In vitro Investigation of Dipeptide Repeat Proteins in C9ORF72-associated Motor Neuron Disease and Frontotemporal Dementia

Mona Radwan

ORCID ID: 0000-0002-3110-4684

Submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

June 2020

Supervised by Professors **Danny M. Hatters** & **Gavin E. Reid**

Department of Biochemistry and Molecular Biology

Bio21 Molecular Science and Biotechnology Institute

The University of Melbourne

Abstract

Mutations that expand a hexanucleotide (GGGGCC) tandem repeat sequence (HTRS) in intron-1 of the C9ORF72 gene is the major cause of familial motor neuron disease MND and frontotemporal dementia (FTD). The expanded HTRSs trigger abnormal bi-directional transcription, which generates both sense and antisense RNA transcripts that form foci. The subsequent translation of all reading frames of HTRSs generates five types of dipeptide repeat proteins (DRPs) via Repeat-Associated Non-ATG (RAN) translation. The sense (G_4C_2) RNA encodes poly-glycine-alanine (poly-GA), poly-glycine-proline (poly-GP), and poly-glycine-arginine (poly-GR), while the antisense (C_4G_2) RNA generates poly-GP, poly-proline-arginine (poly-PR), and poly-alanine-proline (poly-AP). The presumably toxic RAN proteins aggregate in neurons suggesting that they play a crucial role(s) in NMD and FTD pathogenesis. The mechanism underlying the DPR protein-induced neuronal toxicity have not been fully elucidated.

The current thesis has employed a quantitative proteomics approach aimed at understanding how these DRPs, as putative “junk” proteins, confer damage to the cellular biochemistry via first identifying the protein interactome of each DRP and secondly by determining the impact of these DRP sequences on the global proteome. To this end, we have developed artificial cDNA constructs engineered by a randomized codon strategy to express short (10x) and long (101x) DPR proteins avoiding GC repeat sequences. When transiently expressed in mouse neuroblastoma cells, the soluble poly-PA and poly-GP were well tolerated at the lengths tested. Clear length-dependent neurotoxicity was seen for poly-GA that formed cytoplasmic aggregates only at the long length. When compared to 97Q-expanded huntingtin exon 1 aggregates, GA_{101} exhibited faster aggregation kinetics and stronger proteasomal recruitment as revealed in the aggregate proteome. Conversely, the arginine-rich DPRs (Poly-PR and Poly-GR) formed discrete nuclear aggregates and were profoundly toxic at both short and long repeat lengths.

Data presented within this thesis has also provided evidence that the arginine valency in poly-PR and poly-GR triggers their pernicious binding to the proteome, compared to the sparse binding of the more inert poly-AP and poly-GA. The poly-PR and poly-GR interactome included proteins involved in ribosome biogenesis, translation and RNA splicing machinery. Using stalling reporters, long R-rich DPRs stalled translation, through a putative electrostatic jamming mechanism. These findings also indicate that Arg-rich dipeptide polymers are mimetics of endogenous arginine methylase substrates and recruit these

methylase proteins. A consequence is hypomethylation of the proteome, including key proteins known to be involved in ALS mechanisms. The mechanism explains the previously reported abnormalities in arginine methylation patterns in the brains of ALS patients. Finally, this work has unravelled how Arg-rich dipeptide polymers impair the assembly of the actin cytoskeleton. This finding provides important mechanistic insight into the previously reported actin polymerization defects in multiple forms of amyotrophic lateral sclerosis.

Declaration

This is to certify that

i. the thesis comprises only my original work towards the Doctor of Philosophy except where indicated in the Preface;

ii. due acknowledgement has been made in the text to all other material used; and

iii. the thesis is fewer than 100,000 words in length, exclusive of tables, figures and bibliographies, and complies with the stipulations set out for the degree of Doctor of Philosophy by the University of Melbourne.

Mona Radwan (B. Pharm., M.Sc.)
Department of Biochemistry & Molecular Biochemistry
Faculty of Medicine, Dentistry & Health Sciences
The University of Melbourne;
Parkville, Victoria 3010, Australia

Date: 01/06/2020

Preface

Chapter 3 of the thesis is presented as a peer-reviewed publication:

> **Radwan M,** Ang CS, Ormsby AR, Cox D, Reid GE, Hatters DM (2020). Arginine in C9ORF72 dipolypeptides mediates promiscuous proteome binding and multiple modes of toxicity. Molecular & Cellular Proteomics. 19 (4) 640-654; doi.org/10.1074/mcp.RA119.001888.

C.S.A. supervised performing the proteomics experiments. A.R.O., J.C.D. and D.C. cloned exon 1 of Huntington (Httex1) with different polyQ expansions (25Q, 72Q and 97Q) into the dual-fluorescence translation stall reporter and performed the flow cytometry analysis presented in Figure 3-4D.

Chapter 4 served as a basis for a manuscript under review with my authorship:

> **Radwan M**, Lilley JD, Ang CS, Reid GE, Hatters DM (2020). Inclusion bodies formed by polyglutamine and poly(glycine-alanine) are enriched with distinct proteomes but converge in proteins that are risk factors for disease and involved in protein degradation. bioRxiv 2020.05.04.076547; doi: https://doi.org/10.1101/2020.05.04.076547.

J.D.L. has participated in purifying protein aggregates using fluorescence-activated cell sorting (FACS) for subsequent aggregate proteome analysis in Chapter 4.

This work was funded by a grant to Danny M. Hatters (Motor Neuron Disease Research Institute, Australia small grant). I also acknowledge support from an Australian Government Research Training Program (RTP) Scholarship and an Egyptian Ministry of Higher Education and Scientific Research PhD scholarship.

Acknowledgements

"Keep up the good work" and "It is part of learning process!" are two supportive messages that ended every one-to-one meeting with my PhD mentor Prof. Danny Hatters whenever the experiments worked or failled. I was so fortunate to have such an ideal mentor who encouraged me throughout my PhD journey and believed in my abilities, even when I don't believe in them myself. I will be forever grateful to you for helping me re-discover my passion for science and with the freedom culture and independence in his lab I can now conquer the world as a research scientist!

I wish to express my heartfelt gratitude to Prof. Gavin Reid, my co-supervisor. Gavin invested a lot of time patiently introducing me to the amazing world of proteomics. I really appreciate his patience, immediate feedback and guidance in helping me overcome challenges risen during sample preparation for Mass spectrometry and data analysis. My sincere thanks go to him also for the excellent questions that greatly improve the quality of this thesis.

I am incredibly thankful to Dr. Ching-Seng Ang who patiently taught me sample preparation for Mass spectrometry and MS data analysis using Proteome Discoverer, Max Quant and Perceus. Associate professor Nick Williamson and Dr. Shuai Nie have been providing a happy MS facility to work at. Dr. Matt Burton helped with cell sorting by flow cytometry described in Chapter 3 and 4.

I thank my distinguished committee members; Dr. Marie Bogoyevitch and Dr. Bradley Turner, for their support of my work with valuable suggestions during each committee meeting.

Thank you to all the lovely past and present members of Hatters lab: Yasmin Ramdzan, Rebecca Wood, Dr. Dezerae Cox, Dr. Xiaojing Sui, Nagaraj Moily, Autumn Bricker, Yoon Hee Choi, Candice Raeburn, Jordan Lilley and Dr. Estella Newcombe. You all have immensely contributed to my personal and professional time throughout my PhD candidature. Special mention must be made to Angelique Ormsby, for her wealth of technical expertise and assistance, particularly with flow cytometry.

My foremost gratitude goes to my dearest parents Radwan Elsayed and Aziza Mousa, to whom this thesis is dedicated. Last but not least, I would like to thank my husband, Khaled Elnahriry, my two kids and my twin sister Asmaa Radwan who have shared every step of this massive rollercoaster ride. Without their endless love, patience and support, none of this would have happened.

Communications

Publications:

1. **Radwan M**, Lilley JD, Ang CS, Reid GE, Hatters DM (2020). Inclusion bodies formed by polyglutamine and poly(glycine-alanine) are enriched with distinct proteomes but converge in proteins that are risk factors for disease and involved in protein degradation. bioRxiv 2020.05.04.076547; doi: https://doi.org/10.1101/2020.05.04.076547.
2. **Radwan M**, Ang CS, Ormsby AR, Cox D, Reid GE, Hatters DM (2020). Arginine in C9ORF72 dipolypeptides mediates promiscuous proteome binding and multiple modes of toxicity. *Molecular & Cellular Proteomics*. 19 (4) 640-654. doi.org/10.1074/mcp.RA119.001888.
3. Wood RJ, Ormsby AR, **Radwan M**, Cox D, Sharma A, Vöpel T, Ebbinghaus S, Oliveberg M, Reid GE, Dickson A, Hatters DM (2018). A biosensor-based framework to measure latent proteostasis capacity. *Nature Communications*, 9(1):287. doi.org/10.1038/s41467-017-02562-5.
4. **Radwan M,** Wood RJ, Sui X, Hatters DM (2017). When proteostasis goes bad: Protein aggregation in the cell. IUBMB Life,69(2):49-54. doi.org/10.1002/iub.1597.
5. Chen MZ, Moily NS, Bridgford JL, Wood RJ, **Radwan M**, Smith TA, Song Z, Tang BZ, Tilley L, Xu X, Reid GE, Pouladi MA, Hong Y, Hatters DM (2017). A thiol probe for measuring unfolded protein load and proteostasis in cells. *Nature Communications*, 8(1): 474. doi.org/10.1038/s41467-017-00203-5.

Table of Contents

List of Figures

List of Tables

Abbreviations

ABCE-1	ATP-binding cassette sub-family E member 1
Adox	Adenosine dialdehyde
ALS	Amyotrophic lateral sclerosis
ALYREF	Aly/REF Export Factor
ASDMA	Asymmetric dimethylated arginine
Asos	Antisense oligonucleotides
ATG101	Autophagy Related 101
BSA	Bovine serum albumin
BvFTD	Behavioural variant frontotemporal dementia
C9-ALS/FTD	C9ORF72-associated amyotrophic lateral sclerosis and frontotemporal dementia
C9ORF72	Chromosome 9 open reading frame 72
CAN	Acetonitrile
CSF	Cerebrospinal fluid
DDR	DNA damage response
DENN	Differentially expressed in normal and neoplastic cells
DM1	Myotonic dystrophy type 1
DM2	Myotonic dystrophy type 2
Dprs	Dipeptide repeat proteins
Dsbs	Double-strand breaks
EAAT2	Excitatory amino acid transporter 2
EMC	Endoplasmic reticulum membrane protein complex
ER	Endoplasmic reticulum
F-actin	Filamentous actin
FACS	Fluorescence-activated cell sorting
Fals	Familial amyotrophic lateral sclerosis
FCS	Foetal bovine serum
FDA	Food and Drug Administration
FG	Phenylalanine: glycine
FRET	Fluorescence resonance energy transfer
FSC	Forward scatter
FTD	Frontotemporal dementia
FTLD	Frontotemporal lobar degeneration
FUS	Fused in sarcoma
FXTAS	Fragile X-associated tremor/ataxia syndrome
G-actin	Globular actin
GAR	Gly-Arg rich
GC	Granular component
GO	Gene Ontology
GRN	Granulin Precursor
GWAS	Genome-wide association studies
HBSS	Hank's Balanced Salt Solution
HCD	High energy collision
Hnrnps	Heterogeneous nuclear ribonucleoproteins
HSP	Hereditary spastic paraplegia
HTRS	Tandem repeat sequence
Httex1	Huntington exon 1
INFK	Interferon kappa

iPSCs	Induced pluripotent stem cells
KPNB1	Karyopherin Subunit Beta 1
LC-nESI-MS/MS	Liquid chromatography-nano electrospray ionization-tandem mass spectrometry
LCR	Low complexity sequence region
LLP	Liquid-liquid phase separation
Lmns	Lower motor neurons
MAPT	Microtubule Associated Protein Tau
MATR3	Matrin 3
Mb	Megabases
MND	Motor neuron disease
MOBKL2B	Mps One Binder Kinase Activator-Like 2B
Mrna	Messenger ribonucleic acid
NPC	Nuclear pore complex
NPM1	Nucleophosmin
PBP	Progressive bulbar palsy
PBS	Phosphate-buffered saline
PERK	Protein kinase RNA-like endoplasmic reticulum kinase
PFN1	Profilin1
PLS	Primary lateral sclerosis
PMA	Progressive muscular atrophy
PNFA	Progressive non-fluent aphasia
Poly-AP	Poly-alanine-proline
Poly-GA	Poly-glycine-alanine
Poly-GP	Poly-glycine-proline
Poly-GR	Poly-glycine-arginine
Poly-PR	Poly-proline-arginine
PolyQ	Polyglutamine
PRMT	Protein arginine methyltransferase
PulSA	Pulse shape analysis
RAN	Repeat-associated non-ATG initiated
ROS	Reactive oxygen species
RQC	Ribosome-associated quality control
rRNA	Ribosomal ribonucleic acid
sALS	Sporadic amyotrophic lateral sclerosis
SD	Semantic dementia
SDMA	Symmetric dimethylated arginine
SGs	Stress granules
SMA	Spinal muscular atrophy
SMCR8	Smith-Magenis syndrome chromosomal region candidate gene 8
SMN	Survival of motor neuron
SNP	Single nucleotide polymorphism
SOD1	Superoxide dismutase 1
SSC	Side scatter
SURF6	Surfeit locus protein 6
TBK1	TANK binding kinase 1
TCEP	Tris(2-carboxyethyl)phosphine
TDP-43	DNA-binding protein 43
TEAB	Triethylammonium bicarbonate
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
TUDCA	Tauroursodeoxycholic Acid

U2 snrnp	U2 small nuclear ribonucleoprotein particle complex
UBQLN2	Ubiquitin 2
ULK1	Unc-51-like kinase 1
Umns	Upper motor neurons
UPS	Ubiquitin-proteasome system
VCP	Valosin-containing protein
VHP	Villin headpiece
WDR41	WD Repeat Domain 41
XPC	Xeroderma pigmentosum C
H2AX	H2A histone family member X

CHAPTER 1

General Introduction

1 General Introduction

1.1 Overview

1.1.1 Thesis statement

C9ORF72-associated Motor Neuron Disease patients feature abnormal expression of five dipeptide repeat proteins (DPRs). This thesis examined the proteome interactions of these DPRs by quantitative proteomics and their impact on cell biology in a Neuro2a cell model. The work found two of these, poly-PR and -GR, to be promiscuous binders that stalled ribosomes, bound arginine methylases, and destabilized the actin cytoskeleton. It was also found that aggregates of poly-GA co-aggregated distinct proteins to another protein that forms aggregates in Huntington Disease. In this view, novel explanations are offered for the pathologic signatures previously reported in patient brains.

1.1.2 The role of dipeptide repeat proteins in C9ORF72-linked MND/FTD

Motor neuron disease (MND) and Frontotemporal dementia (FTD) are two clinically, pathologically and genetically heterogeneous neurodegenerative diseases that adversely impact the quality of life for the patients and their carers as well. Both disorders impose an economic burden on the community in the form of health and social care. For instance, MND affects 1 in 11,434 Australians at a cost of $1.13 million per person (MND-Australia, 2015). Similar to other neurodegenerative diseases, senility is the key risk factor for MND and FTD. Therefore, the socio-economic impact of both disorders is expected to strike as the population worldwide grows and ages resulting in a dramatic rise in MND/FTD patients percentage. Despite global efforts to bolster NMD/FTD research, there are currently no effective treatments or cures and available treatment options mainly focus on relieving symptoms and do little to delay disease progression.

A landmark discovery in MND and FTD research occurred in 2011 when two independent groups identified a mutation in the C9ORF72 gene, that is now known to be the most common genetic cause of both these diseases (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Although several hypotheses have been proposed, the molecular mechanism of the disease remains unknown and this has led to a flurry of exciting research within the FTD/ALS field in recent years. To add to this, the work in this thesis describes the generation and use of cellular models and post-mortem brain tissue to investigate molecular mechanisms underlying C9orf72-linked FTD and ALS, specifically focusing on the unique dipeptide repeat protein (DPR) pathology associated with the disease.

1.1.3 Overview of general introduction

This introductory chapter aims to put into context the dissertation work on the pathogenesis of C9ORF72-derived dipeptide repeat proteins described in subsequent chapters. This review will initially discuss the key overlapping clinical and genetic features between MND and FTD. Next, we will discuss C9ORF72 repeat expansion mutation as the most common cause of both diseases. Last but not least, the current knowledge of mechanisms of C9ORF72-related neurodegeneration is reviewed, with special focus on aspects of C9ORF72-associated dipeptide repeat proteins pathogenesis. The final section of this chapter will be devoted to providing a rationale for the molecular mechanism of DPRs toxicity.

1.2 Clinical features of FTD and MND

1.2.1 Frontotemporal dementia

Frontotemporal dementia (FTD) refers to a clinically and pathologically heterogeneous group of progressive neurodegenerative disorders characterized by frontotemporal lobar degeneration (FTLD); selective atrophy of the frontal and/or temporal cortices, with a non-Alzheimer's disease pathology (Warren et al., 2013). FTD disease accounts for 5 – 15 % of all cases of dementia making it the second most common form of young-onset dementia, after Alzheimer's disease (Rademakers et al., 2012) with an equal incidence between men and women (Hogan et al., 2016). The disease onset often occurs between the ages of 45-65 years, with a median duration of 6-8 years (Thompson et al., 2005). The population prevalence of FTD ranges from 2.7 to 15.1 per 100,000 in different epidemiological studies (Rabinovici and Miller, 2010).

Up to 50% of FTD cases have a heritable form of the disorder due to mutations in one of the FTD-linked genes (familial FTD), whilst the remaining cases are sporadic with no currently known genetic cause (sporadic FTD) (Rademakers et al., 2012). Clinically, FTD patients undergo a progressive decline in behavioural and/or language faculties. Based on the clinical presentation of the disease, FTD can be classified into three syndromes; behavioural variant frontotemporal dementia (bvFTD), semantic dementia (SD) and progressive non-fluent aphasia (PNFA) (Woollacott and Rohrer, 2016). The bvFTD is the most common FTD subtype accounting for approximately 60% of cases, followed by the SD subtype and PNFA form where each represents 20-25% of diagnosed cases (Onyike and Huey, 2013). Also, these subtypes differ according to the area affected in the brain. The bvFTD subtype typically involves marked degeneration of frontal and anterior temporal lobe that appears to asymmetrical between hemispheres. SD is associated with a pronounced anteroinferior temporal lobe cortical atrophy which is asymmetric

and frequently worst in the left hemisphere. Progressive non-fluent aphasia (PNFA) patients exhibit left-sided frontal and insular atrophy **(Figure 1-1)** (Meeter et al., 2017).

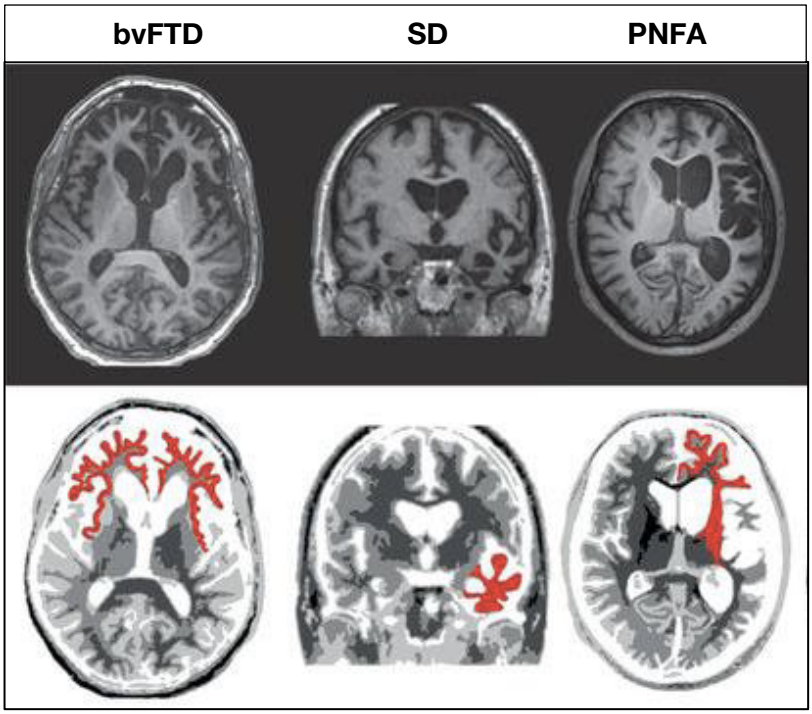

Figure 1-1 | Grey matter atrophy in the three FTD clinical variants.

Characteristic patterns of grey matter atrophy (highlighted in red) in different clinical and genetic subtypes of frontotemporal dementia (FTD). Patients with behavioural variant FTD (bvFTD) exhibit prominent frontal, insular and anterior cingulate atrophy. Typical temporal atrophy in semantic demenia (SD) is asymmetrical (most often left-sided). Patients with progressive nonfluent aphasia (PNFA) exhibit left frontal and insular atrophy. Adapted from Meeter et al. (2017)

Patients with bvFTD manifest with a loss of executive and interpersonal skills with the consequent emotional blunting and the appearance of unusual behaviours including confrontation seeking, personality changes, social withdrawal, a loss of personal hygiene, overeating, obsessions, rituals and even sociopathic behaviours (Lashley et al., 2015). The cognitive decline in bvFTD is more subtle, with patients most commonly showing deficits in executive dysfunctions including attentiveness, judgement, planning and organisation (Rabinovici and Miller, 2010). The SD subtype, on the other hand, predominantly involves language problems, in particular a difficulty understanding the meaning of words, objects and concepts (Rabinovici and Miller, 2010; Warren et al., 2013). Speech is fluent but patients are unable to retrieve words, show an impaired knowledge of familiar objects and are commonly unaware of their lack of comprehension (Ghosh and Lippa, 2013). PNFA sufferers show severe language

deficits characterised by deterioration in language output and slow, non-fluent speech (Warren et al., 2013). Patients present with agrammatism, speech apraxia and a difficulty in understanding complicated sentences (Ghosh and Lippa, 2013; Warren et al., 2013).

Survival varies across the three clinical FTD subtype with median survival being greatest for SD (12 years), whilst mean survival being longest in bvFTD and PNFA (8 years) (Kansal et al., 2016). Heritability also varies across the FTD spectrum with bvFTD having a stronger familial basis compared to both SD and PNFA and the majority of SD cases being sporadic (Rohrer et al., 2009).

1.2.2 Motor neuron disease

The motor neuron disease (MND) refers to a group of neurodegenerative disease that arises from dystrophy of the upper motor neurons (UMNs) of the primary motor cortex and/or the lower motor neurons (LMNs) of the brainstem and spinal cord (Hardiman et al., 2017). The clinical presentation of MND depends on whether the UMNs or LMNs are affected. Examples of the cortical upper motor neurons dysfunction include muscle weakness, hypertonia, hyperreflexia and spasticity. Degeneration of lower motor neuron leads to symptoms as muscle weakness and atrophy, cramps, hyporeflexia, hypotonia and fasciculation (Lee et al., 2016b). Accordingly, the MND diseases are subdivided into 6 variants. MND impacting primarily the UMNs include hereditary spastic paraplegia (HSP) and primary lateral sclerosis (PLS). The LMN clinical spectrum includes progressive bulbar palsy (PBP), progressive muscular atrophy (PMA), and spinal muscular atrophy (SMA). The last MND subtype is amyotrophic lateral sclerosis (ALS) that involves dystrophy of both UMNs and LMNs (van Es et al., 2017). The remainder of this thesiswill focus on ALS and may use the term MND to refers to ALS, unless otherwise specified.

1.2.2.1 Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), as also known as Lou Gehrig's disease in the US, is a progressive fatal MND caused by degeneration of both the cortical UMNs and bulbar and spinal LMNs (Swinnen and Robberecht, 2014; van Es et al., 2017). ALS was first described in 1869 by a French neurologist, Jean-Martin Charcot (Charcot and Joffory, 1869). ALS is the most prevalent MND subtype, accounting for more than 75% of all diagnosed cases (Lattante et al., 2015). The incidence rate per 100,000 in Europe is about 2.16 with more prevalence in men compared to women (60% and 40% respectively) (Chiò et al., 2013). ALS typically manifests as muscle stiffness, muscle wasting and progressive paralysis. ALS patients die of respiratory failure

secondary to progressive weakness of respiratory muscles, with an average prognosis of 3 to 5 years after the first onset of symptom (Kiernan et al., 2011).

Traditionally, ALS are categorized into two subtypes that are indistinguishable clinically; familial ALS (fALS) and sporadic ALS (sALS). The familial ALS compromise around 10% of ALS cases and have been linked to an autosomal dominant pattern of inheritance (Renton et al., 2014). The advancements in genetic studies and sequencing technologies have resulted in identifying≥30 genes mutations underlying numerous ALS cases (Hardiman et al., 2017). The vast majority (around 90%) of ALS cases are sporadic with no family history of the disease. The causes of sALS cases are unknown, but it is possible to have both a genetic and environmental aetiology (Swinnen and Robberecht, 2014).

