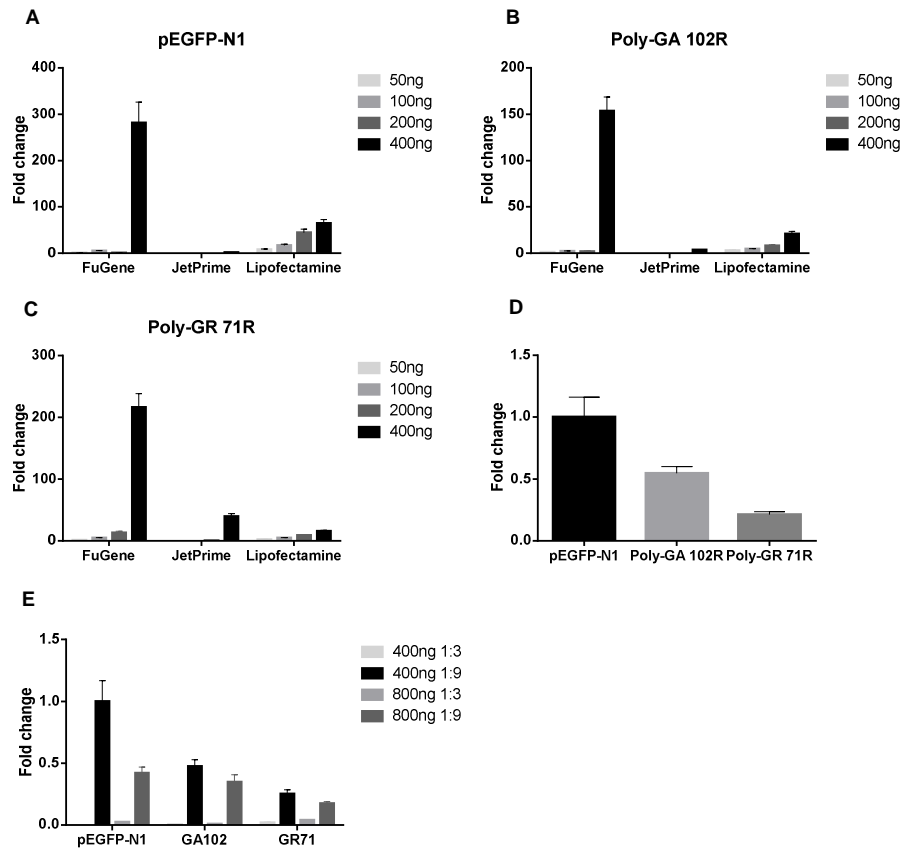


Figure 4.4: Optimisation of DPR transfection using qRT-PCR to quantify and compare mRNA expression levels in HeLa cells 24h post-transfection. All data is presented as fold changes relative to a reference. Poly-GA 102 repeats is used as representative of the alanine-rich DPRs, and poly-GR 71 repeats is used as representative of the arginine-rich DPRs. pEGFP-N1 was also transfected since the empty plasmid will be used as a GFP-only control throughout this project. All values represent the mean of 4 technical replicates. Error bars indicate standard deviation between technical replicates. **(A-C)** Comparison of mRNA expression for the 3 plasmids using 3 different transfection reagents, FuGene ($n=1$), JetPrime ($n=1$) and Lipofectamine ($n=1$). The amount of plasmid used to transfect a single well of a 12-well plate was also varied from 50-400ng. A DNA-to-reagent ratio of 1:3 was maintained with the exception of the 400ng treatments, for which a 1:9 ratio was used in error. The 50ng values were used as references for each plasmid. **(D)** Comparison of mRNA expression of the 3 plasmids following the most efficient transfection method, using 400ng of DNA, FuGene and a 1:9 ratio. pEGFP-N1 was used as a reference. **(E)** Further optimisation to determine whether the most efficient transfection method was due to high DNA concentration or high FuGene concentration. mRNA expression was quantified following transfection with either 400 or 800ng of plasmid, at 1:3 and 1:9 DNA-to-FuGene ratios ($n=1$). The value for pEGFP-N1 with 400ng DNA and a 1:9 ratio was used as a reference.

4.3.3 DPR protein expression *in vitro*

The optimisation of DPR transfection described in Section 4.3.2 produced a suitable protocol to express dipeptide constructs as efficiently as possible in HeLa cells. However, this was measured only by quantification of mRNA expression, and therefore Western blotting was also performed to confirm successful expression of each DPR at the protein level. All 5 DPRs at short lengths below 100 repeats were detected using an antibody against GFP (Figure 4.5A). However, Western blotting confirmed observations from qRT-PCR that DPR expression levels reduced as repeat-length increased, despite optimisation of the protocol (data not shown). This reduction in DPR expression was also observed subjectively by fluorescence microscopy; the GFP signal reduced as repeat-length increased, particularly for the arginine-rich DPRs (data not shown). The reduction in DPR expression appeared to be due to a decrease in the number of successfully transfected cells, as opposed the entire cell population expressing the peptides at a reduced level. Peptides above ~100 repeats in length therefore became difficult to detect by Western blot, since a large proportion of the cell population remained untransfected.

To confirm successful expression of longer DPRs, large volumes of cell lysate expressing 2 example DPRs (GA_{1020} and GR_{1136}) were concentrated using 100kDa filters and blotted. GA_{1020} was detectable by loading filtered lysate from a whole 6-well plate into a single well of an acrylamide gel (Figure 4.5B). The removal of proteins below 100kDa meant that no GFP-only control could be performed, however, a band was detected at the predicted size of ~150kDa and therefore can reasonably be assumed to be GA_{1020} . GR_{1136} was more difficult to detect, and therefore an additional sample was prepared containing twice as much filtered lysate. GR_{1136} was then detected as a smear rather than a band, running from the wells where a small amount of protein remained trapped, to a clear band at 37kDa which most likely corresponded to cleaved GFP (Figure 4.5C). GR_{1136} was also detected by dot blot (Figure 4.5D). Therefore it is clear that the DPRs are successfully expressed at the protein level at longer lengths of ~1000 repeats, if only in a small proportion of cells.

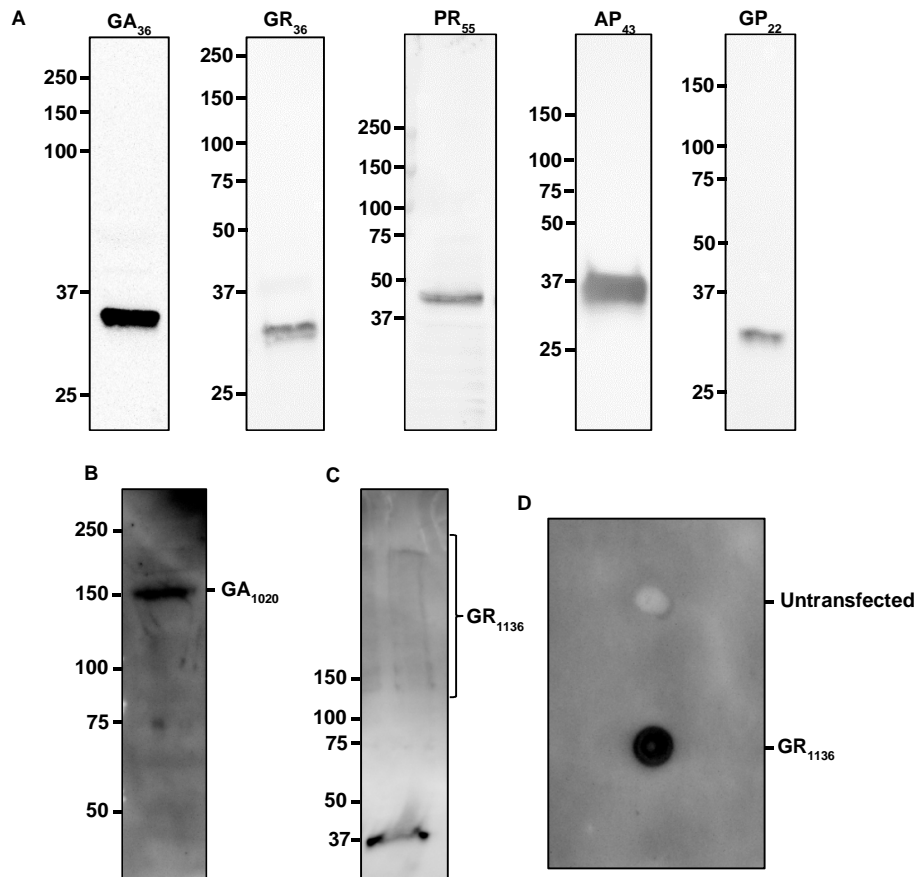


Figure 4.5: Confirmation of GFP-tagged DPR protein expression in HeLa cells 24h post-transfection. All blots labelled with an anti-GFP antibody (Santa Cruz). Numbers indicate protein size in kDa. **(A)** Representative Western blots demonstrating expression of all DPRs at short repeat-lengths ($n=3$). Predicted sizes of GFP-tagged peptides are: GA₃₆ 32kDa; AP₄₃ 34kDa; GP₂₂ 30kDa; GR₃₆ 35kDa; PR₅₅ 41kDa. **(B)** Western blot using concentrated lysate from a whole 6-well plate showing expression of GA₁₀₂₀ ($n=2$). **(C)** Western blot using concentrated lysate from 2 6-well plates showing expression of GR₁₁₃₆ ($n=1$). Predicted sizes of GA₁₀₂₀ and GR₁₁₃₆ including GFP tag are 157.6kDa and 269.2kDa respectively. **(D)** Dot blot further confirming expression of GR₁₁₃₆ in concentrated lysate from transfected, but not untransfected, cells ($n=1$).

4.3.4 Subcellular localisation of DPRs *in vitro*

Further confirmation that each DPR was successfully expressed in HeLa cells was provided by immunofluorescence imaging. The subcellular localisation of each DPR was investigated at a range of lengths and at different time-points post-transfection. Representative images of each DPR compared to GFP-only control (pEGFP-N1) at 24h and 72h post-transfection are shown in Figures 4.6 and 4.7, respectively. Each DPR was expressed at both long and short lengths (~1000 or 36-43 repeats) with the exception of poly-GP, which was only expressed at 22 repeats since technical difficulties with the cloning meant that no longer GFP-tagged constructs were available.

Differential distribution patterns were observed for each of the 5 DPRs, and in some cases differences were also observed between long and short repeat-lengths. Poly-GA consistently formed cytoplasmic inclusions which increased in size with longer repeat-lengths. Poly-GA peptides containing shorter repeats commonly formed 2 or more small inclusions within 1 cell, whereas GA₁₀₂₀ tended to form a single, large aggregate with a fibrous, star-shaped structure. Conversely, poly-AP was diffusely expressed at shorter repeat-lengths, with a similar appearance to vehicle. AP₁₀₂₄, however, formed a characteristic pattern in many cells, with abundant expression surrounding the nucleus and filamentous structures connecting the perinuclear net to the cell membrane. GP₂₂ was diffusely expressed and did not appear significantly different to GFP-only controls. No differences were observed in distribution of poly-GA, -GP or -AP at 24h and 72h.

Both arginine-rich DPRs formed small, punctate inclusions within the nucleus. Poly-GR was also diffusely expressed throughout the cytoplasm. Some cells were present with poly-GR expression solely in the cytoplasm, whilst others contained both diffuse cytoplasmic expression and nuclear inclusions. At 72h post-transfection, some cells expressing GR₁₁₃₆, but not GR₃₆, contained cytoplasmic inclusions. These inclusions maintained the smooth, rounded structure of those found in the nucleus. At 24h post-transfection, poly-PR was primarily found in the nucleus. However, at 72h, some diffuse cytoplasmic expression of poly-PR was observed only at the longer length of 1100 repeats. The distribution patterns of poly-GR and -PR were otherwise unaffected by repeat-length, although the GFP-signal became considerably weaker at lengths of ~1000, and fewer cells containing nuclear inclusions were observed.

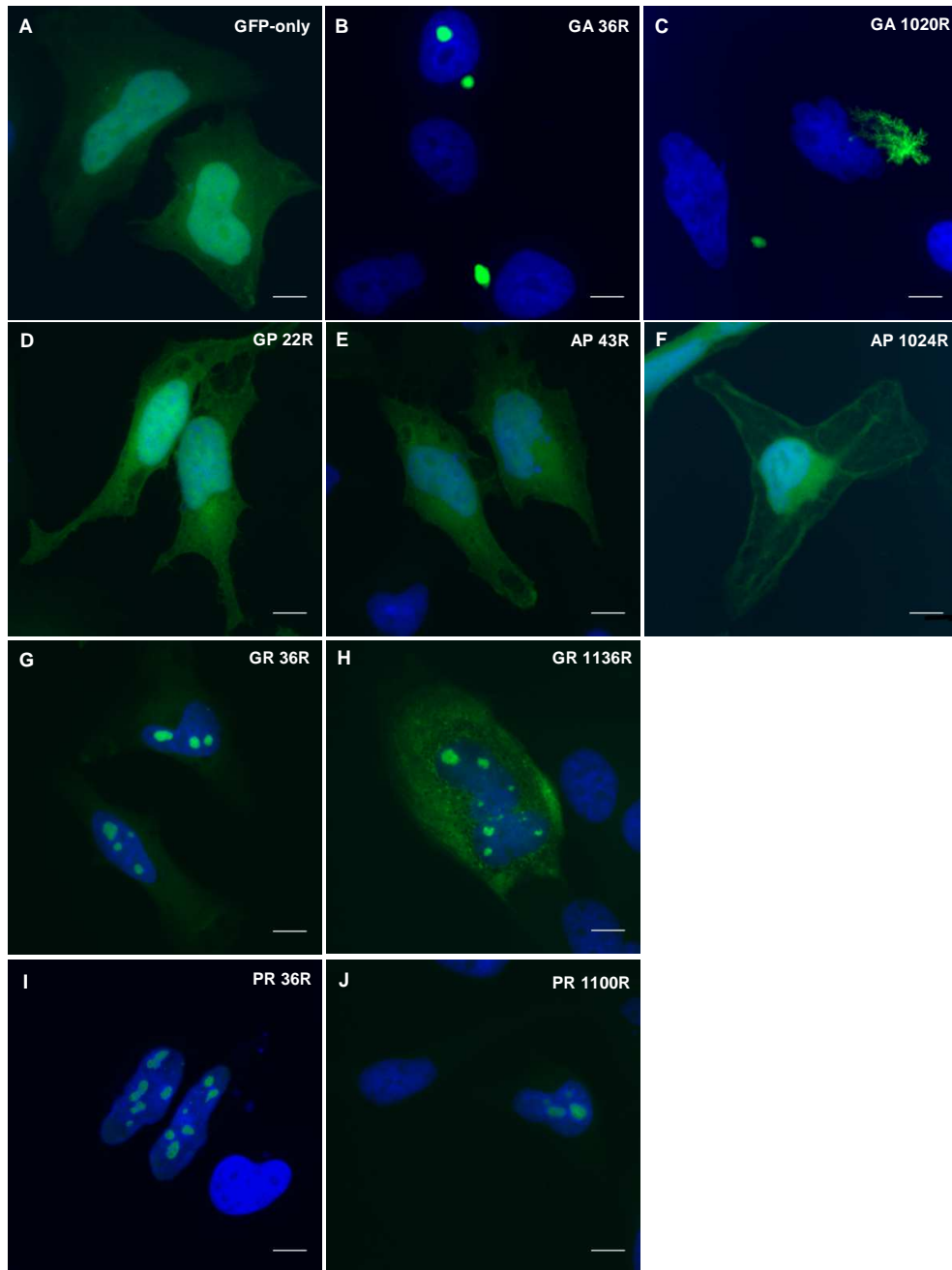


Figure 4.6: Immunofluorescence images showing differential distribution of GFP-tagged DPRs (green) in HeLa cells at 24h post-transfection, or pEGFP-N1 expression as a GFP-only control (A). Representative images are shown at short and long repeat-lengths for Poly-GA (B-C), poly-AP (E-F), poly-GR (G-H) and poly-PR (I-J). Poly-GP is shown at the only length available, 22 repeats (D). Repeat lengths are denoted by the letter “R”. Nuclei are labelled in blue (DAPI). Images captured at x60 magnification. Scale bars represent 15 μ m. n=3 for all transfections.

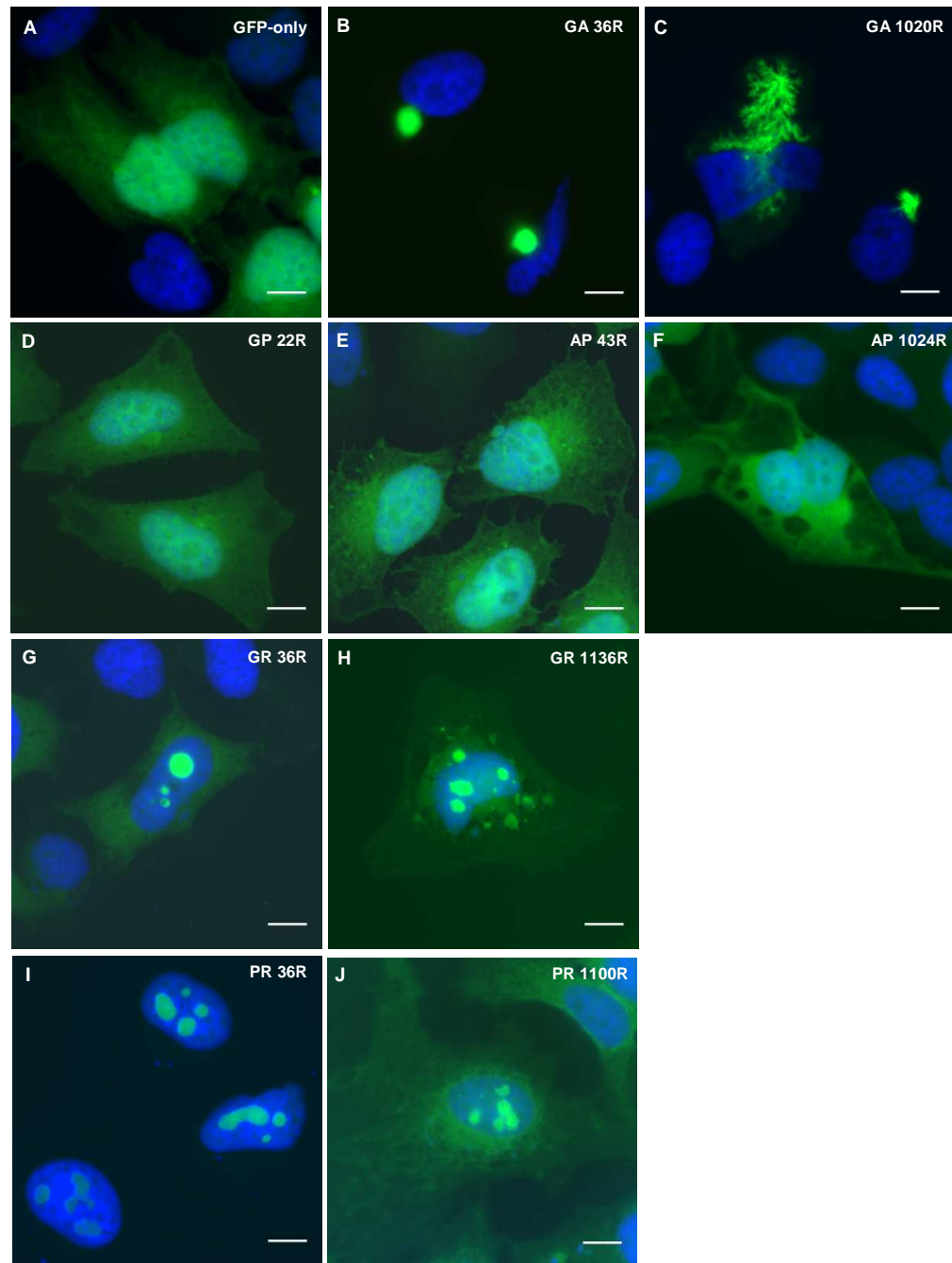


Figure 4.7: Immunofluorescence images showing differential distribution of GFP-tagged DPRs (green) in HeLa cells at 72h post-transfection. Representative images are shown at short and long repeat-lengths for Poly-GA (**B-C**), poly-AP (**E-F**), poly-GR (**G-H**) and poly-PR (**I-J**). Poly-GP is shown at the only length available, 22 repeats (**D**). Repeat lengths are denoted by the letter “R”. Nuclei are labelled in blue (DAPI). Images captured at x60 magnification. Scale bars represent 15 μ m. n=2 for all transfections.

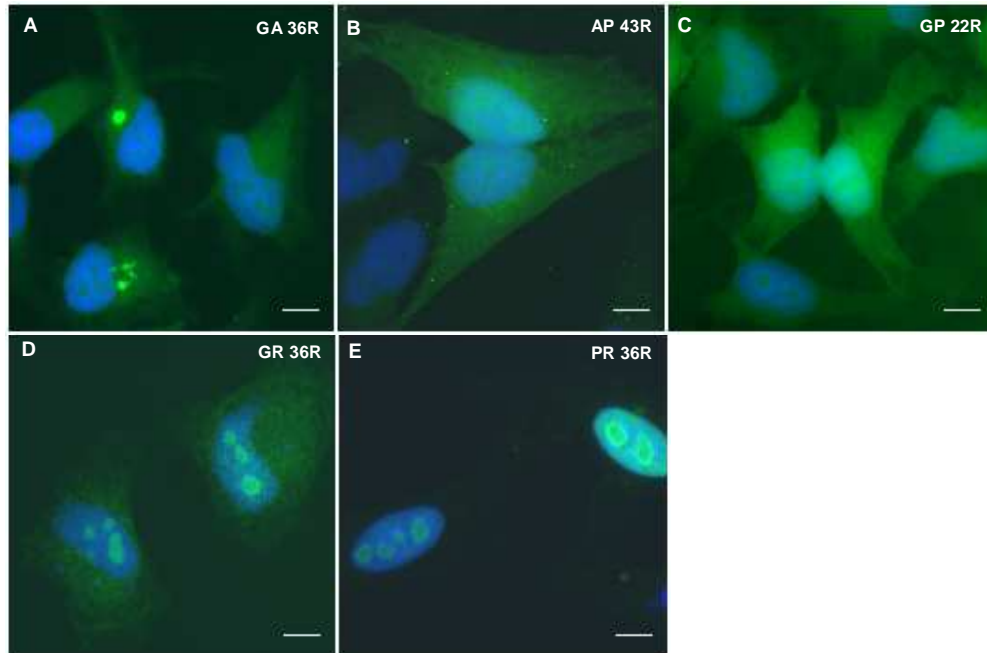


Figure 4.8: Immunofluorescence images showing distribution of untagged poly-GA (**A**), poly-AP (**B**), poly-GP (**C**), poly-GR (**D**), and poly-PR (**E**) using specific antibodies against each peptide (ProteinTech or Genscript; green) in HeLa cells, 48h post-transfection. Repeat lengths denoted by the letter “R”. Nuclei are labelled in blue (DAPI). Images captured at x60 magnification. $n=1$ for all transfections. Scale bars represent $15\mu\text{m}$.

In order to confirm that the GFP-tag did not affect subcellular localisation of DPRs, untagged constructs were also expressed using the same transfection protocol, and labelled with selective antibodies against the relevant DPR (ProteinTech). No differences were observed in expression the untagged constructs (Figure 4.8) and those expressing a GFP-tag.

4.4 Discussion

4.4.1 Generation of cellular models of DPR pathology

This chapter describes the generation of a large library of constructs to express DPRs at a range of lengths in mammalian cell culture. Furthermore, we have optimised expression of DPRs *in vitro* to create and characterise a series of cellular models of DPR pathology. Successful DPR expression has been demonstrated by qRT-PCR, Western blotting and immunofluorescence. Importantly, the DPR library includes constructs longer than 1000 repeats, which is comparable to the expected length of DPRs found in disease; the expansion length in C9FTLD/ALS patients is at least 1000 repeats in almost all cases (DeJesus-Hernandez et al., 2011), and Western blotting of peptides extracted from patient tissue suggest that long peptides do indeed arise from the expansion (Ash et al. 2013; Mori et al. 2013). Furthermore, antibodies specifically targeting the C-terminal of poly-GA successfully labelled inclusions in patient neurons, confirming that translation continues until the end of the expansion (Mori et al. 2013). Therefore our cellular models of dipeptide pathology at ~1000 repeat in length may be considered physiologically relevant. Several independent groups have recently published details of DPR models generated in a similar fashion to ours, which will be discussed in detail in Chapter 5. However, the longest repeat-lengths expressed in these studies are ~100-175 (Mizielinska et al. 2014; May et al. 2014). As such, the constructs of disease-relevant lengths developed during the course of this project will provide an important improvement to the currently existing models.

Poly-GP presents an exception to this, since the cloning to elongate poly-GP constructs was unsuccessful beyond 43 repeats. Thus far, no research group has reported successful generation of a poly-GP construct longer than 80 repeats (May et al. 2014). Since each group designed their alternative codon sequences independently, this suggests the protein itself is problematic, perhaps through leaky expression and excessive toxicity when transformed into competent *E. coli*, preventing growth of successfully cloned plasmids. To overcome this potential issue, transformation was attempted following cloning using several different types of competent *E. coli*, including strains marketed as useful for reducing leaky transcription of toxic plasmids, or for handling repetitive sequences. However, none were successful. The following chapters will therefore focus on the remaining 4 DPRs, for which physiologically relevant repeat-lengths were successfully generated.

The major limitation of the longer length DPR models is that expression levels reduce as repeat-length increases, particularly with the arginine-rich peptides. Imaging data suggests that this is due to a decrease in the number of cells transfected, rather than a reduction in the amount of protein produced within individual cells, although this is also the case to a certain extent with poly-GR and –PR. It is unclear whether this is due to poor transfection efficiency, excessive toxicity of DPRs or both. It is possible that protective mechanisms such as methylation of the transfected plasmid are activated within the cell in response to the introduction of such long, repetitive plasmids, preventing significant DPR expression in many cells. It is also possible that inability to express plasmids above 10kb in length which contain a ~70% GC-rich insert is a technical limitation of the transfection reagents tested. Conversely, the peptides may express for short periods of time post-transfection but cause severe toxicity, preventing survival of transfected cells for analysis.

Despite the low expression level, GA₁₀₂₀ and GR₁₁₃₆ were both possible to detect by Western blot following protein concentration from large amounts of lysate. However, these blots are crude, particularly for GR₁₁₃₆ which formed a smear from the well, where a small amount protein remained trapped, down to a low molecular weight band presumably indicating cleavage of GFP during sample preparation. This type of smear is common when analysing pathological proteins linked to neurodegenerative disease by Western blot, and may be due to post-translational modifications such as methylation and ubiquitination. In the case of poly-GR, the peptides may form secondary structures through cross-linking which were not completely denatured by boiling in SDS buffer, preventing them from running normally on an acrylamide gel. Since part of the smear is below the predicted molecular weight of poly-GR and a distinct GFP band is visible at ~37kDa, some degradation of the peptide is also likely to have occurred. This may explain the difference between the appearances of poly-GA and –GR on Western blot; since poly-GA appears to form large, insoluble aggregates, degradation of the peptide during lysate preparation is less likely to occur.

Since the filtration step removes any proteins smaller than 100kDa, the use of Western blotting to investigate the functional impact of DPRs on cell signalling is severely limited. Furthermore, the filtration step prevents the use of a standard GFP-only control for these types of experiments. However, the blots shown in Figure 4.5 are useful for confirmation that peptides of the expected size are indeed produced in our cellular models. Whilst experimental tools requiring analysis of whole populations of cells may not be useful for investigation of cellular function in the current expression system, immunofluorescence

imaging remains an informative method, since individual cells expressing the DPR can be selected for analysis from a population containing many untransfected cells.

4.4.2 Differential distribution of DPRs *in vitro*

Example images showing the subcellular localisation of each dipeptide at long and short repeat-lengths are shown in Figures 4.6 and 4.7. Interestingly, the distribution patterns differ considerably between different DPRs, highlighting the requirement to separately study the properties of each in this type of model. Furthermore, several differences were noted between cells expressing long and short repeat-lengths of the same peptide, demonstrating the importance of modelling DPR pathology at physiologically relevant lengths. It is also important to note that no differences were observed in localisation of DPRs in the presence or absence of a GFP-tag, confirming that the tag did not affect the distribution of the peptides.