ALS is heterogenous phenotypically with patients showing significant variance in the specific area of symptom onset (Swinnen and Robberecht, 2014). Approximately 70% of ALS cases have a spinal-onset that commences with asymmetric weakness in limbs. Spinal-onset ALS patients experience classic signs of UMN dystrophy including hyperreflexia and hypertonia and LMN dysfunction such as muscle weakness and atrophy (Mejzini et al., 2019). Bulbar-onset form develops in approximately 25% of ALS patients and begins with a weakness in the bulbar muscles resulting in swallowing and speech difficulties and wasting of tongue. The remainder of ALS cases (5%) have a respiratory-onset form that involves severe breathing difficulties that needs an immediate respiratory intervention (Hardiman et al., 2017). ALS patients with spinal-onset form have longer mean survival of 2.6 years compared to those with bulbar-onset form (2 years mean survival). Patients respiratory-onset ALS have the bleakest prognosis with a mean survival of only 1.4 years (Chiò et al., 2013).

ALS is a fatal disease with no cure to date. Management of ALS patients depends entirely on symptomatic treatments including the use of muscle relaxants for spasticity and speech therapy for dysarthria. Out of >50 drugs tested for treating ALS, two disease-modifying drugs are the sole US Food and Drug Administration (FDA)-approved medications for ALS[5]. Riluzole is speculated to reduces glutamate-mediated excitotoxicity but can merely extend patients' survival for 2 to 3 months (Lacomblez et al., 2002). The antioxidant edaravone has shown to slow disease progression in early-onset fast-progressing disease form (Abe et al., 2017). The use of edaravone for all ALS cases is still under debate (Hardiman and van den Berg, 2017). One reason for this apparent unmet therapeutic need may be the genetic heterogeneity of the ALS population. With a better understanding of the genes associated with ALS and the molecular pathways responsible for neurodegeneration, more effective therapies can be developed.

1.2.3 ALS and FTD as a Continuum

Multiple lines of evidence have supported the idea that MND and FTD are at the opposite ends of a single spectrum disorder, termed MND-FTD (Burrell et al., 2016). An overlap between many clinical, genetic and pathological characteristics has been recorded. Several clinical studies have observed the frequent occurrence of ALS and FTD in the same family. In addition, cognitive impairments have been developed in around 20-50% of MND patients that was accompanied by dystrophy of frontotemporal cortices, reminiscent of FTD, and around 15% of FTD patients have developed motor neuron dysfunction that is indistinguishable from ALS (Ringholz et al., 2005; Wheaton et al., 2007; Geser et al., 2010; Ambikairajah et al., 2014). In terms of pathology, frequent TDP-43 protein inclusions have been observed in both MND (~97%) and both FTD (~45%) (Neumann et al., 2006; Geser et al., 2010; Ling et al., 2013). Strikingly, TDP-43 proteinopathies in ALS patients were frequently detected in prefrontal and temporal cortices (Lowe, 1994; Nishihira et al., 2008), the brain regions most affected by FTD.

The close relationship between the two disorders was further strengthened following the discovery of mutations in several genes that contribute to both MND and FTD. Mutations in genes such as fused in sarcoma (FUS) (Kwiatkowski et al., 2009; Neumann et al., 2009), transactive response DNA-binding protein 43 (TDP-43) (Sreedharan et al., 2008; Benajiba et al., 2009; Borroni et al., 2009), ubiquitin 2 (UBQLN2) (Deng et al., 2011; Synofzik et al., 2012), TANK binding kinase 1 (TBK1) (Freischmidt et al., 2015), Matrin 3 (MATR3) (Millecamps et al., 2014), valosin-containing protein (VCP) (Forman et al., 2006; Johnson et al., 2010), and chromosome 9 open reading frame 72 (C9ORF72) (DeJesus-Hernandez et al., 2011; Renton et al., 2011) cause both ALS and FTD. However, mutations in superoxide dismutase 1 (SOD1) gene (Rosen et al., 1993) are exclusively linked to MND and genes such as progranulin (PGRN) (Baker et al., 2006) or microtubule-associated protein Tau (MAPT) (Hutton et al., 1998) for FTD. Altogether, these findings indicate that FTD and ALS are closely related diseases with considerably clinical, pathological and genetic overlaps, thereby strongly supporting the idea of the MND-FTD continuum.

1.3 MND-FTD with C9ORF72 mutation

1.3.1 The discovery of the C9ORF72 repeat expansion mutation

In an attempt to unravel new hereditary causes of MND and FTD, several groups have conducted extensive sequencing and genome-wide linkage analysis in MND and FTD families with unknown causal mutations (Morita et al., 2006; Vance et al., 2006; Luty et al., 2008; Le Ber et al., 2009; Gijselinck et al., 2010). Thesis studies have observed a linkage region of 9.8

megabases (Mb) located at chromosome 9p13.3-p21.3 within European MND and FTD patients (Morita et al., 2006; Vance et al., 2006). Linkage analysis has then narrowed the linkage region to 3.7 Mbs without identifying any specific disease-causing mutations (Boxer et al., 2011; Pearson et al., 2011). The same region has been identified by genome-wide association studies (GWAS) in 25% of ALS patients and 47.3% of fALS patients in Finnish and Dutch cohorts (Laaksovirta et al., 2010)(van Es et al., 2009). A larger GWAS study including sALS patients from several countries has also observed an association with the same chromosome 9 locus (Shatunov et al., 2010). These GWAS studies have reported a single nucleotide polymorphism (SNP), rs3849942, on this locus, hence further narrowed the associated signal on chromosome 9p to a block of linkage disequilibrium ~106.5Kb in size that includes three genes; Mps One Binder Kinase Activator-Like 2B (MOBKL2B), interferon kappa (INFK) and C9ORF72 (Laaksovirta et al., 2010; Shatunov et al., 2010).

To identify potential pathogenic mutations, following sequencing studies of MOBKL2B, INFK and C9ORF72 have focused initially on the protein-coding region of these genes and found no exonic mutations linked to disease (Pearson et al., 2011). Subsequent sequencing of the non-coding regions led two groups to independently identify a massive hexanucleotide repeat expansion of the sequence G_4C_2 within a non-protein-coding region of C9ORF72 as a genetic cause of fFTD and fALS (DeJesus-Hernandez et al., 2011; Renton et al., 2011). The inadequacy of standard PCR sequencing techniques in identifying the GC-rich sequencing has led to difficulty in observing the GC rich nature of the repeats. The DeJesus-Hernandez et al. (2011) and Renton et al. (2011) studies overcome this problem by utilizing repeat-primed PCR and next-generation sequencing to identify the presence of the G_4C_2 repeats. Across different studies conducted in several populations, this repeat expansion in C9ORF72 has accounted for fALS (3-46%), sALS (0.4%-21%), fFTD (3-48%) and sFTD (2-23%), fALS-FTD (10-87.5%) and sALS-FTD (6.3%-22.2%) (Cruts et al., 2013). Another study that included a large cohort of 4448 ALS patients and 1425 FTD patients has reported C9ORF72 mutation as the most common cause of ALS and FTD, accounting in total for around 5·0–7·0% of cases in white Europeans, Americans and Australians (Majounie et al., 2012) **(Figure 1-2)**. The expansion was particularly frequent in the Scandinavian countries Finland, Sweden and Denmark whereas was rare in Asian populations; suggesting a northern European founder (Cruts et al., 2013).

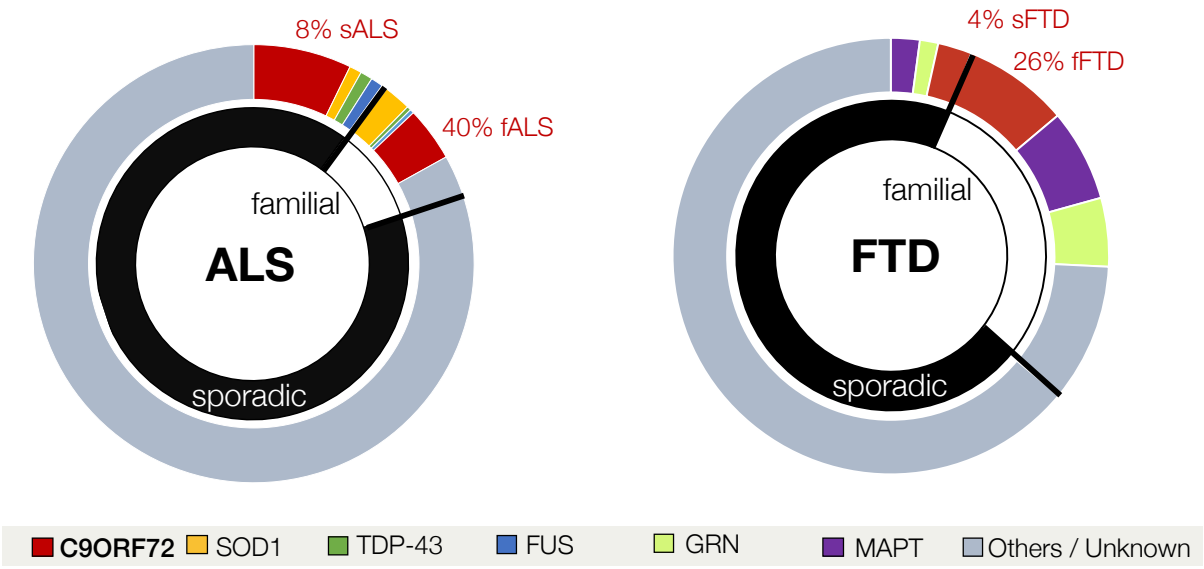

Figure 1-2 | Frequency of the main genes linked to familial and sporadic ALS and FTD.

Modified from Turner et al. (2017)

1.3.2 The C9ORF72 gene structure

The C9orf72 gene comprises 12 exons; 10 coding exons (2-11) and 2 non-coding exons (1a and 1b) **(Figure 1-3)** (Stepto et al., 2014). Based on transcriptional start and termination sites, the gene is alternatively spliced into three variant pre-mRNAs. The three transcripts are V1 (NM_145005.5), V2 (NM_0183525.3) and V3 (NM_001256054.1). These transcripts differ in the incorporation of either the non-coding exon 1a or 1b; transcripts V1 and V3 utilise exon 1a whilst V2 utilises 1b (Haeusler et al., 2016). Since the GGGGCC repeat expansion is located in intron 1 between exons 1a and 1b, transcripts V1 and V3 contain the G_4C_2 sequence within intron 1 whilst the repeats are located in the promoter region of V2. Both fibroblasts derived from C9orf72-ALS patients and motor neurons derived from induced pluripotent stem cells (iPSCs) of C9orf72-ALS individuals suggest selective use of exon 1a compared to exon 1b (Gijselinck et al., 2018). Together, the three transcripts code for 2 different C9ORF72 protein isoforms – the short 24 kDa isoform (C9S) derived from V1 (exons 2-5) and the long 54kDa isoform derived from both V2 and V3 (exons 2-11) (Stepto et al., 2014; Haeusler et al., 2016).

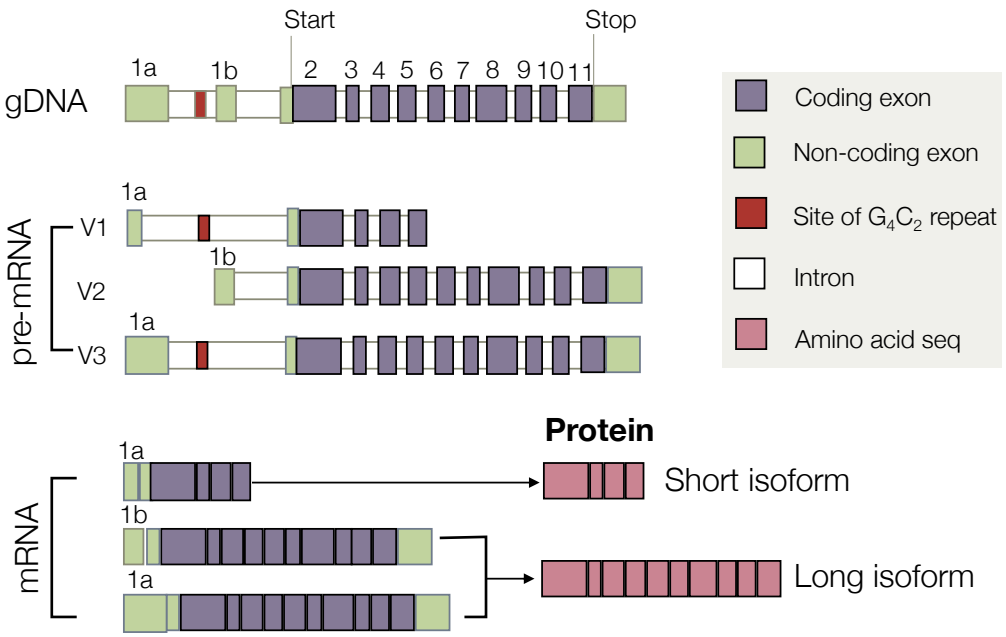

Figure 1-3 | Structure of the C9orf72 gene, transcripts and protein isoforms.

1.3.3 The pathogenic threshold of the C9ORF72 repeat expansion

A crucial question that followed the discovery of C9ORF72 mutation is what the minimum hexanucleotide repeat length required for pathogenicity. Healthy individuals usually harbour <33 repeats with 2 repeats being the most frequent length (van der Zee et al., 2013; Rohrer et al., 2015). ALS/FTD patients harbour much larger repeat expansions that range from hundreds to thousands (Rohrer et al., 2015; Van Mossevelde et al., 2017) **(Figure 1-3)**. However, the exact threshold for a pathogenic expanded repeat is at present unclear. The apparent intermediate-range between 30 to few hundreds are rarely detected in both patients and healthy controls, hence confer unknown disease risk (Rohrer et al., 2015). Hexanucleotide expansions of 20 – 22 repeats have been observed in FTD patients (Gomez-Tortosa et al., 2013) and 24 – 28 repeats have been detected in ALS patients expansions (Millecamps et al., 2012; Garcia-Redondo et al., 2013). Healthy controls with a repeat length range of 20 to a few hundred repeats have been reported in several studies (Cooper-Knock et al., 2012; Majounie et al., 2012; Simon-Sanchez et al., 2012; Beck et al., 2013; Smith et al., 2013; Gami et al., 2015). Also, no clear association between the intermediate repeat length and risk of developing ALS, FTD or ALS-FTD is established yet. One study has reported an 84-year-old-women carrying a 30-unit repeat expansion but didn't develop ALS or FTD (Gami et al., 2015), whereas another study has reported an increase in the risk of developing ALS or FTD with an intermediate 24-30 repeats expansions (Chen et al., 2016). What adds to the complexity is that due to the germline and somatic instability of the GGGGCC sequence, the repeats expand or shrink between generations and show even variability between tissues (Beck et al., 2013; van Blitterswijk et al., 2013). For instance, a family

in which a 90-year-old neurologically healthy father carried a 70-repeat allele that has expanded to ≈1,750 repeats in his offsprings, causing ALS (McGoldrick et al., 2018). Hence, more efforts are required to clarify this ambiguous correlation between the repeat lengths and the risk to develop clinical disease.

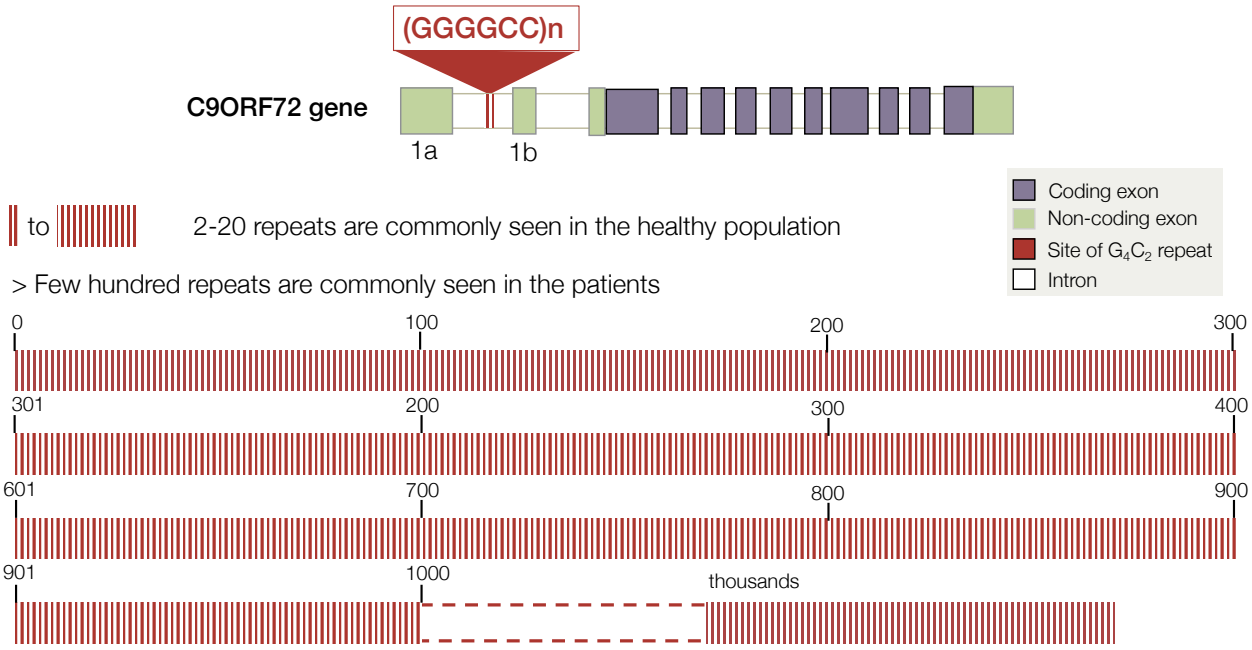

Figure 1-4 | Repeat lengths in C9orf72 associated with disease risk.

A firm correlation between repeat length and disease severity and survival does not currently exist. Repeat length in ALS patients did not affect bulbar or spinal presentations of the disease (Dols-Icardo et al., 2013). However, some authors have observed correlations between repeat length and patient survival. Suh et al. (2015) have reported a reduction in FTD duration in patients with large G4C2 repeat alleles. Post mortem analysis of ALS and FTD patient brains have shown a link between repeat length in the cerebellum and disease duration and a significant reduction in survival rates in FTD patients with repeats over 1467 in length (van Blitterswijk et al., 2013). On the other hand, other groups saw no correlation between length of repeat in the spinal cord and disease duration after onset for either ALS or FTD (Dols-Icardo et al., 2013; Nordin et al., 2015). An inverse correlation has been even recorded where larger repeat sizes have correlated to increased FTD duration (Russ et al., 2015). The authors have also detected hypermethylation of the CpG island located 5' of the G4C2 repeats with the larger allees suggesting a possible protective role due to epigenetic silencing which may cause a reduction in mutant C9ORF72 mRNA, and as a consequence potentially attenuating the toxic effects of the gain of function species derived from the G4C2 expansion (Russ et al., 2015). However, a

conflicting study has reported increased ALS pathology with increased methylation of the 5' CpG island (Xi et al., 2013).

1.4 Hypotheses of C9orf72-mediated pathogenesis

Since its discovery in 2011, scientists have been hunting for clues as to how C9ORF72 expanded G_4C_2 repeats may cause pathogenesis. Currently, three non-mutually exclusive hypotheses have been postulated for C9ORF72-mediated toxicity (Ling et al., 2013; Gitler and Tsuiji, 2016) **(Figure 1-5)**. Firstly, the presence of expanded G_4C_2 repeats within C9ORF72 disrupts the gene transcription leading to a reduction in C9ORF72 protein levels. The second hypothesis is that the expansion causes disease via RNA-mediated toxicity via the formation of nuclear foci of G_4C_2 RNA that sequester essential RNA binding proteins. The third proposed mechanism for toxicity is the translation of the expansion from both the sense and anti-sense strand leading to the production of the dipeptide repeat proteins (DPRs) which aggregate in patient brain and spinal cord. A surge of studies has aimed at defining which of these mechanisms are the driving force behind the disease. In this section, each of these mechanisms will be reviewed in detail with reference to pathological and clinical data, as well as evidence from *in vitro* and *in vivo* disease models.

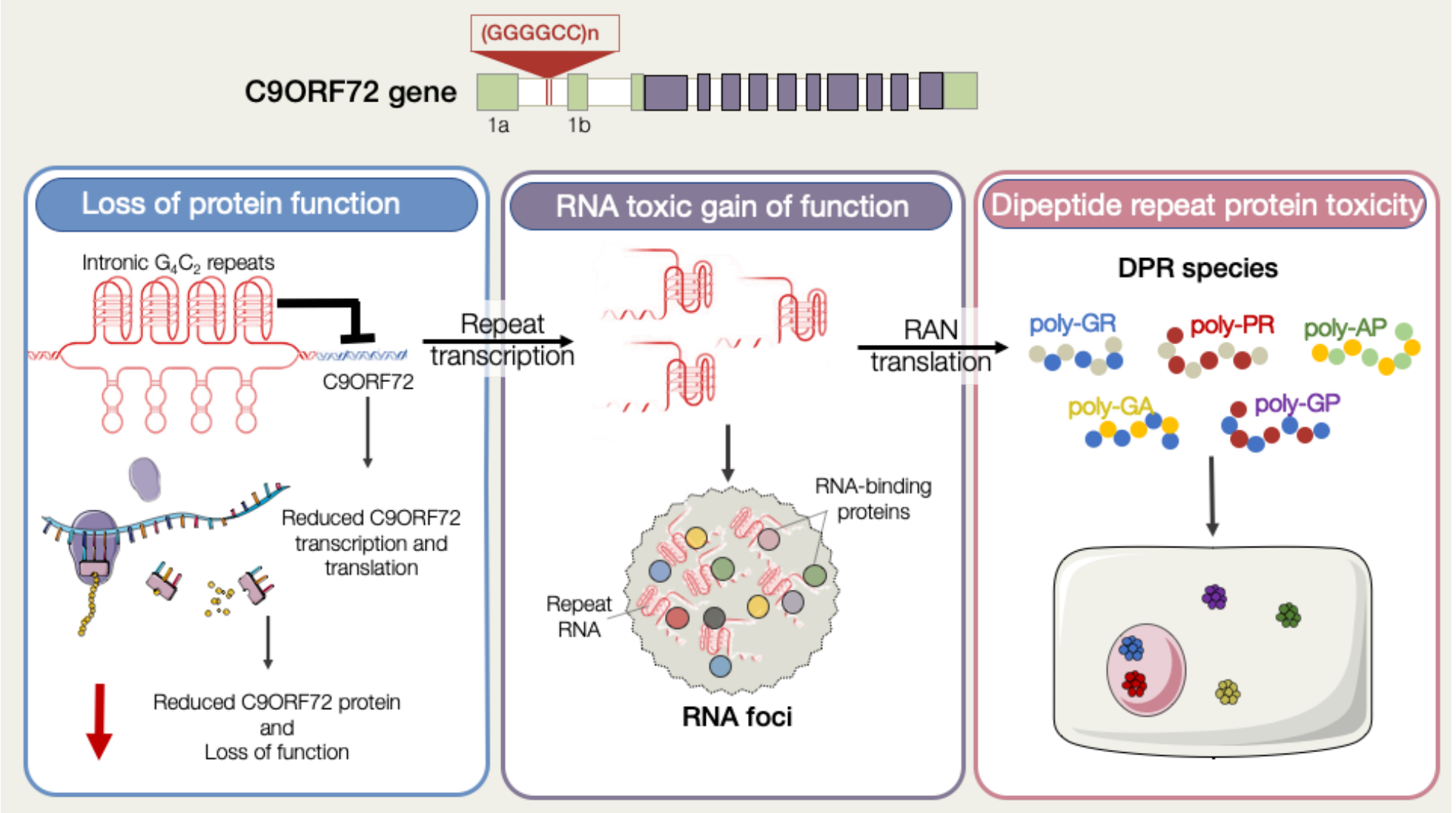

Figure 1-5 | Pathogenic mechanisms associated with C9orf72 repeat expansion toxicity.

1.4.1 Loss of C9ORF72 protein function

1.4.1.1 The C9ORF72 protein function

The precise function of C9ORF72 remains unclear. In humans, C9ORF72 transcripts are expressed widely including the brain. As discussed in section 1.3.2, C9ORF72 has short and long isoforms. Antibodies specific to these two isoforms have shown that the differences in cellular localization within Purkinje neurons and spinal motor neuron suggesting distinct functions. The long C9ORF72 protein isoform has shown diffuse cytoplasmic staining and predicted to be involved in autophagy (Sellier et al., 2016; Sullivan et al., 2016; Xiao et al., 2016; Yang et al., 2016). On the other hand, the short isoform localizes to the nuclear membrane in patient brain tissue and has been shown to interact with the nucleocytoplasmic transport proteins; Ran-GTPase and Importin β hence implicating the short C9ORF72 isoform in nucleocytoplasmic transport (Xiao et al., 2015). Intriguingly, loss of C9ORF72-nuclear membrane localization has correlated with TDP-43 mislocalization in spinal motor neurons of C9or72-ALS cases (Xiao et al., 2015). However, the differential role of each protein isoform in the disease still ambiguous.

Bioinformatic analyses have inferred the C9orf72 protein to function as a DENN (differentially expressed in normal and neoplastic cells) domain protein, regulating membrane trafficking (Zhang et al., 2012; Levine et al., 2013). Recently, C9orf72 has been confirmed to form a complex with Smith-Magenis syndrome chromosomal region candidate gene 8 (SMCR8), WD Repeat Domain 41 (WDR41) and Autophagy Related 101 (ATG101) that function as a GDP-GTP exchange factor for Rab8a and Rab39b; small GTPases regulating vesicle trafficking in the autophagy pathway (Yang et al., 2016). Moreover, this protein complex is phosphorylated by the autophagy-related TBK1 and interacts with the autophagy initiation complex Unc-51-like kinase 1 (ULK1) thus further supporting a role in autophagy regulation (Shannon et al., 2003; Sullivan et al., 2016; Yang et al., 2016).

C9ORF72 has recently been linked to actin dynamics. Immunocytochemical and proteomic studies have detected the interaction of C9ORF72 with cofilin and other actin-binding proteins (Sivadasan et al., 2016). C9ORF72 depletion in cultured motor neurons has reduced actin filament production, and truncated axons and growth cones, accompanied by enhanced cofilin phosphorylation. This observation replicates what has been reported in human iPSC-derived motor neurons and post mortem tissue from patients with C9ORF72 expansion (Sivadasan et al., 2016). shRNA knockdown of C9ORF72 has also impaired the activity of small GTPases, Rac1 and Arf6 and thus actin dynamics (Sivadasan et al., 2016).

1.4.1.2 Evidence supporting loss of protein function

Several studies have probed the mechanisms behind C9ORF72 mutation pathogenesis and have provided support for C9ORF72 haploinsufficiency theory. Both C9ORF72 transcripts (DeJesus-Hernandez et al., 2011; Gijselinck et al., 2012; Belzil et al., 2013; Ciura et al., 2013; Fratta et al., 2013; van Blitterswijk et al., 2013; Xi et al., 2013; Waite et al., 2014) and protein (Xiao et al., 2015; Sivadasan et al., 2016) levels were decreased in post-mortem C9ORF72-ALS/FTD brain samples. Since C9ORF72 plays a role in autophagy and trafficking, loss of protein function is predicted to impair these cellular functions. Marked impairment of vesicle trafficking and a dysfunctional trans-Golgi network have been reported in C9ALS/FTD iPSC motor neurons and C9ORF72 overexpression has rescued this phenotype (Aoki et al., 2017).

C9ORF72 mutation pathology has been characterized by the accumulation of unique p62 positive inclusions (Al-Sarraj et al., 2011). The p62 protein, also known as SQSTM-1, is an autophagy receptor that targets specific cargoes for autophagic degradation (Rusten and Stenmark, 2010). C9ORF72 knockdown has resulted in the accumulation of p62 positive puncta in HeLa cells and primary mouse cortical neurons (Sellier et al., 2016; Webster et al., 2016). The authors have also noticed attenuation in the autophagy basal level in cell lines from C9ALS/FTD patients compared to controls. Re-expressing C9ORF72 mRNA fully rescued autophagy dysfunction and reduced the p62 puncta to the control levels (Sellier et al., 2016; Webster et al., 2016). Seller et al. (2016) have noticed that only the C9L isoform and not the C9S isoform rescued autophagy defects. Reduced C9ORF72 expression has interestingly promoted the aggregation of TDP-43 in mouse embryonic cortical neurons. C9ORF72 overexpression has decreased the aggregation propensity of the D196G mutant of TDP-43 (Sellier et al., 2016). The reduced levels of short C9ORF72 isoform (C9-S) correlate to nucleocytoplasmic transport deficits; C9-S was found at the nuclear membrane in healthy control spinal motor neurons, whilst it has mislocalized to the plasma membrane in C9ORF72- and sALS patient brains (Xiao et al., 2015). Loss of C9-S isoform was associated with mislocalisation of Karyopherin Subunit Beta 1 (KPNB1), RAN and TDP-43 in C9ALS spinal motor neurons (Xiao et al., 2015).

C9ORF72 knockdown in zebrafish model has disrupted arborisation and shortened motor neuron axons compared to controls leading to motor deficits. These observed defects have been attributed to the role of C9ORF72 in actin dynamics (Sivadasan et al., 2016). Co-expression of C9ORF72 mRNA has rescued these motor deficits. Motor impairments have also reported in *C. elegans* C9ORF2 knockout model (Therrien et al., 2013). A C9ORF72 knockout mouse model

showed rapid progressive lethargy and decreased survival compared to control. Mouse embryonic fibroblasts developed from these animals have shown impaired mTOR signalling (Ugolino et al., 2016). Autophagy and lysosomal deficits has been observed in C9ORF72-deficient mice (Sullivan et al., 2016).

1.4.1.3 Evidence against loss of protein function

Several pieces of evidence have excluded a significant role of C9ORF72 Haploinsufficiency in C9ALS/FTD neurodegeneration. To date, no disease-causing mutation has been observed in the coding region of C9ORF72 gene, with one exception, implying that the repeat expansion, not the C9ORF72 protein itself, is crucial to the disease pathogenesis (Harms et al., 2013). The rare exception was a single sALS case that carried a splice-site mutation in exon 5 of the C9orf72 gene and lacked the repeat expansion (Liu et al., 2016). Another case study has not reported more severe clinical or pathological features in patients who are homozygous for the C9orf72 repeat expansion compared to heterozygote cases, which would be expected if the pathological mechanism of disease were purely loss of protein function (Fratta et al., 2013).

Further evidence to support this notion is that most of the C9ORF72-deficient models rarely demonstrated neurodegeneration, mislocalisation of TDP-43 or any motor or behavioural phenotypes associated with C9ALS-FTD. Instead, almost all C9orf72 knockout mouse models displayed immune-system related pathology including lymphadenopathy, splenomegaly and elevated levels of autoantibodies (Atanasio et al., 2016; Burberry et al., 2016; Jiang et al., 2016; O'Rourke et al., 2016; Sudria-Lopez et al., 2016; Sullivan et al., 2016; Ugolino et al., 2016). These models imply a loss of C9ORF72 is not sufficient to cause ALS/FTD phenotypes or neurodegeneration, albeit its effect on the immune system is interesting particularly with reported immunological dysfunction in a range of neurodegenerative disorders (Doty et al., 2015).

Taken together, these studies indicate a loss of function of the protein is not the primary disease mechanism in C9ALS/FTD. Instead, it is likely to play a modulatory role in the disease pathogenesis. For instance, defective autophagy induced by C9ORF72 may enhance DPR accumulation and thereby toxicity (Webster et al., 2016). Indeed, the DPRs are mainly cleared via autophagy (Cristofani et al., 2017); hence a loss of C9ORF72 protein function may enhance DPR accumulation. Additionally, a recent link between C9ORF72 and Stress granule formation (Maharjan et al., 2017) is intriguing particularly with the recent findings that the arginine-rich DPRs perturb SG dynamics (Lee et al., 2016a; Boeynaems et al., 2017).

1.4.2 Mechanism of toxicity: RNA toxic gain of function

Analogues to other neurodegenerative disorders caused by expanded nucleotide sequences in the non-coding regions, it is expected that the C9orf72 hexanucleotide repeat expansion can be bi-directionally transcribed in both the sense and antisense direction (DeJesus-Hernandez et al., 2011; Gendron et al., 2013; Mizielinska et al., 2013; Zu et al., 2013). Since the repeat expansions are GC-rich sequences, the DNA and RNA tend to form a highly stable secondary structure called a G-quadruplex (Fratta et al., 2013; Haeusler et al., 2016). The DNA forms an anti-parallel G-quadruplex structure, whilst RNA mostly forms a parallel structure (Reddy et al., 2013; Haeusler et al., 2016).

Indeed, RNA foci corresponding to both sense and antisense RNA transcripts from repeat expanded C9ORF72 have been detected in patient brain tissue and patient-derived cell lines **(Figure 1-6)** (DeJesus-Hernandez et al., 2011; Donnelly et al., 2013; Gendron et al., 2013; Lee et al., 2013; Mizielinska et al., 2013; Cooper-Knock et al., 2015). These RNA foci are predominantly localised to the nucleus and a lesser extent in the cytoplasm (Lagier-Tourenne et al., 2013; Mizielinska et al., 2013) and have been detected in both neurons and glial cells in patient tissue (Gendron et al., 2013; Lagier-Tourenne et al., 2013; Mizielinska et al., 2013).

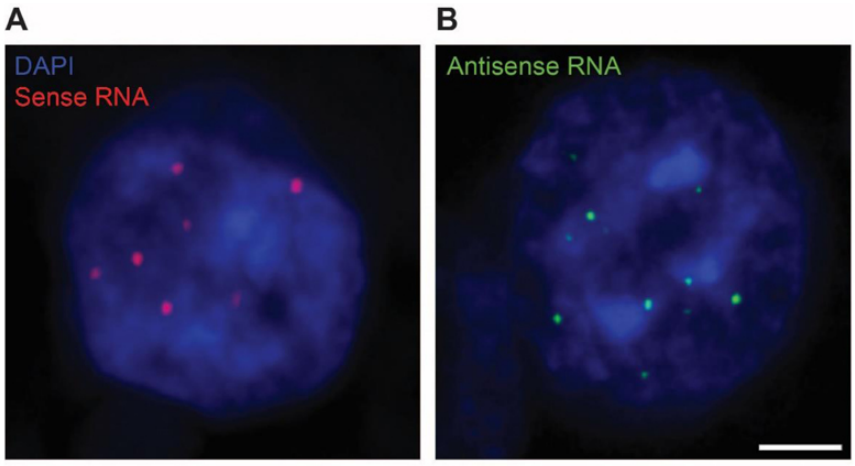

Figure 1-6 | Sense and antisense RNA foci in the frontal cortex of C9FTD/ALS patients.
Scale bar: 2.5 μm. Adapted from Freibaum and Taylor (2017).

Toxicity secondary to sequestration of essential RNA-binding proteins into RNA foci has been proposed as a possible mechanism of C9orf72 FTD and ALS pathogenicity, similar to some other noncoding repeat expansion disorders, such as myotonic dystrophy type 1 (DM1) and fragile X-associated tremor/ataxia syndrome (FXTAS) (Shi et al., 2017b). Various biochemical techniques, including RNA pulldowns, proteome arrays and co-localisation studies have been employed to identify the foci-interacting proteins in multiple *in vitro* and *in vivo* models (Donnelly

et al., 2013; Xu et al., 2013; Cooper-Knock et al., 2014; Haeusler et al., 2016; Fay et al., 2017). The RNA interactome was predominated by proteins having RNA recognition motifs. These proteins are involved in splicing, mRNA nuclear export and/or translation. Many members of the heterogeneous nuclear ribonucleoproteins (hnRNPs), a group of proteins that regulate pre-mRNA splicing have been identified among RNA foci sequestered proteins. Specifically, hnRNP A1, A3, H/F, K, U have bound to C9ORF72 repeat RNA (Lee et al., 2013; Mori et al., 2013b; Sareen et al., 2013; Cooper-Knock et al., 2014; Cooper-Knock et al., 2015; Haeusler et al., 2016). The proteins involved in nuclear mRNA trafficking and nucleocytoplasmic transport are among those sequestered by C9orf72 RNA foci, specifically Ran-GAP1, Aly/REF Export Factor (ALYREF) and Pur-α (Sareen et al., 2013; Cooper-Knock et al., 2014; Rossi et al., 2015; Zhang et al., 2015).

1.4.2.1 RNA toxicity models

Several *in vivo* models have been generated to investigate C9orf72 repeat RNA toxicity. Two independent *Drosophila* models have shown that the formation of sense RNA foci was not sufficient to cause neurodegeneration (Mizielinska and Isaacs, 2014; Tran et al., 2015). These studies have observed that although overexpression of the expanded G4C2 repeat in Drosophila adult neurons or eyes has resulted in neurodegeneration, this phenotype was abolished once the repeats were interrupted with stop codons to prevent the translation of DPR proteins implying that it is DPR proteins that mediate this neurodegeneration (Mizielinska et al., 2014a). This was attributed to efficient splicing of intronic repeat RNA preventing its nuclear export and translation (Tran et al., 2015). In line with the previous findings for the sense RNA, neuronal expression of antisense repeat RNA that was interrupted with stop codons has led to the frequent formation of RNA foci but did not significantly alter lifespan or neurodegeneration (Simone et al., 2018).

Few studies have supported the significant contribution of C9orf72 repeat RNA to disease pathogenesis. Motor axonopathy was induced when *in vitro* transcribed sense and antisense C9orf72 repeat RNA injected into developing zebrafish embryos (Swinnen et al., 2018). This phenotype was still observed when stop-codon-interrupted RNA was injected suggesting independence of DPR proteins. Transfection with 42 GGGGCC repeats located inside an artificial intron of eGFP has resulted in degeneration of rat primary cortical and motor neurons (Wen et al., 2014). These neurons contained abundant RNA foci in the absence of detectable DPR protein expression indicating the reduced survival was RNA mediated. Similarly, Drosophila model with neuronal expression of 30 C9orf72 hexanucleotide repeats interrupted by a 6-base pair restriction endonuclease cut site has shown age-dependent disruption in eye morphology (Xu et al., 2013; Zhang et al., 2015). Yet, lack of DPR detection does not rule out the possibility that

they are being produced. Indeed, upon strong overexpression, DPR proteins were demonstrated in this Drosophila model, thus suggesting that the phenotypes observed may not be independent of DPR protein-induced toxicity (Zhang et al., 2015).

Taken together, these findings indicate that the expanded repeat RNA alone is not likely the main pathogenic cause of disease, albeit a contributor to toxicity. Worth-mentioning, these models have analysed much shorter RNA repeats lengths compared to those observed in patients. Moreover, few studies have assessed the impact of antisense RNA (Simone et al., 2018; Swinnen et al., 2018). Therefore, the generation of newer models expressing more disease-relevant repeat length is crucial to fully understand the role of C9ORF72 RNA toxicity in disease pathogenesis.

1.4.3 Mechanisms of toxicity: Dipeptide repeat protein toxicity

The third mechanism of toxicity relates to the production and accumulation of unique proteins that are generated from the C9orf72 expanded repeat RNA transcripts.

1.4.3.1 RAN translation

A study in 2011 reported that the CAG repeat expansion seen in spinocerebellar ataxia type 8 and DM1 was unconventionally translated in all reading frames despite the absence of an upstream ATG start codon. The method of protein translation was discovered in a study investigating the CAG repeat expansion that occurs in DM1 and spinocerebellar ataxia type 8 (Nichols et al., 2011). The authors named this non-canonical mechanism as repeat-associated non-ATG initiated (RAN) translation. Later, this mechanism was shown in other repeat expansion disorders such as Huntington's disease, myotonic dystrophy type 2 (DM2) and FXTAS (Todd et al., 2013; Cleary and Ranum, 2014; Wojciechowska et al., 2014; Bañez-Coronel et al., 2015).

Several groups have studied RAN translation in the C9orf72 hexanucleotide repeat expansion and discovered the bidirectional transcription and translation of the repeat by RAN translation in all reading frames from both the sense and antisense transcripts (Figure 1.6) (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013a; Zu et al., 2013). This translation generates five different DPR proteins. The sense transcript is translated into glycine-alanine (poly-GA), glycine-arginine (poly-GR) and glycine-proline (poly-GP), while the translation of the antisense transcript produces alanine-proline (poly-AP), proline-arginine (poly-PR), and poly-GP **(Table 1-1)**.

Table 1-1 | Putative protein products from the bidirectional RAN translation of the C9orf72 repeat expansion.

G_4C_2 (sense strand)	
Sense frame 1 (GP_S)	*GPGPGPGPGPGPGPGPGP(GP)$_{EXP}$GPGPGRGRGGPGGGPGAGLRLRCL RPRRRRRRRWRVGE*
Sense frame 2 (GR_S)	*RLTRRKQGGKQPQPVASSGTQESRARGRGRGRGRGRGRGRGRGR(GR)$_{EXP}$GR GRGVVGAGPGAGPGRGCGCGACARGGGGAGGGEWVSEEAASWRVAVWGSAA GKRRG*
Sense frame 3 (GA_S)	*QALELRSRALGAGAGAGAGAGAGAGAGAGA(GA)$_{EXP}$GAGAWSGRARGRARG GAAVAVPAPAAAEAQAVASG*
G_2C_4 (antisense strand)	
Antisense frame 1 (PA_{AS})	*GEPPLLPAPLPGSRTPNSHPPGCRLLTHPLATACASAAAGAGTATAAPPRA RPRARPDHAPAPAPAPAPAPAPAPAPA(PA)$_{EXP}$PAPAPSARLLSSRACYRLRL FPSLFSSG*
Antisense frame 2 (PR_{AS})	MQAIPPVARGESPTPSFGQRNERESKNASSSEESPRFYPRLFPAAEPQTATR QDAASSLTHSPPPAPPPPRAQAPQPQPRPGPAPGPAPTTPRPRPRPRPRPR PRPRPR(PR)$_{EXP}$PRPRPLARDS*
Antisense frame 3 (GP_{AS})	MRGKVKMRRALRRAPASTRASSRQPNPKQPPARMPPPHSPTRHRLRLRRR GRRHRNRSPAPGPPPGPPRPRPGPGPGPGPGPGPGPGPGP(GP)$_{EXP}$GPGP*

Note: Repeat motifs are highlighted in red. Possible upstream AUG start sites (M) are shown in red for antisense frame 2 and frame 3. Stop codons are indicated by an asterisk. Adapted from Zu et al. (2013).

Numerous studies have investigated the mechanism(s) of RAN translation initiation. These studies have postulated that at least in the CCG repeat in FXTAS and the C9orf72 repeat, RAN translation initiates at a near-cognate start codon upstream of the repeat sequence (Kearse et al., 2016; Green et al., 2017; Sellier et al., 2017; Tabet et al., 2018). For instance, RAN translation of the sense C9orf72 RNA is predicted to initiate at a CUG codon located upstream of the repeat that is in frame with poly-GA. Nevertheless, the peptide sequence predicted to lie upstream of the DPR proteins has not yet been seen in patient post-mortem tissue. On the other hand, translation of poly-GR and poly-GP is proposed to depend on ribosomal frameshifting due to formation of stable RNA G-quadruplexes; similar to what seen in other repeat expansion disorders (Wojciechowska et al., 2014; Yu et al., 2014). RAN translation has also been shown to depend on the ribosomal translational machinery and mRNA 5'7-methylguanosine (m7G) capping (Kearse et al., 2016; Green et al., 2017; Sellier et al., 2017; Tabet et al., 2018). Conversely, one study has demonstrated independently of C9ORF72-associated RAN translation on m7G transcript capping despite being less efficient compares to the cap-dependent translation (Cheng et al., 2018).

1.4.3.2 Poly-GA toxicity

Poly-Glycine-Alanine (poly-GA) is the most abundant among the five DPR species in the patients' brains (Schludi et al., 2015). It was first discovered when a subset of ALS/FTD patients with TDP-43 proteinopathy showed abundant and unusual p62-positive, TDP-43 negative inclusions in the frontal neocortex, hippocampus and cerebellum (Pikkarainen et al., 2010). Reportedly, these patients were also negative for fused in sarcoma (FUS), optineurin, alpha-internexin and neurofilament (King et al., 2009; King et al., 2011). The identity of the protein in the p62-positive, TDP-43-negative inclusions remained obscure until the discovery of C9orf72 repeat expansion mutation in familial and sporadic ALS and FTD cases (DeJesus-Hernandez et al., 2011; Renton et al., 2011). Nearly all TDP-43-negative inclusions contained C9ORF72-derived poly-GA peptide. These neuronal poly-GA inclusions were associated with neurodegeneration in this patients cohort (Mackenzie et al., 2015).

Poly-GA has strong aggregation properties due to the biophysical properties of the peptide (Freibaum and Taylor, 2017). Since the peptide is composed of small hydrophobic residues which are uncharged, it is expected to collapse into poorly soluble globules with strong aggregation tendencies (Lee et al., 2016a; Freibaum and Taylor, 2017). Indeed poly-GA forms amyloidogenic fibrils which are positively stained by Congo red and thioflavin T (Chang et al., 2016; Freibaum and Taylor, 2017) and is predominantly found in the urea soluble fraction when expressed in cells (Lee et al., 2016a). The poly-GA amyloidogenic fibrils form a parallel β-sheet structure with similar structural properties to the amyloid-beta protein in Alzheimer's disease (Lee et al., 2016a).

1.4.3.2.1 UPS dysfunction and ER stress associated with poly-GA

The contribution of the aggregation-prone poly-GA protein to disease pathogenesis has been the focus of many studies. Poly-GA has shown to be toxic in many *in vitro* and *in vivo* model systems (May et al., 2014; Mizielinska et al., 2014a; Yamakawa et al., 2014; Zhang et al., 2014; Schludi et al., 2015; Khosravi et al., 2017; Lee et al., 2017). Initial studies have shown that expression of poly-GA using alternative codon constructs induced cellular toxicity (May et al., 2014; Zhang et al., 2014). Both studies have reported also that poly-GA formed frequent cytoplasmic and nuclear aggregates in transfected HEK293T cells. These inclusions were dot-like and star-shaped and co-localised with ubiquitin and p62, but not with TDP-43; features reminiscent of the poly-GA aggregates seen in C9ORF72-ALS/FTD cases. The same phenotype was observed in cortical and hippocampal neurons upon poly-GA expression. Besides, neurons had reduced neurite outgrowth and increased caspase3 activation (May et al., 2014; Zhang et

al., 2014). The two studies have posited that the accumulation of ubiquitinated poly-GA aggregates indicates an impairment of the ubiquitin-proteasome system (UPS). May et al. (2014) have identified members of the UPS as top hits in the interactome of soluble polyGA peptide. The authors have continued to confirm the colocalization of Unc119, a myristoyl-binding protein within polyGA inclusions in both rat primary neurons and post-mortem patient tissue. The endoplasmic reticulum (ER) stress is usually a consequence of UPS dysfunction. Indeed, Zhang et al. (2014) have noted the up-regulation of multiple ER stress markers in the PERK-CHOP pathway in poly-GA transfected neurons and patient tissue. Moreover, pharmacological inhibition of ER stress inhibitors using salubrinal or tauroursodeoxycholic Acid (TUDCA) has rescued poly-GA-induced toxicity. Unc119 pathology has been frequently detected in the C9ALS/FTD frontal cortex and cerebellum and was more severe in the C9FTD cerebellum compared to C9ALS or C9ALS/FTD cerebellar tissues (Schludi et al., 2015).

The poly-GA toxicity was also observed *in vivo*. A poly-GA mouse model has developed brain atrophy as well as motor and cognitive deficits and was accompanied by astrogliosis and degeneration in the hippocampus and the cortex (Zhang et al., 2016). Both UPS-related proteins HR23A and HR23B aggregated into cytoplasmic inclusion and were also recruited into poly-GA inclusions and in the brains of poly-GA mice suggesting the involvement of UPS dysfunction in poly-GA induced toxicity (May et al., 2014; Zhang et al., 2014; Zhang et al., 2016). The HR23 inclusions were also observed in post-mortem human hippocampus where the inclusions again co-localised with poly-GA (Chew et al., 2015). HR23 stabilises the DNA-binding protein xeroderma pigmentosum C (XPC). In poly-GA mice, the poly-GA expression has reduced XPC levels via enhancing its degradation in addition to coaggregation of XPC with poly-GA in cortical and hippocampal neurons. Additionally, co-immunoprecipitation revealed the specific interaction of only poly-GA and not poly-GP or poly-GR with HR23B. This interaction was confirmed in the brains of 66 repeat G_4C_2 mice that showed around 85% of HR32B inclusions were positive for poly-GA pointing toward poly-GA and not the other sense DPRs being the main cause of H32 dysfunction. Zhang et al. (2016) also reported sequestration of the nuclear pore proteins RanGAP1 and Pom121 into poly-GA implying that poly-GA can also disrupt the nuclear pore complex leading to HR23 mislocalisation. Another poly-GA mouse model was developed by Schludi et al. (2017) and showed co-aggregation of the Hsp-70 associated protein, Mlf2 with poly-GA in the spinal cord of mice. Mlf2 pathology was also seen in patient post-mortem brain tissue. Additionally, markers of neuroinflammation were upregulated in the poly-GA mice spinal cord. Microglial activation has correlated to the progression of disease in ALS, and C9ORF72 patients showed higher levels of microglial activation compared to non-C9ORF72 mutation

carriers (Brettschneider et al., 2012). A *Drosophila* (Mizielinska et al., 2014b) and Zebrafish (Ohki et al., 2017) poly-GA models showed the poly-GA to cause a mild reduction in lifespan in comparison to controls (Mizielinska et al., 2014b).

Different studies have investigated the contribution of aggregation to poly-GA toxicity and made different conclusion. Two studies have predicted the soluble form of polyGA to be the toxic culprit and the polyGA aggregation to be cell protective. A Zebrafish model of poly-GA_{80} has found that the muscle structure was not affected by the frequent poly-GA inclusions seen in the tissue despite the overall toxicity of poly-GA to the fish (Ohki et al., 2017). Another supporting study has observed that poly-GA has efficiently sequestered the highly toxic poly-GR into inert cytoplasmic inclusions when co-expressed together. On the contrary, other studies report that the accumulation of insoluble poly-GA inclusions correlates with the increased cytotoxicity. Mouse poly-GA model has observed that poly-GA aggregates sequester the proteasomal degradation protein HR23 and impair nucleocytoplasmic transport proteins leading to neurodegeneration (Zhang et al., 2016). Another study has generated a mouse model noticed expressing a mutant form of $(GA)_{50}$ that contains 50 GA dipeptide repeats interrupted by a proline every 5 repeats. The mutant poly-GA has failed to aggregate and no signs of neurodegeneration or neuropathological and behavioural deficits were observed in this mice compared to mice expressing the aggregating poly-GA_{50} (Zhang et al., 2016). Dual expression of poly-GA and poly-PA inhibited poly-GA aggregation and hence ameliorated poly-GA toxicity both *in vitro* and *in vivo* in the chick embryonic spinal cord (Lee et al., 2017), suggesting that the insoluble poly-GA form is the more likely effector. These discrepancies in results necessitate more research to determine the nature of the neurotoxic poly-GA species.