Our data shows that poly-GA consistently forms cytoplasmic inclusions, with short repeats forming 1 or more small punctate inclusions within an individual cell, whereas GA₁₀₂₀ formed single, larger aggregations with a fibrous, star-shaped structure. Poly-AP was also primarily cytoplasmic, with short repeats expressing diffusely and AP₁₀₂₄ forming structures which appeared perinuclear in nature and which connected to the cell membrane via multiple fibrous, linear structures. This distribution pattern is particularly unusual when compared to other proteins known to aggregate in neurodegenerative disease, and appears to resemble a cytoskeletal structure. Poly-GP was also diffusely expressed at the short length of 22 repeats, however conclusions must be drawn with caution from this data since the repeat length is not physiologically relevant, and the difference between subcellular localisation of AP₄₃ and AP₁₀₂₄ strongly highlights the importance of expressing DPRs at longer repeat lengths.

In contrast to the 2 alanine-rich DPRs, both arginine-rich peptides formed small, rounded inclusions in the nucleus. Interestingly, poly-GR was also expressed diffusely throughout the cytoplasm, and some cells contained diffuse cytoplasmic poly-GR but no nuclear inclusions. Poly-PR, conversely, was primarily nuclear. No cells were found which contained nuclear poly-GR inclusions in the absence of cytoplasmic staining. This may reflect a time-course of inclusion formation, with GR initially expressing in the cytoplasm and forming nuclear inclusions over time. GR₁₁₃₆ also formed cytoplasmic inclusions in some cells at 72h post-transfection, but not at 24h. These inclusions differed from cytoplasmic poly-GA inclusions, more closely resembling the rounded

nuclear poly-GR inclusions observed at 24h time-points. The formation of cytoplasmic inclusions only at later time-points, along with the reduction in number of cells containing nuclear inclusions at 72h, suggests that nuclear inclusions may form initially and translocate to the cytoplasm over time. However, this is a preliminary observation and cannot be confirmed from the images presented in this chapter. Video microscopy to monitor the production and localisation of DPRs over time would therefore be an interesting future study.

In order to validate our cellular models of DPR pathology, it is important to compare the data presented here to observations from patient tissue, and consider possible reasons for any differences noted. The first observations of DPR pathology in patient tissue were reported by 2 key studies which generated antibodies against the sense strand peptides, poly-GA, -GR and -GP (Ash et al. 2013; Mori et al. 2013). Ash *et al.* developed 2 polyclonal antibodies, both of which detected all 3 sense strand DPRs. Immunohistochemical analysis using these antibodies identified a combination of small, round nuclear inclusions and cytoplasmic star-shaped inclusions in brain tissue from expansion-carriers but not non-C9FTLD/ALS cases. Inclusions were neuronal, with no glial staining observed, and found in the cerebellum, hippocampus, motor cortex, frontal and temporal cortices and amygdala. Particular abundance of inclusions was noted in the granular cells of the hippocampus and cerebellum, pyramidal cells of the cortex and cerebellar purkinje cells (Ash et al. 2013).

Whilst it is not possible to determine which of the 3 sense strand DPRs was present in each inclusion using these antibodies, based on the data presented here it is likely that poly-GA may have been present in the cytoplasmic inclusions, and poly-GR in the nucleus. Mori and colleagues then developed more selective antibodies to specifically detect either poly-GA, -GR or -GP, with no cross-reactivity (Mori et al. 2013). This study confirmed the presence of p62-positive poly-GA cytoplasmic inclusions in neurons in the cerebellum and hippocampus of expansion-carrying patient tissue. A mixture of small spherical inclusions and larger star-shaped aggregates were observed. This is comparable to our observation that poly-GA may form either small, rounded inclusions or larger, fibrous structures resembling the star-shaped inclusions frequently found in patient tissue. More recent pathological studies using a range of antibodies confirm these early observations in patient tissue. Poly-GA is consistently found to form cytoplasmic inclusions in areas affected by C9FTLD/ALS such as the cerebellum, hippocampus and frontal, temporal and motor cortices (Mackenzie et al. 2013; Mann et al. 2013; Davidson et al. 2014). Less abundant GA-positive inclusions have also been identified in subcortical regions including the striatum and substantia nigra, and in lower

motor neurons (Mackenzie et al. 2013). Poly-GA inclusions are p62-positive and most commonly exhibit the typical star-shaped structure which is comparable to our observations of GA₁₀₂₀-expressing cells (Figure 4.9). Importantly, this provides validation of our GA₁₀₂₀-expressing cells as a relevant model of poly-GA pathology in C9FTLD/ALS.

In contradiction to observations from our cellular model, poly-GR and -GP staining is also observed in p62-positive, cytoplasmic star-shaped inclusions in patient tissue, although less commonly than poly-GA (Mori et al. 2013; Mann et al. 2013). However, inclusions containing any of these 3 DPRs are very similar in appearance (Figure 4.9D), and it is possible that when all 5 peptides are present in patient tissue, poly-GA inclusions sequester other DPRs present in the cytoplasm, leading to the formation of larger aggregates containing 2 or more peptides. This would explain the contradiction between poly-GR localisation *in vitro* compared to patient tissue. Indeed, diffuse cytoplasmic expression of poly-GR is noted in our cellular model, as well as occasional cytoplasmic inclusions with a more rounded structure than those observed in patient tissue, indicating that the peptide is available for sequestration by poly-GA. This hypothesis is supported by a recent study which co-expressed short (88 repeat) alternative-codon poly-GA and poly-GR sequences similar to our own in *Drosophila*. When expressed separately, poly-GA formed cytoplasmic inclusions but poly-GR did not, whereas when both peptides were co-expressed, poly-GR was sequestered into the poly-GA aggregates (Yang et al. 2015). An interesting future avenue of research using our longer repeat constructs will be to co-express multiple DPRs using alternative fluorescent tags to allow visualisation of both peptides. Co-staining of patient tissue with multiple DPR antibodies would also be of value.

Antibodies targeting the DPRs arising from the antisense strand (poly-AP and -PR, as well as poly-GP) are now also available, and have been used in several immunohistochemical studies of patient tissue. Positive poly-AP staining has been reported to very rarely occur in p62-positive star-shaped cytoplasmic inclusions in the CA4 region of the hippocampus (Mann et al. 2013). However, this staining was reportedly very weak and no further poly-AP-positive inclusions were observed in the cerebellum or other hippocampal regions (Mann et al. 2013). Like poly-GR, it is possible that poly-AP is only present in these inclusions due to sequestration by poly-GA. No observations of cytoskeletal-like distribution patterns similar to those found in our AP₁₀₂₄-expressing cells have been reported, although this may be due to difficulties detecting the signal from these structures, which is considerably lower than the signal from GA₁₀₂₀ aggregates in our model. It is unclear whether the weak staining was because poly-AP

was not abundantly present in patient tissue or because the antibodies to detect poly-AP are less efficient than those available for poly-GA. As such, it is difficult to draw reliable conclusions about the relevance of our AP₁₀₂₄ model to human disease at this stage.

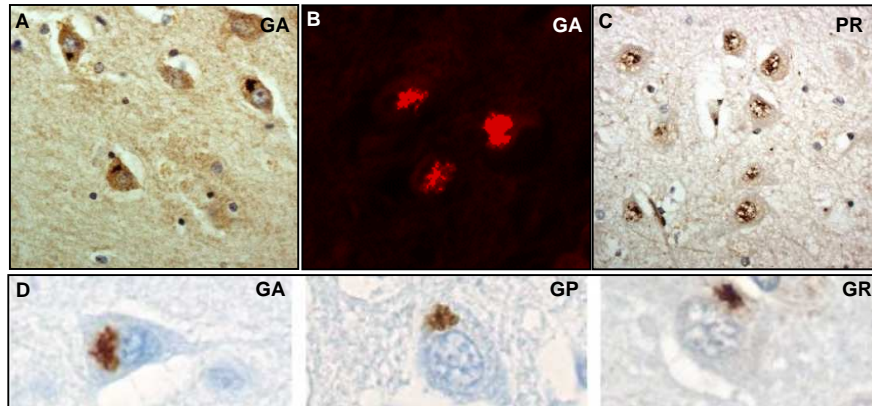


Figure 4.9: Example images of DPR pathology in patient brain tissue, taken from published literature. Poly-GA labelling of pyramidal cells in the CA4 region using 2 different antibodies; images adapted from Mann et al., 2013 (A) and Mackenzie et al., 2013 (B). (C) Poly-PR staining in pyramidal cells of CA4; image adapted from Mann et al., 2013. (D) Example images taken from Mori et al., 2013 showing similarity in the structure of inclusions in neurons of the CA3 region labelled with separate antibodies for either poly-GA, -GP or -GR, from left to right.

Finally, poly-PR has been found to form small, nuclear inclusions in neurons of patient tissue (Figure 4.9C), which are comparable to those observed in our cellular poly-PR model (Mann et al. 2013). Particular abundance of nuclear poly-PR was noted in cerebellar purkinje cells and pyramidal cells in the CA4 region of the hippocampus (Mann et al. 2013). No cytoplasmic inclusions containing poly-PR were observed in patient tissue (Mann et al. 2013), which further supports the hypothesis that DPRs contained within the cytoplasm are sequestered by poly-GA aggregates; poly-PR staining was predominantly nuclear in our cellular model, and therefore has no opportunity to be sequestered into cytoplasmic inclusions. These observations from patient tissue also validate our cellular model of poly-PR pathology as being directly relevant to the human disease condition.

4.4.3 Summary

This chapter describes the generation and characterisation of the first models of DPR pathology expressing physiologically relevant repeat-lengths of over 1000. The importance of repeat-length is evidenced by the differential distribution and aggregation patterns of DPRs at different lengths, particularly poly-GA and poly-AP. Therefore our models have potential to add great value to the field, by providing a significant improvement to currently available DPR models. Our poly-GA and poly-PR models in particular are highly comparable to observations of patient tissue, validating our approach as representative of DPR pathology in the human disease condition. Furthermore, the noted differences in subcellular localisation of poly-GR and poly-AP between patient tissue and our *in vitro* models have led to formation of a hypothesis regarding the interaction of DPRs in patient neurons, which is supported by data from a *Drosophila* model and which may prove a useful line of future research.

There are two major limitations of our *in vitro* models. Firstly, whilst DPR localisation in our models mostly correlates with observations from patient tissue, the impact of DPR pathology on cellular function may differ between HeLa cells and neurons *in vivo*. However, *in vitro* models may still be of great value to research at least in the preliminary stages, particularly given that any phenotypic changes observed from immunofluorescence imaging can be similarly investigated in patient tissue, in order to validate the findings. HeLa cells were chosen for this work for ease of transfection and suitability for DPR imaging. The second limitation of these models is that the low expression of longer DPRs prevents use of techniques which require analysis of whole populations of cells, such as Western blotting. Despite this, the current expression system remains a valuable tool for imaging studies to investigate the impact of DPRs on cellular function, which will be the primary focus of the remainder of this project.

Chapter 5: Impact of DPRs on cellular function

5.1 Introduction

The previous chapter described the generation and characterisation of a series of cellular models, each expressing one of the five DPRs found in C9FTLD/ALS. Each dipeptide exhibited a unique subcellular distribution pattern, with the alanine-rich DPRs (poly-GA and –AP) primarily located in the cytoplasm and the arginine-rich DPRs (poly-GR and –PR) primarily nuclear. These differences in subcellular localisation suggest that each dipeptide is likely to differentially impact cellular function. Our models expressing each DPR individually are therefore useful tools with which to elucidate the specific contribution of each peptide to neurotoxicity in C9FTLD/ALS.

This chapter aimed to further characterise our DPR models, and investigate the impact of DPR expression on cellular function using immunofluorescence imaging. Since poly-GA is known to co-localise with UPS components in patient tissue (Mann et al. 2013), the subcellular localisation of p62, ubiquitin and ubiquilin-2 were assessed in HeLa cells expressing poly-GA. The distribution pattern of AP₁₀₂₄ shown in Figure 4.6F suggested that poly-AP may bind to cytoskeletal components, and therefore actin microfilaments and microtubules were stained in cells expressing AP₁₀₂₄. Finally, the smooth, rounded structure of nuclear poly-GR and –PR inclusions shown in Figure 4.6 suggested that arginine-rich DPRs may be recruited to nuclear sub-organelles such as the nucleolus. Cells expressing poly-GR or –PR were therefore labelled using antibody markers for the nucleolus and nucleolar stress.

5.2 Materials and Methods

5.2.1 Immunofluorescence

Immunofluorescence was performed according to the standard protocol in Section 2.2.4. Primary antibodies and stains used in this section are listed in Table 5.1. Secondary antibodies used were either donkey anti-rabbit IgG or goat anti-mouse IgG, conjugated to Alexa Fluor-594 (Life Technologies; 1 drop antibody solution in 500 μ l PBS).

One exception to the standard protocol was in preparation of coverslips for microtubule imaging, since the α -tubulin antibody required methanol fixation of cells instead of PFA. Coverslips were washed once in warm PBS and fixed in 100% methanol at -20°C for ~2min. Cells were then washed, solubilised and labelled according to the standard protocol. Methanol fixation was found to reduce the intensity of GFP signal, and therefore cells were additionally labelled with an antibody against GFP, followed by incubation with a mouse anti-rabbit IgG secondary antibody conjugated to Alexa Fluor-488 (Life Technologies; 1 drop in 500 μ l PBS).

All antibody incubations were performed for ~30min at room temperature, protected from light. Cells labelled with the microfilament stain, phalloidin, were incubated in the same way, however, no secondary antibody was required. NucBlue DAPI stain (Life Technologies) was either combined with the secondary antibody or incubated separately for ~10min at room temperature.

In order to compare the intensity of staining for the nucleolar protein, fibrillarin, at different points within the cell, heat-maps were produced. Cells expressing GFP or nucleolar poly-GR were imaged according to the standard protocol (Section 2.2.4). Cells were selected for imaging without viewing the red fibrillarin channel, to avoid bias. The exposure time was maintained at 500ms for all fibrillarin images. Fibrillarin images were converted into heat-maps using ImageJ (<http://rsb.info.nih.gov/ij>), depicting the relative signal intensities throughout the cell as colour. Areas of the cell which appeared yellow indicated high intensity of fibrillarin staining at that point, ranging through to red, purple and blue as the signal decreased.

Initial staining with an antibody against the survival motor neuron protein (SMN) suggested there may be a difference in the number of nuclear SMN puncta in cells expressing nucleolar poly-GR. Semi-quantitative analysis of SMN images was therefore performed by manually counting the number of nuclear SMN puncta in cells expressing

nucleolar poly-GR or pEGFP-N1. Cells were selected for imaging without viewing the red SMN channel, to avoid bias. The exposure time was maintained at 15ms for SMN images. For control cells, the number of SMN puncta present in the nucleus was counted in all cells within the field that expressed GFP. For poly-GR, only cells expressing nucleolar poly-GR were analysed (minimum of 1 per image). The mean number of nuclear SMN puncta per cell from a minimum of 10 images was considered as 1 independent replicate. The number of nuclear SMN puncta in poly-GR and control cells was compared by unpaired t-test (n=3) using GraphPad Prism. A graph showing the mean number of nuclear puncta per cell was also produced using Prism.

Table 5.1: List of primary antibodies and stains used for immunofluorescence in this chapter, together with source and dilution factor in PBS.

Target	Host organism/ stain	Manufacturer	Concentration (dilutions in PBS)
F-actin	Phalloidin	Invitrogen	1:100
Coilin	Rabbit	ProteinTech	1:200
EWS	Rabbit	ProteinTech	1:50
Fibrillarlin	Mouse	Abcam	1:1000
FUS	Rabbit	Abcam	1:50
GFP	Rabbit	Santa Cruz	1:200
Lamin B1	Rabbit	Abcam	1:500
α -tubulin	Mouse	Sigma	1:500
Nuclei	NucBlue (DAPI)	Life Technologies	1 drop in 500 μ l
Nucleolin	Rabbit	ProteinTech	1:200
p62	Mouse	Abcam	1:50
SMN	Mouse	Thermo Fisher	1:400
TDP-43	Rabbit	ProteinTech	1:50
Ubiquilin-2	Mouse	Abnova	1:50
Ubiquitin	Rabbit	Abcam	1:400

5.3 Results

5.3.1 Effects of DPRs on the FTLD/ALS-related proteins, TDP-43, FUS and EWS

In order to further characterise the pathological hallmarks of DPR-expressing cells, the impact of full-length (>1000 repeats) dipeptides on FTLD/ALS-related FET proteins was investigated. Cells were labelled with antibodies against TDP-43, EWS or FUS, 48h following transfection with GA₁₀₂₀, AP₁₀₂₄, GR₁₁₃₆, PR₁₁₀₀ or a GFP-only control (Figures 5.1-5.3). No TDP-43 inclusions were observed in DPR-expressing cells. TDP-43 staining was primarily nuclear in all cells including GFP-only controls, with some diffuse expression in the cytoplasm. It was noted that nuclei of some cells expressing poly-GA (Figure 5.1B) and poly-GR (Figure 5.1D) were misshapen, larger than those of surrounding cells or elongated. In these cells, a slight reduction in nuclear TDP-43 signal compared to cytoplasmic was occasionally observed, however this was not found in the majority of cells and therefore considered to most likely be an artefact or feature of apoptotic cells, rather than TDP-43 mislocalisation caused by DPR expression.

EWS staining was also primarily nuclear with some diffuse cytoplasmic expression, although EWS also formed small structures with an aggregate-like appearance in some cells, including GFP-only controls (Figure 5.2A). Very weak co-localisation was observed between EWS and poly-AP, although this did not occur in all cells expressing AP₁₀₂₄ (Figure 5.2C). Misshapen nuclei were again observed in poly-GA and poly-GR cells, as well as in poly-PR cells, however this did not notably affect EWS distribution.

Like TDP-43 and EWS, FUS staining was also primarily nuclear. However, DPR-expressing cells with misshapen nuclei frequently exhibited FUS mislocalisation, with regions of dense FUS-positive staining found in the cytoplasm. This mislocalisation was predominantly observed in cells expressing poly-GA (Figure 5.3B) and poly-GR (Figure 5.3D), and only in cells with misshapen nuclei. As such, FUS mislocalisation is unlikely to be directly caused by expression of DPRs, but rather as a consequence of the impact of DPRs on cell nuclei.

In order to further assess the effect of DPRs on nuclear structure, cells were labelled with an antibody against the nuclear membrane protein, lamin B1 (Figure 5.4). In control cells, lamin B1 formed a clear, continuous membrane around the nucleus, with no visible breaks or pores. However, in cells expressing poly-GA, poly-GR and poly-PR, multiple disruptions to the nuclear membrane were observed, in addition to the presence of

misshapen nuclei in many cells (Figure 5.4). No changes in lamin B1 distribution were observed in cells expressing poly-AP (data not shown). This supports the hypothesis that FUS mislocalisation is caused by disruption to the nuclear membrane integrity, perhaps allowing FUS protein to leak into the cytoplasm.

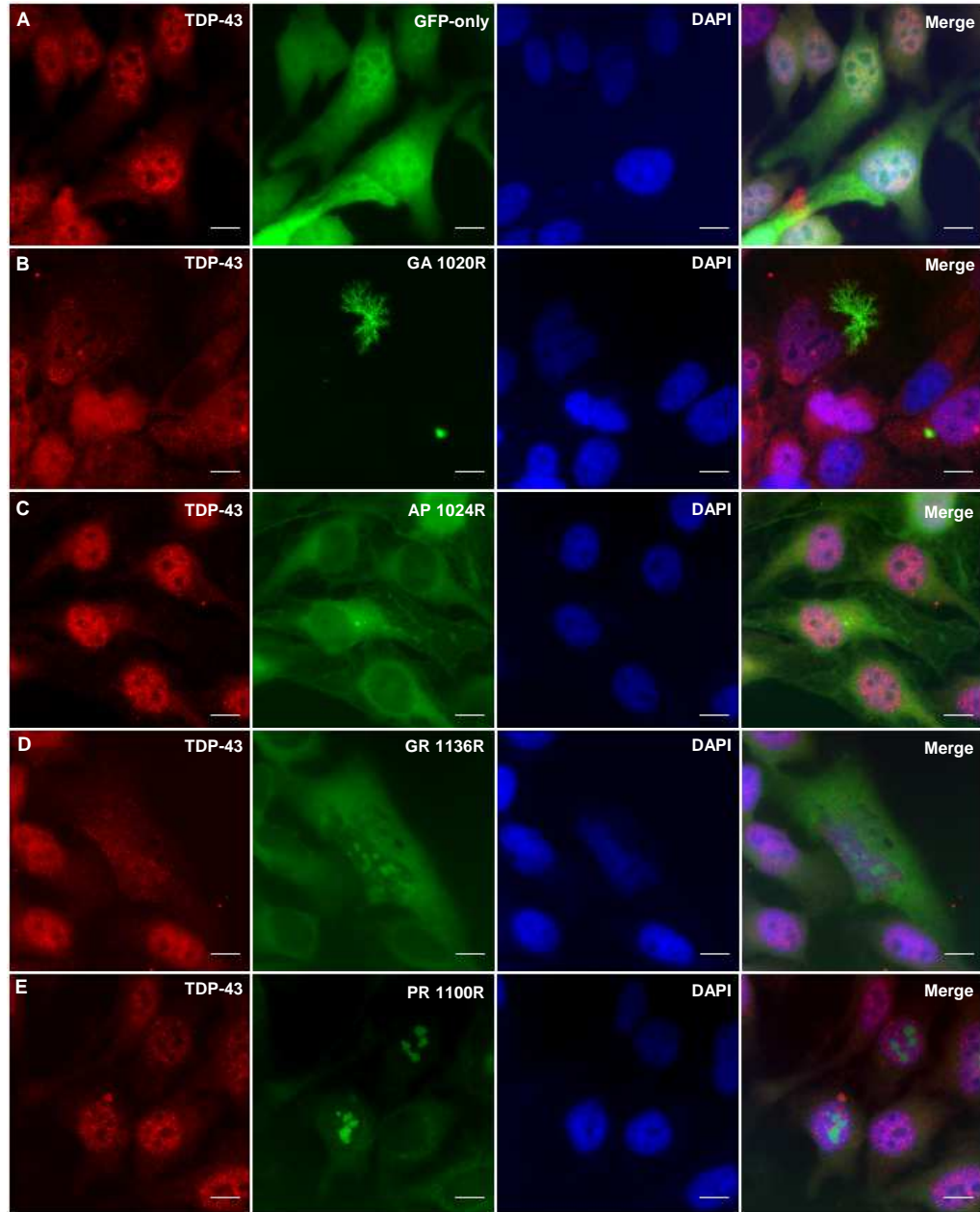


Figure 5.1: Impact of DPR expression (green) on TDP-43 distribution (red) in HeLa cells, 48h post-transfection. pEGFP-N1 expression shown as a GFP-only control (A). All DPRs are expressed at the longest lengths available: GA₁₀₂₀ (B), AP₁₀₂₄ (C), GR₁₁₃₆ (D) and PR₁₁₀₀ (E). Nuclei are labelled in blue (DAPI). Images captured at x60 magnification. Scale bars represent 15µm. n=2 for all transfections

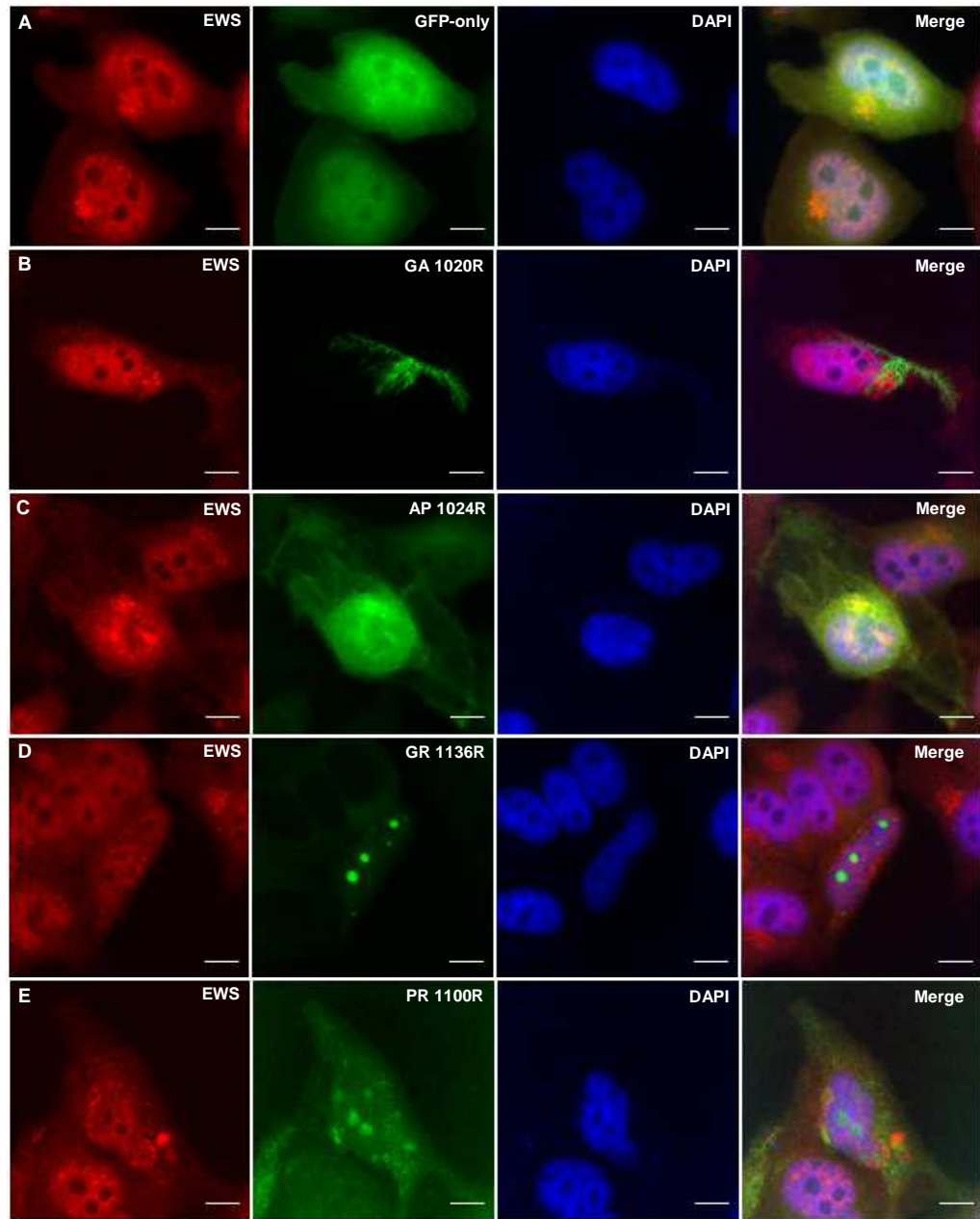


Figure 5.2: Impact of DPR expression (green) on EWS distribution (red) in HeLa cells, 48h post-transfection. pEGFP-N1 expression shown as a GFP-only control (**A**). All DPRs are expressed at the longest lengths available: GA₁₀₂₀ (**B**), AP₁₀₂₄ (**C**), GR₁₁₃₆ (**D**) and PR₁₁₀₀ (**E**). Nuclei are labelled in blue (DAPI). Images captured at x60 magnification. Scale bars represent 15µm. n=2 for all transfections