1.4.3.3 Poly-GR and Poly-PR toxicity

The arginine-rich DPRs, poly-PR and poly-GR, are highly polar and positively charged due to the presence of arginine residues (Freibaum and Taylor, 2017). Transmission electron microscopy revealed a synthesized $(GR)_3$ peptide forms spherical aggregates (Flores et al., 2016). Secondary structure predicted from circular dichroism spectra indicates 55% of the protein has turn and/or random coil conformation and the other 41% of the protein is β-sheet (Flores et al., 2016). Poly-GR peptide has very poor stability *in vitro* than polyPR (<30 min for poly-GR compared to ~72h for poly-PR) (Kwon et al., 2014). Poly-PR and poly-GR were highly toxic compared to other DPRs and caused severe eye degeneration and a dramatic reduction in lifespan of Drosophila model (Mizielinska et al., 2014a). The Arg-rich DPRs high toxicity has been replicated in various cellular models and human iPSCs (Kwon et al., 2014; May et al., 2014; Wen

et al., 2014; Zhang et al., 2014; Freibaum et al., 2015; Tao et al., 2015; Lee et al., 2016a). This section will discuss in details the various mechanisms of poly-PR and poly-GR toxicity unravelled to date.

1.4.3.3.1 Poly-GR and poly-PR associated nucleolar dysfunction

The Arg-rich DPRs localize to the nucleolus when expressed from alternative codon ATG constructs in cells (Kwon et al., 2014; May et al., 2014; Wen et al., 2014; Zhang et al., 2014; Tao et al., 2015; Lee et al., 2016a). The nucleolus is important for ribosomal RNA (rRNA) transcription, pre-rRNA processing and ribosome subunit assembly (Boisvert et al., 2007). Kwon et al. (2014) have noticed a dramatic reduction in the ability of nucleoli to synthesize mature rRNA upon poly-PR expression. Genes involved in rRNA processing were identified among the potent poly-PR toxicity modifiers in yeast and Drosophila screens (Jovičić et al., 2015; Boeynaems et al., 2016). For instance, nucleolin was one of the strongest modifiers of poly-PR toxicity identified in both screens and restoring the function of rRNA processing machinery has mitigated the poly-PR toxicity. Poly-PR and poly-GR both cause the mislocalization of the nucleolar stress sensor protein B23 to the nucleoplasm and nucleolar swelling (Tao et al., 2015). Altogether, these findings imply poly-GR and poly-PR induced nucleolar stress may contribute to C9ORF72 pathogenesis.

The clinical relevance of nucleolar stress in C9ALS/FTD pathogenesis was recently investigated in patient brains (Mizielinska et al., 2017). This study noted smaller neuronal nucleoli in the C9FTD brain compared to controls, yet the nucleolar volume in neurons with poly-GR inclusions was significantly larger compared neurons lacking a poly-GR inclusion. This nucleolar enlargement was also observed in a Drosophila model expressing poly-GR. The same nucleolar phenotype has been also associated with repeat RNA toxicity (Haeusler et al., 2014). The authors have also observed the preferential binding of nucleolin to the sense G4C2 RNA G-quadruplex *in vitro* and detected the nucleolin sequestration into RNA foci in the motor cortex of expansion carriers. Mizielinska et al. (2017) also noticed a significant nucleolar enlargement in C9FTD frontal cortex neurons containing sense RNA foci, hence suggesting a synergism between RNA and Arg-rich DPR toxicity with regards to nucleolar dysfunction.

1.4.3.3.2 Poly-PR and poly-GR disrupt the nucleocytoplasmic transport

Deficits in nucleocytoplasmic transport have been identified among the pathogenic mechanisms linked to C9ORF72-associated ALS/FTD (C9ALS/FTD). Zhang et al. (2015) have detected the impairment of nucleocytoplasmic transport in Drosophila expressing 30 G_4C_2

repeats. The authors have attributed this defect to the G_4C_2 RNA rather than DPRs. Nevertheless, poly-GR, like others DPRs, can be translated from the repeats expressed in the transgenic flies, thus it cannot be ruled out the contribution of the DPRs in this phenotype. A parallel Drosophila study has detected both poly-GP and poly-GR in transgenic flies harbouring 58 G_4C_2 repeats and reported similar defective nucleocytoplasmic defects (Freibaum et al., 2015). A chromosomal deficiency screen spanning the entire 2nd and 3rd chromosomes has identified numerous proteins involved in nucleocytoplasmic transport, including Crm1, Nup153, Nup107, Nup106 Transportin, Gle1 and Ran, as enhancers and suppressors of the 58 G_4C_2 repeat eye degeneration (Freibaum et al., 2015). The nuclear pore protein Nup50 was among the strongest enhancer whilst the strongest suppressor was Ref1, the Drosophila homologue of ALYREF, an RNA binding protein which acts in a complex to deliver processed mRNAs to the nuclear pore (Freibaum et al., 2015). Morphological deficits in the nuclear envelope were visualised by Lamin C staining, in which an abnormal wrinkled and jagged nuclear envelope in addition to Nup107 inclusions near the envelope was detected in over 40% of cells expressing the 58 G_4C_2 repeats. This study was ambivalent whether the observed nuclear pore deficits are caused by G4C2 RNA or the DPRs.

The link between the arginine DPRs and nucleocytoplasmic transport was first elucidated in a yeast screen by Jovičić et al. (2015). A yeast screen using poly-PR only alternative codon constructs has observed that genes involved in nucleocytoplasmic transport were overrepresented among the poly-PR toxicity modifiers. For instance, yeast karyopherin genes were the top suppressors of poly-PR toxicity. Overexpression of KPNA3 has doubled the survival of poly-PR-expressing rodent cortical neurons. Notably, the toxicity modifier RCC1 was depleted from the nucleus in 70-80% of C9ORF72 iPSCs. The same findings have been reported in a Drosophila screen by (Boeynaems et al., 2016). Nuclear pore complex (NPC) proteins, importins, exportins, regulators of the Ran-GTP cycle and arginine methylases were all modifiers of poly-PR retinal degeneration. Trn, the Drosophila homologue of human importin TNPO1, was the strongest poly-PR toxicity enhancer. The authors rationalized for such finding that poly-PR may compete with Trn for Trn cargoes, and indeed the Trn cargo Elav, a neuronal RNA binding protein, was depleted from the nucleus and accumulated in the cytosol in retinas of flies expressing poly-PR. Moreover, aggregates of hnRNP A3, a TNPO1 cargo, was seen in patient C9ORF72-FTD brain. A following study has provided mechanistic insight into how poly-PR cause deficits in nucleocytoplasmic transport (Shi et al., 2017a). The authors postulated that the phenylalanine: glycine (FG) repeats of nuclear pore proteins exist in equilibrium between the unpolymerized and polymerized state. Thus binding of poly-PR binds to the FG repeats stabilize them in the polymerized state. This disruption in the equilibrium of FG repeat polymerization

makes the barrier of central channel less permeable thereby impairing transport through the nuclear pore.

1.4.3.3.3 Poly-GR and poly-PR associated mis-splicing

Genome-wide mis-splicing reported in C9ORF72 mutation carriers have been attributed to both the G_4C_2 RNA (Cooper-Knock et al., 2014; Haeusler et al., 2014; Cooper-Knock et al., 2015; Conlon et al., 2016) and the Arg-rich DPRs (Kwon et al., 2014; Yin et al., 2017). RNA sequencing analysis of cultured human astrocytes incubated with a synthesized poly-PR has shown splicing errors including exon skipping, different 5' UTRs and intronic retention. The aberrant processing of the excitatory amino acid transporter 2 (EAAT2) mRNA has been reported in ALS patients (Lin et al., 1998). The same disrupted EAAT2 splicing alterations have been detected in human astrocytes exposed to poly-PR in a peptide concentration-dependent manner (Kwon et al., 2014).

Recent studies have provided a mechanistic insight of arginine DPR induced mis-splicing (Yin et al., 2017). The proteomic analysis of poly-GR and poly-PR pulldowns revealed that both poly-GR and poly-PR specifically associate with the U2 small nuclear ribonucleoprotein particle complex (U2 snRNP) disrupting their normal function. The authors went on to demonstrate that poly-GR blocked the spliceosome resulting in mis-splicing. The U2 snRNP-associated proteins in C9ORF72 iPSC motor neuron were mislocalized to the cytoplasm compared to their normal nuclear localization. Mis-splicing secondary to RNA repeat toxicity (Conlon et al., 2016) was significantly less frequent than that detected in the C9-ALS frontal cortex and cerebellum. This observation strongly a more crucial role of DPR in the mis-splicing observed in C9 patient brains compared to that resulting from the sequestration of RNA binding proteins into G_4C_2 foci.

1.4.3.3.4 DNA damage and oxidative stress associated with poly-PR and poly-GR

Spinal cord motor neurons differentiated from multiple iPSC have produced both RNA foci and detectable levels of DPRs. Intriguingly, these motor neurons have shown a significant increase in the expression of γH2AX, a marker of DNA double-strand breaks (DSBs) (Lopez-Gonzalez et al., 2016). Transfection of poly-GR into control iPSC motor neurons has increased the number of DNA strand breaks, increased levels of γH2AX and activated the p53 pathway. Poly-GR toxicity was mitigated by lowering p53 levels. The production of reactive oxygen species (ROS) was significantly increased in C9 iPSC motor neurons compared to controls suggesting this is how poly-GR might be causing DNA damage. Poly-GR pulldown has recovered many mitochondrial ribosomal proteins crucial for the translation of 13 mitochondrial complexes

subunits. This may explain the elevated mitochondrial membrane potential detected in both control iPSCs transfected with poly-GR and C9 iPSC motor neurons. Clinically, multiple markers of DNA damage as γH2AX and p-ATM were significantly upregulated in the spinal cord neurons of C9ALS patients compared to controls (Farg et al., 2017). Taken as a whole, it is believed that poly-GR disrupts the neuronal mitochondrial function and induces DNA damage potentially via elevated oxidative stress.

1.4.3.3.5 Poly-GR and poly-PR associated ER Stress and perturbed cell signalling pathways

ER stress was identified as a potent modifier of arginine DPR toxicity in a CRISPR-Cas9 screen in primary neurons (Kramer et al., 2018). Two proteins of the endoplasmic reticulum membrane protein complex (EMC), namely TMX2 and CANX, and Endosomal trafficking genes such as Rab7a were strong modifiers of poly-GR and poly-PR toxicity. These findings indicate that poly-PR and poly-GR induce ER stress, analogous to what observed before for poly-GA (May et al., 2014; Zhang et al., 2014). RNA sequencing of primary neurons expressing poly-PR has revealed the 'apoptotic signalling pathway in response to endoplasmic reticulum stress' was the most significantly enriched Gene Ontology term for upregulated genes. The ATF4, an ER stress gene was upregulated in cells expressing poly-PR and treating these cells with an ER stress inhibitor have alleviated poly-PR toxicity. Zhang et al. (2014) have reported a protein kinase RNA-like endoplasmic reticulum kinase (PERK)-Atf4 mediated ER stress response in C9ALS/FTD patient brains.

1.4.3.3.6 Poly-GR and poly-PR associated disruptions in global protein translation

Another common mechanism hypothesised for poly-GR and poly-PR toxicity is impairment of protein synthesis. Several studies have published the interactomes of DPR proteins which have observed the ribosomal proteins and translation initiation and elongation factors as the most abundant interactors of the Arg-rich DPRs (Kanekura et al., 2016; Lee et al., 2016a; Lin et al., 2016; Lopez-Gonzalez et al., 2016; Boeynaems et al., 2017; Yin et al., 2017; Hartmann et al., 2018; Shi et al., 2018), indicating poly-PR and poly-GR may influence protein translation. A recent poly-GR mouse model that exhibited behavioural deficits and neurodegeneration, showed colocalization of poly-GR with ribosomal subunits and altered expression of genes encoding ribosomal subunits. The authors of the study also demonstrated that poly-GR impaired protein translation in HEK293 cells and neurons containing poly-GR in the cortex of the poly-GR mice (Shi et al., 2018).

The mechanistic isight into how the Arg-rich DPRs impair protein translation is still indefinite. Poly-PR and poly-GR expression have inhibited protein translation, albeit without altering the phosphorylation levels of eIFs, the phosphorylation status of which regulates protein translation (Kanekura et al., 2016). *In vitro*, the synthetic poly-PR peptide has potentially formed complexes with RNA and RNA has induced aggregation of poly-PR and poly-GR that were disassembled by RNase treatment. Poly-PR/RNA complexes were recognised as misfolded proteins *in vitro* by HSP70. These data have prompt the authors to postulate that the Arg-rich DPRs form complexes with mRNA and prevents the access of translation factors to mRNA; thereby blocking protein synthesis. Since nucleolus is the primary site of ribosome biogenesis and assembly and several studies have reported nucleolar stress in poly-GR and poly-PR models (Kwon et al., 2014; May et al., 2014; Wen et al., 2014; Zhang et al., 2014; Tao et al., 2015; Lee et al., 2016a), it follows that nucleolar stress seen in poly-GR and poly-PR models could be a cause or consequence of ribosomal impairment.

1.4.3.3.7 Disruption of liquid-liquid phase separation dynamics of membrane-less organelles by Arg-rich DPRs

Proteomic analyses of poly-GR and poly-PR interactomes have a strong interaction with proteins that harbour a low complexity sequence region (LCR) (Lee et al., 2016a; Lin et al., 2016). LCRs are evolutionarily conserved amino acid segments, typically 75–300 amino acids that are compositionally biased in amino acid representation. These domains are present in one-third of the human proteome and are expected to be unstructured (Freibaum and Taylor, 2017). LCRs mediate multivalent, low-affinity protein-protein and protein-nucleic acid interactions that govern many cellular functions such as the dynamic assembly of membrane-less organelles (Aguzzi and Altmeyer, 2016; Freibaum and Taylor, 2017). Membrane-less organelles are viscous liquid-like structures formed of RNA-protein complexes. These organelles include nucleolus, nuclear pore complex, nuclear speckles, stress granules and Cajal bodies (Brangwynne et al., 2015). Recently, it has been postulated that these organelles may be formed via the liquid-liquid phase separation (LLP) of protein-laden RNAs from the surrounding nucleoplasm or cytoplasm (Taylor et al., 2016b). A recent study has shown the Arg-rich DPRs to phase separate *in vitro* into liquid droplets (Boeynaems et al., 2017). Poly-PR and poly-GR have been proved to directly bind to LCRs of many RNA-binding proteins, some of them are linked to ALS such as TDP-43, FUS, hnRNPA2B1, Ataxin-2, and Matrin-3 (Lee et al., 2016a; Lin et al., 2016). Knockdown of these LCR interacting proteins revealed 35 suppressors and 21 enhancers of poly-GR toxicity (Lee et al., 2016a). These toxicity modifiers were shown to modulate the rough eye phenotype of $(G_4C_2)_{58}$-expressing flies with detectable poly-GR levels (Freibaum et al., 2015). Both poly-PR and

poly-GR also interacted with nucleophosmin (NPM1) that contains 3 LCRs and normally organise the liquid-like granular component (GC) of the nucleolus (Lee et al., 2013). These interactions have altered the ability of NMP1 to phase separate with one of its native nucleolar binding partners SURF6. Poly-GR and poly-PR were found to outcompete surfeit locus protein 6 (SURF6) for NPM1 binding, thereby disrupt nucleolar dynamics, nucleolar morphology and nucleolar function.

Stress granules (SGs) are another example of membraneless organelle that poly-GR and poly-PR interfere with. SGs are cytoplasmic assemblies of RNA and protein formed following translational arrest via dynamic LLP (Jain et al., 2016). Both poly-PR and poly-GR were identified to interact with several SG components such as G3BP1, G3BP2 and Caprin1; their binding partner (Lee et al., 2016a). Also, both peptides have induced spontaneous assembly of SGs that were poorly dynamic. The same study has also identified poly-PR and poly-GR induced perturbation of Cajal bodies assembly and nuclear speckles dynamics (Lee et al., 2016a). Taken together, the Arg-rich DPRs alter the assembly, dynamics, and function of membrane-less organelles leading to the widespread cellular abnormalities.

1.4.3.4 Poly-GP toxicity

Poly-Glycine-Proline (poly-GP) is the second most abundant DPR in the human C9 ALS/FTD brain following poly-GA (Schludi et al., 2015). Uniquely, poly-GP is translated from both the sense and antisense strand, with the anti-sense poly-GP species surprisingly being the more abundant species (Zu et al., 2013). Poly-GP peptide is uncharged and forms a compact flexible coil structure (Freibaum and Taylor, 2017). Cellular models using poly-GP only constructs have observed that poly-GP is not aggregation-prone and do not associate with markers of amyloid aggregates (May et al., 2014; Wen et al., 2014; Yamakawa et al., 2014; Zhang et al., 2014; Tao et al., 2015; Flores et al., 2016). Proteomic studies have shown poly-GP peptide to interact with few cellular proteins (Lee et al., 2016a).

Poly-GP peptide is unlikely to be the primary pathogenic species in C9ALS/FTD with no toxicity reported in several studies (May et al., 2014; Wen et al., 2014; Zhang et al., 2014; Tao et al., 2015) and mild toxicity observed in other studies (Zu et al., 2013; Yamakawa et al., 2014). Yamakawa et al. (2014) have found that poly-GP impaired UPS degradation and enhanced cell death following exposure to the proteasome inhibitor MG132. In human brains, poly-GP level was significantly higher in the frontal cortex of ALS/FTD cases compared to FTD frontal cortex (Gendron et al., 2015), whilst was significantly lower in the cerebellum of ALS cases compared

to either ALS/FTD or FTD cerebellum. Poly-GP levels in the frontal cortex, motor cortex, cerebellum and the hippocampus did not correlate with cognitive impairment and also did not associate with the age of disease onset or survival after onset in any of the previously mentioned brain regions.

Poly-GP does appear to represent a useful pharmacodynamic biomarker for C9ORF72-ALS/FTD (Su et al., 2014; Gendron et al., 2017; Lehmer et al., 2017). A poly-GP immunoassay was developed by Su et al. (2014) to measure the soluble peptide levels in cerebrospinal fluid (CSF). Poly-GP was only detectable in CSF of C9ORF72-ALS, but not sALS or control CSFs. Two following studies have replicated the previous findings from Su et al. (2014) and went further to detect poly-GP in CSF of both symptomatic and asymptomatic ALS and ALS/FTD who are C9ORF72 expansion carriers and referred to the stability of CSF poly-GP over time (Gendron et al., 2017; Lehmer et al., 2017). Poly-GP was also detected in the media of the lymphoblastoid cell lines from C9 expansion carriers indicating the active secretion of poly-GP from cells. It is important to highlight that the poly-GP CSF levels did not correlate with age at disease onset, survival after disease onset, onset site or disease group and hence it cannot be employed as a prognostic marker for either C9ALS or C9ALS/FTD (Gendron et al., 2017).

Poly-GP immunoassay was also useful for testing the efficacy of antisense oligonucleotides (ASOs) targeting the G_4C_2 RNA. Both extracellular and intracellular poly-GP levels have been reduced following treating in C9ORF72 iPSC neurons with G_4C_2 ASO. So, these studies strongly support the use of poly-GP as a diagnostic marker for C9ALS/FTD and a measure of the efficacy of therapies targeting the G_4C_2 RNA in C9orf72 mutation carriers.

1.4.3.5 Poly-AP toxicity

Poly-Alanine-Proline (poly-AP) is one of the DPRs translated from the anti-sense C_4G_2 strand. Only very rare Poly-AP inclusions are reported in the motor neurons or cortex of C9ALS/FTD patients (Mackenzie et al., 2015; Schludi et al., 2015), with one study of exception that has observed poly-AP aggregates to be more prevalent than poly-GA inclusions in motor neurons (Cooper-Knock et al., 2015). Poly-AP is uncharged with a compact flexible coil structure and seemingly inert (Freibaum and Taylor, 2017). No evidence of poly-AP aggregation or toxicity has been reported in studies utilizing alternative codons (Freibaum and Taylor, 2017), instead, the peptide had a diffuse cytoplasmic localization in cellular models (May et al., 2014; Wen et al., 2014; Yamakawa et al., 2014; Zhang et al., 2014). Interestingly, a recent study has suggested a protective role of poly-AP in ameliorating poly-GA toxicity (Lee et al., 2017). These authors have

noticed that when poly-AP and poly-GA are co-expressed *in vitro* and *in vivo* in chick embryos, poly-AP binds to poly-GA and reduce its aggregation propensity and inturn its toxicity.

1.5 Hypothesis and main aims

Mutations that expand a hexanucleotide (GGGGCC) tandem repeat sequence (HTRS) in intron-1 of the C9ORF72 gene is the major cause of familial MND and FTD. The transcription and subsequent translation of all reading frames of HTRSs generate five types of dipeptide repeat proteins (DRPs) via Repeat-Associated Non-ATG (RAN) translation. The presumably toxic RAN proteins aggregate in neurons suggesting that they play an important role(s) in the pathogenesis of NMD. A surge of MND/FTD research in recent years have proposed several hypotheses for the molecular mechanism of these abnormal proteins. To add to this, the work contained in this thesis describes the generation of a cellular model to investigate the possible contribution of the different DPR proteins to the pathophysiology of c9orf72-ALS/FTD. I hypothesize that DPRs, and maybe specific DPR species, induce gain of toxic function through aberrant interactions with cell components.

The main aims of this study were to:

1. Probe how short and long polymers of DPRs differently engage with the proteome when expressed in a simple cellular model. and examine the toxic consequences of these interactions.
2. Conduct hypothesis-free analyses to determine how different DPRs could change the protein abundances and to examine the toxic consequences of the interactions emerging from findings generating from aims 1.
3. Define the proteinaceous composition of polyGA inclusions and elucidate how poly-GA aggregation impact the proteome.

By answering these questions, it is anticipated that the developed knowledge will uncover novel molecular underpinnings disrupted by C9ORF72-derived DPRs that may ultimately unravel novel drug targets for effective treatments for these hitherto incurable neurodegenerative diseases.

CHAPTER 2

General Methods

2 General Methods

2.1 DNA vectors and constructs

2.1.1 DPR constructs

Synthetic genes for short (10×) and (101×) dipeptide repeats were synthesized by GeneArt (Life Technologies, Regensburg, Germany). **Table 2-1** lists the full sequence information. These constructs were flanked 5' by XhoI and PstI recognition sites and 3' by BclI and BamHI recognition sites. XhoI and BamHI enzymes (NEB) were used to introduce the gene cassette into the mammalian expression vector pEGFP-C2 vector so the GFP tag was on the N-terminus of the 10× repeat sequence. The 101× DPR sequences were inserted between PstI and BclI sites in the backbones of the previously developed pEGFP-C2-10× DPR via classical restriction digest cloning following manufacturers' recommendations. Transformations were performed using recombination-deficient Stbl3 *E. coli* (Life Technologies) at 30 °C to minimize retraction of repeats. DNA was extracted using PureYield™ Plasmid Miniprep and Midiprep kits (Promega), following the manufacturer's instructions. Constructs were screened using standard restriction enzyme digest, agarose gel electrophoresis, and were sequence-verified to confirm the integrity of all constructs.

Table 2-1 | Sequence of the open reading frames from the synthetic DPR expression constructs.

Construct	Sequence
	10x DPRs
10xGA	5'-GGCGCTGGCGCTGGGGCAGGCGCAGGGGCTGGCGCAGGCGCTGGGGCTGGGGCTGGGGCA-3'
10xGR	5'-GGCAGAGGAAGAGGCAGGGGACGCGGAAGGGGGAGAGGACGCGGCAGAGGCCGGGGAAGA-3'
10xAP	5'-GCACCAGCTCCAGCCCCTGCTCCTGCTCCCGCCCCAGCACCCGCCCCTGCCCCAGCCCCA-3'
10xPR	5'-CCCAGACCTAGACCTCGGCCTAGACCAAGACCCAGGCCAAGGCCACGGCCAAGACCTAGA-3'
	101x DPRs
101xGA	5'GCCGGCGCTGGCGCTGGGGCAGGCGCAGGGGCTGGCGCCGGGGCCGGGGCCGGCGCTGGC GCAGGGGCTGGGGCTGGCGCAGGCGCTGGGGCAGGGGCTGGCGCTGGGGCTGGCGCAGGCGC AGGCGCTGGCGCTGGCGCAGGGGCTGGCGCAGGCGCTGGGGCTGGCGCTGGGGCAGGGGCAG GGGCAGGGGCTGGGGCAGGCGCTGGCGCAGGCGCAGGCGCAGGGGCAGGCGCTGGGGCTGGG GCTGGGGCTGGCGCAGGGGCCGGGGCCGGGGCAGGGGCAGGCGCAGGCGCAGGGGCTGGGGC AGGGGCAGGCGCAGGGGCTGGCGCTGGCGCTGGCGCCGGGGCCGGGGCCGGCGCAGGGGCTG GCGCTGGGGCAGGCGCTGGCGCAGGGGCAGGGGCAGGCGCTGGCGCTGGGGCAGGGGCTGGGG CCGGGGCCGGCGCAGGCGCTGGGGCAGGCGCAGGCGCAGGCGCTGGGGCCGGGGCCGGGGCT GGCGCTGGCGCTGGCGCAGGCGCTGGGGCTGGGGCAGGCGCCGGGGCCGGGGCCGGGGCAGG CGCTGGGGCTGGGGCTGGCGCAGGGGCAGGCGCTGGGGCAGGCGCAGGGGCC-3'

101xGR	5’CGGGGCAGAGGCCGGGGAAGAGGCAGAGGACGCGGAAGGGGAAGGGGGAGAGGAAG AGGGCGGGGACGCGGCCGGGGCCGCGGGCGGGGCCGGGGCCGGGGGCGCGGCCGGG GGCGGGGGCGCGGGCGCGGCCGGGGCCGGGGCCGCGGGCGGGGGCGCGGCCGGGG CCGGGGGCGGGGGCGCGGGCGGGGGCGGGGAAGAGGCAGGGGCAGAGGAAGAGGAA GAGGACGGGGGAGGGGCAGAGGAAGGGGACGCGGCAGAGGCAGGGGACGGGGACGC GGAAGAGGAAGAGGCAGAGGCAGGGGGAGAGGACGGGGCAGGGGGAGGGGCAGGGG ACGGGGGAGAGGACGCGGACGCGGACGCGGCAGGGGAAGAGGAAGGGGAAGGGGAAG AGGACGCGGCAGAGGACGGGGCAGAGGAAGAGGCCGGGGAAGAGGACGCGGAAGGGG GAGAGGCCGCGGAAGAGGAAGGGGAAGAGGGCGGGGACGGGGAAGGGGACGGGGACG GGGCAGAGGCAGAGGAAGGGGCAGGGGCAGGGGAAGGGGCAGAGGAAGAGGAAGGGG AAGAGGCAGAGGAAGAGGACGCGGCAGAGGCCGCGGAAGAGGACGG-3’
101xAP	5’CCCGCCCCTGCCCCTGCTCCAGCTCCTGCACCAGCACCCGCTCCAGCCCCAGCCCCCG CACCCGCCCCTGCACCCGCACCAGCACCAGCCCCTGCTCCAGCCCCCGCTCCTGCTCCTG CCCCAGCCCCAGCTCCCGCTCCAGCTCCTGCTCCCGCTCCAGCCCCTGCACCAGCCCCTG CCCCCGCTCCCGCACCAGCTCCAGCACCAGCTCCCGCCCCTGCTCCAGCACCAGCCCCA GCACCAGCACCAGCTCCCGCACCAGCCCCTGCCCCAGCTCCTGCTCCCGCCCCTGCTCC TGCCCCCGCACCCGCACCCGCTCCCGCCCCAGCTCCAGCTCCTGCCCCTGCCCCCGCTC CAGCACCCGCCCCAGCTCCCGCTCCTGCACCCGCTCCCGCTCCCGCACCCGCCCCAGCC CCTGCACCAGCTCCAGCTCCAGCTCCTGCTCCTGCACCAGCTCCCGCACCCGCCCCAGCT CCAGCTCCCGCTCCAGCACCAGCCCCTGCCCCTGCCCCAGCCCCTGCCCCTGCACCAGC ACCAGCACCCGCTCCAGCACCCGCACCAGCCCCAGCACCCGCCCCTGCACCAGCTCCCG CCCCAGCCCCCGCTCCT-3’
101xPR	5’CGGCCCAGACCCCGGCCTAGACCAAGACCCAGACCAAGGCCAAGACCTCGGCCTCGCC CTAGGCCACGCCCTCGCCCCAGACCTAGACCTAGGCCTCGGCCAAGACCTAGGCCCAGGC CTAGACCCCGCCCACGCCCTAGACCACGGCCAAGGCCTAGACCTAGACCAAGGCCCAGGC CCAGACCAAGGCCTCGCCCCAGGCCAAGACCACGGCCAAGACCAAGACCACGCCCCAGAC CCAGACCCAGACCTAGACCACGCCCTAGGCCAAGGCCTCGGCCACGGCCTCGGCCTAGAC CCAGGCCAAGACCTAGACCTCGGCCTAGACCACGGCCTCGCCCACGCCCAAGGCCAAGAC CAAGACCTAGACCCCGGCCTCGCCCAAGGCCCAGACCTCGCCCTAGACCTCGCCCAAGAC CAAGGCCTAGACCTCGCCCACGGCCCAGACCTAGACCAAGACCACGGCCACGCCCTAGAC CTAGGCCCAGACCTCGGCCCAGACCCAGACCACGGCCTAGACCTAGGCCAAGGCCACGCC CAAGGCCCAGGCCAAGGCCCAGACCAAGACCAAGACCCAGACCTCGGCCAAGGCCAAGGC CAAGACCAAGG-3’

2.1.2 Other vectors used in the study

pT-REx vector expressing exon 1 of Huntington (Httex1) with polyQ sequence length of 97 and C-terminal mCherry or GFP fluorescent tags and pT-REx-mCherry were prepared as previously described (Tsvetkov et al., 2010; Ramdzan et al., 2017).

2.2 Cell lines and transfection

Mouse neuroblastoma cell line (Neuro-2a cells) and Human embryonic kidney 293 cell line (HEK293T cells), obtained originally from the American Type Culture Collection, were routinely cultured in a reduced Serum Media Opti-MEM (Thermo Fischer Scientific, MA, USA) supplemented with 10% (v/v) foetal bovine serum (FCS), 1 mM glutamine, 100 Unit mL^{-1} penicillin and 100 µg mL^{-1} streptomycin in a humidified incubator under 5% v/v atmospheric CO2 at 37 °C. For microscopy experiments, $9x10^4$ Neuro2a cells were seeded per well into 8-well chamber µSlides (Ibidi, Martinsried, Germany). For flow cytometry, $4x10^5$ Neuro-2a cells were plated per well on 12-well plates. For proteomics experiments, Neuro2a cells were seeded on T75 flask at a density of $6x10^6$ cells. Cells were transfected using Lipofectamine 2000 (Thermo Fischer Scientific, Waltham, USA) as per the manufacturer instructions with 0.5 µg DNA for the microscopy experiments, 1.6 µg DNA for the flow cytometry and 24 µg DNA for proteomic experiments. The next day, the media were changed to Opti-MEM and for the time course, the media were refreshed daily.

2.3 Confocal Imaging

Cells expressing GFP-tagged DPRs were fixed 48 h after transfection in 4 % paraformaldehyde for 15 min at room temperature. Nuclei were counterstained with Hoechst 33342 at 1:200 dilution (Thermo Fisher Scientific) for 15 min at 37 °C, followed by a wash in Hank's Balanced Salt Solution (HBSS). Cells were fixed in 4 % paraformaldehyde and incubation was carried out in the dark for 15 min at room temperature. Afterwards, cells were washed with HBSS two times. Ibidi chambers were wrapped with foil and stored at 4 °C till imaging.

Fixed cells were imaged on a Leica SP5 confocal microscope using HCX PL APO CS 40× or 63× oil-immersion objective lens (NA 1.4) at room temperature. The Hoechst 33342 channel was collected with an excitation wavelength of 405 nm and emission wavelengths of 445–500 nm; EGFP was collected by excitation at 488 nm and emission from 520–570 nm; and mCherry was collected by excitation of 561 nm and emission range of 600–700 nm. Single colour controls were used to establish and adjust to remove bleed-through of the emission filter bandwidths. FIJI version of ImageJ (Abramoff et al., 2004) and Inkscape were used for image processing.

2.4 Survival analyses

2.4.1 Longitudinal live-cell imaging

Survival and risk of death analyses were performed on images acquired by an automated JuLI-stage imaging system (NanoEnTek) fitted into cell culture incubator. Stage and shutter movements, focusing and image acquisition are fully automated and controlled by PC running JuLI-stage imaging software. For automated imaging, neuro2a cells in 12-well plate format were co-transfected with individual GFP-tagged DPRs or GFP-tagged Httex1-97Q along with mCherry in a pT-Rex vector. mCherry fluorescence was employed to reveal cells independently of DPR expression. The media was refreshed 24 hours after transfection. Cells were then imaged longitudinally with the automated microscope at regular 15 min intervals for 96 hours. Channels used: GFP for EGFP (Excitation: 466/40, Emission: 525/50), RFP for mCherry (Excitation: 525/50, Emission: 580 LP).

2.4.2 Image analysis

Measurements of DPR expression, time of inclusion formation and cell survival were extracted from files generated with automated imaging by the FIJI version of ImageJ (Abramoff et al., 2004) and visual inspection. Death was recorded as the time points at which mCherry fluorescence was lost. This event corresponded to the loss of membrane integrity and cell death and was found to be a highly sensitive and specific assay of cell death through different pathways and in different types of cell (Strebel et al., 2001; Arrasate et al., 2004). Cells that drifted from focus were censored. The expression of GFP-tagged versions of DPRs was estimated by measuring the fluorescence intensity of co-transfected mCherry over a region of interest that corresponded to the cytosol of each cell. These mCherry intensity values were background-subtracted by using an adjacent area of the image.

For statistical analysis, survival time was defined as the imaging time point at which a cell was last seen alive. Kaplan–Meier curves were used to estimate survival and hazard functions with Prism software. Differences in Kaplan–Meier curves were assessed with Log-rank (Mantel-Cox) test. Correlations between DPR expression and survival or IB formation were made with Cox proportional hazard analysis. Differences in mean measurements were compared by analysis of variance or t-test.

2.5 Flow cytometry

2.5.1 Cell preparation for flow cytometry

Forty-eight hours following transfection, media was aspirated, and adherent cells were detached by gentle agitation and pipetting in 1 mL of phosphate-buffered saline (PBS) and then sedimented (120 xg for 6 min at room temperature). The supernatant was discarded and cell pellets were resuspended in 500 µL PBS followed by 0.5 µL of 5 µM SYTOX™ Blue dead cell stain (Invitrogen). Cell suspensions were kept on ice until analysis by flow cytometry (which was all completed within 1 h after labelling).

2.5.2 Flow cytometry

Cells were analysed using LSR Fortessa flow cytometer, equipped with 405-nm, 488-nm, 561 nm and 640-nm lasers (BD Biosciences). A minimum of 100,000 events per sample was collected at a high flow rate. Forward scatter and side scatter were collected using a linear scale. Fluorescent emissions were collected as area (log scale), pulse height (log scale), and pulse width (linear scale) for each channel. Data were collected in pulse height, area and width parameters for each channel. For DAPI, data were collected with the 405-nm laser and DABI (450/50) filter. For GFP, data were collected with the 488-nm laser and FITC (530/30) filter. For mCherry, data were collected with the 561 nm laser and PE-Texas Red (621/20) filter. For SYTOX Red Dead stain was collected using 640-nm laser and APC (670/14) filter. Flow cytometric gating and data analysis was performed using Flow Jo software (Tree Star Inc.) and graphs were analysed in Prism 8.

2.6 Sample preparation for proteome analysis of GFP immunoprecipitated samples.

An initial step to understand how abnormal polydipeptides exert toxicity is to identify their interacting partners. Interactomics have unravelled interesting insights into toxicity mechanisms of other neurodegenerative diseases (Olzscha et al., 2011; Hipp et al., 2014; Ripaud et al., 2014). We employed a quantitative proteomics approach using a dimethyl-labelling and LC-MS/MS based experiment with immunoprecipitation of the GFP tagged DPRs from neuro2a cells expressing with either GFP-DPR or GFP alone. **Figure 2-1** shows a scheme of the affinity purification-mass spectrometry and whole proteome analysis workflow employed in the current study.

6×10^6 Neuro-2a cells were seeded into 75cm^2 flasks and transfected the following day with either GFP-tagged DPRs or GFP-only constructs (24µg DNA and 60µL Lipofectamine 2000) according to the manufacturer's instructions (Life Technologies). The experiment was designed

as 3 or 4 biological replicates. Media was refreshed 24h after transfection. At 48h post-transfection, cells were gently rinsed with PBS and harvested in PBS by gently pipetting. Cells were pelleted (120g; 6min; room temperature) and resuspended in 1mL PBS and pelleted again (400g; 6min; room temperature). The pellet was resuspended in 200µL ice-cold lysis buffer (10mM Tris-HCl, pH 7.4; 150mM NaCl; 0.5mM EDTA; 0.5% v/v NP-40; 1 mM PMSF; 10 units/ml DNase I) supplemented with EDTA-free Complete protease inhibitor cocktail (Roche Diagnostic). The cell suspensions were then passed through a 27 Gauge syringe needle 25 times, followed by a 31 Gauge needle 10 times and incubated on ice for 30 min. The resultant lysates were clarified by centrifugation (21000 xg; 10 min; 4°C). Protein concentrations were quantified by the Pierce BCA Protein Assay (Catalogue Number: 23225, Thermo Fischer Scientific, MA, USA) using bovine serum albumin (BSA) as the mass standard. 0.5 mg of cellular protein was added to 25 µL of GFP-Trap MA beads (ChromoTek GmbH, Germany) pre-washed and equilibrated in the wash buffer (10 mM Tris/Cl pH 7.4; 150 mM NaCl; 0.5 mM EDTA; 1 mM PMSF; EDTA-free Protease inhibitor cocktail). The solution was incubated for 2 hr at 4°C (end-over-end rotation). Magnetically separated beads were then washed 3 times with wash buffer and 2 times more with 25mM triethylammonium bicarbonate (TEAB) buffer. The immunoprecipitated proteins were eluted by the addition of 100 µl of 50% v/v aqueous 2,2,2-Trifluoroethanol (TFE), 25mM TEAB. The supernatant was collected after pelleting (2000g; 2min; room temperature) and adjusted t

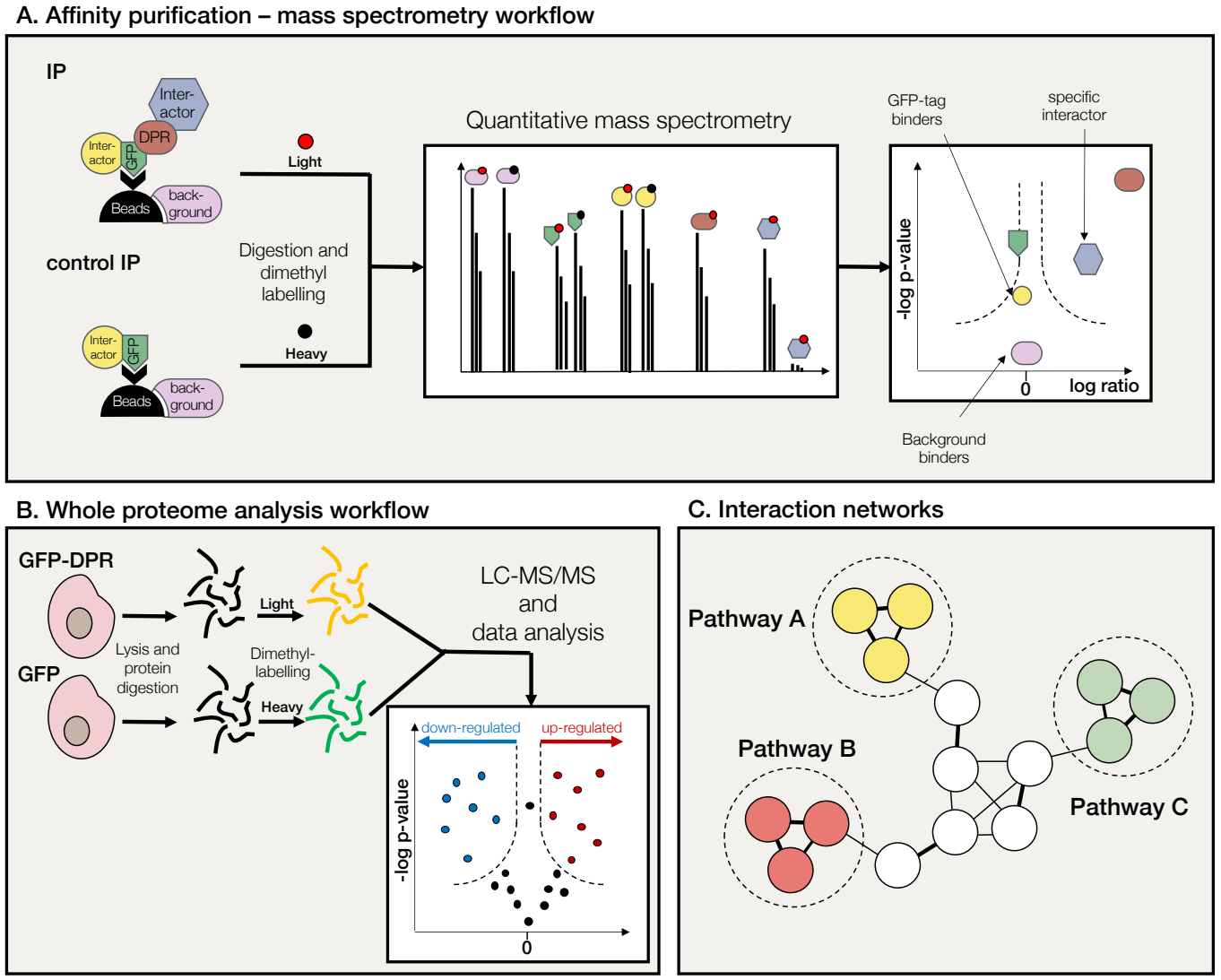

Figure 2-1 | A scheme of the affinity purification-mass spectrometry and whole proteome analysis workflow employed in the current study.

a final concentration of 100mM TEAB by addition of 1M stock solution and the pH was validated to be approximately 7 after this treatment). The samples were further processed for mass spectrometry analysis **(Section 2.9).**

2.7 Sample preparation for whole proteome analysis

2.7.1 Fluorescence-activated cell sorting (FACS) of live cells

Neuro2a cell lines of early passage numbers (18-20) were used in the study. Cells were collected at 48 h post-transfection by first rinsing in PBS, followed by resuspension in PBS with a cell scraper and gentle pipetting. Cells were pelleted at 120 xg for 6 min and resuspended in 2 ml PBS supplemented with 10 units/ml DNase I and filtered through 100-µm nylon mesh before analysis by flow cytometry. Just before sorting, 2 µL of the nuclear marker DAPI (1:1000) was spiked into cell suspensions to stain dead cells.

Cells were fluorescence-activated cell sorted using a FACS ARIA III cell sorter (BD Biosciences) equipped with 405-nm, 488-nm, 561-nm and 640-nm lasers with using a 100-µm nozzle. Gating was done using the BD FACS Diva™ software (Becton, Dickinson Biosciences). Cells (1,000,000) of each population of interest were sorted at a speed of 1500 cells/s. Side scatter (SSC) and forward scatter (FSC) height, width, and area were collected to gate for single-cell population. DAPI area was collected to gate-out dead cells. Data were also collected for pulse height, width, and area of GFP with FITC filter. To match for expression, cells were further gated to the same median GFP intensity of 2200 fluorescence units by varying the window of expression. Pulse shape analysis (PulSA) by flow cytometry was utilized to distinguish the 'ni' (non-inclusion) and 'i' (inclusion) population of 101xGA as described (Ramdzan et al., 2012). PulSA relies on standard pulse height and width parameters of a fluorescence channel where diffuse protein generates a wider and shorter pulse shape, whereas punctate protein as seen in inclusions generates a narrow and higher pulse shape. Since cells with i tend to be higher expresses then ni, the gates for both i and ni were not identical. Cells were sorted in parallel across three days and performed as three matched replicates. Cells were kept on ice for all steps of the sorting preparation and handling. The targeted population was directly sorted into PBS. Cells were pelleted at 120 g for 6 min, and the resultant pellet was snap-frozen in liquid nitrogen and kept at - 80 °C until use.

2.7.2 Preparing whole-cell protein extracts

Cell pellets were thawed then lysed in 100 µL RIPA lysis buffer (25mM Tris-HCl, pH 7.4, 150mM NaCl, 1% NP40, 0.1%SDS, 1% Sodium deoxycholate, 1x complete mini-protease cocktail; Roche), and incubated for 30 min. The concentration of proteins was measured by the

Pierce™ BCA Protein Assay according to the manufacturer's instruction (Thermo Fisher Scientific). To remove any residual detergent that can interfere with subsequent MS analysis, equal amounts of protein for each sample were precipitated with six volumes of pre-chilled (-20°C) acetone and the precipitation was allowed to proceed overnight. Next day, samples were spun at 13,000 rpm, 4°C, 10 min. Acetone was decanted without disturbing the protein pellet. The pellets were washed once with pre-chilled acetone then allowed to dry for 10 minutes. The protein precipitates were resuspended in 100 µL 0.1M TEAB and were vortexed and sonicated 3 times for 30 sec to help solubilize the pellet. The samples were further processed for mass spectrometry analysis **(Section 2.9).**

2.8 Sample preparation for poly-GA and poly-Q aggregates proteome analysis

Neuro-2a cells expressing either GFP-tagged GA_{101} or $Httex1Q_{97}$ in 3 replicates were harvested by pelleting (200 g; 5 min; 24 °C) 24 h post-transfection. Cell pellets were resuspended in lysis buffer (20 mM Tris, pH 8.0; 2 mM $MgCl_2$; 150mM NaCl; 1% (w/v) Triton X-100; 20 Units/mL Benzonase, Novagen; 1× complete mini-protease cocktail; Roche) and then incubated for 30 min on ice. Lysates were diluted 2 times with PBS supplemented with protease inhibitor and aggregates were pelleted at 1000 g for 6 minutes. The aggregates were washed twice with 1 mL PBS, then resuspended in 1 ml PBS and sorted on a BD FACS Aria III instrument with an outlet nozzle of 100 µm in diameter. The flow rate was adjusted to ~500 events/min, and EGFP fluorescence was monitored for sorting. Sorted aggregates were pelleted (12,000 g; 5 min; 4 °C), resuspended in PBS and washed 3 times by pelleting as above and resuspension in PBS. The final pellets were harvested by pelleting (21,000 g, 6 min, 4 °C) and dissolved in 10 µL neat formic acid for 30 min at 37 °C, vortexed for 20 seconds and sonicated for 1 min three times then incubated in a shaking microfuge tube incubator (30 min, 37 °C). Samples were neutralized to pH 7.0 by titration with unbuffered 3 M Tris. The protein concentration in the sample was determined by a Bradford assay using bovine serum albumin as mass standard. A total protein of 200 µg was further processed for mass spectrometry analysis **(Section 2.9).**

2.9 Mass spectrometry

2.9.1 Protein digestion

Proteins were subjected to reduction with 10mM tris(2-carboxyethyl)phosphine (TCEP), pH 8.0, and alkylation with 55 mM iodoacetamide for 45 min, followed by trypsin digestion (0.25µg, 37°C, overnight). The resultant peptides were then desalted by solid-phase extraction following acidification in 1% v/v formic acid; the cartridge (Oasis HLB 1 cc Vac Cartridge, product

number 186000383, Waters Corp., USA) was pre-washed with 1mL of 80% v/v acetonitrile (ACN) containing 0.1% v/v trifluoroacetic acid (TFA) and equilibrated with 1.2mL of 0.1% v/v TFA three times. Samples were then loaded on the cartridge and washed with 1.5mL of 0.1% v/v TFA before being eluted with 0.8ml of 80% v/v ACN containing 0.1% v/v TFA and collected in 1.5mL microcentrifuge tubes. Peptides were then lyophilized by freeze-drying (Virtis, SP Scientific). The peptides were resuspended in 100µL distilled water and quantified using microBCA assay (Catalogue Number: 23235, Thermo Fischer Scientific) with BSA as the mass standard.

2.9.2 Peptide labelling

10 µg of each sample (in a volume of 50µL containing 100 mM TEAB) were differentially labelled by reductive dimethyl labelling using equal volumes (2 µL) of 4% light formaldehyde (CH_2O), 4% medium formaldehyde (CD_2O, 98% D) or heavy formaldehyde ($^{13}CD_2O$, 99% ^{13}C, 98% D) and 0.6 M Sodium cyanoborohydride ($NaCNBH_3$, for light and medium label) or Sodium cyanoborodeuteride ($NaCNBD_3$, 96% D, for heavy label) added in sequence. The peptide solutions were incubated on an Eppendorf Thermomixer at room temperature for 1 h. After quenching with 8µL of 1% v/v ammonium hydroxide followed by 8µL of neat formaldehyde, dimethyl-labelled peptides were mixed up in equal volumes. The labelling strategy used for the different proteomic experiments is shown in **Table 2-2**. The peptide mixtures were then analysed using liquid chromatography-nano electrospray ionization-tandem mass spectrometry (LC-nESI-MS/MS) analysis.

Table 2-2 | Dimethyl labelling strategy employed in the proteomic experiments*.

Light	Medium	Heavy
GFP-immunoprecipitation experiment		
101PR	10PR	GFP
GFP	10GR	101GR
GFP	101GA	10GA
GFP	101AP	10AP
Whole proteome analysis		
101PR		GFP
---------	101GR	GFP
101GA-ni	----------	GFP
101AP	----------	GFP
101GA-ni	101GA-i	----------
Aggregate proteome analysis		
Htt-97Q	101GA	----------

* n = 3 biological replicates for all analyses except for PR and PA interactomes where n = 4 biological replicates.

2.9.3 NESI-LC-MS/MS analysis

Peptides were analyzed by LC-nESI-MS/MS using an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) fitted with nanoflow reversed-phase-HPLC (Ultimate 3000 RSLC, Dionex, Thermo Fisher Scientific). The nano-LC system was equipped with an Acclaim Pepmap nano-trap column (Dionex - C18, 100 Å, 75 μm × 2 cm) and an Acclaim Pepmap RSLC analytical column (Dionex - C18, 100 Å, 75 μm × 50 cm, Thermo Fisher Scientific). For each LC-MS/MS experiment, 1 μg (whole proteome) or 0.135 μg (aggregate proteome) of the peptide mix was loaded onto the enrichment (trap) column at a flow of 5 μl/min in 3% CH_3CN containing 0.1% v/v formic acid for 6 min before the enrichment column was switched in-line with the analytical column. The eluents used for the LC were 5% DMSO/0.1% v/v formic acid (solvent A) and 100% CH_3CN/5% DMSO/0.1% formic acid v/v. The gradient used was 3% v/v B to 20% B for 95 min, 20% B to 40% B in 10 min, 40% B to 80% B in 5 min and maintained at 80% B for the final 5 min before equilibration for 10 min at 3% B prior to the next analysis.