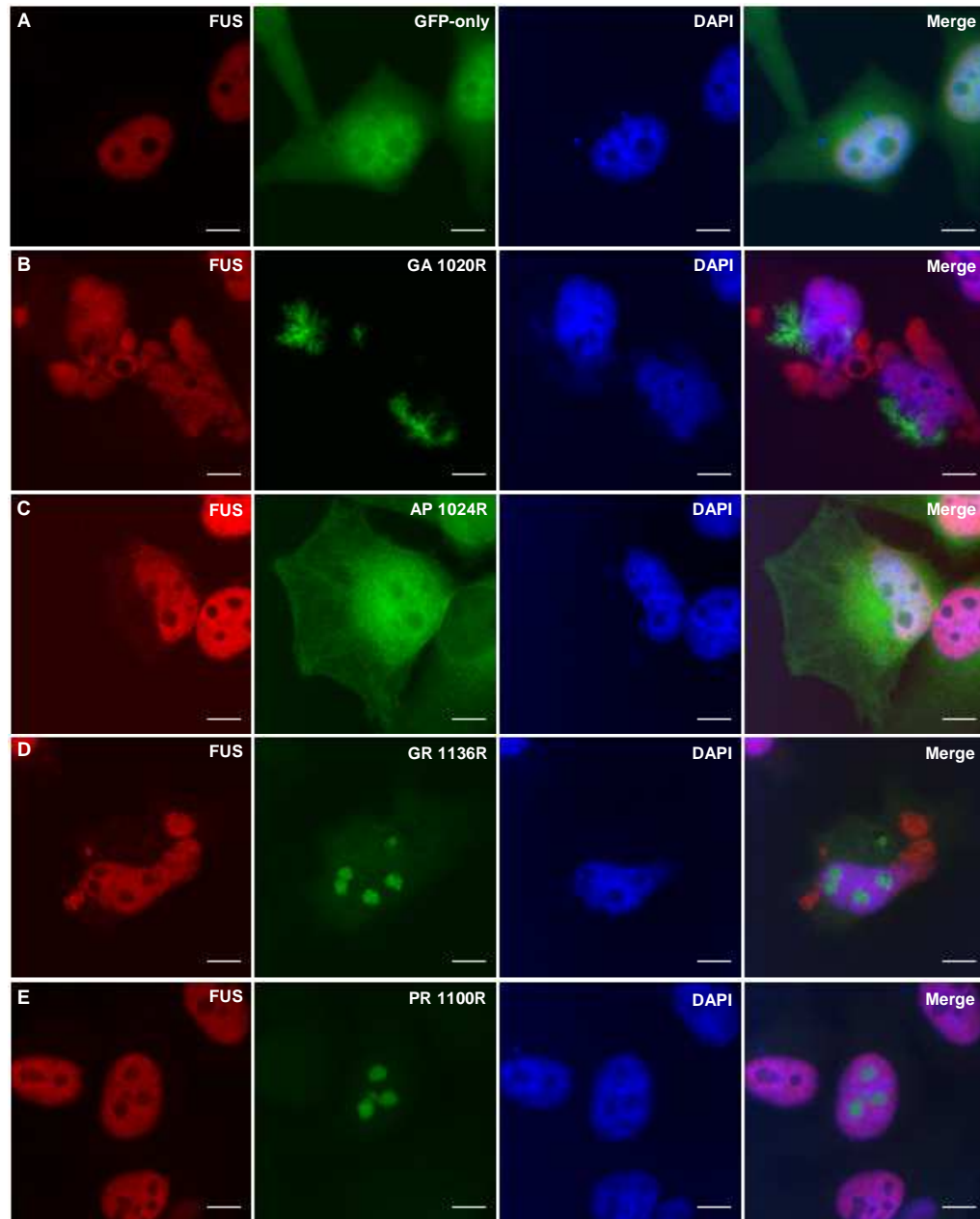


Figure 5.3: Impact of DPR expression (green) on FUS distribution (red) in HeLa cells, 48h post-transfection. pEGFP-N1 expression shown as a GFP-only control (**A**). All DPRs are expressed at the longest lengths available: GA₁₀₂₀ (**B**), AP₁₀₂₄ (**C**), GR₁₁₃₆ (**D**) and PR₁₁₀₀ (**E**). Nuclei are labelled in blue (DAPI). Images captured at x60 magnification. Scale bars represent 15 μ m. n=2 for all transfections

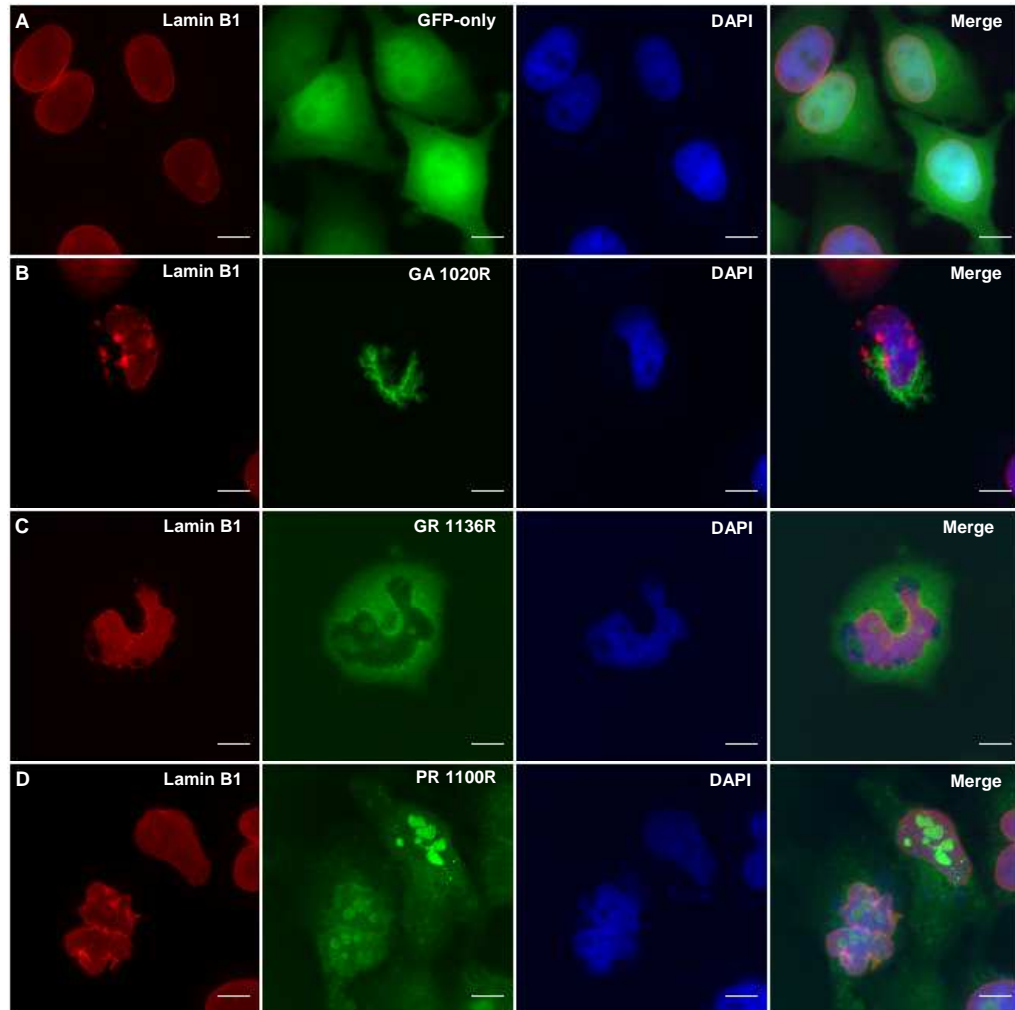


Figure 5.4: Disruption of the nuclear membrane in HeLa cells expressing DPRs. Lamin B1 (red) was used as a marker of the nuclear membrane. **(A)** GFP-only control cells expressing pEGFP-N1. Disruption to the nuclear membrane was observed in cells expressing **(B)** poly-GA_{1020^r} **(C)** poly-GR_{1136^r} and **(D)** poly-PR_{1100^r}. Images captured at x60 magnification, 48h post-transfection. $n=1$ for all transfections Scale bars represent 15 μ m.

5.3.2 Impact of alanine-rich DPRs on cytoplasmic proteins

Since poly-GA and –AP are primarily localised in the cytoplasm and poly-GR and –PR primarily in the nucleus, it is likely that alanine-rich and arginine-rich DPRs may differentially impact cellular processes and function. Furthermore, the common amino acids within these repetitive peptides may cause similarities in structural and functional properties between the two alanine-rich and two arginine-rich DPRs. The remainder of this chapter will therefore report separately on the impact of alanine- and arginine-rich dipeptides on cellular function.

Poly-GA inclusions are known to co-localise with several components of the ubiquitin-protease system (UPS) in C9FTLD/ALS patient tissue (Mackenzie et al. 2013; Mann et al. 2013; Mori et al. 2013). Therefore cells expressing poly-GA were labelled for p62, ubiquitin and ubiquilin-2 (Figure 5.5). Since poly-AP is also cytoplasmic, and UPS proteins are predominantly found in the cytoplasm in patient tissue, poly-AP cells were also investigated. Poly-GA strongly co-localised with p62 (Figure 5.5B) and ubiquilin-2 (Figure 5.5E), but not ubiquitin (Figure 5.6). Poly-AP also co-localised with ubiquilin-2 (Figure 5.5F), but not p62 (Figure 5.5C) or ubiquitin (Figure 5.6).

Since AP₁₀₂₄ forms structures which appear cytoskeletal in nature, cells expressing poly-AP were also labelled for the cytoskeletal proteins α -tubulin and F-actin. However, no co-localisation was observed between poly-AP and either cytoskeletal protein (data not shown). Therefore the cause of the distribution pattern of poly-AP remains unclear.

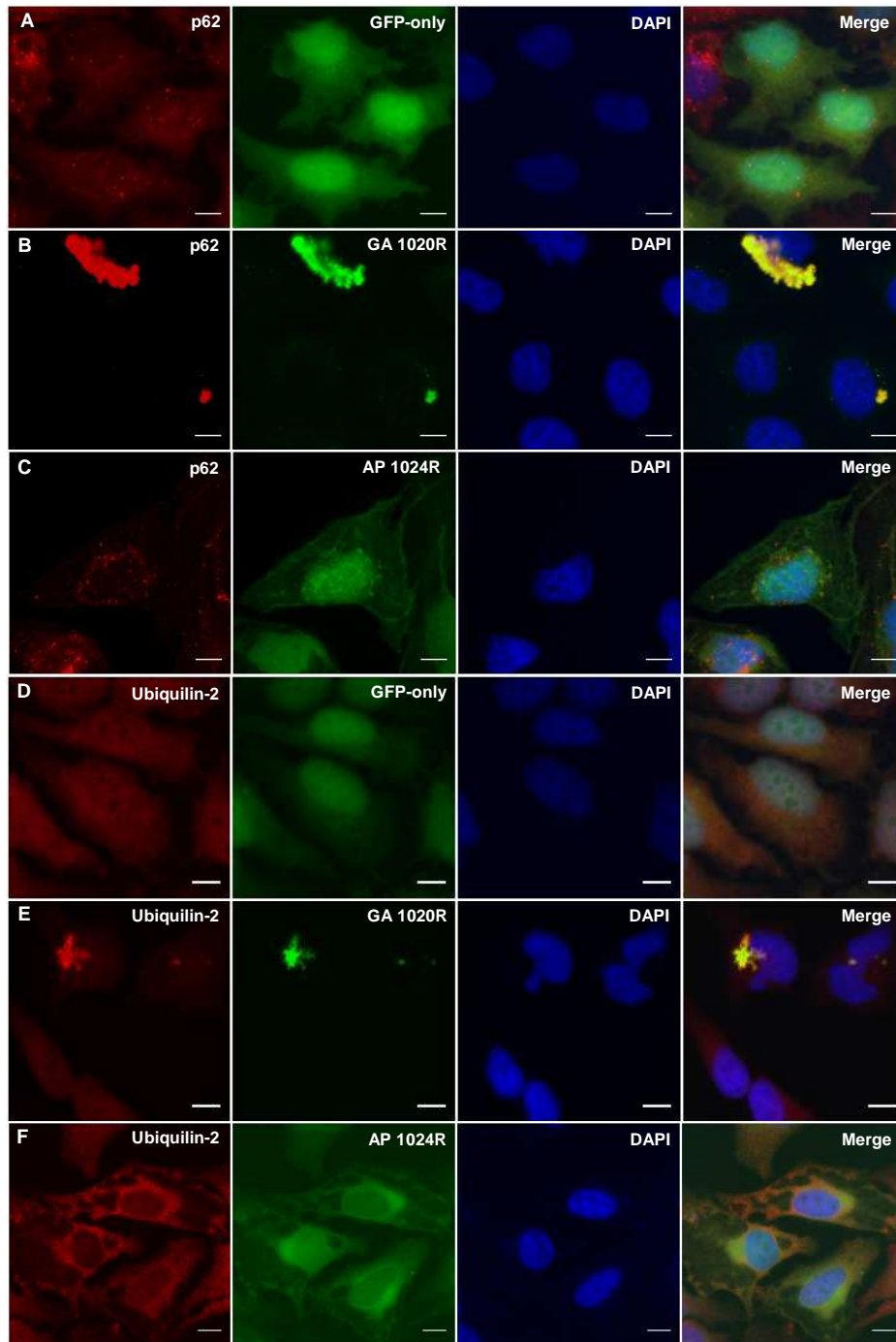


Figure 5.5: Co-localisation of alanine-rich DPRs with components of the ubiquitin-protease system in HeLa cells, 48h post-transfection. GFP-tagged poly-GA at 1020 repeats (labelled GA 1020R; green) co-localised with **(B)** p62 (red; $n=3$) and **(E)** ubiquilin-2 (red; $n=2$). Poly-AP at 1024 repeats (labelled AP 1024R) also co-localised with ubiquilin-2 **(F)** ($n=2$), but not p62 ($n=3$) **(C)**. The distribution of p62 and ubiquilin-2 in GFP-only control cells transfected with pEGFP-N1 are shown in panels **A** and **D** respectively. Nuclei are labelled in blue (NucBlue). Images captured at x60 magnification. Scale bars represent 15 μ m.

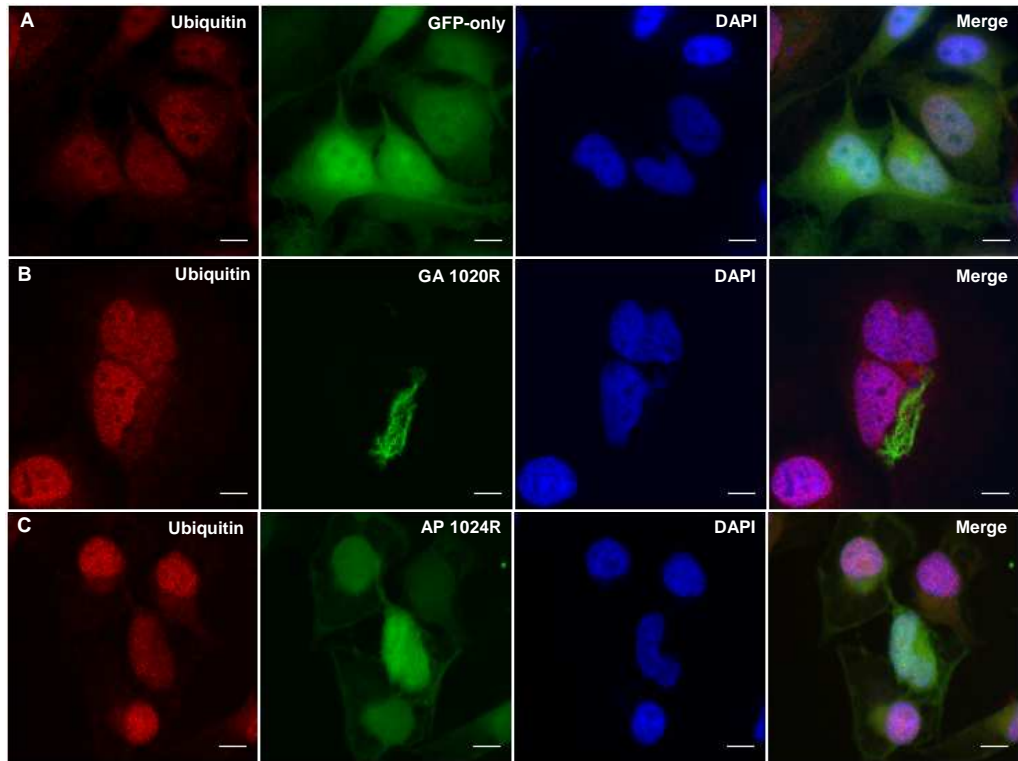


Figure 5.6: Alanine-rich DPRs did not co-localise with ubiquitin in our cellular model. HeLa cells expressing pEGFP-N1 as a GFP-only control (**A**), GA₁₀₂₀ (**B**) or AP₁₀₂₄ (**C**) were labelled with an antibody against ubiquitin (red). n=3. Nuclei are labelled in blue (NucBlue). Images captured at x60 magnification, 48h post-transfection. Scale bars represent 15µm.

5.3.4 Arginine-rich DPRs translocate to the nucleolus, causing nuclear stress

It was reported in Chapter 4 that the arginine-rich dipeptides, poly-GR and poly-PR, form punctate nuclear inclusions when expressed in HeLa cells. Further investigation showed that these inclusions were immunopositive for the nucleolar marker, nucleolin (Figure 5.7). Strong co-localisation was observed between nucleolin and poly-GR at the long and short repeat lengths of 1136 and 36, and between nucleolin and poly-PR at 1100 and 36 repeats in length. However, images captured at different time-points post-transfection (24, 48 or 72h) show that GR₁₁₃₆ and PR₁₁₀₀ also formed punctate structures within the cytoplasm, and that these more frequently occurred at later time-points (example images for poly-GR shown in Figure 5.8). It is possible that the nucleolar DPR structures are exported from the nucleus over time. However, this is a preliminary, subjective observation and further work is required to test this hypothesis.

Several endogenous nucleolar proteins contain repetitive arginine-rich motifs known as arginine-glycine and arginine-glycine-glycine boxes (RG/RGG boxes). Arginine methylation within these motifs acts as a nucleolar localisation signal, perhaps explaining the translocation of poly-GR and poly-PR to the nucleolus. The introduction of a novel poly-GR repeat may therefore impact the function of endogenous proteins containing RG/RGG boxes in a number of ways. To assess this, the effect of poly-GR expression on the nucleolar RG box-containing protein, fibrillarin, was therefore investigated. Cells expressing poly-GR at various repeat-lengths or GFP-only control were labelled with an antibody against fibrillarin, and cells selected for imaging based on the presence of nuclear poly-GR, or diffuse GFP expression for control cells. Fibrillarin images were captured using a Texas Red filter with constant exposure time. Heat-maps showing relative intensity of fibrillarin staining throughout the cell were then produced (Figure 5.9). Areas of relatively high fibrillarin expression are shown in yellow, ranging through to red, purple and finally blue where no fibrillarin signal was detected. GFP-only control cells contained several small yellow “hot-spots” of high intensity fibrillarin staining in the nucleus as expected, presumably corresponding to the nucleoli. Expression of poly-GR at the short repeat length of 36 caused some disruption to these structures, since the yellow areas of high intensity staining became more diffuse but remained nuclear. As repeat-length was increased, the yellow regions became yet more diffuse. In cells expressing GR₁₁₃₆, no yellow “hot-spots” remained. Therefore poly-GR caused loss or re-distribution of fibrillarin in a length-dependent manner. This change in fibrillarin distribution is typically considered a sign of nucleolar stress, and may be detrimental to cellular function and survival.

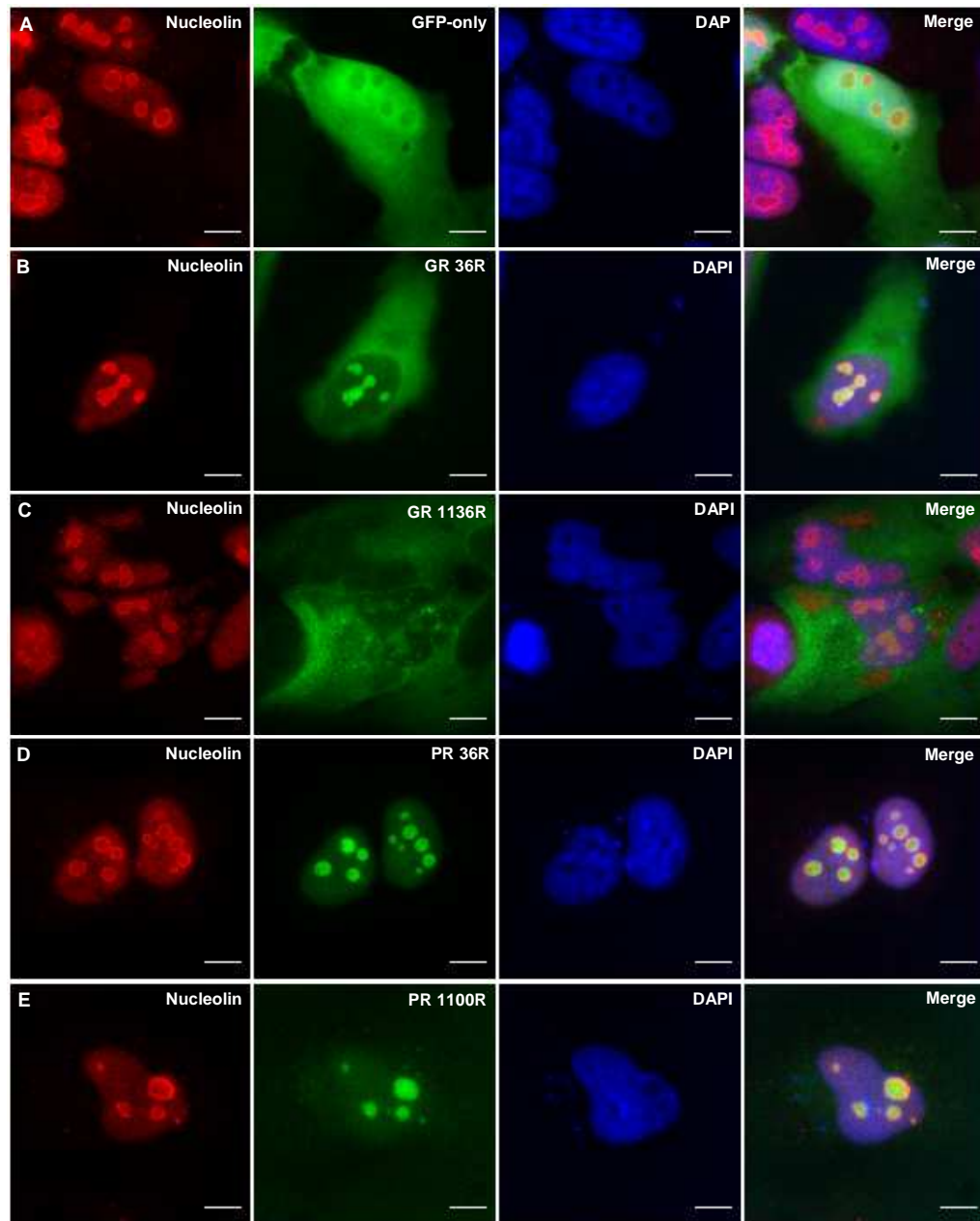


Figure 5.7: Arginine-rich DPRs translocate to the nucleolus in HeLa cells, as demonstrated by co-localisation with the nucleolar marker protein, nucleolin (red). Nucleolin staining in GFP-only control cells expressing pGFP-N1 (green) is shown in **A**. Poly-GR (green) co-localises with nucleolin at the short ($n=3$) and long ($n=1$) lengths of 36 and 1136 repeats (**B** and **C** respectively). Poly-PR (green) also co-localises with nucleolin at the short ($n=3$) and long ($n=3$) lengths of 36 and 1100 repeats (**D** and **E**). Images captured at x60 magnification, 24h post-transfection. Scale bars represent $15\mu\text{m}$.

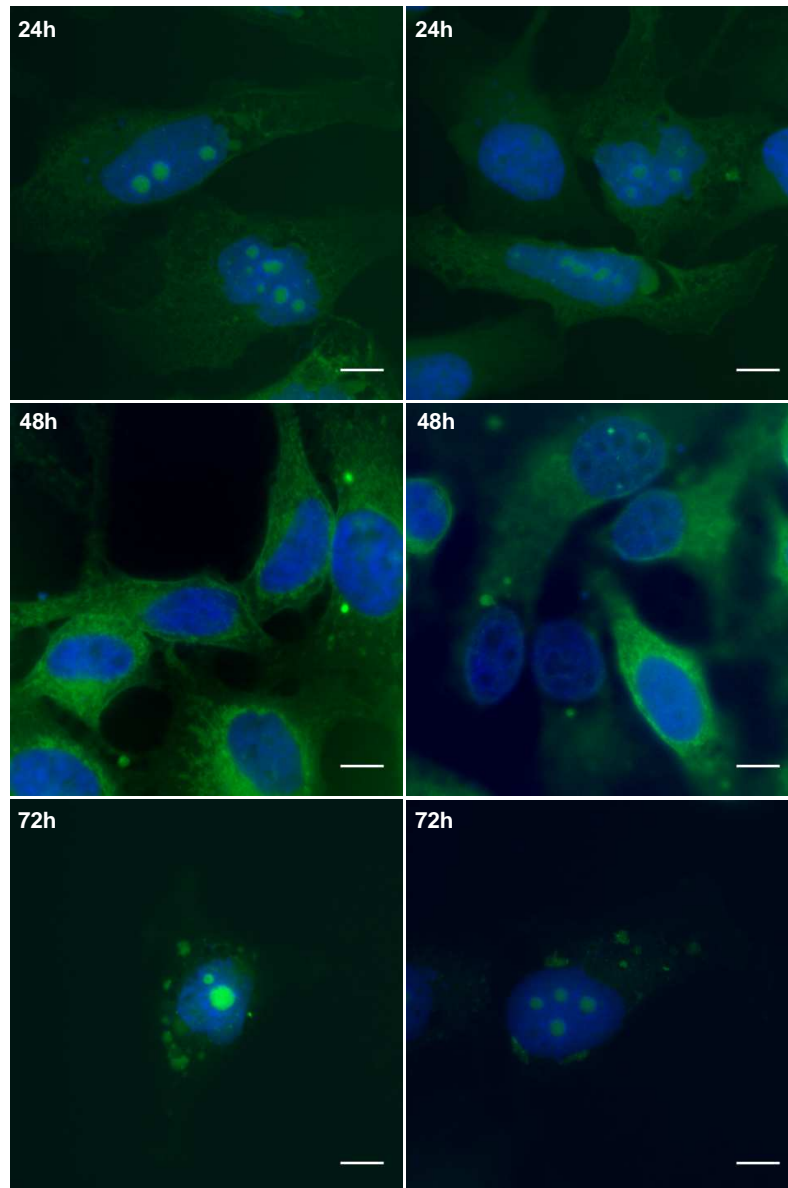


Figure 5.8: Preliminary evidence suggesting arginine-rich DPRs may be exported from the nucleus over time. Example images shown for GR_{1136} (green) at 24, 48 or 72h post-transfection. In addition to diffuse cytoplasmic poly-GR expression, poly-GR forms punctate cytoplasmic inclusions at later time-points. These inclusions are structurally similar to the nucleolar inclusions observed from 24h onwards, suggesting poly-GR may form inclusions in the nucleoli and translocate to the cytoplasm over time. Images captured at x60 magnification in HeLa cells. Scale bars represent $15\mu\text{m}$.

Whilst our DPR constructs do not contain the G₄C₂ sequence and corresponding RNA found in C9FTLS/ALS, their sequences remain somewhat repetitive, with a high percentage of guanine and cytosine pairs. As such, it is possible that RNA transcripts arising from our constructs may have detrimental effects on cellular function. To ensure that the changes observed in fibrillarlin distribution were caused by the poly-GR peptide and not RNA, an “RNA-only” poly-GR construct was generated by removing the start codon from the beginning of the repeat sequence. A start codon in the GFP gene was retained, allowing diffuse GFP expression to be utilised as a marker of successfully transfected cells. Fibrillarlin staining and imaging was performed as before. As expected, yellow “hot-spots” of high fibrillarlin signal were observed in the nucleus in cells expressing the RNA-only construct (Figure 5.9B), and no differences were noted compared to the GFP-only control. Therefore the nucleolar stress phenotype was caused by poly-GR peptide, and not RNA.

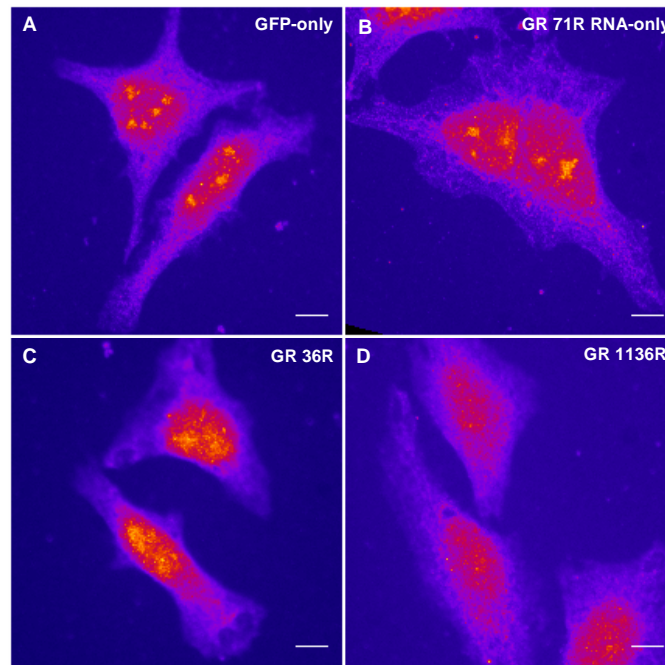


Figure 5.9: Poly-GR peptide, not RNA, causes a nucleolar stress phenotype. Heat-maps showing loss or re-distribution of fibrillarlin in cells expressing nucleolar GR₁₁₃₆ compared to control (n=2). Yellow “hot-spots” indicated areas of high intensity fibrillarlin staining, with lower signal areas displayed in red, purple and blue, in decreasing order. (A) Fibrillarlin forms punctate nuclear “hot-spots” corresponding to the nucleoli in cells expressing pEGFP-N1. Yellow hot-spots become more diffuse with increasing repeat-length of poly-GR from 36 (C) until no yellow is visible at GR₁₁₃₆ (D). (B) Expression of an RNA-only poly-GR construct did not affect fibrillarlin distribution. Images captured in HeLa cells at x60, 48h post-transfection. Scales bars represent 15µm.