The mass spectrometer was operated in positive-ionization mode with spray voltage set at 1.9 kV and source temperature at 275 °C. Lockmass of 401.92272 from DMSO was used. The mass spectrometer was operated in the data-dependent acquisition mode, with MS spectra acquired by scanning from m/z 400–1500 at 120,000 resolution with an AGC target of 5e5. For MS/MS, the “top speed” acquisition method mode (3 s cycle time) on the most intense precursor was used whereby peptide ions with charge states ≥2 were isolated with an isolation window of 1.6 m/z and fragmented with high energy collision (HCD) mode, with a stepped collision energy of 30 ± 5%. Product ion spectra were acquired in the Orbitrap at 15,000 resolution. Dynamic exclusion was activated for 30s.

2.9.4 Proteomic data analysis

For GFP-immunoprecipitated samples, raw MS data were analysed using Proteome Discoverer (version 2.3.0.81; Thermo Fisher Scientific) with the Mascot search engine (Matrix Science version 2.4.1). Data were searched against the SwissProt *Mus Musculus* database (version 2016_07; 16794 proteins) combined with common contaminant proteins. GFP sequence (UniProt ID: P42212) was also added to the database. For protein identification, the search was conducted with 20ppm MS tolerance, and 0.8Da MS/MS tolerance. The enzyme specificity was set as trypsin. The maximum number of missed cleavage sites permitted was two, and the minimum peptide length required was six. The following modifications were allowed: Oxidation (M), Acetylation (Protein N-term), Dimethylation (K), Dimethylation (N-Term), Dimethylation: 2H(4) (K), Dimethylation 2H(4) (N-term), 2H(6)13C(2) Dimethylation (K),

2H(6)13C(2) Dimethylation (N-term) (Variable); Carbamidomethyl (C) (Fixed). The false discovery rate (FDR) was calculated by Percolator node in Proteome Discoverer v 2.3.0.81 and peptide identification accepted with a cut-off of 0.01 and a FDR cut-off value of 0.05 was applied for protein identification. Proteins were filtered for those identified by at least two peptides, one of which was unique, in all three replicates. The common contaminant, Keratin, was excluded from the dataset.

Peptide quantitation was performed in Proteome Discoverer v.2.3.0.81 using the precursor ion quantifier node. Dimethyl labelled peptide pairs (between two comparisons of light, medium or heavy) were established with a 2ppm mass precision, and a signal to noise threshold of 3. A retention time tolerance of isotope pattern multiplets was set to 0.8min. Three single peak or missing channels were allowed for peptide identification. The protein abundance in each replicate was calculated by summation of the unique peptide abundances that were used for quantitation (light, medium and-or heavy dimethyl derivatives). Missing quantitation values were replaced with a constant (zero-filling). The peptide group abundance and protein abundance values were normalized to account for sample loading. In brief, the total peptide abundances for each sample was calculated and the maximum sum for all files was determined. The normalization factor was the factor of the sum of the sample and the maximum sum in all files. After calculating the normalization factors, the Peptide and Protein Quantifier node normalized peptide group abundances and protein abundances by dividing abundances with the normalization factor over all samples. The normalized protein abundances were imported into Perseus software (v 1.6.5.0). Protein abundances were transformed to $\log_2$ scale. The samples were then grouped according to the replicates. For pairwise comparison of proteomes and determination of significant differences in protein abundances Welch's t-test based on permutation-based FDR statistics was then applied (250 permutations; FDR=0.01; S0=1). This, and all other t-tests below, are justified on the basis the proteomics abundance data is normally distributed.

For whole proteome data analysis, the conditions were similar to those above, but with the following differences: Raw MS data were analysed using Proteome Discoverer (version 2.2.0.388; Thermo Fisher Scientific). For protein identification, the search was conducted with 20ppm MS tolerance, and 0.6Da MS/MS tolerance. The maximum number of missed cleavage sites permitted was three. Additional variable modifications for mono- and dimethylation of arginine were included. Peptide quantitation was performed in Proteome Discoverer v.2.2.0.388, and the retention time tolerance of isotope pattern multiplets was set to 0.6min. After grouping samples, proteins with at least four valid values across all groups were retained for the

subsequent analysis. Missing values were imputed from distribution of all other log_2-transformed protein values from that sample, using the default settings in Perseus (1.8 standard deviation downshift, 0.3 mean downshift). For pairwise comparison of proteomes and determination of significant differences in protein abundances, a two-sample Students t-test based on permutation-based FDR statistics was applied (250 permutations; FDR=0.05; S_0=0.1).

For the aggregate proteome analysis, raw data were analysed using Proteome Discoverer (version 2.3; Thermo Scientific) with the Mascot search engine (Matrix Science version 2.4.1). Database searches were conducted against the Swissprot Mus musculus database (version 2016_07; 16794 proteins) combined with common contaminant proteins. GFP sequence (UniProt ID: P42212) was also concatenated to the $Httex1Q_{97}$ and $PolyGA_{101}$ sequences. The search was conducted with 20 ppm MS tolerance, 0.2 Da MS/MS tolerance and 2 missed cleavages allowed. Variable modifications were used for all experiments: oxidation (M), acetylation (Protein N-term), dimethylation (K), dimethylation (N-Term), 2H(4) dimethylation: (K) and 2H(4) dimethylation (N-term). A fixed modification used for all experiments was carbamidomethyl (C). The false discovery rate (FDR) was calculated by the Percolator node in Proteome Discoverer v 2.3.0.81 and was set to 0.5 % at the peptide identification level and 1 % at the protein identification level. Proteins were filtered for those containing at least two unique peptides in 3 biological replicates. Peptide quantitation was performed in Proteome Discoverer v.2.3 using the precursor ion quantifier node. Dimethyl labelled peptide pairs were established with a 2 ppm mass precision and a signal to noise threshold of 3. The retention time tolerance of isotope pattern multiplex was set to 0.6 min. Two single peak or missing channels were allowed for peptide identification. The protein abundance in each replicate was calculated by summation of the unique peptide abundances that were used for quantitation (light or medium derivatives). Missing quantitation values were replaced with a constant (zero-filling). The peptide group abundance and protein abundance values were normalized to account for sample loading. In brief, the total peptide abundances for each sample was calculated and the maximum sum for all files was determined. The normalization factor was the factor of the sum of the sample and the maximum sum in all files. After calculating the normalization factors, the Peptide and Protein Quantifier node normalized peptide group abundances and protein abundances by dividing abundances with the normalization factor over all samples. The normalized protein abundances were imported into Perseus software (v 1.6.5.0). Protein abundances were transformed to log2 scale. The samples were then grouped according to the replicates. For pairwise comparison of proteomes and determination of significant differences in protein abundances, paired Student's t-test based on permutation-based FDR statistics was then applied (250 permutations; FDR =

0.05; S0 = 0.1). This was justified on the basis the proteomics abundance data is normally distributed.

2.10 Bioinformatic analysis and visualization using Cytoscape

Protein interaction networks were generated using Cytoscape 3.7.1 (Shannon et al., 2003) built-in String (v11.0) (Szklarczyk et al., 2017) using active interaction sources parameters on for Experiments, Databases, Co-expression neighbourhood, Gene Fusion and Cooccurrence. The minimum required interaction score setting was 0.9 for interactome data or 0.7 for whole proteome and aggregate proteome data. Protein abundance changes were used as node color attributes and the node size reflected the p-value. The corresponding enriched GO annotation terms were determined by calculating their enrichment P-value, which we compute using a Hypergeometric test, as explained in (Rivals et al., 2006). The P-values are corrected for multiple testing using the method of Benjamini and Hochberg (Benjamini and Hochberg, 1995). Selected GO terms were used to manually re-arrange nodes and were added on the protein interaction network using Inkscape.

Cytoscape 3.7.1 plugin BiNGO was used to perform GO term analyses performed on proteins significantly enriched in the long versus short Arg-rich DPRs immunoprecipitates. Overrepresented categories were chosen after Benjamini & Hochberg False Discovery Rate correction. A hypergeometric statistical test was used to ascertain the p-value of GO term enrichment. The annotation plots show the GO terms comprising at least three proteins with Benjamini-Hochberg adjusted $p < 0.05$ and the $\log_2$ fold enrichment of their associated proteins were significantly different from zero as assessed by one-sample student T-test.

A Python algorithm developed by Mitrea *et al*. (2016) was employed in this study to identify proteins exhibiting multivalent arginine-rich motifs (R-motifs) with the sequence pattern, $RX_{n1}RX_{n2}RX_{n3}R$, where n1, n3≤2 and n2≤20. This algorithm was applied to the DPR-interacting proteins as well as proteins up- or downregulated upon DPR expression. The background proteins observed in the whole proteome analysis and which were not deemed significantly affected in abundance by DPR expression was used as a control. For determining significant differences, Fisher exact test was used.

IUPred is a predictor that assigns each amino acid an energetic contribution to the stability of the protein; this contribution is a function of the frequencies of all other amino acids in the sequence. IUPred-L predicts global structural disorder that encompasses at least 30 consecutive residues of the protein (Dosztányi et al., 2005). This algorithm was applied to predict the intrinsically unstructured/disordered regions of proteins significantly enriched in polyGA or

Httex1Q_{97} aggregates. Glutamine content was analyzed with the web-server COPid (Kumar et al., 2008) (http://www.imtech.res.in/raghava/copid/). A control set of 100 random proteins **(Table 4-S1)** was generated from a list of the mouse proteome obtained from the Uni-ProtKB database (http://www.uniprot.org/uniprot/?query=reviewed:yes+AND+organism:10090&random=yes). The Mann-Whitney- Wilcoxon test was employed to determine significant differences.

2.11 Data availability

The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium database via the PRIDE (Vizcaino et al., 2016) partner repository program with the following dataset identifier: PXD015177 (for interactome data), PXD015180 (for the whole proteome data), PXD018505 (aggregate proteome data) and PXD018824 (polyGA aggregation impact on the whole proteome data).

2.12 Statistical Analysis

The details of the tests were reported in the Figures and corresponding Figure Legends. All statistical analyses were performed with GraphPad Prism v 7.05 (Graphpad Software Inc., San Diego, CA). Significant results were defined on the figures for $P < 0.05$. P values lower than 0.05 are coded as *, $P<0.05$; **, $P<0.01$; ***, $P<0.001$; ****, $P< 0.0001$.

CHAPTER 3

Arginine in C9ORF72 Dipolypeptides Mediates Promiscuous Proteome Binding and Multiple Modes of Toxicity

Published April 2020 in Molecular & Cellular Proteomics

3 Chapter 3: Arginine in C9ORF72 Dipolypeptides Mediates Promiscuous Proteome Binding and Multiple Modes of Toxicity*

3.1 Highlights

- Quantitative proteome interactions with 5 different C9ORF72 dipolypeptides (DPRs).
- The Arg-rich DPRs promiscuously bound to the proteome compared with the other DPRs.
- Long repeat lengths of Arg-rich DPRs, but not short lengths, stalled ribosomes.
- The Arg-rich DPRs also reduced arginine methylation and actin cytoskeleton assembly.

3.2 Abstract

C9ORF72-associated Motor Neuron Disease patients feature abnormal expression of 5 dipeptide repeat (DPR) polymers. Here we used quantitative proteomics in a mouse neuronal-like cell line (Neuro2a) to demonstrate that the Arg residues in the most toxic DPRS, PR and GR, leads to a promiscuous binding to the proteome compared to a relative sparse binding of the more inert AP and GA. Notable targets included ribosomal proteins, translation initiation factors and translation elongation factors. PR and GR comprising more than 10 repeats appeared to robustly stall on ribosomes during translation suggesting Arg-rich peptide domains can electrostatically jam the ribosome exit tunnel during synthesis. Poly-GR also recruited arginine methylases, hypomethylation of endogenous proteins, and induced a profound destabilization of the actin cytoskeleton. Our findings point to arginine in GR and PR polymers as multivalent toxins to translation as well as arginine methylation that may explain the dysfunction of biological processes including ribosome biogenesis, mRNA splicing and cytoskeleton assembly.

3.3 Introduction

The major genetic cause of motor neuron disease (MND) (also known as amyotrophic lateral sclerosis (ALS)) and frontotemporal dementia (FTD) is an expansion in the number of GGGGCC hexanucleotide repeats in C9ORF72 from less than 15 in the general population to

* This chapter represents the following manuscript:

Radwan M, Ang CS, Ormsby AR, Cox D, Daly JC, Reid GE, Hatters DM. Arginine in C9ORF72 Dipolypeptides Mediates Promiscuous Proteome Binding and Multiple Modes of Toxicity. Mol Cell Proteomics. 2020;19(4):640-654. doi:10.1074/mcp.RA119.001888

over 20 (and typically hundreds) (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Majounie et al., 2012; Gomez-Tortosa et al., 2013). Toxicity has been proposed to arise through multiple mechanisms including C9ORF72 haploinsufficiency (DeJesus-Hernandez et al., 2011), the formation of C9ORF72 mRNA foci that sequesters critical RNA binding proteins (Donnelly et al., 2013; Gendron et al., 2013; Lee et al., 2013; Mori et al., 2013a; Sareen et al., 2013; Cooper-Knock et al., 2014) and the production of abnormal translation products by non-AUG-initiated translation (RAN translation) (Zu et al., 2011). RAN translation occurs from expanded hexanucleotide repeat lengths in both sense and anti-sense transcripts resulting in abnormal expression of 5 dipeptide repeats (DPRs) in neurons of patient brains: poly-GP, poly-GA, poly-GR, poly-AP and poly-PR (Ash et al., 2013; Gendron et al., 2013; Mori et al., 2013a; Zu et al., 2013).

Experimental animal and cell culture models expressing the DPRs have revealed poly-GR and poly-PR to be particularly toxic, with the others being comparatively inert (May et al., 2014; Mizielinska et al., 2014a; Wen et al., 2014; Lee et al., 2016a). Furthermore, while all DPRs are widely distributed in human brain of patients with ALS, only poly-GR is correlated to clinically related regions (Saberi et al., 2018). Various interactome studies have indicated that the poly-GR and poly-PR DPRs engage with RNA binding proteins, ribosome machinery and proteins with low complexity domains, which mediate the formation of membrane-less organelles by phase separation (Lee et al., 2016a; Zhang et al., 2018; Moens et al., 2019). These interactions have been proposed to negatively impact on the functioning of ribosome biogenesis (Kwon et al., 2014), ribosome activity (Kanekura et al., 2016), nucleolus function (Kwon et al., 2014; Tao et al., 2015), nucleocytoplasmic transport (Zhang et al., 2015; Freibaum and Taylor, 2017) and stress granule dynamics (Tao et al., 2015; Lee et al., 2016a; Zhang et al., 2018).

Here, the role of the poly-GR and poly-PR DPRs expressed in a simple cell model was examined by defining what they interact with using quantitative proteomics and examining how short DPR lengths (10× repeats) differed to longer lengths (101× repeats). This proteomics data suggested a potent engagement of poly-PR and poly-GR to the ribosome and translational machinery. Given that poly-GR and poly-PR suppress protein translation, role of the interactions of ribosomes was further explored (Kanekura et al., 2016; Lee et al., 2016a; Zhang et al., 2018; Moens et al., 2019). Here evidence is provided that disruption of protein translation may arise by Arg-rich peptides stalling on ribosomes during their synthesis. other key mechanisms mediating poly-PR and poly-GR toxicity are also revealed. This includes destabilization of the actin cytoskeleton and proteome arginine hypomethylation. These findings point to the repetitive

arginine sequences in the DPRs promoting promiscuous binding to the proteome that in turn enact multiple modes of toxicity.

3.4 Specialized methods

3.4.1 Dual fluorescence translation stall assay

Genes were synthesized to produce the dual-fluorescence translation stall reporter as described previously (Juszkiewicz and Hegde, 2017), except we used mCherry as the red fluorescent protein. In these translation stall reporter plasmids, GFP and mCherry fluorescent protein are expressed in-frame from a single mRNA that is separated by a linker sequence. Two Ribosomal 2A skipping sequences flank the linker region, allowing GFP and ChFP to be independent markers of translation prior to or after the linker region. The linker sequence contained the cytosolic domain of Sec61b containing the autonomously folding villin headpiece (VHP) domain inserted between residues 14 and 15 (Shao et al., 2016) and was utilized as a negative control. A stretch of 20 consecutive lysine codons (encoded by AAA) that have been proved to robustly induce ribosome stalling was fused downstream of VHP domain and was utilized as a positive control. DPRs constructs or exon 1 of Huntington (Httex1) with different polyQ expansions (25Q, 72Q and 97Q) were cloned to replace the linker region using PstI and BAMHI restriction sites. Frame shifting was corrected using standard PCR-based strategies and was confirmed by sequencing.

Dual fluorescence reporter plasmids were transfected into cells using Lipofectamine 2000 (Life Technologies) according to manufacturer guidelines. Two days after transfection, cells were harvested in PBS containing SYTOX Blue dead cell stain (S34857, Thermo Fischer Scientific) to exclude dead cells from subsequent analysis. Cellular fluorescence of 100,000 events per sample was analysed on a LSRFortessa X-20 flow cytometer (BD Biosciences) using the 488-nm laser and 530/30 filter for GFP and the 561nm laser and 610/20 bandpass filter for mCherry. Subsequent analysis of flow cytometry data was done using FlowJo software (v10.5.3) and graphs were analysed in GraphPad Prism 7.05.

3.4.2 Flow Cytometry analysis of G- and F-actin.

F- and G-actin levels in Neuro2a cells were measured as described (Grosse et al., 2003). Briefly, Neuro2a cells seeded into 12-well plates were harvested 48 h following transfection of different GFP-tagged DPRs. Cells were treated with 2 μM cytochalasin-D in 0.1% DMSO/PBS for 1 h for the positive control. Cells incubated with 0.1% DMSO for 1 h were used as a negative control (untreated cells). Cells were fixed with 4% paraformaldehyde for 15 min, then

permeabilized with 0.2% Triton X-100 for 5 min. After washing in PBS, cells were blocked with 1% BSA in 0.1% v/v Triton X-100/PBS solution for 15min and then incubated in the dark at room temperature for 30 min with Alexa Flour 594 deoxyribonuclease1 (DNase1) conjugate (10 µg mL^{-1}, Invitrogen Molecular Probes, D12372) for G-actin detection and Alexaflour-405 Phalloidin (1:1000, Invitrogen Molecular Probes, A30104) for F-actin detection. After thorough washing with 1× PBS, fluorescence was measured using blue (BV421), green (FITC) and red (PE-Texas Red) channels on a FACSCanto flow cytometer (BD Bioscience). A total of 1×10^4 cells were analysed per acquisition. Unstained cells were used to set the baseline. G- and F-actin contents were determined from the respective fluorescence and the ratio of F/G calculated from the mean values as determined with FlowJo software.

3.4.3 Confocal Microscopy for F-actin analysis.

Cells were grown in 8–well ibidi culture chambers (Sarstedt, Nümbrecht, Germany). Cells were washed with PBS once before being fixed with 4% paraformaldehyde for 15 min at room temperature, then washed with PBS 3 times and permeabilized in 0.2% w/v Triton X-100/PBS for 5 min. The cells were then blocked with 1% w/v BSA in 0.1% w/v Triton X-100/PBS for 15 min at room temperature. F-actin was stained with Alexa Fluor 594 phalloidin (1:1000, Invitrogen, A12381) for 30 min and with Hoechst 33342 (1:200, Thermo Fisher Scientific) for nuclei staining for 30 min in the dark. Cells were imaged on a Leica SP5 confocal microscope using HCX PL APO CS 40× or 63× oil-immersion objective (NA 1.4) at room temperature. Fluorescence intensity profiles of phalloidin in >50 cells expressing GFP-tagged DPRs and untransfected cells were analysed using ImageJ software.

3.5 Results

3.5.1 Characterization of the cell model expressing the DPRs

Expression constructs of the DPRs were synthesized using mixed codons and an ATG-start codon, designed to minimize the influence of RNA-mediated effects from the repeat sequences. 10× and 101× repeat lengths of the toxic poly-PR and poly-GR, as well as the less toxic polyGA and polyAP, were prepared.

These constructs displayed patterns of localization and toxicity in the mouse neuroblastoma cell line Neuro2a similar to that described previously in other models (May et al., 2014; Mizielinska et al., 2014a; Wen et al., 2014; Lee et al., 2016a; Hartmann et al., 2018; Suzuki et al., 2018) and observed *in vivo* for cases of MND with C9ORF72 mutation (Gendron et al., 2013; Mori et al., 2013c; Zu et al., 2013; Schludi et al., 2015). This included a predominately

nuclear punctate pattern of localization for PR_{10} and PR_{101} with residual expression in the cytosol **(Figure 3-1A)**. The longer repeat length accentuated the localization into the nuclear foci. GR_{10} also formed nuclear foci and appeared almost identical to PR_{10}, however, GR_{101} was excluded from the nucleus and had a mostly diffuse cytoplasmic distribution. By contrast, GA_{10}, AP_{10} and AP_{101} were slightly enriched in the nucleus without forming distinct puncta. Similarly, GA_{101} was enriched in the nucleus but also formed large cytosolic (and less commonly nuclear) inclusions in many cells **(Figure 3-1A & Table 3-S1)**. All DPR constructs expressed at levels comparable to that of GFP except for PR_{101} and GR_{101} that have very weak expression levels; about 30% lower than that of either GFP or other long DPRs **(Figure 3-1B)**. According to cell survival rates when expressing the DPRs, the poly-PR and poly-GR constructs were most toxic and this was true in both 10× and 101× lengths **(Figure 3-1C)**. In contrast, poly-GA was only toxic in 101× lengths and the other DPRS were not toxic.

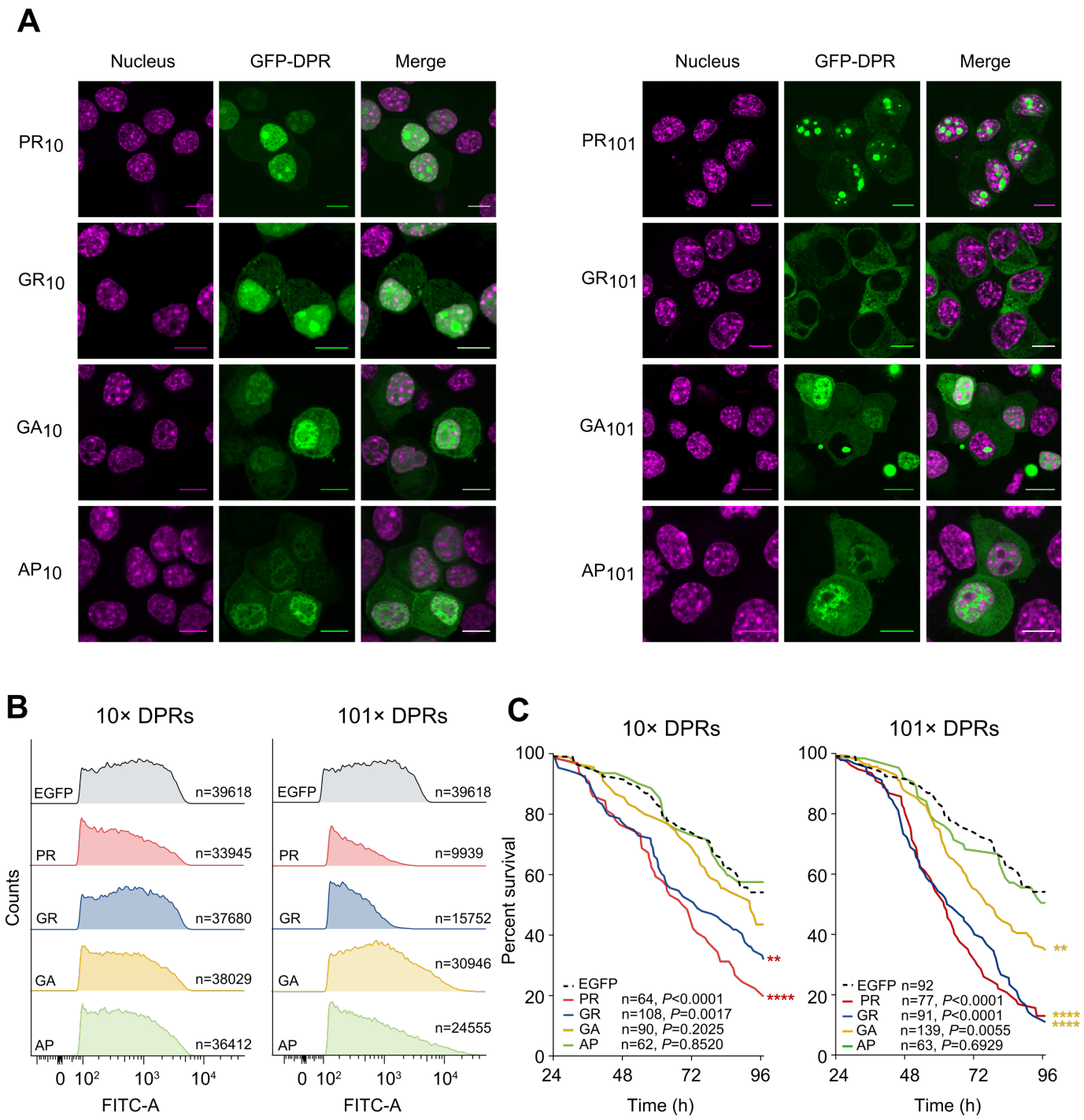

Figure 3-1 | DPR toxicity and cellular localization.

A, Confocal micrographs of Neuro2a cells 48 h after transfection with GFP-tagged DPRs. The nucleus was stained with Hoechst 33258. Scale bars represent 10 µm. B, Expression levels of the DPRs as assessed by flow cytometry, 48 h following transfection. FITC-A channel tracks the GFP fluorescence. Shown are histograms of cell populations. C, Kaplan-Meier survival analysis of Neuro-2a transfected with the GFP-tagged DPRs. Cells were tracked from 24 h post-transfection. p values correspond to a log-rank (Mantel-Cox) test of each sample versus EGFP control.

3.5.2 Arginine-rich DPRs attracts promiscuous proteome interactions

With the model system established, we next sought to investigate how the toxic 101× length DPRs (poly-GR, poly-PR and poly-GA) interacted with the proteome. To do this we transiently transfected cells with GFP or the 101× DPRs fused to GFP and then captured proteins that bound to the DPRs by immunoprecipitation with GFP trap. Proteins were assessed by quantitative proteomics by comparing the pulldowns to GFP-only transfected cells, with protein levels normalized to protein mass recovered from the immunoprecipitation. Under these conditions, GFP levels were anticipated to differ in the pulldown due to different expression levels **(Figure 3-1B)**. Indeed, the amount of GFP appeared heavily enriched in the GFP-only control for the toxic DPRs (PR_{101}, GR_{101} and to a lesser extent GA_{101}) **(Figure 3-2A)**. Yet, despite the larger amount of GFP coming from the control GFP-only transfected cells (which would enrich for non-specific interactors to GFP), we observed many proteins strongly enriched to the Arg-rich DPRs and comparatively few for GA_{101} and AP_{101} **(Figure 3-2A; Table 3-S1)**. The result suggested two conclusions. One was that the Arg appears to mediate promiscuous binding to the proteome and the second was that these interactions are responsible for mitigating toxicity. The Arg-rich DPRs in particular enriched for proteins involved in ribosome biogenesis and RNA splicing machinery, which is consistent with prior findings (Lee et al., 2016a; Lin et al., 2016; Boeynaems et al., 2017; Hartmann et al., 2018). However, we also found additional novel interactions with proteins involved in ribosome-translation, cytoskeleton and chromatin machineries **(Figure 3-2B)**. GR_{101} also enriched specifically for methylosome proteins and PR_{101} with mitochondrial proteins. GA_{101}, which is the only DPR that formed large cytosolic inclusions, was enriched for a very distinct proteome. These interactions may be indicative of a distinct set of molecular mechanisms involved in its more modest toxicity. It will be interesting in the future to investigate the target of ubiquitination since all different DPRs lack lysine. In our model, the GFP tag fused to the DPRs may be ubiquitinated. diseased brains, the DPRs are translated with additional endogenous N- and C-terminal sequences specific for each reading frame as indicated in **Table 1-1**. The antibodies raised against these sequences were capable to specifically probe each DPR species in patient tissues (Zu et al., 2013). So, it is possible that these N- and C-terminal sequences are the target of ubiquitination in vivo. Another possibility is that ubiquitin is attracted to proteins that bind to the DRPs.

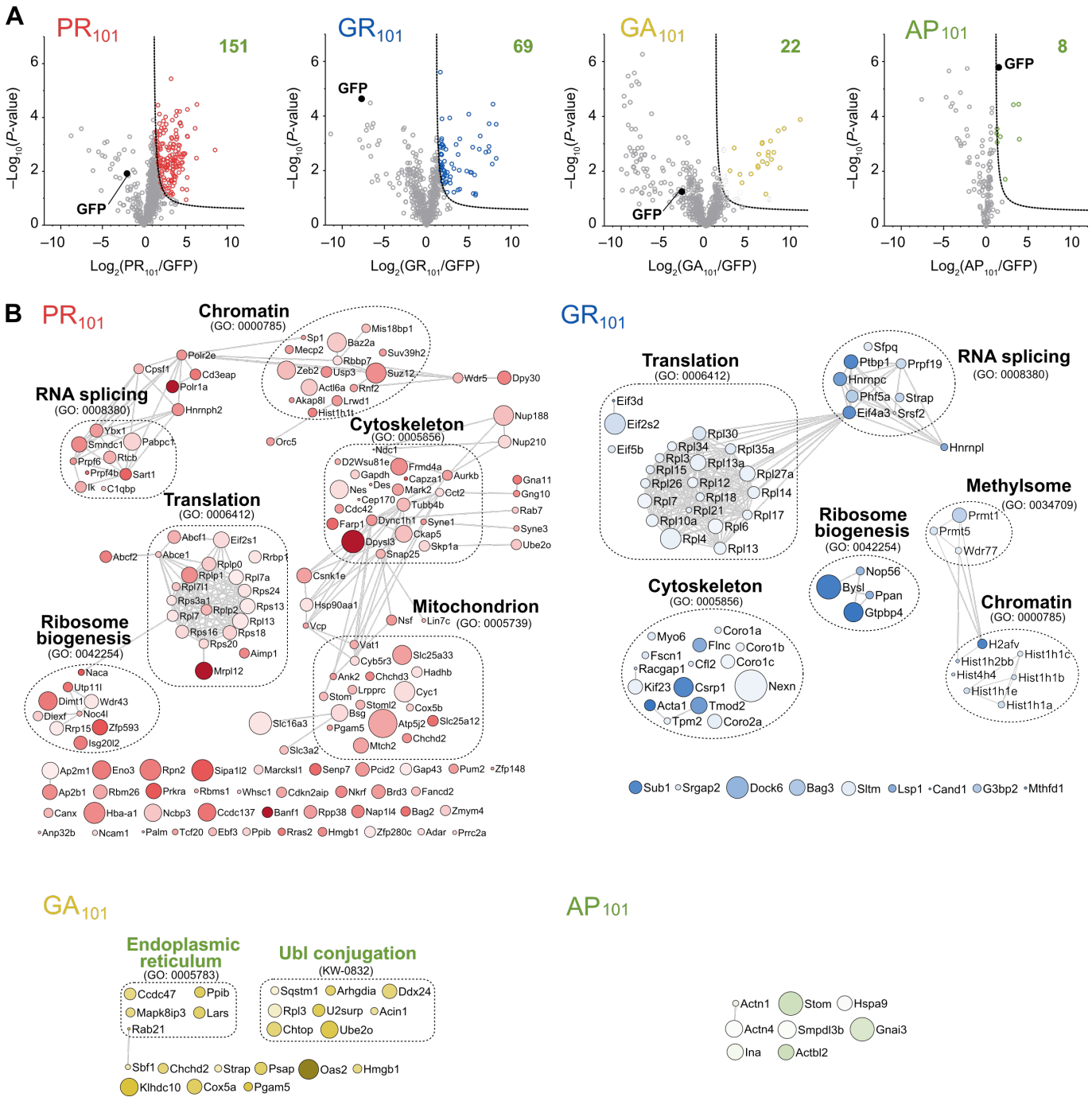

Figure 3-2 | Interactome analysis of the DPR_{101} variants.

A, Volcano plots of each DPR_{101}-GFP versus GFP-only control from quantitative proteomic analysis of GFP-Trap immunoprecipitates of DPR_{101}-GFP transfected in Neuro2a cells harvested 48h after transfection. Significant binders (shown in colored circles) were classified with False Discovery Rate of ≤ 0.01 (dotted lines). The number of interactors is indicated. n = 4 biological replicates for PR and AP and n = 3 biological replicates for GR and GA. B, STRING (v10) interaction maps for proteins significantly enriched with confidence set at 0.9 (highest stringency). Circle sizes are proportional to –log10 (*P*-value). The color intensity is proportional to the log2 (fold change). Selected significantly enriched GO terms (GOCC, GOPB, and UniProt keywords) are displayed.

3.5.3 Arginine-rich DPRs lead to ribosome stalling

It was previously reported that Arg-rich DPRs can cause translational suppression although a mechanism that remains undetermined (Moens et al., 2019). Our proteomics data revealed the ATP-binding cassette sub-family E member 1 (ABCE1) was enriched in the PR_{101} interactome. ABCE1 is involved in translation termination ribosomal recycling (Pisarev et al., 2010), which led us to wonder whether synthesis of the Arg-rich DPRs impairs translation. To examine this possibility, we employed a previously established assay to measure the ability of a protein sequence to stall or delay protein synthesis rates (Juszkiewicz and Hegde, 2017). This assay involves a cassette containing two fluorescent reporters on each side of the peptide sequence to be tested for stalling (GFP at the N-terminus and mCherry at the C-terminus) **(Figure 3-3A)**. Each construct is encoded in frame without stop codons. However, the sequence to be tested; e.g. DPR-coding sequence, is flanked by viral P2A sequences, which causes the ribosome to skip the formation of a peptide bond but otherwise continue translation elongation uninterrupted. This means that complete translation of the cassette from one ribosome will generate three independent proteins (GFP, protein tested, and mCherry) in an equal stoichiometry. However, should the ribosome stall during synthesis, mCherry is produced at lower stoichiometries than the GFP **(Figure 3-3A)**.

It has been previously validated that stalling reporters that have 21 AAA codons as a sequence to be tested (which therefore encodes poly-lysine) stall the ribosome while the reporters with no AAA codons allow read-through (Juszkiewicz and Hegde, 2017). When the different DPR-coding sequences was cloned inbetween the P2A sequences to replace the previous tested sequences, GR_{101} and PR_{101} constructs induced marked stalling **(Figure 3-3 B&C and Figure 3-S3)**. This was DPR-length-dependent in that the 10× repeats showed no stalling **(Figure 3-3C)**. In addition, the AP_{101} and GA_{101} did not lead to stalling **(Figure 3-3C)**. Indeed, there was a significantly increased ratio of mCherry to GFP **(Figure 3-3C)**. A likely explanation for this increased ratio comes from intermolecular fluorescence resonance energy transfer (FRET) arising from low rates of translational readthrough of the stall construct and self-association of these read-through protein products. This was more evident in another control construct of mutant Huntington exon 1 (Httex1), which when containing a polyglutamine (polyQ) expansion above 36 glutamines causes Huntington Disease and becomes highly aggregation-prone (Duyao et al., 1993; Scherzinger et al., 1999). The expanded polyQ forms of Httex1 did not cause stalling **(Figure 3-3D)**. However, increased pathological lengths of polyglutamine increased the ratio of GFP to mCherry. Microscopic images confirmed the presence of GFP and mCherry in aggregates in cells expressing the GA_{101} and Httex1 reporters **(Figure 3-S2)**.

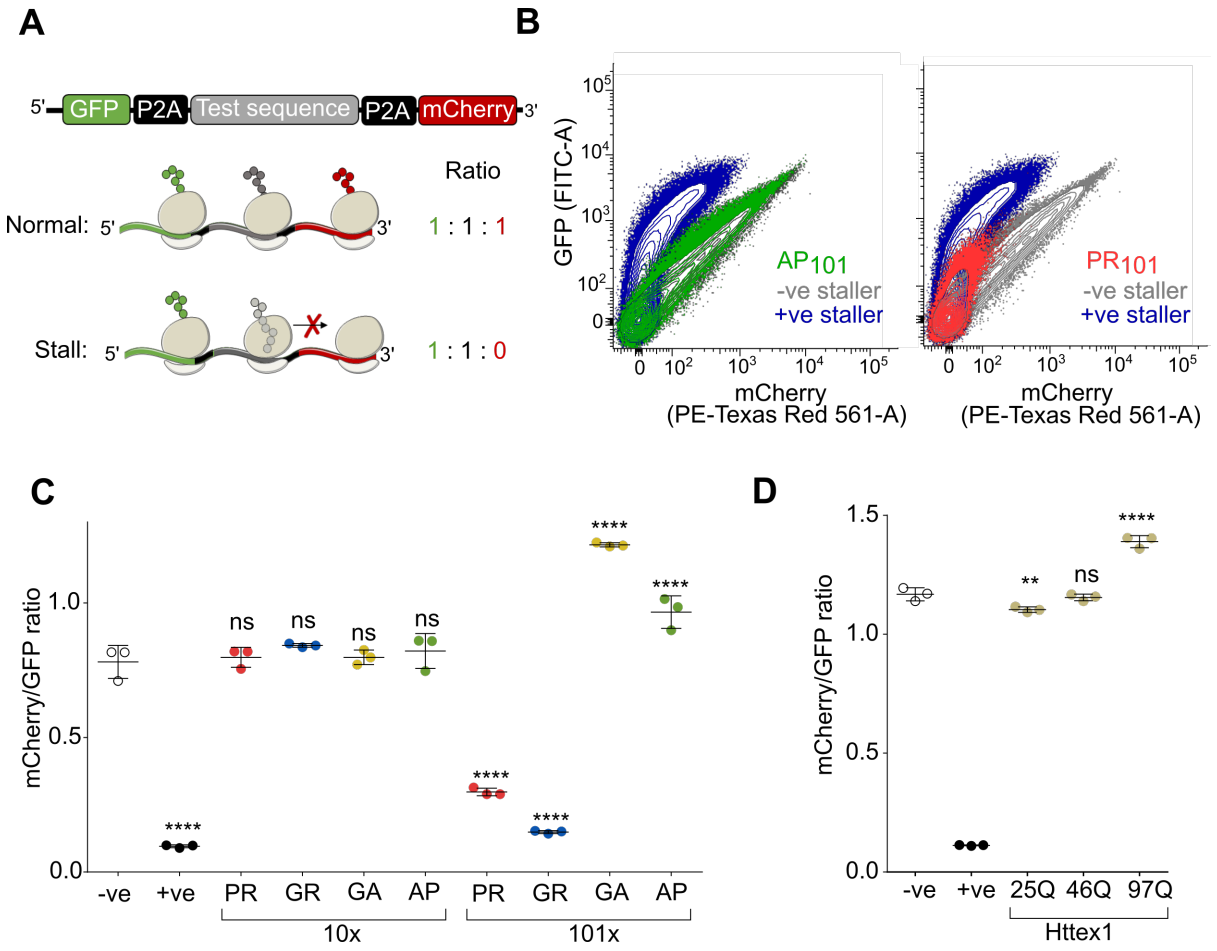

Figure 3-3 | Long Arg-rich DPRs stall ribosomes during translation.

A, Schematic of the reporter construct design. The P2A sequence causes the ribosome to skip the formation of a peptide bond but otherwise continue translation elongation uninterrupted. Complete translation of the cassette from one ribosome will generate three independent proteins (GFP, test protein, and mCherry). However, should the ribosome stall during synthesis (such as through the previously established positive stall reporter sequence with poly-lysine (K20; +ve sequence. The –ve sequence is stall reporter sequence without lysine residues. B, Flow cytograms of Neuro2a cells 48 h after transfection with the indicated test sequences inserted into the reporter. C, Median mCherry:GFP ratios calculated from transfected cells population (from 100,000 analyzed cells). Error bars indicate standard deviations from three independent transfections and flow cytometry measurements. p values determined for one-way ANOVA and Dunnett's post hoc test using the –ve as the control. ****, $p < 0.0001$; ***, $p < 0.001$; ns, $p > 0.05$. D, the same assay (and statistical tests) using Huntington exon 1 (Httex1) transfected in HEK293 cells with the indicated polyglutamine (polyQ) lengths.

3.5.4 Arginine-rich DPRs alter the state of the actin cytoskeleton

We next sought to examine whether the interactomes of the Arg-rich DPRs were influenced by the length of the repeat sequence **(Figure 3-4A)**. For this analysis, we measured the relative enrichment of the interacting proteins (101× versus 10×) in each gene ontology (GO) term significantly associated with the Arg-rich DPRs of both lengths. (The interactomes of 10× DPRs compared to GFP-control are shown in **Figure 3-S4**). The majority of GO terms were significantly enriched to the longer Arg-rich DPRs **(Figure 3-4A)**. These data, therefore, suggest arginine content generally mediates the binding, which may arise through increased charge per molecule (i.e. arg valency) or greater aggregation capacity. We also saw enrichment patterns consistent with different cellular localizations of 10× and 101× PR constructs (as shown in **Figure 3-1**). In particular, the PR_{101} DPR revealed a substantial enrichment of proteins in GO terms relevant to nucleus localization (including DNA replication, positive regulation of transcription) which was consistent with the enhanced localization of PR_{101} to nucleolar substructures compared to PR_{10} **(Figure 3-4A)**. While the actin-related GO terms were enriched also for the Arg-rich DPRs, the enrichment correlated with the DPR localization in the cytosol where most actin is expected to reside. Namely, there was a lesser enrichment for PR_{101} compared to PR_{10}, in accordance with the shift from more diffuse nuclear and cytoplasmic localization into the nucleolar substructures. GR_{101} was also excluded from the nucleus compared to GR_{10}, and this was reflected in the greater enrichment of GR_{101} with actin-related GO terms.

Next, we examined whether the enrichment with actin GO terms was indicative of changes in the actin cytoskeleton. The Arg-rich DPRs except for GR_{10} had a significant impact on the formation of filamentous (F) actin compared to the other DPRs and GFP-alone control using a flow cytometry protocol for measuring filamentous (F) and globular (G) actin ratios **(Figure 3-4B)**. There was no apparent colocalization of actin to the DPRs, certainly not to punctate structures of the DPRs **(Figure 3-4C)**. However, it was clear that F actin was reduced in individual cells expressing the Arg-rich DPRs **(Figure 3-4C)**. Therefore, the results collectively suggested that the Arg-rich DPRs, either directly or indirectly, leads to destabilization of machinery involved in actin filament assembly without binding to the actin filaments directly.

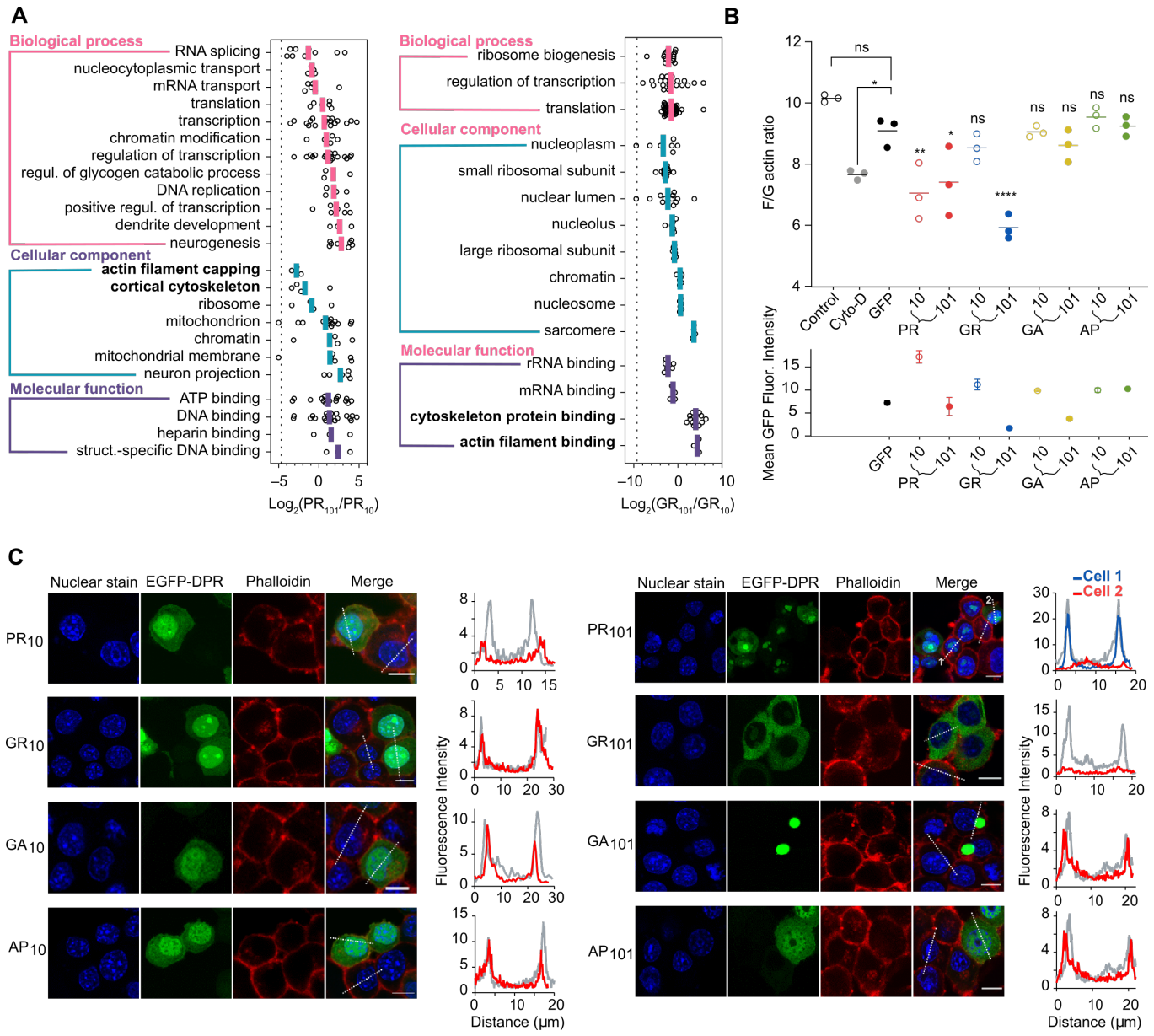

Figure 3-4 | Arg-content of the DPRs and subcellular location influence actin cytoskeleton assembly.

A, Gene Ontology (GO) enrichment analysis of proteins significantly associated with the long and short Arg-rich DPRs in the immunoprecipitation. GO terms are shown that contain at least three proteins, and where the $\log_2$ ratios of DPR length-enrichment values (DPR_{101}/DPR_{10}), shown as open circles, were deemed different to zero by a Student t-test and Benjamini-Hochberg-adjusted $p < 0.05$. The colored lines indicate the means of the ratios. The dashed line represents the abundance ratio of GFP, which therefore provides an estimate of the bias of each DPR in the immunoprecipitation. GO terms related to cytoskeleton are shown in bold. B, Flow cytometry analyses of population ratios (1×10^4 Neuro2a cells) of F-actin stain (stained with Phalloidin) and G-actin stain (stained with DNase I) 48 h after transfection with the GFP-tagged DPRs. Cyto-D cells represent 1 h Cytochalasin-D treatment in untransfected cells; Control indicates vehicle control treatment (0.1% DMSO) for Cyto-D sample. The lower graph shows the corresponding mean GFP fluorescence intensities of the samples in the upper graph. Data represent mean ± S.D., n = 3

with *p* values determined from a one-way ANOVA and Bonferroni's multiple comparisons post hoc-test with the GFP-alone sample as the control: *, $p < 0.05$; **, $p < 0.01$; ****, $p < 0.0001$; ns, $p > 0.05$. C, Intensity distance trajectories of F-actin (stained with Phalloidin) in Neuro2a cells transfected with GFP-tagged DPRs and nuclei stained with Hoechst 33258. Trajectories are shown on the cells as dotted white lines. Comparisons are shown in the graphs of representative cells (from over 50 measured cells) expressing GFP-tagged DPRs (red or blue lines) and untransfected cells in adjacent cells (grey lines). Scale Bars, 10 μm.

Next, we examined how proteome abundances changed in cells expressing the 100× DPRs. Because the DPRs were toxic and had variable expression levels, we sorted transfected cells into similar levels of expression by virtue of GFP levels before analysis. Cell lysates were collected and quantitatively analyzed by a reductive dimethyl labelling proteome analysis approach **(Figure 3-5A; Table 3-S2)**. The poly-GR and poly-PR DPRs resulted in the most changes to the proteome in accordance with them having more potent toxicity and promiscuous pattern of interactions. Poly-GR and poly-PR also had many overlapping GO terms annotated consistent with the Arg content driving a similar pathological consequence **(Figure 3-5B)**. Conversely, the non-toxic poly-AP resulted in few changes to the proteome abundance and poly-GA had a distinct signature consistent with a different mechanism of toxicity.