Another nuclear protein which contains an endogenous RG box is the Cajal body (CB) marker, coilin. CBs may also be affected by cellular stress (Cioce et al. 2006; Dahl et al. 2008), and therefore the impact of expressing a novel poly-GR repeat peptide on coilin expression was next investigated (Figure 5.10). In control cells, CBs were easily visible since coilin formed very small puncta within the nucleus. The number of coilin-containing bodies varied, but several were visible in almost all cells. In cells expressing nucleolar GR_{1136} , however, the number of coilin-containing bodies was notably reduced or completely depleted. Interestingly, no change in coilin expression was observed in cells exhibiting cytoplasmic but not nucleolar poly-GR expression. Poly-PR did not affect the number of coilin-containing nuclear bodies (data not shown).

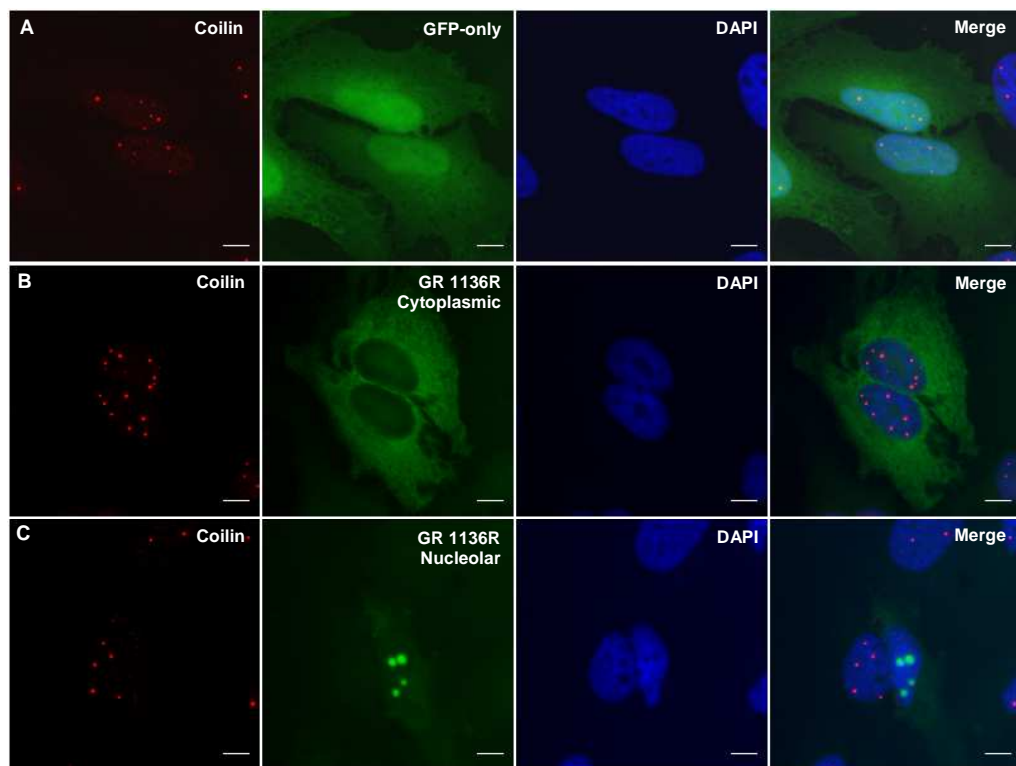


Figure 5.10: Nucleolar, but not cytoplasmic, poly-GR (green) causes a reduction in staining for the Cajal body marker, coilin (red; $n=1$). **(A)** Coilin-positive nuclear bodies in GFP-only control cells expressing pEGFP-N1. **(B)** Coilin staining is unaffected in cells expressing GR_{1136} in the cytoplasm only. **(C)** The number of coilin-positive nuclear bodies is reduced in cells expressing GR_{1136} in the nucleolus. HeLa cells, 48h post-transfection. All images captured at x60 magnification. Scale bars represent $15\mu\text{m}$.

Since this data clearly demonstrates that expression of nucleolar poly-GR causes loss or re-distribution of nuclear proteins containing RG boxes, the impact of this on proteins known to interact with the RG motifs were next considered. Survival motor neuron protein (SMN) is a primarily nuclear protein known to interact with the RG boxes of both coilin and fibrillarin (Tapia et al. 2014; Jones et al. 2001). As the name suggests, SMN is vital for the function and survival of motor neurons, and loss of SMN is known to cause the severe degenerative motor neuron disease, spinal muscular atrophy (SMA; Tisdale & Pellizzoni 2015). As such, the impact of poly-GR expression on SMN function is of particular interest to C9ALS. Cells expressing GR₁₁₃₆ were labelled with an antibody

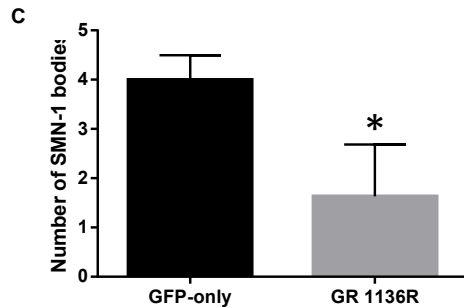
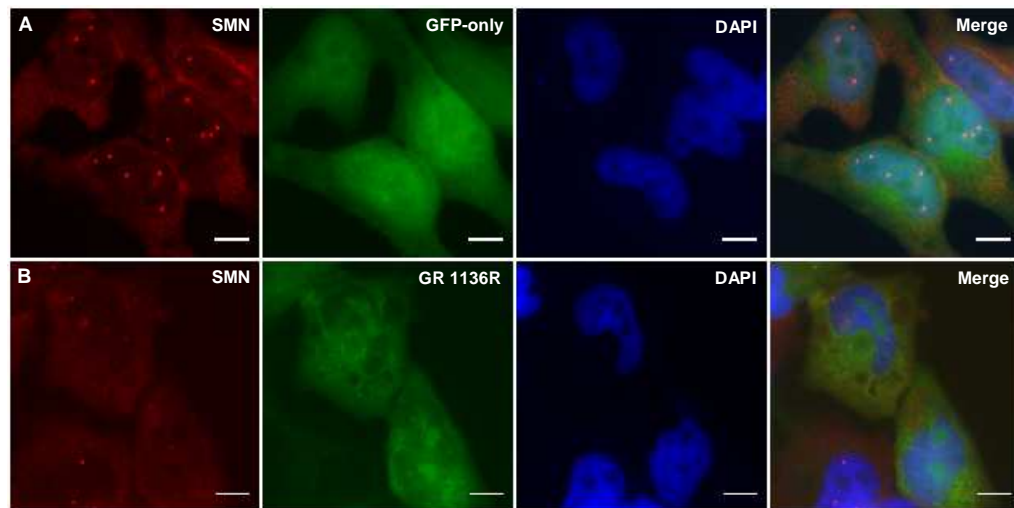


Figure 5.11: Nucleolar poly-GR causes loss of nuclear SMN-containing bodies (red). **(A)** SMN forms small, punctate nuclear bodies which vary in number in cells expressing pEGFP-N1. **(B)** Loss of punctate SMN in cells expressing nucleolar poly-GR. **(C)** Semi-quantitative analysis of SMN images showing significant reduction in number of visible SMN-containing bodies in nucleus. Mean values of at least 10 technical replicates were used as independent replicates for analysis ($n=3$) and compared by unpaired t -test ($P=0.4$). Error bars indicate standard deviation. Images captured at $\times 60$ in HeLa cells, 48h post-transfection. Scale bars represent $15\mu\text{m}$.

against SMN, and those exhibiting nucleolar poly-GR expression were selected for imaging (Figure 5.11). SMN is abundant in CBs, where it co-localises with coilin. Therefore GFP-only control cells contained varying numbers of small, punctate, SMN-containing nuclear bodies. The number of SMN-containing nuclear bodies in cells expressing nucleolar poly-GR also varied, however, a clear reduction was observed compared to control cells (Figure 5.11B and C). Semi-quantitative analysis of SMN images demonstrated that this reduction was statistically significant ($P=0.04$; $n=3$). Therefore the presence of nucleolar poly-GR caused a protein known to be crucial for the survival of motor neurons to be depleted from the nucleus.

5.4 Discussion

This chapter used immunofluorescence imaging to identify a number of important effects of DPR expression on cellular function *in vitro*. Poly-GA, -GR and -PR were of particular interest, since all 3 peptides caused disruption to the nuclear membrane, presumably leading to mislocalisation of FUS from the nucleus to the cytoplasm. Conversely, TDP-43 and EWS did not mislocalise to cytoplasm in DPR-expressing cells. Poly-GA and poly-AP were found to co-localise with important components of the UPS, suggesting a potential role of impaired protein degradation in C9FTLD/ALS. Finally, poly-GR and poly-PR translocated to the nucleolus, where poly-GR in particular caused nucleolar stress and loss of the nuclear proteins, fibrillarin, coilin and SMN from the nucleus. This section will discuss the implications of these findings in context of the recent literature regarding the impact of DPRs on cellular function and survival.

5.4.1 Existing models of DPR pathology

Since this project commenced, several independent groups have also described models of dipeptide pathology generated in a similar way to ours, using alternative codons in a randomised fashion. Mizielińska *et al.* (2014) generated constructs containing pure G₄C₂ repeats, “RNA-only” constructs containing G₄C₂ interrupted by stop codons at regular intervals in all frames, and “protein-only” constructs similar to those reported here. Expression of pure G₄C₂ constructs of 36 or 103 repeats in length in *Drosophila* caused degeneration of the eye and reduced life-span in a repeat length-dependent manner. RNA-only constructs did not cause a degenerative phenotype or affect survival even when expressed at the longer length of 288 repeats, suggesting that the toxic effects were due to RAN-translation of the pure repeats. Inhibition of translation using low dose cycloheximide partially rescued toxicity, further implying that DPR production rather than repetitive RNA was the primary cause of lethality. DPR constructs of 36 or 100 repeats were then expressed in fly. Interestingly, poly-PR and -GR were found to be highly toxic, causing degeneration of the eye and reduced survival, whereas poly-GA only exerted a mild toxic effect at 100 repeats, and poly-AP was not toxic at either length. The toxicity of arginine-rich repeats was again length-dependent, with 100 repeat constructs causing a more severe phenotype than 36 repeats (Mizielińska *et al.* 2014). The finding that arginine-rich DPRs are most severely toxic, with poly-GA also exhibiting toxic effects albeit to a lesser extent, have been replicated by several studies. Yang *et al.* also demonstrated that expression of poly-GR and -PR at 80 repeats caused a

degenerative phenotype in the fly eye, as well as impaired motility and survival (Yang et al. 2015). In cell culture, expression of poly-GR and -PR has been shown to be toxic in HEK293 cells, mouse spinal cord neuroblastoma cells and mouse primary cortical, hippocampal and motor neurons, whereas poly-GA or -AP had no significant effect in these cell types (Tao et al. 2015; Wen et al. 2014). In addition, external application of synthetic poly-GR and -PR peptides was toxic to human osteosarcoma cells *in vitro* (Kwon et al. 2014). Other studies have focused primarily on poly-GA as the most easily detectable DPR in patient tissue; Poly-GA was found to modestly increase caspase-3 activation and toxicity in HEK293 cells and mouse primary cortical neurons (May et al. 2014; Zhang et al. 2014).

Taken together, these studies provide strong evidence for a toxic effect of arginine-rich DPRs on multiple cell types, including neurons. A less significantly toxic role of poly-GA is also likely. The major limitation of these studies is the relatively short repeat length assessed, since the severity of toxic phenotypes appears to be length-dependent (Mizielinska et al. 2014). Quantification of toxicity in our models expressing DPRs at physiologically-relevant repeat-lengths is therefore an urgent priority. However, our finding that poly-GA, -GR and -PR caused structural abnormalities in the nucleus tentatively supports the conclusion that these peptides are the most toxic. Poly-GR, -PR and -GA were all reported to cause abnormalities in nuclear structure, with many nuclei appearing elongated or “U-shaped”. Furthermore, immunolabelling of lamin B1 demonstrated that all 3 DPRs caused fragmentation of the nuclear membrane, and the FTL/ALS-related protein FUS was mislocalised to the cytoplasm in many cells most likely as a result of this. It is important to note that fragmentation of the nucleus and FUS mislocalisation were also observed in some control cells, although less frequently. Therefore it is likely that these cells are undergoing apoptosis or cell-cycle arrest, and the increased incidence of nuclear abnormalities in cells expressing poly-GR, poly-PR and poly-GA most likely represents increased toxicity in these populations.

5.4.2 Impact of cytoplasmic alanine-rich DPRs on cellular function

In Chapter 4, alanine-rich DPRs were found to be primarily cytoplasmic, with poly-GA forming star-shaped inclusions, and poly-AP forming an unusual fibrous pattern resembling the cytoskeleton. The work presented in this chapter aimed to further characterise these pathological features, by immunolabelling proteins predicted to co-localise with either peptide. In order to determine whether the cytoskeleton-like expression pattern of poly-AP is caused by binding to microtubules or microfilaments,

cells expressing AP₁₀₂₄ were labelled for α -tubulin or F-actin. However, no co-localisation was observed between poly-AP and either cytoskeletal protein (data not shown). It is possible that poly-AP binds to a cytoskeletal protein more specifically expressed in HeLa cells, such as keratin. Alternatively, the observed distribution pattern may simply be an artefact; no observations of a poly-AP distribution pattern similar to that shown in Figure 4.6-7 have yet been reported in patient tissue. This may be difficult to detect *ex vivo* if the peptide is expressed at low levels or also diffusely present in the cytosol. Therefore an important future experiment will be to express AP₁₀₂₄ in neuronal cell culture, in order to determine whether the distribution pattern is cell-specific.

It is well documented that poly-GA co-aggregates with components of the UPS in patient tissue, most notably p62, ubiquilin-2 and ubiquitin (Mackenzie et al. 2013; Mann et al. 2013; Mori et al. 2013). Indeed, prior to the discovery of RAN-translated DPRs, poly-GA aggregates were first observed as insoluble neuronal cytoplasmic inclusions which were p62 positive, TDP-43-negative and present only in C9FTLD/ALS cases (Al-Sarraj et al. 2011; Bigio et al. 2013; Mahoney et al. 2012). Co-localisation of poly-GA and p62 has also been demonstrated in HEK293 cells and rodent primary neuronal cultures (May et al. 2014; Zhang et al. 2014). Furthermore, co-immunoprecipitation of poly-GA in primary cortical neurons followed by mass spectrometry identified p62, ubiquilin-1, -2 and -4 and the proteasomal subunit PSMB6 as poly-GA interactors (May et al. 2014).

The finding that poly-GA also co-localised with p62 and ubiquilin-2 in HeLa cells provided further validation of our *in vitro* model as relevant to human disease. However, ubiquitin did not co-localise with GA₁₀₂₀ in HeLa cells at 48h post-transfection. The reason for this discrepancy with patient tissue is unclear. Earlier work in the Pickering-Brown lab showed that ubiquitin did co-localise with poly-GA at the shorter repeat length of 102, when expressed in H4 cells (data not shown). GA₁₀₂ also did not co-localise with ubiquitin in HeLa cells (data not shown), suggesting that this may be a cell-specific difference. Despite this, poly-GA aggregates in HeLa do encapsulate many features of those observed in patient tissue, including their fibrous, star-like structure, cytoplasmic localisation and co-localisation with p62 and ubiquilin-2. Interestingly, poly-AP also co-localised with ubiquilin-2 in our model. As discussed above, however, it is important to note that until the distribution pattern of AP₁₀₂₄ is validated in a neuronal cell type, conclusions regarding the impact of poly-AP on cellular function must be drawn with caution.

The sequestration of UPS proteins by poly-GA, and perhaps poly-AP, implicates a role of impaired protein degradation systems in C9FTLD/ALS pathogenesis. Impairments in

both the UPS and autophagy have been previously implicated in many neurodegenerative diseases, including FTLD/ALS (reviewed by Chen et al. 2012; Dantuma & Bott, 2014; Majcher, Goode, James, & Layfield, 2015; Wolfe et al., 2013). This is particularly well studied in other repeat expansion disorders such as Huntington's disease, where long repetitive peptides are thought to overwhelm protein degradation systems (reviewed by Li & Li, 2011; Ortega & Lucas, 2014). The identification of mutations in the genes encoding ubiquilin-2, p62 and VCP in FTLD/ALS provide further evidence for a role of UPS dysfunction (Deng et al. 2011; Johnson et al. 2010; Schröder et al. 2005; Kaleem et al. 2007; Rubino et al. 2012; Miller et al. 2015). Compromised protein degradation systems may contribute to neuronal dysfunction and toxicity by preventing various regulatory and signalling proteins from normal degradation, affecting numerous important pathways and cellular functions.

Interestingly, Zhang *et al.* (2014) found that poly-GA expression in mouse primary cortical neurons caused a reduction in proteasome activity as well as endoplasmic reticulum (ER) stress, which is known to be a downstream consequence of proteasomal inhibition, leading to apoptosis (Park et al. 2011; Obeng et al. 2006; Fribley et al. 2004). Inhibition of ER-triggered cell death pathways by treatment with salubrinal partially rescued toxicity caused by poly-GA expression in these cells (Zhang et al. 2014). The authors therefore proposed that poly-GA inhibits proteasome function by sequestration of key UPS proteins, leading to ER stress and premature apoptosis. iPSC-derived motor neurons from C9FTLD/ALS patients have previously been reported to exhibit increased toxicity in response to treatment with tunicamycin, an agent known to induce ER stress (Haeusler et al. 2014), further supporting this hypothesis. Further work to assess UPS function and ER stress in our poly-GA expressing cells would therefore be of interest. However, since the majority of studies assessing DPR toxicity to date conclude that arginine-rich DPRs are the most toxic, it is likely that poly-GA-induced ER stress is not the primary toxic event in C9FTLD/ALS pathogenesis.

5.4.3 Arginine-rich DPRs translocate to the nucleolus, causing nuclear stress

In Figure 5.7, nuclear poly-GR and –PR inclusions were found to co-localise with the nucleolar marker protein, nucleolin, indicating that both arginine-rich DPRs translocated to the nucleolus. This finding has been replicated in other *in vitro* DPR models, including human osteosarcoma and astrocytes (Kwon et al. 2014), mouse primary cortical neurons and human iPSC-derived neurons (Wen et al. 2014), and mouse spinal cord neuroblastoma cells and HEK293 cells (Tao et al. 2015) expressing poly-GR or –PR at

varying repeat-lengths. Interestingly, many nuclear proteins contain one or more glycine-arginine-rich (GAR) domains known as RGG/RG boxes, and the methylation status of arginine within these regions is considered vital for localisation of such proteins to the nucleus, and commonly the nucleolus specifically (Thandapani et al. 2013). Examples of such proteins include RNA-binding proteins such as FUS and EWS, as well as the CB marker, coilin, and nucleolar proteins fibrillarin and nucleolin (Thandapani et al. 2013). Therefore it is likely that arginine-rich DPRs mimic RGG/RG box-containing proteins, resulting in their nucleolar localisation.

The effects of abnormal arginine-rich peptides on endogenous GAR domain-containing proteins were next considered. Figure 5.9 demonstrates that the RGG box-containing protein, fibrillarin, was depleted from the nucleolus following poly-GR expression. This finding is complimented by several studies demonstrating the loss or mislocalisation of other nucleolar proteins such as nucleolin and nucleophosmin in cells expressing arginine-rich DPRs (Wen et al. 2014; Tao et al. 2015). In addition, cells expressing nuclear poly-PR were found to exhibit nucleolar swelling, as well as an increase in cytoplasmic P-bodies and stress granules (Wen et al. 2014; Tao et al. 2015). Furthermore, B-lymphocytes and iPSC-derived motor neurons from C9FTLD/ALS patients also exhibited a more diffuse pattern of nucleolin expression, validating this observation as highly disease-relevant. Interestingly fibrillarin knockdown has previously been reported to cause abnormalities in nuclear shape in HeLa cells (Amin et al. 2007), which may partially explain the structural changes observed in poly-GR expressing cells (Figure 5.4). Furthermore, depletion of nucleolin has recently been reported to suppress poly-GR and –PR toxicity in yeast (Jovičić et al. 2015), suggesting that displacement of nucleolin may exert a gain of function toxicity mechanism in the cytoplasm or nucleoplasm. The small amount of nucleolin found in the nucleoplasm and cytoplasm under normal conditions plays an important role in the stabilisation of anti-apoptotic proteins (Berger et al. 2015; Thandapani et al. 2013). However, re-distribution of both nucleolin and nucleophosmin from the nucleolus has been previously reported as an early event in apoptosis following various toxic insults *in vitro*, suggesting that the amount of these proteins present in the cytoplasm may be crucial in determining their role in pro- or anti-apoptotic pathways (Lindenboim et al. 2010). Taken together, these findings strongly suggest that arginine-rich DPRs cause nucleolar stress, which may trigger apoptotic pathways and could explain the increased toxicity of arginine-rich compared to alanine-rich DPRs.

In contradiction of these findings, Schludi *et al.* (2015) recently reported that poly-GR and –PR did not co-localise with nucleolin or fibrillarin, but that the majority of nuclear

DPR inclusions were immediately adjacent to nucleoli. However, this was not convincingly demonstrated since very faint poly-GR staining which appeared to co-localise with fibrillarin was visible, adjacent to a brighter poly-GR inclusion in the Supplemental Figures (Schludi et al. 2015). It is possible that nucleolar DPRs may be more difficult to detect in patient tissue, given their short half-lives (demonstrated *in vitro* by Kwon et al., 2014). Furthermore, nuclear DPRs are more likely to be soluble and form an inclusion-like appearance through interaction with other proteins rather than aggregation, since this is how nuclear bodies such as nucleoli remain intact. Since post-translational modifications such as arginine methylation, which are required for this type of interaction, may not be well-preserved in post-mortem tissue, data from small sample sizes such as this (tissue was obtained from 3 patients) must be analysed with caution. Finally, it is possible that cells expressing nucleolar poly-GR and –PR are least likely to survive, and therefore difficult to observe in patient tissue samples which represent the end-stage of disease. Further analysis of patient tissue with larger sample size would therefore be of great value to determine whether arginine-rich DPRs are present in, or merely adjacent to, nucleoli in C9FTLD/ALS brain.

Many proteins containing RGG/RG boxes are found in other nuclear bodies besides the nucleolus. The CB marker, coilin, is one such protein. The data presented in Figure 5.10 demonstrated that the number of coilin-positive CBs in the nucleus is dramatically reduced in cells expressing nucleolar poly-GR. Therefore poly-GR-induced loss or redistribution of GAR domain-containing proteins extends beyond the nucleolus. This implies that the downstream effects of poly-GR expression may be extensive and varied. Interestingly, loss of CBs was not observed in cells expressing nucleolar poly-PR (data not shown), suggesting that despite the structural similarities and subcellular localisation shared by the two arginine-rich DPRs, poly-GR and poly-PR may differentially impact cellular function.

In order to investigate the downstream effects of poly-GR-induced loss of nuclear proteins, interactors of fibrillarin and coilin were next considered. SMN is an important nuclear protein which is crucial for the survival of motor neurons. It contains a tudor domain which interacts with symmetrical dimethylated arginine residues in the RGG/RG boxes of both coilin and fibrillarin, amongst other proteins (Pellizzoni et al. 2001; Thandapani et al. 2013; Boisvert et al. 2005; Jones et al. 2001). Therefore the impact of poly-GR expression on SMN was investigated. Figure 5.11 shows that the presence of nucleolar poly-GR reduced the number of SMN-containing nuclear bodies in HeLa cells. This is of particular interest to C9ALS, since it is well established that haploinsufficiency

of SMN causes degeneration of motor neurons in SMA (reviewed by Anderton et al. 2013; Tisdale & Pellizzoni, 2015).

The SMN-containing nuclear bodies lost in poly-GR expressing cells are most likely CBs, since SMN is primarily located in CBs under normal conditions, where it interacts with coilin. SMN may also be present in coilin-negative nuclear puncta known as gem bodies, and is present at lower levels in the cytoplasm and nucleolus. Whilst the functions of SMN are not entirely understood, the protein appears to act as a shuttle to transport proteins between the cytoplasm and nucleus, in order to assemble the ribonucleoprotein complexes required for pre-mRNA splicing (Thandapani et al. 2013). Furthermore, SMN is required for CB formation, and depletion of SMN leads to subsequent depletion of coilin-positive CBs in transgenic mice and cell culture (Turner et al. 2008; Girard et al. 2006). The SMN tudor domain is required for the formation of CBs, indicating that the interaction between SMN and the RGG/RG domain in coilin is critical (Tapia et al. 2014). Therefore it is possible that loss of SMN from the nucleus is the primary event, and coilin bodies are depleted as a result of this.

The nucleolar localisation of poly-GR indicates that it is likely to contain dimethylated arginine residues, thus rendering it capable of interacting with the SMN tudor domain. This may create competition for normal binding partners of SMN such as coilin, preventing formation of CBs and maintaining a more diffuse distribution of both SMN and coilin throughout the nucleoplasm. Conversely, poly-GR may create competition for binding to methyltransferase enzymes, preventing methylation of RGG/RG boxes in other proteins and thus interrupting protein-protein interactions and the formation of nuclear bodies. Finally, it is possible that by preventing SMN from binding to coilin and forming CBs, poly-GR triggers stress pathways resulting in downregulation of both proteins. The data presented here is insufficient to determine whether the changes in coilin and SMN staining are due to mislocalisation or overall depletion of either protein.

Nonetheless, our findings strongly suggest that dysfunction of SMN and CBs, as well as nucleolar stress, are implicated in poly-GR-mediated cytotoxicity. The CB is vital for regulation of pre-mRNA splicing, and the nucleolus required for ribosomal assembly. Therefore poly-GR may induce neurotoxicity by impairing control of RNA processing. Aberrant RNA processing has been previously linked to FTL/ALS pathogenesis; mutations in the RNA-binding proteins TDP-43 and FUS are causative in some cases (Borroni et al. 2009; Kabashi et al. 2008; Benajiba et al. 2009; Yan et al. 2010; Vance et al. 2009), and TDP-43 is the most common form of insoluble protein aggregate found in patient tissue (Neumann et al. 2006; Mackenzie et al. 2007; Arai et al. 2006).

Furthermore, RNA sequencing in cells expressing poly-PR identified abnormal splicing of a number of genes (Kwon et al. 2014), and qRT-PCR has demonstrated that both poly-PR and –GR caused a reduction in expression levels of several ribosomal proteins (Tao et al. 2015). Finally, RNA sequencing of patient tissue recently identified a range of splicing defects and other RNA processing abnormalities in C9ALS brain, which were not present in non-C9ALS or healthy controls (Prudencio et al. 2015). Therefore strong evidence exists supporting a role of aberrant RNA processing in C9FTLD/ALS pathogenesis, and the data presented in this chapter suggests that arginine-rich DPRs, particularly poly-GR, may contribute to this dysfunction through loss of key proteins.

Large-scale screening for genetic modifiers of G₄C₂-induced toxicity has recently identified aberrant nucleocytoplasmic transport of RNA and proteins as key to C9FTLD/ALS pathogenesis (Jovičić et al. 2015; Zhang et al. 2015; Freibaum et al. 2015). Freibaum and colleagues expressed G₄C₂ repeats in *Drosophila* and observed a severe toxicity phenotype. These flies were then crossed into 372 different transgenic strains, each exhibiting a different chromosomal deficiency, in order to identify genes which modified G₄C₂-induced toxicity when suppressed. Partial loss of two proteins involved in nuclear export of mRNA, NFX1 and CHTOP, was found to increase toxicity, suggesting that mRNA export may be impaired following G₄C₂ expression (Freibaum et al. 2015).

CHTOP, along with the nuclear export factor, REF/ALY, forms part of the transcription export complex (TREX), which mediates nuclear export of spliced mRNAs. CHTOP and REF/ALY selectively bind to spliced mRNAs. Interestingly, the RNA-binding domains of these proteins contain methylated GAR domains which are required for the interaction, and therefore arginine-rich DPRs may impact their function. CHTOP and REF/ALY then bind to NFX1, facilitating transfer of the mRNA to the methylated GAR domain in NFX1. NFX1 interacts with the nuclear pore via an intermediary protein, NXT1, and the mRNA is transported through the pore.

In contradiction to the hypothesis that this mRNA export pathway is impaired following G₄C₂ expression, depletion of REF/ALY was found to reduce toxicity in fly (Freibaum et al. 2015). However, expression of G₄C₂ increased sensitivity of fly motor neurons to leptomycin B, an inhibitor of nuclear export. Furthermore, an increase in the amount of RNA present in the nucleus compared to the cytoplasm was observed in HeLa and HEK293 cells expressing G₄C₂ repeats, as well as in iPSC-derived cortical neurons from C9FTLD/ALS patients (Freibaum et al., 2015). The authors propose the contradictory finding that REF/ALY depletion suppressed toxicity may be explained by changes in

nuclear exosome function; REF/ALY binding protects mRNAs from exosomal degradation, and depletion of exosomal proteins was found to increase toxicity in the *Drosophila* model (Freibaum et al. 2015). It is possible that impaired nuclear export is toxic via two distinct mechanisms, by loss of mRNAs at cytoplasmic translation sites, and by a gain of function toxicity mechanism caused by excessive nuclear mRNA retention. Therefore downregulation of REF/ALY and exosomal proteins may increase G₄C₂ toxicity by preventing degradation of mRNA species accumulating in the nucleus. These findings further support a role of aberrant RNA processing in C9FTLD/ALS pathogenesis.

Impaired nucleocytoplasmic transport of proteins has also been recently observed in cells and *Drosophila* expressing G₄C₂ repeats. In their large-scale screen using transgenic flies, Freibaum and colleagues also identified several genes required for nuclear import of proteins to act as modifiers of G₄C₂-induced toxicity. An independent group using a similar method of crossing G₄C₂-expressing flies into various mutant strains also identified genes involved in nucleocytoplasmic transport of proteins as modifiers of toxicity. Furthermore, a large-scale screening study recently performed in yeast expressing poly-PR produced similar results (Jovičić et al. 2015).

Like mRNAs, proteins are transported across the nuclear membrane via the nuclear pore complex. However, the mechanisms of transport differ between RNA and protein. The majority of proteins destined for nuclear import contain a nuclear localisation signal (NLS) which is recognised by import factors such as karyopherins, either directly or via adapter proteins. Import factors then utilise the Ran cycle to provide the energy required to shuttle cargo proteins across the membrane. Ran is a small GTPase which is predominantly bound to GDP in the cytoplasm (RanGDP) and GTP in the nucleus (RanGTP). RanGDP binds to cargo-import factor complexes in the cytoplasm, and is transported through the membrane pore. Once in the nucleus, the Ran guanine nucleotide exchange factor (RanGEF) catalyses the phosphorylation of RanGDP to RanGTP, encouraging the detachment of cargo proteins from their import factors. Nuclear pore complex proteins (nucleoporins) such as Nup50 promote disassociation of adapter proteins from cargo (reviewed by Cautain et al. 2015).

Depletion of karyopherin β₂, Ran, RanGEF, Nup50 and a second nucleoporin required for karyopherin docking to the NPC, Nup153, were all found to enhance eye degeneration in flies expressing G₄C₂, suggesting that nuclear import of proteins via the karyopherin/Ran system may contribute to toxicity (Freibaum et al. 2015; Zhang et al. 2015). In support of this hypothesis, overexpression of karyopherin α and RanGTP

reduced G₄C₂ toxicity in flies (Zhang et al. 2015) and overexpression of 6 different members of the karyopherin family reduced toxicity in yeast expressing G₄C₂ repeats (Jovičić et al. 2015). Interestingly, depletion of nuclear export factor, exportin, also reduced toxicity in G₄C₂-expressing flies (Zhang et al. 2015). Therefore toxicity was rescued by either upregulation of nuclear import or downregulation of nuclear export. Immunofluorescence imaging of *Drosophila* S2 cells and iPSC-derived neurons obtained from C9FTLD/ALS patients showed mislocalisation of several proteins involved in nucleocytoplasmic transport; RanGEF was lost from the nucleus, and several nucleoporins as well as the regulator of RanGTP hydrolysis, RanGAP, formed perinuclear inclusions in many cells (Freibaum et al. 2015; Jovičić et al. 2015; Zhang et al. 2015). Taken together, these observations suggest that nuclear import of proteins is impaired in C9FTLD/ALS, and may explain the loss or mislocalisation of nuclear proteins such as FUS, coilin, SMN and fibrillarin reported in this chapter.

The complimentary findings from yeast, *Drosophila*, cultured neurons and patient tissue provide persuasive evidence for impaired nucleocytoplasmic transport of both proteins and RNA in C9FTLD/ALS. Since two of the three studies described above were performed in models expressing pure G₄C₂ repeats, or in cells and tissue obtained from C9FTLD/ALS patients, the respective roles of DPRs and expanded RNA remain unclear. However, the data provided from yeast cells expressing poly-PR strongly suggests that arginine-rich DPRs play a major role in disruption of nucleocytoplasmic transport (Jovičić et al. 2015). Analysis of nucleocytoplasmic transport pathways in our DPR models would therefore be of value.

In addition to the effects of poly-GR expression reported here, our finding that poly-GR severely impacts the localisation of RGG/RG box-containing proteins has much wider implications for cellular functions which might also be affected. The RGG/RG domain is a key signalling motif with important roles in regulation of transcription and translation, nucleocytoplasmic transport, response to DNA damage and splicing control. RGG/RG boxes are also involved in apoptotic signalling; The GAR domain in nucleolin is required for stabilisation of the anti-apoptotic protein, Bcl-xl, for example (Thandapani et al. 2013). Therefore the downstream effects of poly-GR expression are likely to be multifaceted, and may contribute to neurotoxicity in C9FTLD/ALS via a number of distinct mechanisms.

5.4.4 Summary

This chapter aimed to investigate the impact of DPRs of physiologically-relevant repeat-length on cellular function. Poly-GA, -GR and -PR were found to exhibit abnormalities in nuclear structure, implying increased cell death or cell-cycle arrest. Disruption to the nuclear membrane was accompanied by mislocalisation of FUS to the cytoplasm, although the implications of this remain unclear. Poly-GA was found to sequester important components of the UPS. When considered alongside recent literature, it is proposed that this may inhibit proteasomal function, leading to ER stress and activation of apoptotic pathways. Arginine-rich DPRs have been consistently reported as the most toxic in other cellular and *in vivo* models. Poly-PR and -GR translocate to the nucleolus and cause nucleolar stress in *in vitro* models, which may account for their toxicity. Our findings highlight a potential mechanism for poly-GR-induced toxicity in particular, in which arginine residues are methylated and act as RGG/RG boxes, disrupting nuclear function in a number of ways. Of particular interest is the loss of SMN-positive nuclear bodies following poly-GR expression, which may account for the selective vulnerability of motor neurons in C9ALS and C9FTLD-ALS. Finally, the noted disruptions to nucleoli and CBs strongly implicate a role of aberrant RNA processing triggered by poly-GR expression in C9FTLD/ALS.

Chapter 6: General discussion

6.1 Summary of main findings

The aim of this project was to generate a series of models which encapsulated different features of C9FTLD/ALS, and which could be utilised to investigate mechanisms of toxicity caused by the C9orf72 repeat expansion. C9orf72 mRNA expression is reduced in expansion-carrying patients, suggesting that loss of C9orf72 function may cause neurodegeneration in C9FTLD/ALS. Chapter 3 therefore described the development and characterisation of a transgenic *C. elegans* strain which did not express the worm orthologue of C9orf72. Behavioural analysis showed that no motility or survival impairments were present in this null mutant worm, demonstrating that loss of C9orf72 function was insufficient to cause a phenotype reminiscent of ALS in an *in vivo* model. This finding was consistent with the overwhelming body of literature emerging to support a toxic gain of function mechanism as causative in C9FTLD/ALS. The respective roles of expanded C9orf72 RNA and DPRs arising from the expansion remain unclear, however. The relevance of loss of function mechanisms of toxicity in other repeat expansion disorders was touched upon in Chapter 3, however, the remainder of this thesis will focus specifically on C9FTLD/ALS.

Chapter 4 discussed the generation of *in vitro* models separately expressing the five DPRs, four of which were expressed at physiologically-relevant lengths of over 1000 repeats. The importance of repeat-length has been repeatedly demonstrated, both within this thesis and in the published literature; the appearance of poly-GA and poly-AP differed between cells expressing long and short repeat-lengths in Chapter 4, and several studies have reported length-dependent toxicity of DPRs in both cell culture and *in vivo* models (Mizielinska et al. 2014; Wen et al. 2014; Jovičić et al. 2015). Length-dependent toxicity has also been reported in models expressing pure G₄C₂ repeats at various repeat-lengths (Freibaum et al. 2015). As the first models expressing DPRs at physiologically-relevant repeat-lengths, our cells therefore offer an important improvement to existing models of DPR pathology.

In agreement with observations from patient tissue, differential subcellular localisations were noted for each DPR in our cellular models. Poly-GA formed large, fibrous, star-shaped cytoplasmic inclusions, which co-localised with p62 and ubiquilin-2. Poly-AP exhibited an unusual distribution pattern which resembled the cytoskeleton but did not co-localise with actin microfilaments or microtubules. It is possible, however, that the

localisation of poly-AP is caused by binding to a HeLa-specific cytoskeletal protein which is not expressed in neurons, and therefore further validation of this model as disease-relevant is required before reliable conclusions may be drawn from it. Finally, the arginine-rich DPRs, poly-GR and –PR, translocated to the nucleolus, where poly-GR in particular was found to cause nucleolar stress and loss or re-distribution of nuclear proteins known to be important for RNA processing and neuronal survival. Therefore the work presented in Chapter 5 strongly supports a key role of DPRs in C9FTLD/ALS pathogenesis. This chapter will discuss the wider implications of these findings, as well as suggestions for further research directed by this thesis.

6.2 Evidence for RNA vs. DPR toxicity in C9FTLD/ALS

The role of RNA transcripts arising from the C9orf72 expansion in neurodegeneration is yet to be established. Whilst numerous studies report a toxic effect of G₄C₂ repeat expression *in vitro* and *in vivo*, these data are insufficient to distinguish between potentially toxic effects of RNA and DPRs. As discussed in Chapter 1, RNA FISH has demonstrated the presence of nuclear RNA foci containing both sense and antisense transcripts as a consistently present feature of C9FTLD/ALS patient tissue (DeJesus-Hernandez et al. 2011; Mizielinska et al. 2013; Lagier-Tourenne et al. 2013; Lee et al. 2013). Nuclear RNA foci have also been observed in patient iPSC-derived neurons (Almeida et al. 2013; Donnelly et al. 2013; Sareen et al. 2013), as well as cultured cells expressing pure G₄C₂ repeats (Gendron et al. 2013; Lee et al. 2013; Rossi et al. 2015). Finally, nuclear foci were present in the cortex, cerebellum, spinal cord, thalamus and hippocampus of transgenic mice expressing G₄C₂ repeats by 6 months of age, indicating that foci are formed reasonably early in pathogenesis (Chew et al. 2015). Therefore it is possible that expanded RNA plays an important role in C9FTLD/ALS pathogenesis.

A number of RNA-binding proteins have been found to co-localise with RNA foci in cells expressing G₄C₂ repeats and patient tissue. Several members of the heterogeneous ribonucleoprotein (hnRNP) family, including hnRNP A1, H and F, were detected in RNA foci in the cerebellum and motor neurons (Cooper-Knock et al. 2014), as well as in patient iPSC-derived neurons (Sareen et al. 2013). hnRNP A2/B1, H, U and Q were also identified as interactors of G₄C₂ repeat RNA *in vitro* (Rossi et al. 2015). hnRNPs are primarily nuclear proteins which play crucial roles in regulation of transcription, translation, splicing and nucleocytoplasmic RNA transport. ADARB2, a regulatory protein with a role in RNA editing also co-localised with RNA foci *in vitro* and in patient

tissue (Donnelly et al. 2013), as did the serine-arginine rich splicing factors, SF2 and SC35 (Cooper-Knock et al. 2014; Lee et al. 2013). It has been proposed that sequestration of RNA binding proteins into nuclear foci may cause neurodegeneration through a loss of function mechanism leading to aberrant RNA processing. Early work from Xu and colleagues supports this hypothesis; the translational activator, Pura α , was found to interact with G₄C₂ RNA *in vitro* and in *Drosophila*, and overexpression of Pura rescued G₄C₂-induced toxicity (Xu et al. 2013). This suggests that loss of Pura contributed to the toxic phenotype. Finally, the mRNA export factor, REF/ALY, has also been reported to co-localise with nuclear RNA foci in the cerebellum and motor neurons of C9ALS patient tissue (Cooper-Knock et al. 2014). This finding is of particular interest, given the evidence for impaired nucleocytoplasmic transport in C9FTLD/ALS. However, depletion of REF/ALY, rather than overexpression, was found to reduce G₄C₂-induced toxicity in fly (Freibaum et al. 2015). This argues against the hypothesis that sequestration of RNA-binding proteins causes toxic loss of function. Furthermore, Cooper-Knock and colleagues suggested that binding of proteins such as REF/ALY, SF2 and hnRNPs to RNA foci may be transient, since they only observed co-localisation in a small proportion of cells in patient tissue (Cooper-Knock et al. 2014). Therefore the presence of RNA-binding proteins in nuclear foci may not indicate loss of their normal function. Whilst aberrant RNA processing has been strongly implicated in C9FTLD/ALS pathogenesis, the evidence remains insufficient to determine whether this is caused solely by DPRs or by a combination of both DPRs and RNA.

Perhaps the most persuasive evidence arguing against DPR production as the primary toxic event in C9FTLD/ALS is the poor correlation between DPR immunostaining, TDP-43 pathology and neurodegeneration in patient tissue (Mackenzie et al. 2013; Mackenzie et al. 2015). The distribution of DPR pathology across different brain regions appears to be consistent regardless of clinical subtype and associated neurodegeneration patterns, with inclusions most abundantly present in the cerebellum, frontal and motor cortices and hippocampus, and less frequently present in the spinal cord (Mackenzie et al. 2013; Davidson et al. 2014; Gomez-Deza et al. 2015; Mori et al. 2013). Conversely, the distribution of TDP-43 pathology more closely correlated with areas of neuronal loss in patient tissue, and has been reported to be more abundant than DPR pathology (Mackenzie et al. 2013; Gomez-Deza et al. 2015). However, these findings do not necessarily indicate that DPRs are not the primary cause of neurodegeneration in C9FTLD/ALS, for several reasons.

Firstly, many studies investigating the distribution of DPR pathology in patient tissue have focused primarily on p62-positive cytoplasmic poly-GA inclusions, which are

easiest to detect but which may not reflect the distribution pattern of other DPRs. Secondly, the data presented in this thesis and elsewhere suggests that nucleolar, but not cytoplasmic, arginine-rich DPRs may be most toxic. However, as discussed in Section 5.4.3, the presence of nucleolar DPRs may be more difficult to detect in patient tissue. Thirdly, a major limitation of these studies is their inability to detect insoluble forms of each DPR, which may be more toxic than the aggregated, insoluble forms. Indeed, soluble forms of proteins key to other neurodegenerative diseases are widely considered to be more toxic than insoluble protein aggregates, for example soluble amyloid- β oligomers are considered to be the toxic amyloid species in AD (reviewed by Rosenblum, 2014). It is possible that cytoplasmic aggregation of poly-GA is in fact protective, particularly if poly-GR is sequestered by these inclusions as suggested in Chapter 4. Finally, analysis of post-mortem autopsy tissue only provides insight into the end-stage of disease, where the majority of vulnerable neurons have already undergone cell death. It is not possible to determine whether these neurons contained DPR pathology prior to degeneration. Indeed, post-mortem visualisation of DPR pathology may only highlight the cells which were most resistant to DPR toxicity. Therefore poor correlation between DPR aggregation and neuronal atrophy in patient tissue may not reflect the toxicity of DPRs.

Several case studies of expansion-carrying patients who died of unrelated causes have been reported. Proudfoot and colleagues described a C9FTLD patient with a clinical diagnosis of FTD who died aged 53 due to bronchopneumonia. This patient exhibited wide-spread poly-GA inclusions but minimal TDP-43 pathology, suggesting that the formation of DPR inclusions is an early event in disease pathogenesis, which occurs before TDP-43 inclusions form and which is sufficient to cause severe neurological dysfunction (Proudfoot et al. 2014). The patient's son was also described; although he was not diagnosed with FTLTLD, he carried the C9orf72 expansion and displayed bvFTD-like symptoms from an early age, including anxiety, ritualistic behaviour and requirement of full-time care by age 25. The son died of pulmonary embolism at age 26. Post-mortem analysis showed that poly-GA inclusions were present in the cerebellum, cortex and hippocampus, but no TDP-pathology was detected. No significant neuronal loss was evident in this case, again indicating that DPR pathology is an early event in disease pathogenesis (Proudfoot et al. 2014). Three C9FTLD patients who died of other causes within 10 months – 3 years of diagnosis were also found to exhibit an extensive DPR burden but minimal TDP-43 pathology (Baborie et al. 2014). Mori *et al.*, also reported that a single C9FTLD/ALS case in their cohort of 15 displayed DPR, but not TDP-43, pathology at autopsy (Mori et al. 2013). Overall, these findings strongly support

a role of DPRs as causative in the early stages of C9FTLD/ALS pathogenesis, and suggest that DPR production may lead to TDP-43 pathology and neurodegeneration over time. However, the link between DPRs and TDP-43 is inconclusive, and the presence of RNA foci was not investigated in these studies.

Like DPR pathology, RNA foci are commonly found in the frontal and motor cortices, cerebellum and hippocampus (Mizielinska et al. 2013; Cooper-Knock et al. 2015), although foci and DPR inclusions are only occasionally found in the same cell (Gendron et al. 2013; Mizielinska et al. 2013). Importantly, RNA foci did not correlate well with TDP-43 pathology or neurodegeneration either (Mizielinska et al. 2013). However, both RNA foci burden and the extent of DPR pathology have been reported to correlate with age of onset (Mizielinska et al. 2013; Davidson et al. 2014).

No changes in TDP-43 localisation have been reported in cells or flies expressing DPRs without G₄C₂ RNA, including those generated in this project (Figure 5.1). This would imply that DPRs do not trigger the pathological changes observed in TDP-43 in C9FTLD/ALS. However, TDP-43 pathology was also absent from cells and flies expressing pure G₄C₂ repeats. These models involve expression of G₄C₂ repeats or DPR sequences at high levels over relatively short-periods of time, whereas in patient tissue, DPRs and G₄C₂ RNA are present at presumably lower levels over several decades before TDP-43 pathology becomes wide-spread. Therefore it is possible that the changes in TDP-43 solubility and localisation are very gradual, and as such cannot be reproduced in cell culture models. As such, the relationship between RNA foci, DPRs and TDP-43 remains elusive. The role of TDP-43 in neurodegeneration also remains unclear. Interestingly, the recently described transgenic mouse model expressing G₄C₂ repeats exhibited extensive TDP-43 pathology at 6m of age, indicating that this mammalian model more accurately reflects the pathological features of C9FTLD/ALS. Expression of DPR-only constructs such as ours in a mouse model may therefore be a more appropriate method to determine the relationship between DPRs and TDP-43.

Overall, there is a considerable body of evidence emerging which provides compelling evidence for the role of DPRs in neurodegeneration. As discussed in Chapter 5, a number of *in vivo* and *in vitro* models expressing DPRs without G₄C₂ RNA exhibit severe toxicity phenotypes, particularly when expressing arginine-rich DPRs. The evidence supporting a role of G₄C₂ RNA in toxicity is less persuasive. Whilst several models expressing G₄C₂ repeats have been reported to also exhibit a toxic phenotype, it is not possible to determine from this whether the toxicity was caused by RNA or peptides.

Perhaps the most useful study to address this question was the development of an “RNA-only” G₄C₂ construct of ~288 repeats by Mizielinska and colleagues. This construct incorporated regular stop codons in all frames, thus preventing RAN-translation of the sequence. When expressed in *Drosophila*, RNA foci were produced but no toxicity was observed (Mizielinska et al. 2014b). This finding strongly suggests that G₄C₂ RNA is not toxic in the absence of DPRs. However, generation of similar RNA-only G₄C₂ constructs of longer repeat-length would be more suitable to definitively draw this conclusion. Furthermore, the interruption of the G₄C₂ repeat sequence with regular stop codons may prevent toxicity. Therefore it is possible that G₄C₂ RNA is only toxic at long repeat-lengths, when the sequence is 100% GC-rich or that it may require the presence of DPRs to trigger its toxic effects. Interestingly, several of the proteins reported to co-localise with RNA foci in patient tissue contain RGG/RG boxes, and GAR domains are known to preferentially bind to G-quartet nucleic acid structures. This offers an exciting potential link between arginine-rich DPRs and G₄C₂ RNA in disease pathogenesis, and may be an interesting avenue of future research.

6.3 Proposed mechanisms of neurodegeneration in C9FTLD/ALS

The C9orf72 expansion is likely to exert multiple, distinct toxic effects culminating in neurodegeneration in C9FTLD/ALS (summarised in Figure 6.1). The evidence discussed in Chapter 5 strongly implicates a role of DPR induced stress phenotypes, impairments in nucleocytoplasmic transport and aberrant RNA processing in disease pathogenesis. Furthermore, expanded G₄C₂ RNA may play a role in toxicity by sequestering RNA-binding proteins and exacerbating the RNA processing defects proposed to be caused by DPRs. Since GAR domains in RNA-binding proteins are known to interact with G-quartet RNA structures, it is possible that arginine-rich DPRs may interact with G₄C₂ RNA with unknown consequences, although this remains to be tested *in vitro* or *in vivo*.

Poly-GA aggregates sequester UPS proteins, leading to proteasomal inhibition and ER stress. Unresolvable ER stress conditions trigger pro-apoptotic signalling cascades culminating in caspase activation and cell death (reviewed by Szegezdi et al. 2006). Therefore poly-GA is proposed to exert neurotoxicity via ER stress. In addition to the sequestration of UPS proteins, poly-GA was also observed to disrupt the nuclear membrane, which became misshapen and adopted a “frayed” appearance. In Chapter 5, mislocalisation of FUS was also noted in cells expressing poly-GA. This may be caused

by nuclear membrane disruption, allowing certain proteins to leak into the cytoplasm. However, the evidence for G₄C₂-induced impairments in nucleocytoplasmic transport

suggests that this may be caused by defects in nuclear import of proteins such as FUS. Thus far, nucleocytoplasmic transport impairments have only been reported to be caused by G_4C_2 repeats and the arginine-rich DPRs. However, the similarities in nuclear appearance and FUS distribution in cells expressing poly-GA, -GR and -PR suggest that poly-GA may exert a similar effect on nucleocytoplasmic transport. This provides a second potential mechanism of neurodegeneration caused by poly-GA, which is shared by poly-GR and -PR.

In particular, impairments have been observed in nuclear export of mRNA, which would prevent normal translation of many proteins and could therefore impact any number of cellular processes and pathways. The mechanisms through which DPR expression leads to impaired nucleocytoplasmic transport are unknown. One possibility is that arginine-rich DPRs compete with key proteins such as REF/ALY and NXF1 for mRNA binding, which is dependent on methylation of GAR domains. Alternatively, DPRs may compete for methylation by methyltransferase enzymes, preventing methylation of RGG/RG boxes in RNA-binding proteins. Further work is required to determine the precise mechanisms involved.

Defective nuclear import of proteins has also been demonstrated to be caused by DPRs or G_4C_2 RNA. It is possible that this is a secondary event caused by nuclear retention of mRNA and subsequent downregulation of translation. Conversely, this may be a separate phenomenon, occurring independently of the defects in mRNA export. Regardless, the downstream effects of preventing nuclear import of proteins are potentially severe. Since the primary functions of the nucleus are gene expression and RNA processing, loss of nuclear proteins is likely to cause toxicity through aberrant RNA processing, something which has been shown to occur as a result of DPR or G_4C_2 expression.

Arginine-rich DPRs have been found to cause nucleolar stress, a phenomenon which is known to lead to defective ribosome biogenesis and trigger pro-apoptotic pathways or cell-cycle arrest (reviewed by Boulon et al. 2010; James et al. 2014). Poly-GR has also been shown to cause loss of SMN and coilin from Cajal bodies. CBs are the site of spliceosome formation and regulation, suggesting that nucleolar poly-GR is likely to impact splicing regulation. Therefore DPRs may induce RNA processing abnormalities via several convergent mechanisms. It is possible that loss of SMN, coilin and fibrillarin are all consequences of impaired nucleocytoplasmic transport. Alternatively, these changes may occur via distinct, as yet unknown mechanisms. Again, the importance of

RGG/RG boxes in nucleolar and CB function suggests that poly-GR may directly interfere with normal protein interactions in these regions.

Aberrant RNA processing could potentially impact a great number of genes and pathways involved in all aspects of cellular function and survival. Therefore three primary DPR-induced impairments are proposed to lead to neurodegeneration in C9FTLD/ALS: (i) proteasome inhibition leading to ER stress, (ii) impaired nucleocytoplasmic transport and (iii) aberrant RNA processing.

6.4 Convergent mechanisms of neurotoxicity in FTLD/ALS

A major question in FTLD/ALS research is how various genetic mutations lead to similar clinical or pathological outcomes. The *C9orf72* expansion may cause FTLD, ALS or combined FTLD-ALS with TDP-43 pathology. Mutations in *TARDBP* and *FUS* cause FTLD/ALS with TDP-43 or FUS pathology, respectively. However, mutations in *SOD1* cause ALS without TDP-43 pathology, and do not cause FTLD, whereas *MAPT* mutations cause FTLD without TDP-43 pathology, but do not cause ALS. It is possible that the selective neurodegeneration of the specific neuronal subgroups associated with FTLD or ALS is caused by increased vulnerability of these neuronal subtypes to dysfunction in certain cellular processes. Therefore different causative events early in disease pathogenesis may converge on similar mechanisms leading to neurodegeneration via different routes. If convergent mechanisms of neurodegeneration do exist, novel therapeutic interventions targeting these pathways could potentially benefit patients with different forms of FTLD/ALS. In the previous section, three aspects of cellular dysfunction were proposed to be key to neurodegeneration in C9FTLD/ALS: (i) proteasome inhibition leading to ER stress, (ii) impaired nucleocytoplasmic transport and (iii) aberrant RNA processing. This section will briefly discuss the evidence for a role of these three impairments in non-C9FTLD/ALS.

6.4.1 Protein degradation and ER stress

The presence of insoluble or misfolded cytoplasmic protein aggregates is a common feature of all neurodegenerative diseases, and in most cases UPS proteins such as p62 and ubiquitin are sequestered by these inclusions (Kurosawa et al. 2015; Lennox et al. 1988; Manetto et al. 1988). In FTLD/ALS, poly-GA is not the only protein inclusion known to co-localise with UPS proteins; *SOD1*, TDP-43, tau, and *FUS* inclusions are also ubiquitinated and contain p62 (Neumann et al. 2006; Neumann et al. 2009; Shibata et al. 1996; de Silva et al. 2006). This may indicate failure of the UPS to degrade the pathological protein despite ubiquitin-tagging, suggesting impairment of protein degradation systems is a common feature of many neurodegenerative diseases.

The discovery of mutations in *UBQLN2*, *SQSTM1* and *VCP* as causative in FTLD/ALS strongly implicates UPS impairment in non-C9orf72-linked disease (Deng et al. 2011; Johnson et al. 2010; Schröder et al. 2005; Kaleem et al. 2007; Rubino et al. 2012; Miller et al. 2015). This is supported by findings from transgenic mice; reduced proteasomal labelling and activity has been observed in mice expressing mutant *SOD1* or TDP-43

(Caccamo et al. 2015; Cheroni et al. 2009). Furthermore, Cheroni and colleagues generated double-transgenic *SOD1* mice also expressing a ubiquitin-tagged GFP molecule known to be degraded by the UPS. Increased GFP fluorescence was observed in spinal motor neurons of mutant *SOD1* mice compared to those expressing the fluorescent reporter construct alone, indicating an impairment in the ability of the UPS to degrade the reporter protein normally (Cheroni et al. 2009). Moreover, depletion of the proteasome subunit, Rpt3, in mouse motor neurons produced a phenotype comparable to ALS; transgenic mice exhibited degeneration of motor neurons, impaired motor function, gliosis and mislocalisation or pathological aggregation of proteins such as TDP-43, FUS and ubiquilin-2. This provides strong evidence for a causative link between impaired UPS activity and neurodegeneration in ALS (Tashiro et al. 2012). Importantly, examination of patient tissue confirmed that proteasomal activity was reduced in spinal cord lysate from sporadic ALS cases, and that proteasomal markers were downregulated in motor neurons (Kabashi et al. 2012). Therefore UPS impairments are a consistently reported feature of non-C9FTLD/ALS.

Impairments in proteasome function are known to cause ER stress and, if unresolvable, apoptosis (Park et al. 2011; Obeng et al. 2006; Fribley et al. 2004). Accordingly, signs of ER stress have been observed in cells expressing mutant *TARDBP* (Wang et al. 2015; Walker et al. 2013), as well as in *SOD1* transgenic mice (Chen et al. 2015). In addition, expression of mutant *FUS*, *TARDBP* or *SOD1* in mice inhibited transport of newly synthesised proteins away from the ER and Golgi, causing ER stress (Soo et al. 2015). Finally, signs of ER stress were observed in anterior horn motor neurons in sporadic ALS patients (Sasaki 2010). Further work to determine whether ER stress markers are consistently upregulated in patient tissue would be of value.

6.4.2 Nucleocytoplasmic transport

Perhaps the strongest evidence for impaired nucleocytoplasmic transport in non-C9FTLD/ALS is the commonness of TDP-43 and FUS pathology. TDP-43 and FUS are primarily nuclear proteins, and therefore their mislocalisation to cytoplasmic inclusions may indicate impaired nuclear import of these proteins. Indeed, the majority of FTLN/ALS-linked mutations identified in *TARDBP* and *FUS* to date are within or in close proximity to the C-terminal NLS which both proteins possess (Baloh 2012). Interestingly, the nuclear import factor karyopherin β 2, which was found to enhance G₄C₂-induced toxicity when depleted in yeast (Jovičić et al. 2015), has been shown to co-localise with FUS inclusions in FTLN patient tissue (Brelstaff et al. 2011). This implies that the

nuclear import defect may affect other proteins besides FUS, through loss of key import factors. Similarly, the Ran binding protein, RanBP1, was downregulated following TDP-43 knockdown *in vitro*, indicating that loss of TDP-43 causes further loss of proteins required for nucleocytoplasmic transport (Štalekar et al. 2015). Furthermore, the retinal neurodegeneration observed in *GRN* knockout mice was preceded by nuclear depletion of TDP-43 and loss of Ran (Ward et al. 2014). Therefore nuclear depletion of both TDP-43 and FUS is associated with loss of several proteins involved in nuclear import. Interestingly, expression of mutant FUS in motor neurons has been found to cause loss of the arginine methyltransferase, PRMT1, from the nucleus. This implies that defects in arginine methylation of RGG/RG box containing proteins may be a feature of non-C9FTLD/ALS.

Expression of mutant VCP has also been reported to affect nucleocytoplasmic transport. Disruption to the nuclear membrane was observed in cells expressing ALS-linked mutant VCP, and nucleoporins were mislocalised to the ER and golgi (Tran et al. 2012). The authors suggest a role for VCP in trafficking nucleoporins to the nuclear envelope. It is of particular interest that nucleoporins were sequestered to the ER in this study, since this may also trigger ER stress pathways.

Finally, defective nucleocytoplasmic transport has also been implicated in *SOD1*-linked ALS. Nuclear depletion of several karyopherins and nucleoporins has been noted in anterior horn motor neurons of transgenic *SOD1* mice, as well in *SOD1*-linked ALS patient tissue (Kinoshita et al. 2009; Zhang et al. 2006; Nagara et al. 2013). Abnormalities in nuclear structure were also observed in both transgenic mice and patient tissue (Nagara et al. 2013; Kinoshita et al. 2009). Therefore considerable evidence exists supporting a role of impaired nucleocytoplasmic transport in non-C9FTLD/ALS.

6.4.3 Aberrant RNA processing

The final mechanism through which the C9orf72 expansion is proposed to cause toxicity is aberrant RNA processing. Strong evidence exists for a role of aberrant RNA processing in non-C9FTLD/ALS. Firstly, the discovery of *TARDBP* and *FUS* mutations in familial forms of FTL/ALS (Kabashi et al. 2008; Kwiatkowski et al. 2009; Vance et al. 2009; Sreedharan et al. 2008; Yan et al. 2010) and the TDP-43 and FUS pathology observed in both familial and sporadic cases (Neumann et al. 2006; Neumann et al. 2009) strongly suggest that aberrant RNA processing is important in disease

pathogenesis. The causative nature of these mutations is supported by work in transgenic animals; expression of mutant *TARDBP* and *FUS* caused motor impairments and neurodegeneration reminiscent of ALS in rodents and invertebrate models (Huang et al. 2011; Kabashi et al. 2010; Liachko et al. 2010; Murakami et al. 2012; Tatom et al. 2009; Xia et al. 2012; Zhou et al. 2010). TDP-43 and FUS perform many functions relating to RNA processing, including regulation of transcription and splicing. They also play a role in post-transcriptional regulation of gene expression through involvement in microRNA biogenesis and interaction with RNA granules in neurons (reviewed by Baloh, 2012). Therefore loss of TDP-43 and FUS from the nucleus is likely to cause a range of RNA processing defects. Indeed, abnormal gene expression has been demonstrated in cells expressing mutant TDP-43 or FUS, as well as *in vivo* (Coady & Manley 2015; Ling et al. 2015; Yu et al. 2015; Štalekar et al. 2015); of particular note, mislocalisation of several splicing factors and associated splicing abnormalities were recently reported in a transgenic pig model expressing mutant TDP-43 (Wang et al. 2015).

A second phenomenon which may contribute to RNA processing defects in FTL/ALS is oxidative stress. Oxidative stress has been strongly implicated in many neurodegenerative diseases, and in particular in ALS (Nunomura et al. 2012; Niedzielska et al. 2015). Increased oxidation of mRNA species has been observed in ALS post-mortem tissue, as well as in *SOD1* transgenic mice (Chang et al. 2008). This was observed in the motor cortex and spinal cord, but not in areas unaffected by ALS. Furthermore, studies in mutant *SOD1* mice and *in vitro* models of oxidative stress suggest that mRNA oxidation is an early event in disease pathogenesis, and may contribute to neurodegeneration (Shan et al. 2007; Chang et al. 2008). Oxidation reduces translation of mRNA transcripts and impairs translational fidelity (Tanaka et al. 2007), which may affect a wide range of cellular processes key for neuronal survival. Of note, the oxidation of mRNA in *SOD1* mice was selective, and when grouped according to function, mRNA species involved in core cellular processes such as protein folding and degradation, mitochondrial electron transport and cytoskeletal function were most severely affected (Chang et al. 2008). This demonstrates that the downstream effects of changes in RNA processing may be widely varied and detrimental to cellular survival.

One of the main findings presented in Chapter 5 was the observation that nucleolar poly-GR caused loss of the SMA-related protein, SMN, and prevented formation of coilin-positive Cajal bodies. Altered SMN distribution has been previously linked to non-C9ALS. Loss of SMN from CBs has been reported in cultured cells and transgenic mice expressing mutant *SOD1* (Kariya et al. 2012; Turner et al. 2014). Furthermore, overexpression of SMN partially rescued *SOD1*-induced toxicity in both of these models

(Kariya et al. 2012; Turner et al. 2014). Depletion of SMN from CBs has also been demonstrated in cells expressing mutant *FUS*, and direct interaction of *FUS* and *SMN* has been observed *in vitro* (Yu et al. 2015; Yamazaki et al. 2012; Groen et al. 2013; Sun et al. 2015). Importantly, a reduction in *SMN* staining has also been observed in spinal cord tissue from patients with sporadic ALS (Turner et al. 2014). Taken together, these studies provide strong evidence for a role of aberrant *SMN* function in ALS, and implies convergent mechanisms of neurodegeneration may exist between ALS and SMA. *SMN* plays a crucial role in ribonucleoprotein trafficking and spliceosome assembly. Therefore loss of *SMN* function is likely to cause splicing defects. The particular vulnerability of motor neurons to *SMN* deficiency may be a contributing factor to the selectivity of neurodegeneration in ALS and FTL-D-ALS.

Taken together, the literature suggests that convergent mechanisms of neurodegeneration in different forms of FTL-D/ALS are likely, and in particular these mechanisms may involve impaired protein degradation systems, nucleocytoplasmic transport and RNA processing. Therapeutic strategies targeting these impairments may therefore prove beneficial to patients with a range of genetic backgrounds.

6.5 Future work

The generation of constructs for expression of over 1000 repeats for 4 of the 5 DPRs creates many possibilities for further research. This final section will briefly discuss some of these possibilities, and address technical problems encountered during this project which remain to be overcome.

6.5.1 Improvement DPR expression levels in our cellular models

The major limitation of our current HeLa cell DPR models is the low expression level of full-length peptides. This is particularly true of the arginine-rich DPRs, which are difficult to detect even by immunofluorescence; nucleolar GFP expression may only be observed in a small proportion of cells on a 22mm coverslip, and exposure times of up to 2s are required for image capture. It appears that rather than every cell within the population expressing a very low level of peptide, only a small proportion of cells express at a level observable by immunofluorescence imaging. This severely limits the potential for investigation of changes in cell signalling, since Western blot analysis is not possible.

The reason for the dramatic reduction in DPR expression levels as repeat-length increases is unclear. One possible explanation is that PR₁₁₀₀ and GR₁₁₃₆ are severely toxic, meaning that successfully transfected cells are less likely to survive long enough for analysis. This is supported by the overwhelming literature discussed in Chapter 5, which suggests that poly-PR and poly-GR are the most toxic DPRs. Toxicity could also explain the subjective observation from immunofluorescence analysis that expression levels decrease over time. The shortest time-point at which coverslips have been prepared for immunofluorescence is 24h post-transfection. If the low expression levels are indeed due to high toxicity, preparation of samples on the same day as transfection may produce better results.

If toxicity is the reason for low DPR expression, reducing the expression level within individual cells may improve survival, thereby increasing the proportion of DPR-expressing cells within the population. The optimisation data in Chapter 4 showed that reducing the amount of plasmid in the transfection reaction further reduced expression. However, this may be because the transfection reagents used require a minimum concentration of DNA in the reaction to facilitate successful transfection. An alternative way to reduce the amount of DPR plasmid without affecting FuGene efficiency could be

to mix it with a second, well-characterised plasmid which does not affect cellular function in any way, thus increasing the amount of total DNA in the reaction. This method may improve the overall transfection efficiency by driving a lower level of DPR expression in a greater proportion of cells. This is currently being tested using an empty pBluescript vector, and may improve transfection efficiency for future work using the DPR constructs.

If unsuccessful, a second option would be to clone the long repeat sequences into a different vector containing a less efficient promoter. Expression of DPRs in the pEGFP-N1 vector is driven by the CMV promoter, which is highly efficient in mammalian cell culture, whereas a less effective promoter may be more suitable for uniform delivery of lower protein levels across the whole cell population. Finally, work is currently underway to generate inducible DPR models using a tetracycline-controlled system, which may improve tolerance of DPRs by allowing expression to be switched on or off for specific periods of time.

6.5.2 Quantification of DPR toxicity

An important future experiment will be to quantify toxicity in our DPR-expressing cells, and to compare toxicity between peptides and between repeat-lengths. In the current expression system, the high proportion of untransfected cells in each well prevents analysis of whole-cell populations. Therefore in the short-term, toxicity will be assessed by immunofluorescence imaging. Cells will be labelled with antibodies against markers of apoptosis such as activated caspases, and signal intensity compared between DPR-expressing cells and controls. This approach allows for selective analysis of cells containing DPR inclusions. Propidium iodide staining may also be used. If further optimisation of DPR transfection is successful and more uniform populations of DPR-expressing cells are obtained, cell death will be quantified using traditional cytotoxicity assays.

6.5.3 Time-course of DPR pathology

An experiment which would be of great interest using our DPR models in their current form is video microscopy to monitor the formation of inclusions over time. Early work performed by Dr Janis Bennion Callister in the Pickering-Brown group suggested that

GA₁₀₂ initially formed 2 small inclusions within a single cell, which later combined to form 1 large inclusion. Observation of aggregate formation by GA₁₀₂₀ would be of interest.

In particular, however, observation of the arginine-rich DPRs over time would be of great value. Our preliminary, subjective observation that poly-GR and –PR inclusions may be exported from the nucleus, and perhaps even the cell, over time has not yet been tested. Video microscopy of individual cells would determine the fate of nucleolar DPR inclusions over time. In addition to video microscopy, it may also be useful to determine whether poly-GR and –PR are present in the media. Indeed, extracellular inclusions of poly-PR were observed in populations of cultured neurons expressing the peptide, which may imply that they are released from living cells or remain intact following cell death (Wen et al. 2014). If arginine-rich DPRs are present extracellularly *in vivo*, the potential impact of this is great. Novel peptides such as these are likely to trigger an immune response, which may contribute to neuronal damage. Gliosis and neuroinflammation are hallmarks of all neurodegenerative diseases including FTL/ALS, and a growing body of evidence suggests that this may contribute to toxicity (von Bernhardi et al. 2015; Lant et al. 2014; Olejniczak et al. 2015). Furthermore, poly-GR and –PR have been demonstrated to enter cultured cells following external application of the peptides (Kwon et al. 2014). This implies that if arginine-rich peptides were secreted from cells *in vivo*, they could potentially enter neighbouring cells. Propagation of various types of pathological inclusions from cell to cell across neuronal networks has been demonstrated in other proteinopathies, for example propagation of misfolded tau pathology has been observed in AD models (Sydow & Mandelkow 2010; Goedert et al. 2010; Pooler et al. 2015). Therefore in theory, the potential of extracellular DPRs to exacerbate neuronal damage could be great. These hypotheses are purely conjecture at this stage, and video microscopy of cells expressing arginine-rich DPRs over time would be a useful first step in determining whether DPRs are indeed secreted by cells.

6.5.4 Co-expression of DPRs

A second future experiment which would be of interest is co-expression of multiple DPRs in HeLa cells, using alternate fluorescent tags. The cloning required to insert each of the longest DPR constructs into a dsRed vector is currently underway. Therefore one DPR may be expressed with a green GFP tag and another expressed with a red tag. It would be of particular interest to observe the distribution patterns of poly-GA and –GR when co-expressed. Given the similarities in appearance of cytoplasmic poly-GA and –GR aggregates in patient tissue, it is hypothesised that poly-GA aggregates may

sequester poly-GR, which would not otherwise aggregate in the cytoplasm. This was observed in *Drosophila* lines expressing poly-GA and –GR (Yang et al. 2015). Furthermore, co-expression of poly-GA reduced the toxicity of poly-GR, suggesting that by preventing poly-GR from translocating to the nucleolus, its sequestration into poly-GA aggregates was protective. This could partially explain why neuronal dysfunction does not usually become severe enough to cause a clinical phenotype in patients until middle age. If a suitable method for quantification of toxicity is developed, investigating the effects of poly-GA expression on poly-GR toxicity would therefore be of interest.

6.5.5 Mechanisms of DPR toxicity

Whilst evidence for DPR-induced impairments in protein degradation, nucleocytoplasmic transport and aberrant RNA processing has been published using other models, these phenotypes have not been directly assessed in our DPR-expressing cells. In order to investigate these impairments in more detail, confirmation that they occur in our model is first required. In the short-term, immunofluorescence will be a useful tool to quickly address some of these questions. The distribution and number of proteasomes within individual cells containing poly-GA aggregates may be assessed using antibodies targeting proteasomal subunits. Markers of ER stress may also be observed by this method. Co-expression of poly-GA with a ubiquitin-tagged red fluorescent protein would be a useful method to determine whether poly-GA expression slows protein degradation via the UPS. The distribution of proteins involved in nucleocytoplasmic transport such as REF/ALY, nucleoporins and the karyopherin family may also be assessed by immunofluorescence. RNA FISH labelling of poly(A)-capped transcripts would also be an interesting experiment to determine whether nuclear retention of mRNA is a feature of our poly-GR and –PR expressing cells.

Investigating changes in RNA processing using immunofluorescence is more challenging. In the short-term, labelling for proteins associated with ribosome biogenesis and splicing control may prove interesting, given the loss of SMN, coilin and fibrillarin observed in cells expressing poly-GR. In the long term, however, transcriptome analysis is required. As discussed above, inducible cell lines expressing the longer length DPRs are currently being generated in the Pickering-Brown lab. This will be followed by RNA sequencing of the whole transcriptome, in order to identify any changes in transcription or splicing.

Further investigation of the relationship between poly-GR and SMN is an urgent priority. Analysis of SMN distribution in C9FTLD/ALS patient tissue or iPSC-derived neurons is required, in order to confirm the observed effects of poly-GR are not HeLa cell specific. If a suitable method is established for quantification of toxicity, the effects of SMN overexpression on cell death would next be investigated. In particular, the possibility that poly-GR may affect SMN by mimicking the methylated GAR domains of its protein interactors is of interest. Antibodies are available to selectively label proteins containing methylated or dimethylated arginine residues. Immunofluorescence may therefore be used to determine whether nucleolar poly-GR contains methylated arginines. Knockdown or overexpression of arginine methyltransferase enzymes could be employed to determine whether arginine methylation is required for poly-GR to translocate to the nucleolus, causing nuclear depletion of SMN, coilin and fibrillarin.

A final point of interest is the potential relationship between poly-GR and G₄C₂ RNA. Since RGG/RG box-containing proteins are known to preferentially interact with G-quartet structures, it is possible that poly-GR also interacts with G₄C₂ RNA. No observations of co-localisation between poly-GR and RNA foci have been reported, however it is possible that the interaction is transient and does not result in inclusion formation, making it difficult to detect by immunofluorescence. The best method to test this hypothesis is most likely by either protein or RNA pull-down in cells expressing pure G₄C₂ repeats or in iPSC-derived neurons obtained from patients. If there is indeed an interaction, the downstream effects of this are unknown. Poly-GR may play a role in foci formation, or aid in sequestration of RNA-binding proteins. Alternatively, interaction with poly-GR may prevent translational activators from accessing G₄C₂ RNA, thus acting as a negative feedback signal downregulation production of DPRs. Therefore the relationship between G₄C₂ RNA and poly-GR may prove an exciting avenue of future research.

6.5.6 Generation of *in vivo* models of DPR pathology

Perhaps the most important future work to follow on from this project will be the generation of *in vivo* models of DPR pathology. Cloning is underway to insert the longest repeat sequences into a vector suitable for pan-neuronal expression in *C. elegans*, driven by the *rgef-1* promoter. This construct contains a red fluorescent protein tag (RFP-T) on the C-terminal. Transgenic worm strains will then be generated by microinjection of these constructs. A well-characterised construct driving nuclear expression of GFP in all cells will be co-injected, to aid investigation of subcellular location by fluorescence microscopy. Motility and survival assays will be performed on

DPR-expressing worms as described in Chapter 3. Once characterised, our transgenic *C. elegans* models will be utilised to further investigate the impact of DPR expression *in vivo*. Finally, work is currently being planned in the Pickering-Brown lab to generate transgenic mice expressing GA₁₀₂₀ and GR₁₁₃₆. These will be the first mammalian *in vivo* models of DPR pathology in the absence of G₄C₂ RNA. As such, the constructs generated as part of this project will be utilised to address an urgent gap in the literature, providing more disease-relevant tools for use in C9FTLD/ALS research.

6.6 Concluding remarks

This thesis described the generation and characterisation of a series of *in vivo* and *in vitro* models of C9FTLD/ALS, each possessing one specific aspect of disease pathology. Knockdown of the C9orf72 orthologue in *C. elegans* did not cause a motility impairment or reduce survival, suggesting that loss of C9orf72 function is insufficient to cause neurodegeneration. Our cellular models of DPR pathology, however, demonstrated that each dipeptide exhibits a unique subcellular distribution with differential effects on cellular function. When combined with recent literature, the findings presented in this thesis present a strong argument for DPRs as the main driving force behind neurodegeneration in C9orf72. The arginine-rich DPRs have been most consistently linked to toxicity, and poly-GA also appears to contribute to toxicity albeit in a less severe manner. The respective roles of poly-AP and poly-GP remain to be determined. Poly-GA is proposed to cause proteasomal inhibition, leading to ER stress and apoptosis. Poly-GR and –PR, conversely, cause more damage to nuclear function; both peptides have been reported to cause nucleolar stress, impairments in nucleocytoplasmic transport and aberrant RNA processing. Of particular interest, we have shown that poly-GR causes loss of SMN from Cajal bodies, which is likely to be a major contributor to toxicity in motor neurons specifically. Many of the cellular functions and pathways which are reported to be impaired by DPR expression have also been implicated in non-C9FTLD/ALS, suggesting that convergent mechanisms of neurodegeneration may exist between different forms of these diseases.

Overall the work presented in this thesis has significantly contributed to the field of C9FTLD/ALS research, through generation of the first models expressing DPRs at physiologically-relevant repeat-lengths. Our *in vitro* models will allow detailed further study of the molecular and cellular mechanisms involved in DPR toxicity. Furthermore, our DPR constructs may be modified for expression in various *in vivo* models. We intend to express DPRs in *C. elegans* and mice, generating the first *in vivo* models of DPR pathology at physiologically-relevant repeat-lengths. This will address a major gap in our currently available toolkit for the study of C9FTLD/ALS, enabling further research to progress in many directions. Finally, our DPR models may be utilised for genetic and chemical screening on a large scale, in order to identify novel compounds and therapeutic targets which may be beneficial in the treatment of C9FTLD/ALS.

References

- Ahmed, Z. et al., 2010. Accelerated lipofuscinosis and ubiquitination in granulin knockout mice suggest a role for progranulin in successful aging. *The American journal of pathology*, 177(1), pp.311–24.
- Al-Chalabi, A. et al., 2012. The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta neuropathologica*, 124(3), pp.339–52.
- Almeida, S. et al., 2013. Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta neuropathologica*.
- Al-Sarraj, S. et al., 2011. p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTL and MND/ALS. *Acta neuropathologica*, 122(6), pp.691–702.
- Amin, M.A. et al., 2007. Fibrillarin, a nucleolar protein, is required for normal nuclear morphology and cellular growth in HeLa cells. *Biochemical and Biophysical Research Communications*, 360(2), pp.320–326.
- Amos, L.A., 2004. Microtubule structure and its stabilisation. *Organic & biomolecular chemistry*, 2(15), pp.2153–60.
- Anderton, R.S. et al., 2013. Spinal muscular atrophy and the antiapoptotic role of survival of motor neuron (SMN) protein. *Molecular neurobiology*, 47(2), pp.821–32.
- Anon, WormBase : Nematode Information Resource. Available at: <http://www.wormbase.org/#01-23-6> [Accessed September 17, 2015].
- Aoki, N. et al., 2012. Localization of fused in sarcoma (FUS) protein to the post-synaptic density in the brain. *Acta neuropathologica*, 124(3), pp.383–94.
- Arai, T. et al., 2006. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Biochemical and biophysical research communications*, 351(3), pp.602–11.
- Ash, P.E.A. et al., 2010. Neurotoxic effects of TDP-43 overexpression in *C. elegans*. *Human molecular genetics*, 19(16), pp.3206–18.
- Ash, P.E.A. et al., 2013. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron*, 77(4), pp.639–46.
- Ayala, Y.M. et al., 2008. Structural determinants of the cellular localization and shuttling of TDP-43. *Journal of cell science*, 121(Pt 22), pp.3778–85.
- Baborie, A. et al., 2014. Accumulation of dipeptide repeat proteins predates that of TDP-43 in Frontotemporal Lobar Degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathology and applied neurobiology*.
- Baez, S. et al., 2014. Primary empathy deficits in frontotemporal dementia. *Frontiers in aging neuroscience*, 6, p.262.
- Baker, M. et al., 2006. Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature*, 442(7105), pp.916–9.

- Baloh, R.H., 2012. How do the RNA-binding proteins TDP-43 and FUS relate to amyotrophic lateral sclerosis and frontotemporal degeneration, and to each other? *Current opinion in neurology*, 25(6), pp.701–7.
- Bañez-Coronel, M. et al., 2015. RAN Translation in Huntington Disease. *Neuron*, 88(4), pp.667–677.
- Bannwarth, S. et al., 2014. A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain : a journal of neurology*, 137(Pt 8), pp.2329–45.
- Bateman, A. et al., 1990. Granulins, a novel class of peptide from leukocytes. *Biochemical and biophysical research communications*, 173(3), pp.1161–8.
- Beck, J. et al., 2013. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *American journal of human genetics*, 92(3), pp.345–53.
- Belzil, V. V et al., 2013. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta neuropathologica*, 126(6), pp.895–905.
- Benajiba, L. et al., 2009. TARDBP mutations in motoneuron disease with frontotemporal lobar degeneration. *Annals of neurology*, 65(4), pp.470–3.
- Le Ber, I. et al., 2009. Chromosome 9p-linked families with frontotemporal dementia associated with motor neuron disease. *Neurology*, 72(19), pp.1669–76.
- Le Ber, I. et al., 2007. Progranulin null mutations in both sporadic and familial frontotemporal dementia. *Human mutation*, 28(9), pp.846–55.
- Berger, C.M., Gaume, X. & Bouvet, P., 2015. The roles of nucleolin subcellular localization in cancer. *Biochimie*, 113, pp.78–85.
- Von Bernhardi, R., Eugenín-von Bernhardi, L. & Eugenín, J., 2015. Microglial cell dysregulation in brain aging and neurodegeneration. *Frontiers in aging neuroscience*, 7, p.124.
- Bigio, E.H. et al., 2013. Frontotemporal lobar degeneration with TDP-43 proteinopathy and chromosome 9p repeat expansion in C9ORF72: clinicopathologic correlation. *Neuropathology : official journal of the Japanese Society of Neuropathology*, 33(2), pp.122–33.
- Van Blitterswijk, M. et al., 2013. Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *The Lancet. Neurology*, 12(10), pp.978–88.
- Boeve, B.F. et al., 2012. Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. *Brain : a journal of neurology*, 135(Pt 3), pp.765–83.
- Boisvert, F.-M., Chénard, C.A. & Richard, S., 2005. Protein interfaces in signaling regulated by arginine methylation. *Science's STKE : signal transduction knowledge environment*, 2005(271), p.re2.
- Borrioni, B. et al., 2007. Evidence of white matter changes on diffusion tensor imaging in frontotemporal dementia. *Archives of neurology*, 64(2), pp.246–51.

- Borrioni, B. et al., 2014. Heterozygous TREM2 mutations in frontotemporal dementia. *Neurobiology of aging*, 35(4), pp.934.e7–10.
- Borrioni, B. et al., 2010. Is frontotemporal lobar degeneration a rare disorder? Evidence from a preliminary study in Brescia county, Italy. *Journal of Alzheimer's disease: JAD*, 19(1), pp.111–6.
- Borrioni, B. et al., 2009. Mutation within TARDBP leads to frontotemporal dementia without motor neuron disease. *Human mutation*, 30(11), pp.E974–83.
- Borrioni, B. et al., 2010. TARDBP mutations in frontotemporal lobar degeneration: frequency, clinical features, and disease course. *Rejuvenation research*, 13(5), pp.509–17.
- Boulon, S. et al., 2010. The nucleolus under stress. *Molecular cell*, 40(2), pp.216–27.
- Boxer, A.L. et al., 2011. Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *Journal of neurology, neurosurgery, and psychiatry*, 82(2), pp.196–203.
- Braungart, E. et al., 2004. Caenorhabditis elegans MPP+ model of Parkinson's disease for high-throughput drug screenings. *Neuro-degenerative diseases*, 1(4-5), pp.175–83.
- Brelstaff, J. et al., 2011. Transportin1: a marker of FTL-D-FUS. *Acta neuropathologica*, 122(5), pp.591–600.
- Brettschneider, J. et al., 2012. Pattern of ubiquilin pathology in ALS and FTL-D indicates presence of C9ORF72 hexanucleotide expansion. *Acta neuropathologica*, 123(6), pp.825–39.
- Brooks, B.R. et al., 2000. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis and other motor neuron disorders: official publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases*, 1(5), pp.293–9.
- Bruijn, L.I. et al., 1997. ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. *Neuron*, 18(2), pp.327–38.
- Brun, A. et al., 1994. Clinical and neuropathological criteria for frontotemporal dementia. The Lund and Manchester Groups. *Journal of neurology, neurosurgery, and psychiatry*, 57(4), pp.416–8.
- Buchman, V.L. et al., 2013. Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. *Molecular neurodegeneration*, 8, p.12.
- Caccamo, A. et al., 2015. Reduced protein turnover mediates functional deficits in transgenic mice expressing the 25 kDa C-terminal fragment of TDP-43. *Human molecular genetics*, 24(16), pp.4625–35.
- Cairns, N.J. et al., 2007. Neuropathologic diagnostic and nosologic criteria for frontotemporal lobar degeneration: consensus of the Consortium for Frontotemporal Lobar Degeneration. *Acta neuropathologica*, 114(1), pp.5–22.
- Cautain, B. et al., 2015. Components and regulation of nuclear transport processes. *The FEBS journal*, 282(3), pp.445–62.

- Cerami, C. et al., 2014. Neural correlates of empathic impairment in the behavioral variant of frontotemporal dementia. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, 10(6), pp.827–34.
- Chancellor, A.M. et al., 1993. The prognosis of adult-onset motor neuron disease: a prospective study based on the Scottish Motor Neuron Disease Register. *Journal of neurology*, 240(6), pp.339–46.
- Chang, Y. et al., 2008. Messenger RNA oxidation occurs early in disease pathogenesis and promotes motor neuron degeneration in ALS. *PloS one*, 3(8), p.e2849.
- Chausseot, A. et al., 2014. Screening of CHCHD10 in a French cohort confirms the involvement of this gene in frontotemporal dementia with amyotrophic lateral sclerosis patients. *Neurobiology of aging*, 35(12), pp.2884.e1–4.
- Chen, D., Wang, Y. & Chin, E.R., 2015. Activation of the endoplasmic reticulum stress response in skeletal muscle of G93A*SOD1 amyotrophic lateral sclerosis mice. *Frontiers in cellular neuroscience*, 9, p.170.
- Chen, S. et al., 2012. Autophagy dysregulation in amyotrophic lateral sclerosis. *Brain pathology (Zurich, Switzerland)*, 22(1), pp.110–6.
- Chen, Y. et al., 2011. Expression of human FUS protein in Drosophila leads to progressive neurodegeneration. *Protein & cell*, 2(6), pp.477–86.
- Chen-Plotkin, A.S. et al., 2011. Genetic and Clinical Features of Progranulin-Associated Frontotemporal Lobar Degeneration. *Archives of Neurology*, 68(4), p.488.
- Cheroni, C. et al., 2009. Functional alterations of the ubiquitin-proteasome system in motor neurons of a mouse model of familial amyotrophic lateral sclerosis. *Human molecular genetics*, 18(1), pp.82–96.
- Chew, J. et al., 2015. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science*.
- Chiò, A. et al., Prognostic factors in ALS: A critical review. *Amyotrophic lateral sclerosis: official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, 10(5-6), pp.310–23.
- Cioce, M. et al., 2006. UV-induced fragmentation of Cajal bodies. *The Journal of cell biology*, 175(3), pp.401–13.
- Ciura, S. et al., 2013. Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. *Annals of neurology*, pp.1–44.
- Coady, T.H. & Manley, J.L., 2015. ALS mutations in TLS/FUS disrupt target gene expression. *Genes & development*, 29(16), pp.1696–706.
- Cooper-Knock, J. et al., 2015. Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta neuropathologica*.
- Cooper-Knock, J. et al., 2012. Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain: a journal of neurology*, 135(Pt 3), pp.751–64.
- Cooper-Knock, J. et al., 2014. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain: a journal of neurology*, 137(Pt 7), pp.2040–51.

- Corrado, L. et al., 2009. High frequency of TARDBP gene mutations in Italian patients with amyotrophic lateral sclerosis. *Human mutation*, 30(4), pp.688–94.
- Cruts, M. et al., 2006. Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature*, 442(7105), pp.920–4.
- Dahl, K.N., Ribeiro, A.J.S. & Lammerding, J., 2008. Nuclear shape, mechanics, and mechanotransduction. *Circulation research*, 102(11), pp.1307–18.
- Van Damme, P. et al., 2008. Progranulin functions as a neurotrophic factor to regulate neurite outgrowth and enhance neuronal survival. *The Journal of cell biology*, 181(1), pp.37–41.
- Daniel, R. et al., 2000. Cellular localization of gene expression for progranulin. *The journal of histochemistry and cytochemistry: official journal of the Histochemistry Society*, 48(7), pp.999–1009.
- Dantuma, N.P. & Bott, L.C., 2014. The ubiquitin-proteasome system in neurodegenerative diseases: precipitating factor, yet part of the solution. *Frontiers in molecular neuroscience*, 7, p.70.
- Daoud, H. et al., 2012. UBQLN2 mutations are rare in French and French-Canadian amyotrophic lateral sclerosis. *Neurobiology of aging*, 33(9), pp.2230.e1–2230.e5.
- Davidson, Y. et al., 2007. Ubiquitinated pathological lesions in frontotemporal lobar degeneration contain the TAR DNA-binding protein, TDP-43. *Acta neuropathologica*, 113(5), pp.521–33.
- Davidson, Y.S. et al., 2014. Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta neuropathologica communications*, 2, p.70.
- Van Deerlin, V.M. et al., 2010. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nature genetics*, 42(3), pp.234–9.
- Van Deerlin, V.M. et al., 2008. TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis. *The Lancet. Neurology*, 7(5), pp.409–16.
- DeJesus-Hernandez, M. et al., 2011. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron*, 72(2), pp.245–256.
- DeJesus-Hernandez, M. et al., 2011. Novel p.Ile151Val mutation in VCP in a patient of African American descent with sporadic ALS. *Neurology*, 77(11), pp.1102–3.
- Delacourte, A. et al., 1996. Specific pathological Tau protein variants characterize Pick's disease. *Journal of neuropathology and experimental neurology*, 55(2), pp.159–68.
- Delacourte, A. et al., 1998. Vulnerable neuronal subsets in Alzheimer's and Pick's disease are distinguished by their tau isoform distribution and phosphorylation. *Annals of neurology*, 43(2), pp.193–204.
- Deng, H.-X. et al., 2010. FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis. *Annals of neurology*, 67(6), pp.739–48.
- Deng, H.-X. et al., 2011. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature*, 477(7363), pp.211–5.

- Devenney, E. et al., 2014. Frontotemporal dementia associated with the C9ORF72 mutation: a unique clinical profile. *JAMA neurology*, 71(3), pp.331–9.
- Dobson-Stone, C. et al., 2012. C9ORF72 repeat expansion in clinical and neuropathologic frontotemporal dementia cohorts. *Neurology*, 79(10), pp.995–1001.
- Dols-Icardo, O. et al., 2014. Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. *Human molecular genetics*, 23(3), pp.749–54.
- Donnelly, C.J. et al., 2013. RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron*, 80(2), pp.415–28.
- Van Es, M.A. et al., 2009. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nature Genetics*, 41(10), pp.1083–1087.
- Eslinger, P.J. et al., 2012. Apathy in frontotemporal dementia: behavioral and neuroimaging correlates. *Behavioural neurology*, 25(2), pp.127–36.
- Fahey, C. et al., 2014. Analysis of the hexanucleotide repeat expansion and founder haplotype at C9ORF72 in an Irish psychosis case-control sample. *Neurobiology of aging*, 35(6), pp.1510.e1–5.
- Farg, M.A. et al., 2014. C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Human molecular genetics*, 23(13), pp.3579–95.
- Ferrari, R. et al., 2010. Novel Missense Mutation in Charged Multivesicular Body Protein 2B in a Patient With Frontotemporal Dementia. *Alzheimer disease and associated disorders*.
- Fielenbach, N. & Antebi, A., 2008. C. elegans dauer formation and the molecular basis of plasticity. *Genes & development*, 22(16), pp.2149–65.
- Finch, N. et al., 2009. Plasma progranulin levels predict progranulin mutation status in frontotemporal dementia patients and asymptomatic family members. *Brain: a journal of neurology*, 132(Pt 3), pp.583–91.
- Fiszer, A. & Krzyzosiak, W.J., 2013. RNA toxicity in polyglutamine disorders: concepts, models, and progress of research. *Journal of molecular medicine (Berlin, Germany)*, 91(6), pp.683–91.
- Fogel, B.L. et al., 2012. C9ORF72 expansion is not a significant cause of sporadic spinocerebellar ataxia. *Movement disorders: official journal of the Movement Disorder Society*, 27(14), pp.1832–3.
- Fratta, P. et al., 2012. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Scientific reports*, 2, p.1016.
- Fratta, P. et al., 2013. Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. *Acta neuropathologica*.
- Fratta, P. et al., 2015. Screening a UK amyotrophic lateral sclerosis cohort provides evidence of multiple origins of the C9orf72 expansion. *Neurobiology of aging*, 36(1), pp.546.e1–7.
- Freibaum, B.D. et al., 2015. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature*.

- Fribley, A., Zeng, Q. & Wang, C.-Y., 2004. Proteasome inhibitor PS-341 induces apoptosis through induction of endoplasmic reticulum stress-reactive oxygen species in head and neck squamous cell carcinoma cells. *Molecular and cellular biology*, 24(22), pp.9695–704.
- Galimberti, D. et al., 2013. Autosomal Dominant Frontotemporal Lobar Degeneration Due to the C9ORF72 Hexanucleotide Repeat Expansion: Late-Onset Psychotic Clinical Presentation. *Biological psychiatry*.
- Galimberti, D. et al., 2014. C9ORF72 hexanucleotide repeat expansion is a rare cause of schizophrenia. *Neurobiology of aging*, 35(5), pp.1214.e7–1214.e10.
- García-Redondo, A. et al., 2013. Analysis of the C9orf72 gene in patients with amyotrophic lateral sclerosis in Spain and different populations worldwide. *Human mutation*, 34(1), pp.79–82.
- Gass, J. et al., 2006. Mutations in progranulin are a major cause of ubiquitin-positive frontotemporal lobar degeneration. *Human molecular genetics*, 15(20), pp.2988–3001.
- Gass, J. et al., 2012. Progranulin regulates neuronal outgrowth independent of sortilin. *Molecular neurodegeneration*, 7(1), p.33.
- Gellera, C. et al., 2013. Ubiquilin 2 mutations in Italian patients with amyotrophic lateral sclerosis and frontotemporal dementia. *Journal of neurology, neurosurgery, and psychiatry*, 84(2), pp.183–7.
- Gendron, T.F. 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 neuropathologica*, 126(6), pp.829–44.
- Ghidoni, R. et al., 2008. Low plasma progranulin levels predict progranulin mutations in frontotemporal lobar degeneration. *Neurology*, 71(16), pp.1235–9.
- Gijssels, I. 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. *The Lancet. Neurology*, 11(1), pp.54–65.
- Gijssels, I. et al., 2009. Neuronal inclusion protein TDP-43 has no primary genetic role in FTD and ALS. *Neurobiology of aging*, 30(8), pp.1329–31.
- Gijssels, I. et al., 2008. Progranulin locus deletion in frontotemporal dementia. *Human mutation*, 29(1), pp.53–8.
- Girard, C. et al., 2006. Depletion of SMN by RNA interference in HeLa cells induces defects in Cajal body formation. *Nucleic acids research*, 34(10), pp.2925–32.
- Goedert, M. et al., 1989. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. *Neuron*, 3(4), pp.519–526.
- Goedert, M., Clavaguera, F. & Tolnay, M., 2010. The propagation of prion-like protein inclusions in neurodegenerative diseases. *Trends in neurosciences*, 33(7), pp.317–25.
- Gomez-Deza, J. et al., 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 Neuropathologica Communications*, 3(1), p.38.

- Gómez-Tortosa, E. et al., 2013. C9ORF72 hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. *Neurology*, 80(4), pp.366–70.
- Grice, S.J. et al., 2011. Invertebrate models of spinal muscular atrophy: insights into mechanisms and potential therapeutics. *BioEssays: news and reviews in molecular, cellular and developmental biology*, 33(12), pp.956–65.
- Groen, E.J.N. et al., 2013. ALS-associated mutations in FUS disrupt the axonal distribution and function of SMN. *Human molecular genetics*, 22(18), pp.3690–704.
- Grover, A. et al., 1999. 5' splice site mutations in tau associated with the inherited dementia FTDP-17 affect a stem-loop structure that regulates alternative splicing of exon 10. *The Journal of biological chemistry*, 274(21), pp.15134–43.
- Guerreiro, R.J. et al., 2013. Using exome sequencing to reveal mutations in TREM2 presenting as a frontotemporal dementia-like syndrome without bone involvement. *JAMA neurology*, 70(1), pp.78–84.
- Guo, A. et al., 2010. Progranulin deficiency leads to enhanced cell vulnerability and TDP-43 translocation in primary neuronal cultures. *Brain research*, 1366, pp.1–8.
- Haeusler, A.R. et al., 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature*, 507(7491), pp.195–200.
- Hanger, D.P. et al., 2014. Intracellular and extracellular roles for tau in neurodegenerative disease. *Journal of Alzheimer's disease: JAD*, 40 Suppl 1, pp.S37–45.
- Harms, M.B., Cady, J., et al., 2013. Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. *Neurobiology of aging*, 34(9), pp.2234.e13–9.
- Harms, M.B., Neumann, D., et al., 2013. Parkinson disease is not associated with C9ORF72 repeat expansions. *Neurobiology of aging*, 34(5), pp.1519.e1–2.
- Harvey, R.J., Skelton-Robinson, M. & Rossor, M.N., 2003. The prevalence and causes of dementia in people under the age of 65 years. *Journal of neurology, neurosurgery, and psychiatry*, 74(9), pp.1206–9.
- Hasegawa, M. et al., 1999. FTDP-17 mutations N279K and S305N in tau produce increased splicing of exon 10. *FEBS letters*, 443(2), pp.93–6.
- Hern, J. et al., 1992. The Scottish Motor Neuron Disease Register: a prospective study of adult onset motor neuron disease in Scotland. Methodology, demography and clinical features of incident cases in 1989. *Journal of neurology, neurosurgery, and psychiatry*, 55(7), pp.536–41.
- Hicks, G.G. et al., 2000. Fus deficiency in mice results in defective B-lymphocyte development and activation, high levels of chromosomal instability and perinatal death. *Nature genetics*, 24(2), pp.175–9.
- Hirschbichler, S.T. et al., 2015. Classic PD-like rest tremor associated with the tau p.R406W mutation. *Parkinsonism & Related Disorders*, 21(8), pp.1002–1004.
- Hodges, J.R., 2001. Frontotemporal dementia (Pick's disease): clinical features and assessment. *Neurology*, 56(11 Suppl 4), pp.S6–10.
- Hodges, J.R. et al., 2003. Survival in frontotemporal dementia. *Neurology*, 61(3), pp.349–54.

- Hoell, J.I. et al., 2011. RNA targets of wild-type and mutant FET family proteins. *Nature Structural & Molecular Biology*, 18(12), pp.1428–1431.
- Hosler, B.A. et al., 1996. Three novel mutations and two variants in the gene for Cu/Zn superoxide dismutase in familial amyotrophic lateral sclerosis. *Neuromuscular disorders: NMD*, 6(5), pp.361–6.
- Huang, C. et al., 2011. FUS transgenic rats develop the phenotypes of amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *PLoS genetics*, 7(3), p.e1002011.
- Hübbers, A. et al., 2014. Polymerase chain reaction and Southern blot-based analysis of the C9orf72 hexanucleotide repeat in different motor neuron diseases. *Neurobiology of aging*, 35(5), pp.1214.e1–6.
- Huey, E.D. et al., 2013. C9ORF72 repeat expansions not detected in a group of patients with schizophrenia. *Neurobiology of aging*, 34(4), pp.1309.e9–10.
- Huey, E.D. et al., 2012. FUS and TDP43 genetic variability in FTD and CBS. *Neurobiology of aging*, 33(5), pp.1016.e9–17.
- Hukema, R.K. et al., 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 neuropathologica communications*, 2(1), p.166.
- Hutton, M. et al., 1998. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature*, 393(6686), pp.702–5.
- Irwin, D., Lippa, C.F. & Rosso, A., 2009. Progranulin (PGRN) expression in ALS: an immunohistochemical study. *Journal of the neurological sciences*, 276(1-2), pp.9–13.
- Irwin, D.J. et al., 2013. Cognitive decline and reduced survival in C9orf72 expansion frontotemporal degeneration and amyotrophic lateral sclerosis. *Journal of neurology, neurosurgery, and psychiatry*, 84(2), pp.163–9.
- James, A. et al., 2014. Nucleolar stress with and without p53. *Nucleus (Austin, Tex.)*, 5(5).
- Jansen, G. et al., 1996. Abnormal myotonic dystrophy protein kinase levels produce only mild myopathy in mice. *Nature genetics*, 13(3), pp.316–24.
- Jiang, Z. et al., 2003. Mutations in tau gene exon 10 associated with FTDP-17 alter the activity of an exonic splicing enhancer to interact with Tra2 beta. *The Journal of biological chemistry*, 278(21), pp.18997–9007.
- Joachim, C.L. et al., 1987. Tau antisera recognize neurofibrillary tangles in a range of neurodegenerative disorders. *Annals of neurology*, 22(4), pp.514–20.
- Johnson, J.O. et al., 2010. Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron*, 68(5), pp.857–64.
- Johnson, J.O. et al., 2014. Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nature neuroscience*, 17(5), pp.664–6.
- Jokelainen, M., 1977a. Amyotrophic lateral sclerosis in Finland. I: An epidemiologic study. *Acta neurologica Scandinavica*, 56(3), pp.185–93.

- Jokelainen, M., 1977b. Amyotrophic lateral sclerosis in Finland. II: Clinical characteristics. *Acta neurologica Scandinavica*, 56(3), pp.194–204.
- Jones, K.W. et al., 2001. Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *The Journal of biological chemistry*, 276(42), pp.38645–51.
- Jovičić, A. et al., 2015. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nature Neuroscience*, 18(9), pp.1226–1229.
- Joyce, P.I. et al., 2011. SOD1 and TDP-43 animal models of amyotrophic lateral sclerosis: recent advances in understanding disease toward the development of clinical treatments. *Mammalian genome : official journal of the International Mammalian Genome Society*, 22(7-8), pp.420–48.
- Kabashi, E. et al., 2010. Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. *Human molecular genetics*, 19(4), pp.671–83.
- Kabashi, E. et al., 2012. Impaired proteasome function in sporadic amyotrophic lateral sclerosis. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases*, 13(4), pp.367–71.
- Kabashi, E. et al., 2008. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nature genetics*, 40(5), pp.572–4.
- Kaivorinne, A.-L. et al., 2013. Clinical Characteristics of C9ORF72-Linked Frontotemporal Lobar Degeneration. *Dementia and geriatric cognitive disorders extra*, 3(1), pp.251–62.
- Kaleem, M. et al., 2007. Identification of a novel valosin-containing protein polymorphism in late-onset Alzheimer's disease. *Neuro-degenerative diseases*, 4(5), pp.376–81.
- Kariya, S. et al., 2012. Mutant superoxide dismutase 1 (SOD1), a cause of amyotrophic lateral sclerosis, disrupts the recruitment of SMN, the spinal muscular atrophy protein to nuclear Cajal bodies. *Human molecular genetics*, 21(15), pp.3421–34.
- Kelley, B.J. et al., 2009. Prominent phenotypic variability associated with mutations in Progranulin. *Neurobiology of Aging*, 30(5), pp.739–751.
- Kertesz, A. et al., 2013. Psychosis and hallucinations in frontotemporal dementia with the C9ORF72 mutation: a detailed clinical cohort. *Cognitive and behavioral neurology : official journal of the Society for Behavioral and Cognitive Neurology*, 26(3), pp.146–54.
- Kim, H.J. et al., 2013. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature*, 495(7442), pp.467–73.
- Kinoshita, Y. et al., 2009. Nuclear contour irregularity and abnormal transporter protein distribution in anterior horn cells in amyotrophic lateral sclerosis. *Journal of neuropathology and experimental neurology*, 68(11), pp.1184–92.
- Kleinberger, G. et al., 2010. Increased caspase activation and decreased TDP-43 solubility in progranulin knockout cortical cultures. *Journal of neurochemistry*, 115(3), pp.735–47.
- Koppers, M. et al., 2012. VCP mutations in familial and sporadic amyotrophic lateral sclerosis. *Neurobiology of aging*, 33(4), pp.837.e7–13.

- Kovacs, G.G. et al., 2009. TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea. *Movement disorders : official journal of the Movement Disorder Society*, 24(12), pp.1843–7.
- Kraemer, B.C. et al., 2010. Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis. *Acta neuropathologica*, 119(4), pp.409–19.
- Kraemer, B.C. et al., 2003. Neurodegeneration and defective neurotransmission in a *Caenorhabditis elegans* model of tauopathy. *Proceedings of the National Academy of Sciences of the United States of America*, 100(17), pp.9980–5.
- Kurosawa, M. et al., 2015. Serine 403-phosphorylated p62/SQSTM1 immunoreactivity in inclusions of neurodegenerative diseases. *Neuroscience research*.
- Kwiatkowski, T.J. et al., 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 323(5918), pp.1205–8.
- Kwok, C.T. et al., 2015. VCP mutations are not a major cause of familial amyotrophic lateral sclerosis in the UK. *Journal of the neurological sciences*, 349(1-2), pp.209–13.
- Kwon, I. et al., 2014. Poly-dipeptides encoded by the C9ORF72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science (New York, N.Y.)*.
- Laaksovirta, H. et al., 2010. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet neurology*, 9(10), pp.978–85.
- Lagier-Tourenne, C. et al., 2013. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 110(47), pp.E4530–9.
- Van Langenhove, T. et al., 2010. Genetic contribution of FUS to frontotemporal lobar degeneration. *Neurology*, 74(5), pp.366–71.
- Lanson, N. a & Pandey, U.B., 2012. FUS-related proteinopathies: lessons from animal models. *Brain research*, 1462, pp.44–60.
- Lanson, N.A. et al., 2011. A *Drosophila* model of FUS-related neurodegeneration reveals genetic interaction between FUS and TDP-43. *Human Molecular Genetics*, 20(13), pp.2510–2523.
- Lant, S.B. et al., 2014. Patterns of microglial cell activation in frontotemporal lobar degeneration. *Neuropathology and Applied Neurobiology*, 40(6), pp.686–696.
- Larner, A.J., 2013. Delusion of pregnancy: a case revisited. *Behavioural neurology*, 27(3), pp.293–4.
- Lee, Y.-B. et al., 2013. Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell reports*, 5(5), pp.1178–86.
- Lennox, G. et al., 1988. Ubiquitin is a component of neurofibrillary tangles in a variety of neurodegenerative diseases. *Neuroscience letters*, 94(1-2), pp.211–7.
- Levine, T.P. et al., 2013. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics (Oxford, England)*, 29(4), pp.499–503.

- Li, X.-J. & Li, S., 2011. Proteasomal dysfunction in aging and Huntington disease. *Neurobiology of disease*, 43(1), pp.4–8.
- Liachko, N.F., Guthrie, C.R. & Kraemer, B.C., 2010. Phosphorylation promotes neurotoxicity in a *Caenorhabditis elegans* model of TDP-43 proteinopathy. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 30(48), pp.16208–19.
- Lin, K.-P. et al., 2015. Mutational analysis of MATR3 in Taiwanese patients with amyotrophic lateral sclerosis. *Neurobiology of aging*, 36(5), pp.2005.e1–4.
- Lindenboim, L. et al., 2010. Regulation of stress-induced nuclear protein redistribution: a new function of Bax and Bak uncoupled from Bcl-x(L). *Cell death and differentiation*, 17(2), pp.346–59.
- Linds, A.B. et al., 2015. Trajectories of Behavioural Disturbances Across Dementia Types. *The Canadian journal of neurological sciences. Le journal canadien des sciences neurologiques*, pp.1–6.
- Ling, J.P. et al., 2015. TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD. *Science*, 349(6248), pp.650–655.
- Link, C.D., 1995. Expression of human beta-amyloid peptide in transgenic *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America*, 92(20), pp.9368–72.
- Liu, J. et al., 2004. Toxicity of familial ALS-linked SOD1 mutants from selective recruitment to spinal mitochondria. *Neuron*, 43(1), pp.5–17.
- Livak, K.J. & Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods (San Diego, Calif.)*, 25(4), pp.402–8.
- Logroscino, G. et al., 2010. Incidence of amyotrophic lateral sclerosis in Europe. *Journal of neurology, neurosurgery, and psychiatry*, 81(4), pp.385–90.
- Lomen-Hoerth, C., 2004. Characterization of amyotrophic lateral sclerosis and frontotemporal dementia. *Dementia and geriatric cognitive disorders*, 17(4), pp.337–41.
- Love, S. et al., 1988. Alz-50, ubiquitin and tau immunoreactivity of neurofibrillary tangles, Pick bodies and Lewy bodies. *Journal of neuropathology and experimental neurology*, 47(4), pp.393–405.
- Luigetti, M. et al., 2013. Frontotemporal dementia, Parkinsonism and lower motor neuron involvement in a patient with C9ORF72 expansion. *Amyotrophic lateral sclerosis & frontotemporal degeneration*, 14(1), pp.66–9.
- Mackenzie, I.R. et al., 2013. Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta neuropathologica*, 126(6), pp.859–79.
- Mackenzie, I.R., Rademakers, R. & Neumann, M., 2010. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet neurology*, 9(10), pp.995–1007.
- Mackenzie, I.R.A., Neumann, M., et al., 2011. A harmonized classification system for FTLD-TDP pathology. *Acta neuropathologica*, 122(1), pp.111–3.
- Mackenzie, I.R.A., Munoz, D.G., et al., 2011. Distinct pathological subtypes of FTLD-FUS. *Acta neuropathologica*, 121(2), pp.207–18.

- Mackenzie, I.R.A. et al., 2007. Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. *Annals of neurology*, 61(5), pp.427–34.
- Mackenzie, I.R.A. et al., 2015. Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta neuropathologica*.
- Mahoney, C.J., Beck, J., et al., 2012. Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: clinical, neuroanatomical and neuropathological features. *Brain : a journal of neurology*, 135(Pt 3), pp.736–50.
- Mahoney, C.J., Downey, L.E., et al., 2012. Longitudinal neuroimaging and neuropsychological profiles of frontotemporal dementia with C9ORF72 expansions. *Alzheimer's research & therapy*, 4(5), p.41.
- Majcher, V. et al., 2015. Autophagy receptor defects and ALS-FTLD. *Molecular and cellular neurosciences*, 66(Pt A), pp.43–52.
- Majounie, E. et al., 2012. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet neurology*, 11(4), pp.323–30.
- Manetto, V. et al., 1988. Ubiquitin is associated with abnormal cytoplasmic filaments characteristic of neurodegenerative diseases. *Proceedings of the National Academy of Sciences of the United States of America*, 85(12), pp.4501–5.
- Mann, D.M. et al., 2013. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neuron disease associated with expansions in C9ORF72. *Acta neuropathologica communications*, 1(1), p.68.
- Marat, A.L., Dokainish, H. & McPherson, P.S., 2011. DENN domain proteins: regulators of Rab GTPases. *The Journal of biological chemistry*, 286(16), pp.13791–800.
- Masellis, M. et al., 2006. Novel splicing mutation in the progranulin gene causing familial corticobasal syndrome. *Brain : a journal of neurology*, 129(Pt 11), pp.3115–23.
- May, S. et al., 2014. C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta neuropathologica*.
- Merrilees, J. et al., 2013. Characterization of apathy in persons with frontotemporal dementia and the impact on family caregivers. *Alzheimer disease and associated disorders*, 27(1), pp.62–7.
- Merrilees, J. et al., 2010. Cognitive and behavioral challenges in caring for patients with frontotemporal dementia and amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis*.
- Millecamps, S. et al., 2012. Mutations in UBQLN2 are rare in French amyotrophic lateral sclerosis. *Neurobiology of aging*, 33(4), pp.839.e1–3.
- Miller, L. et al., 2015. p62/SQSTM1 analysis in frontotemporal lobar degeneration. *Neurobiology of aging*, 36(3), pp.1603.e5–9.
- Miller, R.G., Mitchell, J.D. & Moore, D.H., 2012. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). *Cochrane database of systematic reviews (Online)*, 3, p.CD001447.

- Mitchell, J.C. et al., 2013. Overexpression of human wild-type FUS causes progressive motor neuron degeneration in an age- and dose-dependent fashion. *Acta neuropathologica*, 125(2), pp.273–88.
- Mizielinska, S. et al., 2013. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta neuropathologica*, 126(6), pp.845–57.
- Mizielinska, S. et al., 2014. C9orf72 repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science (New York, N.Y.)*, p.science.1256800–.
- Mok, K. et al., 2012. Chromosome 9 ALS and FTD locus is probably derived from a single founder. *Neurobiology of aging*, 33(1), pp.209.e3–8.
- Mori, K., Arzberger, T., et al., 2013. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta neuropathologica*, 126(6), pp.881–93.
- Mori, K., Weng, S.-M., et al., 2013. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL/ALS. *Science (New York, N.Y.)*, 339(6125), pp.1335–8.
- Morita, M. et al., 2006. A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology*, 66(6), pp.839–44.
- Morita, M. et al., 1996. A novel two-base mutation in the Cu/Zn superoxide dismutase gene associated with familial amyotrophic lateral sclerosis in Japan. *Neuroscience letters*, 205(2), pp.79–82.
- Mukherjee, O. et al., 2006. HDDD2 is a familial frontotemporal lobar degeneration with ubiquitin-positive, tau-negative inclusions caused by a missense mutation in the signal peptide of progranulin. *Annals of neurology*, 60(3), pp.314–22.
- Murakami, T. et al., 2012. ALS mutations in FUS cause neuronal dysfunction and death in *Caenorhabditis elegans* by a dominant gain-of-function mechanism. *Human molecular genetics*, 21(1), pp.1–9.
- Murayama, S. et al., 1990. Immunocytochemical and ultrastructural studies of Pick's disease. *Annals of neurology*, 27(4), pp.394–405.
- Nagara, Y. et al., 2013. Impaired cytoplasmic-nuclear transport of hypoxia-inducible factor-1 α in amyotrophic lateral sclerosis. *Brain pathology (Zurich, Switzerland)*, 23(5), pp.534–46.
- Neary, D. et al., 1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology*, 51(6), pp.1546–54.
- Neumann, M. et al., 2009. A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain : a journal of neurology*, 132(Pt 11), pp.2922–31.
- Neumann, M. et al., 2006. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science (New York, N.Y.)*, 314(5796), pp.130–3.
- Niedzielska, E. et al., 2015. Oxidative Stress in Neurodegenerative Diseases. *Molecular neurobiology*.
- Noor, R., Mittal, S. & Iqbal, J., 2002. Superoxide dismutase--applications and relevance to human diseases. *Medical science monitor: international medical journal of experimental and clinical research*, 8(9), pp.RA210–5.

- Nordin, A. et al., 2015. Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Human molecular genetics*.
- Nunomura, A. et al., 2012. Oxidative damage to RNA in aging and neurodegenerative disorders. *Neurotoxicity research*, 22(3), pp.231–48.
- O'Dowd, S. et al., 2012. C9ORF72 expansion in amyotrophic lateral sclerosis/frontotemporal dementia also causes parkinsonism. *Movement disorders : official journal of the Movement Disorder Society*, 27(8), pp.1072–4.
- Obeng, E.A. et al., 2006. Proteasome inhibitors induce a terminal unfolded protein response in multiple myeloma cells. *Blood*, 107(12), pp.4907–16.
- Okamoto, K. et al., 1993. Oculomotor nuclear pathology in amyotrophic lateral sclerosis. *Acta neuropathologica*, 85(5), pp.458–62.
- Okamoto, K., Mizuno, Y. & Fujita, Y., 2008. Bunina bodies in amyotrophic lateral sclerosis. *Neuropathology : official journal of the Japanese Society of Neuropathology*, 28(2), pp.109–15.
- Olejniczak, M., Urbanek, M.O. & Krzyzosiak, W.J., 2015. The Role of the Immune System in Triplet Repeat Expansion Diseases. *Mediators of Inflammation*, 2015, pp.1–11.
- Ortega, Z. & Lucas, J.J., 2014. Ubiquitinâ€“proteasome system involvement in Huntingtonâ€™s disease. *Frontiers in Molecular Neuroscience*, 7, p.77.
- Osborne, R.J. & Thornton, C.A., 2006. RNA-dominant diseases. *Human molecular genetics*, 15 Spec No, pp.R162–9.
- Panda, S.K. et al., 2013. Highly Efficient Targeted Mutagenesis in Mice Using TALENs. *Genetics*.
- Park, H.S. et al., 2011. Proteasome inhibitor MG132-induced apoptosis via ER stress-mediated apoptotic pathway and its potentiation by protein tyrosine kinase p56lck in human Jurkat T cells. *Biochemical pharmacology*, 82(9), pp.1110–25.
- Parkinson, N. et al., 2006. ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology*, 67(6), pp.1074–7.
- Pearson, J.P. et al., 2011. Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *Journal of neurology*, 258(4), pp.647–55.
- Pellizzoni, L. et al., 2001. The survival of motor neurons (SMN) protein interacts with the snoRNP proteins fibrillarin and GAR1. *Current biology : CB*, 11(14), pp.1079–88.
- Pereson, S. et al., 2009. Progranulin expression correlates with dense-core amyloid plaque burden in Alzheimer disease mouse models. *The Journal of pathology*, 219(2), pp.173–81.
- Perry, D.C. et al., 2012. Voxel-based morphometry in patients with obsessive-compulsive behaviors in behavioral variant frontotemporal dementia. *European journal of neurology : the official journal of the European Federation of Neurological Societies*, 19(6), pp.911–7.
- Peters, F. et al., 2006. Orbitofrontal dysfunction related to both apathy and disinhibition in frontotemporal dementia. *Dementia and geriatric cognitive disorders*, 21(5-6), pp.373–9.

- Petkau, T.L. et al., 2012. Synaptic dysfunction in progranulin-deficient mice. *Neurobiology of disease*, 45(2), pp.711–22.
- Petkau, T.L. & Leavitt, B.R., 2014. Progranulin in neurodegenerative disease. *Trends in Neurosciences*, 37(7), pp.388–398.
- Philips, T. et al., 2010. Microglial upregulation of progranulin as a marker of motor neuron degeneration. *Journal of neuropathology and experimental neurology*, 69(12), pp.1191–200.
- Pickering-Brown, S. et al., 2000. Pick's disease is associated with mutations in the tau gene. *Annals of neurology*, 48(6), pp.859–67.
- Pickering-Brown, S.M. et al., 2008. Frequency and clinical characteristics of progranulin mutation carriers in the Manchester frontotemporal lobar degeneration cohort: comparison with patients with MAPT and no known mutations. *Brain: a journal of neurology*, 131(Pt 3), pp.721–31.
- Piguat, O. et al., 2011. Behavioural-variant frontotemporal dementia: diagnosis, clinical staging, and management. *The Lancet Neurology*, 10(2), pp.162–172.
- Pijnenburg, Y.A.L. et al., 2004. Initial complaints in frontotemporal lobar degeneration. *Dementia and geriatric cognitive disorders*, 17(4), pp.302–6.
- Pizzuti, A. & Petrucci, S., 2011. Mitochondrial dysfunction as a cause of ALS. *Archives italiennes de biologie*, 149(1), pp.113–9.
- Pokrishevsky, E. et al., 2012. Aberrant localization of FUS and TDP43 is associated with misfolding of SOD1 in amyotrophic lateral sclerosis. *PloS one*, 7(4), p.e35050.
- Pooler, A.M. et al., 2015. Amyloid accelerates tau propagation and toxicity in a model of early Alzheimer's disease. *Acta neuropathologica communications*, 3, p.14.
- Poorkaj, P. et al., 1998. Tau is a candidate gene for chromosome 17 frontotemporal dementia. *Annals of neurology*, 43(6), pp.815–25.
- Proudfoot, M. et al., 2014. Early dipeptide repeat pathology in a frontotemporal dementia kindred with C9ORF72 mutation and intellectual disability. *Acta neuropathologica*, 127(3), pp.451–8.
- Prudencio, M. et al., 2015. Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nature neuroscience*, 18(8), pp.1175–1182.
- Rademakers, R. et al., 2008. Common variation in the miR-659 binding-site of GRN is a major risk factor for TDP43-positive frontotemporal dementia. *Human molecular genetics*, 17(23), pp.3631–42.
- Rademakers, R. et al., 2007. Phenotypic variability associated with progranulin haploinsufficiency in patients with the common 1477C-->T (Arg493X) mutation: an international initiative. *The Lancet. Neurology*, 6(10), pp.857–68.
- Rademakers, R., Neumann, M. & Mackenzie, I.R., 2012. Advances in understanding the molecular basis of frontotemporal dementia. *Nature reviews. Neurology*, 8(8), pp.423–34.
- Ratnavalli, E. et al., 2002. The prevalence of frontotemporal dementia. *Neurology*, 58(11), pp.1615–21.

- Ratti, A. et al., 2012. C9ORF72 repeat expansion in a large Italian ALS cohort: evidence of a founder effect. *Neurobiology of aging*, 33(10), pp.2528.e7–14.
- Reddy, K. et al., 2013. The disease-associated r(GGGGCC)_n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *The Journal of biological chemistry*, 288(14), pp.9860–6.
- Reddy, S. et al., 1996. Mice lacking the myotonic dystrophy protein kinase develop a late onset progressive myopathy. *Nature genetics*, 13(3), pp.325–35.
- Renton, A.E. et al., 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron*, 72(2), pp.257–68.
- Ringholz, G.M. et al., 2005. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology*, 65(4), pp.586–90.
- Rohrer, J.D. et al., 2008. Parietal lobe deficits in frontotemporal lobar degeneration caused by a mutation in the progranulin gene. *Archives of neurology*, 65(4), pp.506–13.
- Rohrer, J.D. et al., 2009. The heritability and genetics of frontotemporal lobar degeneration. *Neurology*, 73(18), pp.1451–6.
- Rohrer, J.D. & Warren, J.D., 2011. Phenotypic signatures of genetic frontotemporal dementia. *Current opinion in neurology*, 24(6), pp.542–9.
- Rollinson, S. et al., 2015. A small deletion in C9orf72 hides a proportion of expansion carriers in FTLD. *Neurobiology of aging*, 36(3), pp.1601.e1–5.
- Rollinson, S., Mead, S., et al., 2011. Frontotemporal lobar degeneration genome wide association study replication confirms a risk locus shared with amyotrophic lateral sclerosis. *Neurobiology of Aging*, 32(4), pp.758.e1–758.e7.
- Rollinson, S., Rohrer, J.D., et al., 2011. No association of PGRN 3'UTR rs5848 in frontotemporal lobar degeneration. *Neurobiology of aging*, 32(4), pp.754–5.
- Rollinson, S. et al., 2007. TDP-43 gene analysis in frontotemporal lobar degeneration. *Neuroscience letters*, 419(1), pp.1–4.
- Rosen, D.R. et al., 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature*, 362(6415), pp.59–62.
- Rosen, H.J. et al., 2002. Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology*, 58(2), pp.198–208.
- Rosenblum, W.I., 2014. Why Alzheimer trials fail: removing soluble oligomeric beta amyloid is essential, inconsistent, and difficult. *Neurobiology of aging*, 35(5), pp.969–74.
- Rossi, S. et al., 2015. Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *Journal of cell science*, 128(9), pp.1787–99.
- Rosso, S.M. et al., 2003. Frontotemporal dementia in The Netherlands: patient characteristics and prevalence estimates from a population-based study. *Brain: a journal of neurology*, 126(Pt 9), pp.2016–22.

- Rubino, E. et al., 2012. SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Neurology*, 79(15), pp.1556–62.
- Russ, J. et al., 2015. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta neuropathologica*, 129(1), pp.39–52.
- Ryan, C.L. et al., 2009. Progranulin is expressed within motor neurons and promotes neuronal cell survival. *BMC neuroscience*, 10(1), p.130.
- Sapp, P.C. et al., 1995. Identification of three novel mutations in the gene for Cu/Zn superoxide dismutase in patients with familial amyotrophic lateral sclerosis. *Neuromuscular disorders : NMD*, 5(5), pp.353–7.
- Sareen, D. et al., 2013. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Science translational medicine*, 5(208), p.208ra149.
- Sasaki, S., 2010. Endoplasmic reticulum stress in motor neurons of the spinal cord in sporadic amyotrophic lateral sclerosis. *Journal of neuropathology and experimental neurology*, 69(4), pp.346–55.
- Schafer, W.R., 2005. Egg-laying. *WormBook : the online review of C. elegans biology*, pp.1–7.
- Schludi, M.H. et al., 2015. Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta neuropathologica*.
- Schröder, R. et al., 2005. Mutant valosin-containing protein causes a novel type of frontotemporal dementia. *Annals of neurology*, 57(3), pp.457–61.
- Schumacher, A. et al., 2009. No association of TDP-43 with sporadic frontotemporal dementia. *Neurobiology of Aging*, 30(1), pp.157–159.
- Schwahnhäusser, B. et al., 2011. Global quantification of mammalian gene expression control. *Nature*, 473(7347), pp.337–42.
- Ségalat, L., 2007. Invertebrate animal models of diseases as screening tools in drug discovery. *ACS chemical biology*, 2(4), pp.231–6.
- Shan, X., Chang, Y. & Lin, C.G., 2007. Messenger RNA oxidation is an early event preceding cell death and causes reduced protein expression. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 21(11), pp.2753–64.
- Shankaran, S.S. et al., 2008. Missense mutations in the progranulin gene linked to frontotemporal lobar degeneration with ubiquitin-immunoreactive inclusions reduce progranulin production and secretion. *The Journal of biological chemistry*, 283(3), pp.1744–53.
- Shatunov, A. et al., 2010. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet neurology*, 9(10), pp.986–94.
- Shibata, N. et al., 1996. Immunohistochemical study on superoxide dismutases in spinal cords from autopsied patients with amyotrophic lateral sclerosis. *Developmental neuroscience*, 18(5-6), pp.492–8.
- De Silva, R. et al., 2006. An immunohistochemical study of cases of sporadic and inherited frontotemporal lobar degeneration using 3R- and 4R-specific tau monoclonal antibodies. *Acta neuropathologica*, 111(4), pp.329–40.

- Skibinski, G. et al., 2005. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nature genetics*, 37(8), pp.806–8.
- Sleegers, K. et al., 2009. Serum biomarker for progranulin-associated frontotemporal lobar degeneration. *Annals of Neurology*, 65(5), pp.603–609.
- Sleigh, J.N. & Sattelle, D.B., 2010. *C. Elegans* Models of Neuromuscular Diseases Expedite Translational Research. *Translational Neuroscience*, 1(3), pp.214–227.
- Smith, B.N. et al., 2013. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *European journal of human genetics: EJHG*, 21(1), pp.102–8.
- Smith, K.R. et al., 2012. Strikingly different clinicopathological phenotypes determined by progranulin-mutation dosage. *American journal of human genetics*, 90(6), pp.1102–7.
- Snowden, J., Neary, D. & Mann, D., 2007. Frontotemporal lobar degeneration: clinical and pathological relationships. *Acta neuropathologica*, 114(1), pp.31–8.
- Snowden, J.S., Rollinson, S., et al., 2012. Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. *Brain: a journal of neurology*, 135(Pt 3), pp.693–708.
- Snowden, J.S. et al., 2006. Progranulin gene mutations associated with frontotemporal dementia and progressive non-fluent aphasia. *Brain: a journal of neurology*, 129(Pt 11), pp.3091–102.
- Snowden, J.S. et al., 2015. Psychosis associated with expansions in the C9orf72 gene: the influence of a 10 base pair gene deletion. *Journal of neurology, neurosurgery, and psychiatry*, p.jnnp-2015-310441-.
- Snowden, J.S., Rollinson, S., et al., 2012. Psychosis, C9ORF72 and dementia with Lewy bodies. *Journal of neurology, neurosurgery, and psychiatry*, 83(10), pp.1031–2.
- Snowden, J.S. et al., 2011. The most common type of FTL-D-FUS (aFTLD-U) is associated with a distinct clinical form of frontotemporal dementia but is not related to mutations in the FUS gene. *Acta neuropathologica*, 122(1), pp.99–110.
- Soo, K.Y. et al., 2015. Rab1-dependent ER-Golgi transport dysfunction is a common pathogenic mechanism in SOD1, TDP-43 and FUS-associated ALS. *Acta neuropathologica*.
- Spillantini, M.G. et al., 1998. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proceedings of the National Academy of Sciences of the United States of America*, 95(13), pp.7737–41.
- Spina, S. et al., 2007. Corticobasal syndrome associated with the A9D Progranulin mutation. *Journal of neuropathology and experimental neurology*, 66(10), pp.892–900.
- Sreedharan, J. et al., 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science (New York, N. Y.)*, 319(5870), pp.1668–72.
- Štálekár, M. et al., 2015. Proteomic analyses reveal that loss of TDP-43 affects RNA processing and intracellular transport. *Neuroscience*, 293, pp.157–70.
- Sumi, H. et al., 2009. Nuclear TAR DNA binding protein 43 expression in spinal cord neurons correlates with the clinical course in amyotrophic lateral sclerosis. *Journal of neuropathology and experimental neurology*, 68(1), pp.37–47.

- Sun, S. et al., 2015. ALS-causative mutations in FUS/TLS confer gain and loss of function by altered association with SMN and U1-snRNP. *Nature communications*, 6, p.6171.
- Suzuki, H. et al., 2009. ALS-linked P56S-VAPB, an aggregated loss-of-function mutant of VAPB, predisposes motor neurons to ER stress-related death by inducing aggregation of co-expressed wild-type VAPB. *Journal of neurochemistry*, 108(4), pp.973–985.
- Suzuki, N. et al., 2013. The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nature neuroscience*, 16(12), pp.1725–7.
- Van Swieten, J.C. et al., 2007. The DeltaK280 mutation in MAP tau favors exon 10 skipping in vivo. *Journal of neuropathology and experimental neurology*, 66(1), pp.17–25.
- Sydow, A. & Mandelkow, E.-M., 2010. “Prion-like” propagation of mouse and human tau aggregates in an inducible mouse model of tauopathy. *Neuro-degenerative diseases*, 7(1-3), pp.28–31.
- Synofzik, M. et al., 2012. Screening in ALS and FTD patients reveals 3 novel UBQLN2 mutations outside the PXX domain and a pure FTD phenotype. *Neurobiology of aging*, 33(12), pp.2949.e13–7.
- Szegezdi, E. et al., 2006. Mediators of endoplasmic reticulum stress-induced apoptosis. *EMBO reports*, 7(9), pp.880–5.
- Tafari, F. et al., 2015. SOD1 misplacing and mitochondrial dysfunction in amyotrophic lateral sclerosis pathogenesis. *Frontiers in cellular neuroscience*, 9, p.336.
- Tanaka, M., Chock, P.B. & Stadtman, E.R., 2007. Oxidized messenger RNA induces translation errors. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), pp.66–71.
- Tanaka, Y. et al., 2013. Exacerbated inflammatory responses related to activated microglia after traumatic brain injury in progranulin-deficient mice. *Neuroscience*, 231, pp.49–60.
- Tanaka, Y. et al., 2014. Possible involvement of lysosomal dysfunction in pathological changes of the brain in aged progranulin-deficient mice. *Acta neuropathologica communications*, 2, p.78.
- Taniguchi, S. et al., 2004. The neuropathology of frontotemporal lobar degeneration with respect to the cytological and biochemical characteristics of tau protein. *Neuropathology and applied neurobiology*, 30(1), pp.1–18.
- Tao, Z. et al., 2015. Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Human molecular genetics*, 24(9), pp.2426–41.
- Tapia, O. et al., 2014. The SMN Tudor SIM-like domain is key to Smd1 and coilin interactions and to Cajal body biogenesis. *Journal of cell science*, 127(Pt 5), pp.939–46.
- Tashiro, Y. et al., 2012. Motor neuron-specific disruption of proteasomes, but not autophagy, replicates amyotrophic lateral sclerosis. *Journal of Biological Chemistry*, 287(51), pp.42984–42994.
- Tatom, J.B. et al., 2009. Mimicking aspects of frontotemporal lobar degeneration and Lou Gehrig’s disease in rats via TDP-43 overexpression. *Molecular therapy: the journal of the American Society of Gene Therapy*, 17(4), pp.607–13.

- Taylor, R.C. et al., 2013. Changes in translational efficiency is a dominant regulatory mechanism in the environmental response of bacteria. *Integrative biology : quantitative biosciences from nano to macro*, 5(11), pp.1393–406.
- Thandapani, P. et al., 2013. Defining the RGG/RG motif. *Molecular cell*, 50(5), pp.613–23.
- Therrien, M. et al., 2013a. Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in *C. elegans*. *PLoS one*, 8(12), p.e83450.
- Therrien, M. et al., 2013b. Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in *C. elegans*. *PLoS one*, 8(12), p.e83450.
- Theuns, J. et al., 2014. Global investigation and meta-analysis of the C9orf72 (G4C2)_n repeat in Parkinson disease. *Neurology*, 83(21), pp.1906–13.
- Tian, Q. et al., 2004. Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Molecular & cellular proteomics : MCP*, 3(10), pp.960–9.
- Ticozzi, N. et al., 2011. Mutational analysis of TARDBP in neurodegenerative diseases. *Neurobiology of aging*, 32(11), pp.2096–9.
- Tiloca, C. et al., 2013. Screening of the PFN1 gene in sporadic amyotrophic lateral sclerosis and in frontotemporal dementia. *Neurobiology of aging*, 34(5), pp.1517.e9–10.
- Tisdale, S. & Pellizzoni, L., 2015. Disease Mechanisms and Therapeutic Approaches in Spinal Muscular Atrophy. *Journal of Neuroscience*, 35(23), pp.8691–8700.
- Todd, P.K. et al., 2013. CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron*, 78(3), pp.440–55.
- Tran, D. et al., 2012. A mutation in VAPB that causes amyotrophic lateral sclerosis also causes a nuclear envelope defect. *Journal of cell science*, 125(Pt 12), pp.2831–6.
- Troakes, C. et al., 2012. An MND/ALS phenotype associated with C9orf72 repeat expansion: abundant p62-positive, TDP-43-negative inclusions in cerebral cortex, hippocampus and cerebellum but without associated cognitive decline. *Neuropathology : official journal of the Japanese Society of Neuropathology*, 32(5), pp.505–14.
- Tsai, C.-P. et al., 2012. A hexanucleotide repeat expansion in C9ORF72 causes familial and sporadic ALS in Taiwan. *Neurobiology of aging*, 33(9), pp.2232.e11–2232.e18.
- Turner, B.J. et al., 2014. Overexpression of survival motor neuron improves neuromuscular function and motor neuron survival in mutant SOD1 mice. *Neurobiology of aging*, 35(4), pp.906–15.
- Turner, B.J. et al., 2008. TDP-43 expression in mouse models of amyotrophic lateral sclerosis and spinal muscular atrophy. *BMC neuroscience*, 9, p.104.
- Valdmanis, P.N. et al., 2007. Three families with amyotrophic lateral sclerosis and frontotemporal dementia with evidence of linkage to chromosome 9p. *Archives of neurology*, 64(2), pp.240–5.
- Vance, C. et al., 2006. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain : a journal of neurology*, 129(Pt 4), pp.868–76.

- Vance, C. et al., 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science (New York, N.Y.)*, 323(5918), pp.1208–11.
- Vijayvergiya, C., 2005. Mutant Superoxide Dismutase 1 Forms Aggregates in the Brain Mitochondrial Matrix of Amyotrophic Lateral Sclerosis Mice. *Journal of Neuroscience*, 25(10), pp.2463–2470.
- Vogel, C. et al., 2010. Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. *Molecular systems biology*, 6, p.400.
- Waite, A.J. et al., 2014. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiology of aging*, 35(7), pp.1779.e5–1779.e13.
- Walker, A.K. et al., 2013. ALS-associated TDP-43 induces endoplasmic reticulum stress, which drives cytoplasmic TDP-43 accumulation and stress granule formation. *PLoS one*, 8(11), p.e81170.
- Wang, G. et al., 2015. Cytoplasmic mislocalization of RNA splicing factors and aberrant neuronal gene splicing in TDP-43 transgenic pig brain. *Molecular neurodegeneration*, 10(1), p.42.
- Wang, J. et al., 2009. An ALS-linked mutant SOD1 produces a locomotor defect associated with aggregation and synaptic dysfunction when expressed in neurons of *Caenorhabditis elegans*. *PLoS genetics*, 5(1), p.e1000350.
- Wang, X. et al., 2015. Activation of ER Stress and Autophagy Induced by TDP-43 A315T as Pathogenic Mechanism and the Corresponding Histological Changes in Skin as Potential Biomarker for ALS with the Mutation. *International journal of biological sciences*, 11(10), pp.1140–9.
- Ward, M.E. et al., 2014. Early retinal neurodegeneration and impaired Ran-mediated nuclear import of TDP-43 in progranulin-deficient FTLD. *The Journal of experimental medicine*, 211(10), pp.1937–45.
- Watanabe, Y. et al., 2000. Frameshift, nonsense and non amino acid altering mutations in SOD1 in familial ALS: report of a Japanese pedigree and literature review. *Amyotrophic lateral sclerosis and other motor neuron disorders: official publication of the World Federation of Neurology, Research Group on Motor Neuron Diseases*, 1(4), pp.251–8.
- Watts, G.D.J. et al., 2004. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nature genetics*, 36(4), pp.377–81.
- Wen, X. 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(6), pp.1213–25.
- White, J.G. et al., 1986. The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 314(1165), pp.1–340.
- Wolfe, D.M. et al., 2013. Autophagy failure in Alzheimer's disease and the role of defective lysosomal acidification. *The European journal of neuroscience*, 37(12), pp.1949–61.
- Worms, P.M., 2001. The epidemiology of motor neuron diseases: a review of recent studies. *Journal of the neurological sciences*, 191(1-2), pp.3–9.

- Wray, S. et al., 2008. Direct analysis of tau from PSP brain identifies new phosphorylation sites and a major fragment of N-terminally cleaved tau containing four microtubule-binding repeats. *Journal of neurochemistry*, 105(6), pp.2343–52.
- Wu, C.-H. et al., 2012. Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature*, 488(7412), pp.499–503.
- Wu, L.-S., Cheng, W.-C. & Shen, C.-K.J., 2012. Targeted depletion of TDP-43 expression in the spinal cord motor neurons leads to the development of amyotrophic lateral sclerosis-like phenotypes in mice. *The Journal of biological chemistry*, 287(33), pp.27335–44.
- Xi, Z. et al., 2013. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *American journal of human genetics*, 92(6), pp.981–9.
- Xi, Z. et al., 2014. Hypermethylation of the CpG-island near the C9orf72 G₄C₂-repeat expansion in FTLD patients. *Human molecular genetics*, 23(21), pp.5630–7.
- Xi, Z. et al., 2012. Investigation of c9orf72 in 4 neurodegenerative disorders. *Archives of neurology*, 69(12), pp.1583–90.
- Xi, Z. et al., 2015. The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta neuropathologica*, 129(5), pp.715–27.
- Xia, R. et al., 2012. Motor neuron apoptosis and neuromuscular junction perturbation are prominent features in a Drosophila model of Fus-mediated ALS. *Molecular neurodegeneration*, 7, p.10.
- Xiao, S. et al., 2015. Isoform Specific Antibodies Reveal Distinct Subcellular Localizations of C9orf72 in Amyotrophic Lateral Sclerosis. *Annals of neurology*.
- Xu, Y.-F. et al., 2010. Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *The Journal of neuroscience: the official journal of the Society for Neuroscience*, 30(32), pp.10851–9.
- Xu, Z. et al., 2013. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19), pp.7778–83.
- Yamazaki, T. et al., 2012. FUS-SMN protein interactions link the motor neuron diseases ALS and SMA. *Cell reports*, 2(4), pp.799–806.
- Yan, J. et al., 2010. Frameshift and novel mutations in FUS in familial amyotrophic lateral sclerosis and ALS/dementia. *Neurology*, 75(9), pp.807–14.
- Yang, D. et al., 2015. FTD/ALS-associated poly(GR) protein impairs the Notch pathway and is recruited by poly(GA) into cytoplasmic inclusions. *Acta neuropathologica*.
- Yoshino, Y. et al., 2015. No abnormal hexanucleotide repeat expansion of C9ORF72 in Japanese schizophrenia patients. *Journal of neural transmission (Vienna, Austria: 1996)*, 122(5), pp.731–2.
- Yu, Y. et al., 2015. U1 snRNP is mislocalized in ALS patient fibroblasts bearing NLS mutations in FUS and is required for motor neuron outgrowth in zebrafish. *Nucleic acids research*, 43(6), pp.3208–18.

- Van der Zee, J. et al., 2013. A pan-European study of the C9orf72 repeat associated with FTL: geographic prevalence, genomic instability, and intermediate repeats. *Human mutation*, 34(2), pp.363–73.
- Van der Zee, J. et al., 2007. Mutations other than null mutations producing a pathogenic loss of progranulin in frontotemporal dementia. *Human Mutation*, 28(4), pp.416–416.
- Zhang, C.-C. et al., 2015. The Role of MAPT in Neurodegenerative Diseases: Genetics, Mechanisms and Therapy. *Molecular neurobiology*.
- Zhang, D. et al., 2012. Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Frontiers in genetics*, 3, p.283.
- Zhang, J. et al., 2006. Altered distributions of nucleocytoplasmic transport-related proteins in the spinal cord of a mouse model of amyotrophic lateral sclerosis. *Acta neuropathologica*, 112(6), pp.673–80.
- Zhang, K. et al., 2015. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature*.
- Zhang, Y.-J. et al., 2014. Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta neuropathologica*, 128(4), pp.505–24.
- Zhou, H. et al., 2010. Transgenic rat model of neurodegeneration caused by mutation in the TDP gene. *PLoS genetics*, 6(3), p.e1000887.
- Zhukareva, V. et al., 2002. Sporadic Pick's disease: a tauopathy characterized by a spectrum of pathological tau isoforms in gray and white matter. *Annals of neurology*, 51(6), pp.730–9.
- Zou, Z.-Y. et al., 2012. Screening of the FUS gene in familial and sporadic amyotrophic lateral sclerosis patients of Chinese origin. *European journal of neurology: the official journal of the European Federation of Neurological Societies*, 19(7), pp.977–83.
- Zu, T. et al., 2011. Non-ATG-initiated translation directed by microsatellite expansions. *Proceedings of the National Academy of Sciences of the United States of America*, 108(1), pp.260–5.