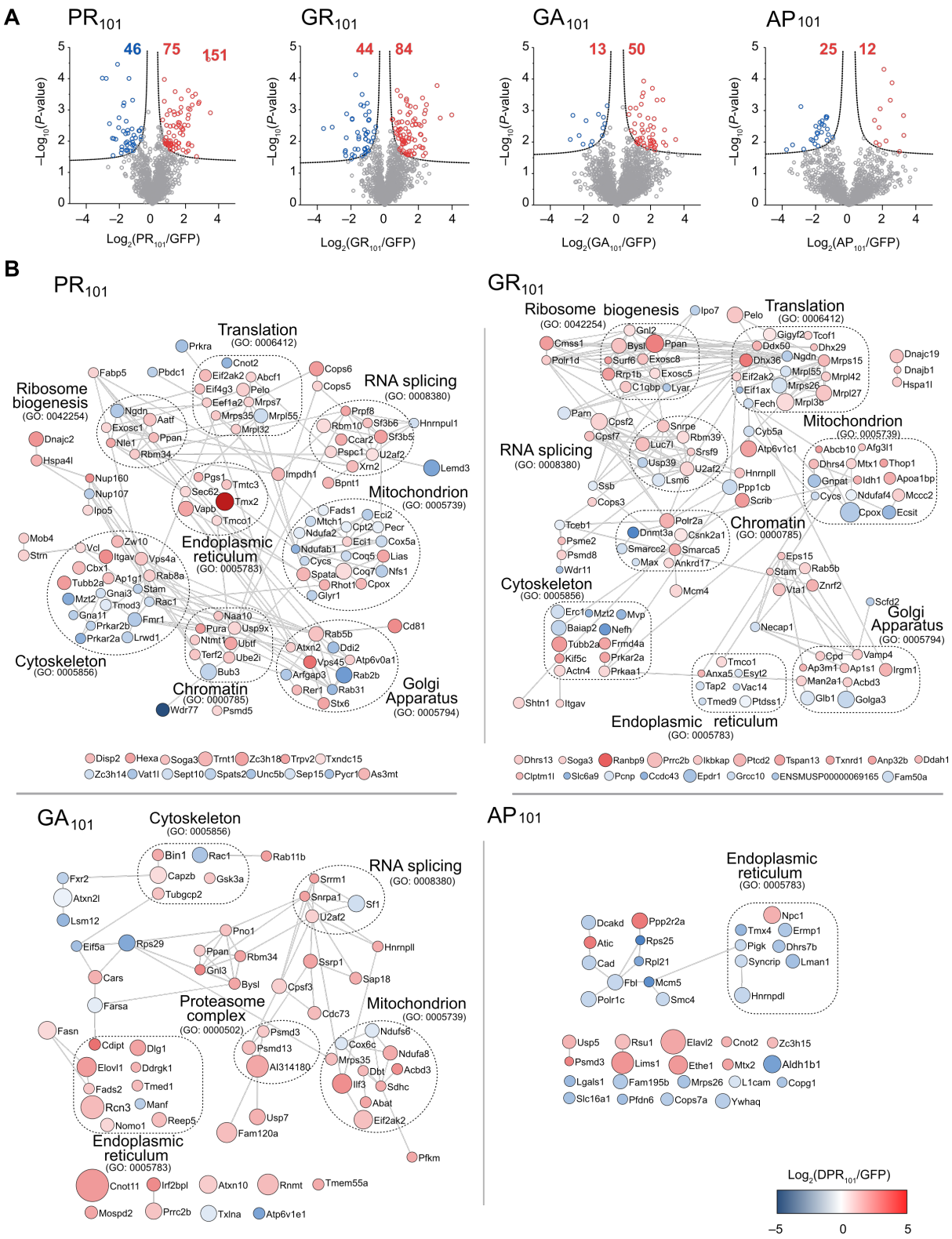

Figure 3-5 | Proteome abundance changes in response to DPR_{101} expression.

A, Shown are protein abundances in Neuro2a cells transfected with each GFP-DPR101 compared with GFP for 48 h, and sorted for matched GFP abundance by flow cytometry. n = 3 biological replicates. p values were determined with a two-sided Welch's t-Test. Significantly changed proteins (red and blue) were defined with FDR ≤ 0.01, S0 = 1 (dashed lines). The total number of proteins changed are indicated.

B, Protein interaction networks for proteins significantly changed in abundance, colour-coded to enrichment using STRING (v10) in Cytoscape (v3.6) at medium confidence. Size of nodes scales proportionally to -$\log_{10}$ (*P*-value). Selected significantly enriched GO terms (GOCC, GOPB, and UniProt keywords) are displayed.

3.5.5 Arginine-rich DPRs lead to altered arginine methylation patterns.

Arginine methylation has been reported to be abnormal in patients with C9ORF72 mutations, including the presence of arginine-dimethylated enriched inclusions (Chitiprolu et al., 2018; Sakae et al., 2018). Arginine residues are commonly methylated to regulate biological activity of many cellular processes and are important in particular in histones which are enriched GO terms in our datasets. In addition, abnormal histone methylation has previously been reported in a mouse model of PR_{50} (Zhang et al., 2019). Furthermore, prior work has suggested that arginine methylase PRMT5 is important in regulating stress granule function in C9ORF72 models of disease and methylates ALS-gene risk factor FUS (Chitiprolu et al., 2018). Here, GR_{101} interactome was found to be significantly enriched for 3 proteins in the Methylosome GO term (GO:0034709) including WDR77 and arginine methylases PRMT1 and PRMT5 **(Figure 3-2B)**. PRMT1 appears particularly important, accounting for 85% of the methylation activity in mammalian cells (Tang et al., 2000). Analysis of the interactome data showed that the Arg-rich DPRs were significantly enriching for other arginine-enriched proteins that are common substrates for methylases **(Figure 3-6A and 3-6B)**. This led us to hypothesize that the Arg-rich DPRs interfere with endogenous arginine methylation activity, that links to the mechanisms of ALS toxicity. To test this hypothesis, we examined the 101× DPRs affected proteome levels and corresponding levels of arginine methylation **(Table 3-S3)**. Because of our reductive dimethylation proteomics workflow, we could only observe a minor fraction of possible arginine methylation patterns that could exist. However, sufficient information was obtained to reveal that GR_{101} leads to a significantly lower level of arginine methylation relative to the GFP only control **(Figure 3-6C; Table 3-S3)**. These same proteins did not show a significant change in abundance **(Figure 3-6C)**. Examination of these peptides revealed that many come from Hnrnp family proteins including Hnrnpa1, Hnrnpab and Hnrnph1 **(Table 3-S3)**. Hnrnpa1 showed a statistically significant reduction in Arg-methylation in the GR_{101} treatment **(Figure 3-6D)**. Hnrnp family proteins are well-known substrates of PRMT family proteins (Dreyfuss et al., 1993). Mutations in Hnrnpa1 also cause ALS (Kim et al., 2013b). It should be noted that the abundances of methylated peptides seen in the whole proteome have not significantly changed in other tested DPRs; PR_{101} and GA_{101} compared to GFP ctrl indicating the specificity of arginine hypomethylation to GR_{101}. Altogether, these data suggest the possibility that the Arg-rich DPRs

act as substrate sinks of arginine methylases that therefore results in a broader deficiency in arginine methylase modification of endogenous proteins.

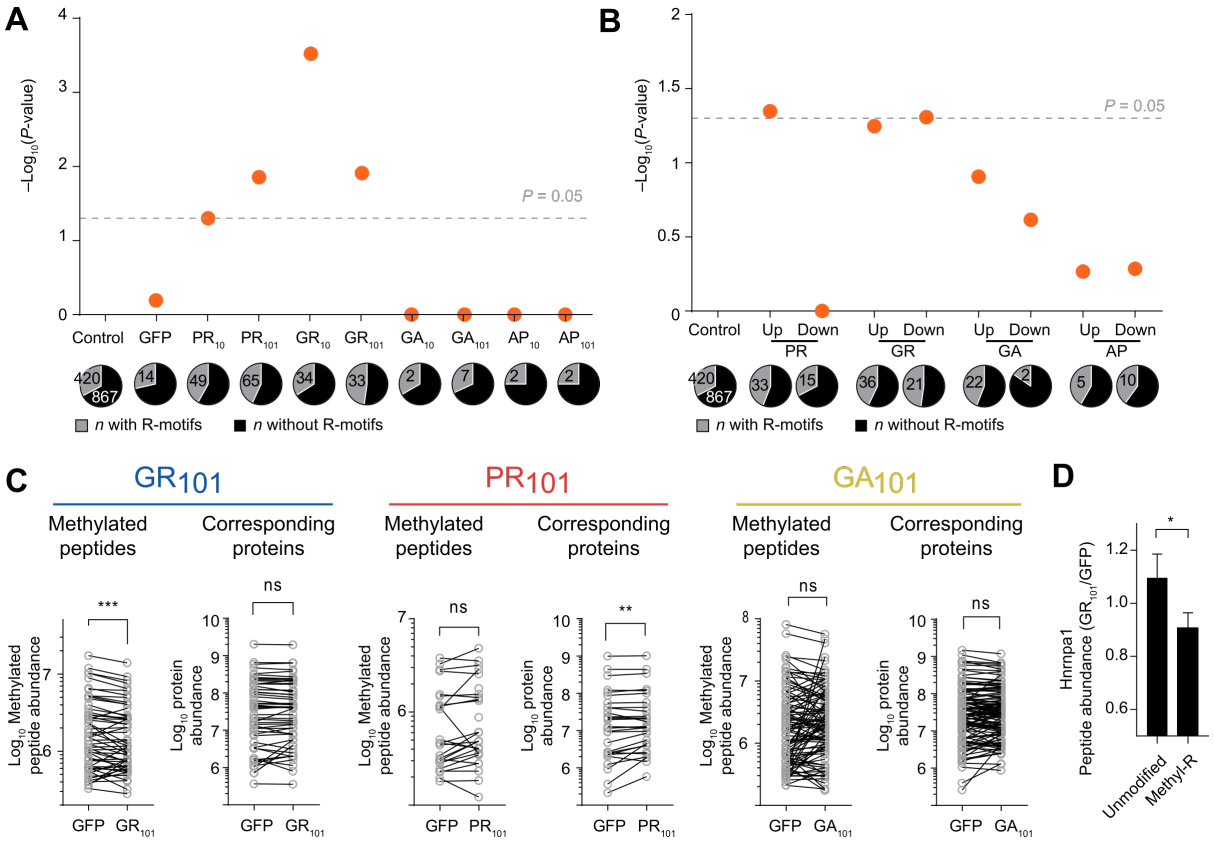

Figure 3-6 | Proteome hypomethylation upon GFP-GR_{101} expression.

A, Shown are *p* values of Fisher exact tests for enrichment of arginine motifs (Mitrea et al., 2016) in the DPR interactomes versus Control, which are the background proteins observed in the whole proteome analysis and which were not deemed significantly affected in abundance by DPR expression. Numbers of proteins are indicated. B, As for A, but for proteins seen significantly upregulated and downregulated in total protein lysate compared with Control. C, Abundances of methylated peptides seen in the whole proteome, and the corresponding protein abundances from which they derive. The data is plotted as matched-pairs of peptides (or proteins) with differences evaluated by 2-tailed Wilcoxon signed-rank test. *p* values are coded as ns > 0.05; **, $p < 0.01$; ***, $p < 0.001$. The mean difference in abundances of matched pairs of methylated peptides (GR_{101} - GFP) is –399,018. D, Shown are the means ± S.E. of peptide abundance ratios of Hnrnpa1 from the PR_{101} and GFP samples. Shown are unmodified (n = 24) and arg-methylated peptides (n = 5). Significance of difference was assessed with an unpaired t-test with Welch's correction. *p*-value is coded as *, $p < 0.05$.

3.6 Discussion

Here we show that the Arg in the Arg-rich DPRs promotes widespread interactions with the proteome relative to the other less toxic DPRs. These interactions centre on various hubs of cellular activity including translation, ribosome biogenesis, chromatin, mitochondria, cytoskeleton, RNA splicing and the methylosome. The effects may be driven by the valency of arginine as well as changes in the cellular localization, and potentially aggregation state, of the different lengths of DPRs. In contrast, the inert and non-toxic AP DPRs showed few interactions. GA also showed relative few interactors and unmasked a distinct interactome, consistent with a different mechanism of toxicity to the potently toxic Arg-rich DPRs.

Overall our data are consistent with the Arg-rich DPRs manifesting toxicity through multiple mechanisms. One mechanism appears to arise by the Arg-rich sequences causing translation to stall. It is notable that we couldn't identify ABCE1 among GR101 interactors despite it stalls the ribosome. This may be attributed to our finding that polyGR interactome is dominated by Arginine methylases and their substrates; a subset of proteins that are completely absent in polyPR interactome. Since we injected the same amount of peptide mix across all the interactive experiments. Our results suggested that lengths longer than at least 10 repeats are needed to induce stalling. Given that the ribosome exit tunnel holds around 33 amino acids and is lined in negative charges, a plausible explanation for stalling is that DPR lengths approaching 16–17 repeat lengths would be sufficient to fill this cavity volume and lock the peptide in place by electrostatic interactions. Previously it was suggested that a canonical polyadenylate tail on mRNA used to stall ribosomes is translated to lysine and that the poly-lysine sequence is recognized as aberrant by ribosomes and results in translation repression (Ito-Harashima et al., 2007). Additional experiments using repeating sequences of Lys and Arg in proteins both slowed translation, which supports the mechanism of electrostatic interactions jamming the emergent poly-basic chain in the negatively charged ribosome exit tunnel (Lu and Deutsch, 2008). Interestingly, antimicrobial peptides enriched for PR-containing motifs have been demonstrated to bind to the ribosomal exit tunnel and inhibit bacterial protein synthesis (Gagnon et al., 2016). This raises the possibility that emergent DPRs can also re-enter and plug the exit tunnel through electrostatic interactions. Further support for this additional mechanism comes from *in vitro* translation assays showing that poly-PR and poly-GR peptides formed insoluble complexes with mRNA, restricted the access of translation factors to mRNA, and blocked protein translation (Kanekura et al., 2016). This study showed that poly-PR and poly-GR inhibit protein translation by binding to the translational complex and ribosomal proteins, leading to neurotoxicity (Kanekura et al., 2016). It should be noted that the overall expression level for PR_{101} reporter are lower

compared to the remaining DPR sequences tested **(Figure 3-S3)**. This may be explained by two possible reasons. Firstly, only live cells are included in our analysis. Since, polyPR peptides are the most toxic DPR species, this increases the possibility of death of cells expressing high levels of the PR_{101} reporters and consequently being excluded from the analysis. Secondly, several studies have reported that poly-PR peptides inhibit the global protein translation (Kanekura et al., 2016; Lee et al., 2016a; Lin et al., 2016; Lopez-Gonzalez et al., 2016; Boeynaems et al., 2017; Yin et al., 2017; Hartmann et al., 2018; Shi et al., 2018). Additionally, antimicrobial peptides enriched for PR-containing motifs have been demonstrated to bind to the ribosomal exit tunnel and inhibit bacterial protein synthesis (Gagnon et al., 2016). It is therefore readily conceivable that emergent DPRs may also re-enter and plug the exit tunnel hence potentiating the translation inhibitory effect of polyPR peptides.

Our data also suggested that the Arg-rich DPRs impede assembly of the actin cytoskeleton. Recently it was observed that promoting actin filament assembly in cell models of ALS can alleviate defects in nuclear-cytoplasmic transport defects (Giampetruzzi et al., 2019a) which supports the conclusion that destabilization of the actin cytoskeleton is pertinent to disease pathomechanisms. We observed a significant reduction in F-actin in cells with Arg-rich DPRs. Whether this effect is a consequence of other cellular effects, such as hypomethylation or ribosome stalling is unclear. Interestingly it has been reported that arginine methylation of an arginine methylase (PRMT2) regulates the activity of actin nucleator protein Cobl, which suggests a possible role for arginine methylation defects being an upstream mediator of effects on the actin cytoskeleton (Hou et al., 2018). In addition, it is thought that dysregulation of actin is a key process in ALS (Hensel and Claus, 2018). Of note is that mutations in profilin-1 protein, which mediates the conversion of G-actin to F-actin, is linked to ALS (Wu et al., 2012). Other cytoskeletal genes have also been linked to ALS including TUBA4A and DCTN1 and KIF5A (Taylor et al., 2016a; Nicolas et al., 2018).

Our mechanism of toxicity attributable to the Arg-rich DPRs was through a broader hypomethylation of the proteome. Previous studies have shown that PRMT1 colocalizes with GR and PR in a Drosophila model and that knockdown of PRMT family members enhanced toxicity (Boeynaems et al., 2016). It was also found that C9ORF72-related brain samples had abundant methylated inclusions (Boeynaems et al., 2016). Thus, the data raises the possibility that the Arg-rich motifs attract and alter the endogenous methylation activity leading to pathological outcomes. The substrate of PRMT family proteins contain glycine- and arginine-rich (GAR) sequences that include multiple arginines in RGG or RXR contexts, which bear resemblance to the Arg-rich DPRs (Blanc and Richard, 2017). It follows that many of the key pathways seen in

our dataset are affected by altered arginine methylase activity – including proteins that are methylated for functional regulation such as histones, proteins involved in mRNA splicing, and ribosomes (Chang et al., 1976; Yang and Bedford, 2013).

Also of note is that other genes that when mutated are risk factors for ALS have activity regulated by arginine methylation and show abnormal methylation patterns in disease. In particular, FUS has been reported to interact with PRMT1 and PRMT8 and undergo asymmetric dimethylation in cultured cells (Scaramuzzino et al., 2013). Importantly, PRMT1 and PRMT8 localized to mutant FUS-positive inclusion bodies in ALS (Scaramuzzino et al., 2013). It has also been reported that arginine methylation modulates the nuclear import of FUS and inclusions in ALS-FUS patients contain methylated FUS (Dormann et al., 2012). We observed the hypomethylation of Hnrnpa1 caused by the Arg-rich DPRs, which indicates a possible link to ALS arising from distinct gene mutations. Hnrnpa1, as well as the Arg-rich DPRs and other ALS-associated proteins, are known to form molecular condensates by phase separation (Taylor et al., 2016a; Boeynaems et al., 2017). Arginine methylation of Hnrnpa1 reduces its ability to phase separate suggesting that an imbalance in molecular condensate mechanisms contributes to the pathogenic response (Ryan et al., 2018). PR DPRs can also promote the aggregation of ALS-related proteins containing prion-like domains, that are involved in mediating phase separation into molecular condensates (Boeynaems et al., 2017). Hence our data indicate a possible convergence of multipronged mechanisms involving methylation, phase separation and cytoskeleton as important contributors to the toxicity of the Arg-rich DPRs.

3.7 Acknowledgements

This work was funded by grants to DMH (National Health and Medical Research Council APP1161803 and Motor Neuron Disease Research Institute, Australia small grant) and to DMH and GER (Australian Research Council DP170103093). MR acknowledges support from an Australian Government Research Training Program (RTP) Scholarship and an Egyptian Ministry of Higher Education and Scientific Research PhD scholarship.

3.8 Contribution to this publication

As first author of this publication, I had major contributions to the experimental work and study design. In detail, I performed all experiments shown in this publication except of the ribosome stalling assay for mutant Huntigton proteins presented in Figure 2-4D. A.R.O., J.C.D. and D.C. cloned exon 1 of Huntington (Httex1) with different polyQ expansions (25Q, 72Q and

97Q) into the dual-fluorescence translation stall reporter and performed the flow cytometry analysis presented in Figure 2-4D.

3.9 Supplemental information

Table 3-S1: Quantitative assessment of DPR pathology in mouse neuroblastoma cell line. Relates to Figure 3-1.

Table 3-S2. Protein interactors to 10× and 101× DPRs. Relates to Figure 3-2 & Figure 3-S2. Separate file accessible at:

https://www.mcponline.org/highwire/filestream/54306/field_highwire_adjunct_files/1/157401_1_supp_477899_q5xsfk.xlsx

Table 3-S3. Cellular abundances of proteins caused by 101× DPR expression. Relates to Figure 3-5. Separate file accessible at:

https://www.mcponline.org/highwire/filestream/54306/field_highwire_adjunct_files/2/157401_1_supp_477900_q5xsfk.xlsx

Table 3-S4: Arginine methylated proteome quantitation by 101× DPR expression. Relates to Figure 3-6. Separate file accessible at:

https://www.mcponline.org/highwire/filestream/54306/field_highwire_adjunct_files/3/157401_1_supp_477901_q5xsfl.xlsx

Table 3-S1: Quantitative assessment of DPR pathology in mouse neuroblastoma cell line*.

DPR species	Souluble (%)	Cytoplasmic inclusion (%)	Intranuclear inclusion (%)
GA_{101}	58.1 ± 7.9	39.6 ± 7.7	2.3 ± 0.3
PR_{101}	55.9 ± 0.16	----------	44.1 ± 0.16
GR_{101}	76.3 ± 7.8	----------	23.7 ± 7.8
AP_{101}	100 ± 0.0	----------	----------
PR_{10}	14.9 ± 2.4	----------	87.0 ± 9.4
GR_{10}	20.1 ± 4.1	----------	78.9 ± 3.4
GA_{10}	100 ± 0.0	----------	----------
AP_{10}	100 ± 0.0	----------	----------

* The numbers represent the percentage of cell having specific phenotype. Three confocal micrographs of Neuro2a cells expressing DPR species were analysed.

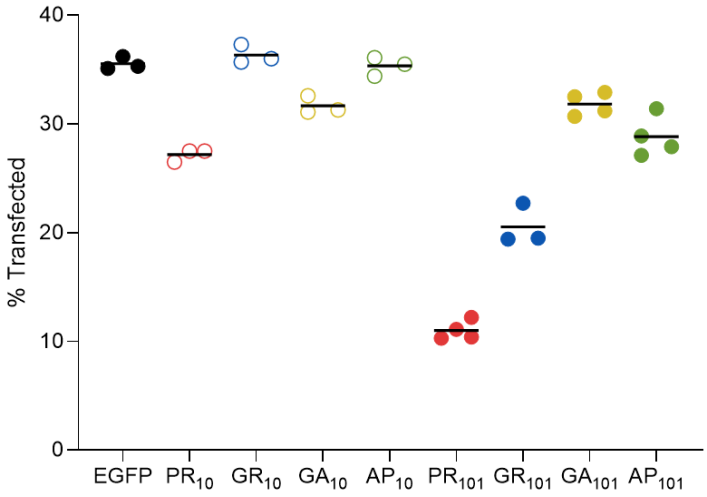

Figure 3-S1 | Transfection efficiency of DPRs as fusions to GFP. Related to Figure 3-1. Neuro2a cells were transfected with GFP-tagged DPRs (or EGFP alone) as labelled. Shown are percentages of cells in each sample with higher GFP fluorescence than background levels. The background levels were derived by reference to untransfected cells. Data were determined by flow cytometry analysis. Shown are biological replicates (circles) and means (bars).

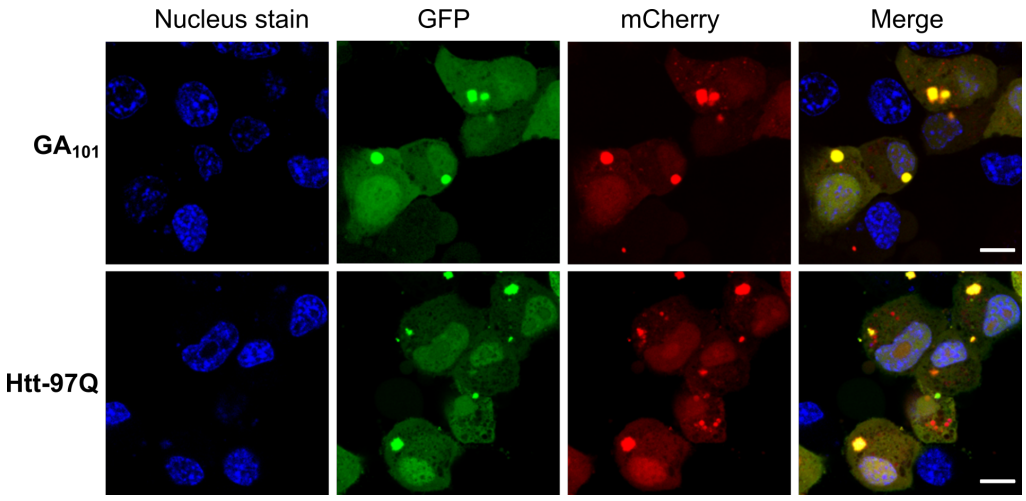

Figure 3-S2 | Expression of DPRs in the ribosome stall reporter constructs. Relates to Figure 3-3. Shown are confocal micrographs of Neuro2a cells overexpressing the reporter construct containing either GA101 (top panel) or Httex1 with 97Q repeats (bottom panel) fixed 48 hr post-transfection and stained with Hoechst 33258 (blue) to visualize nuclei. Scale bars represent 10 μm.

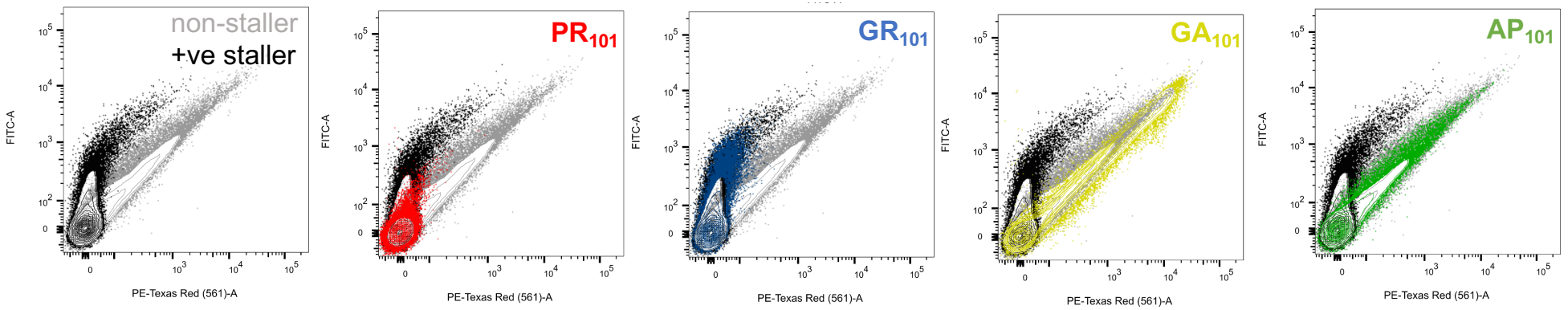

Figure 3-S3 | Flow cytograms of Neuro2a cells 48 h after transfection with different DPRs inserted into the ribosme stalling reporter. Relates to Figure 3-3.

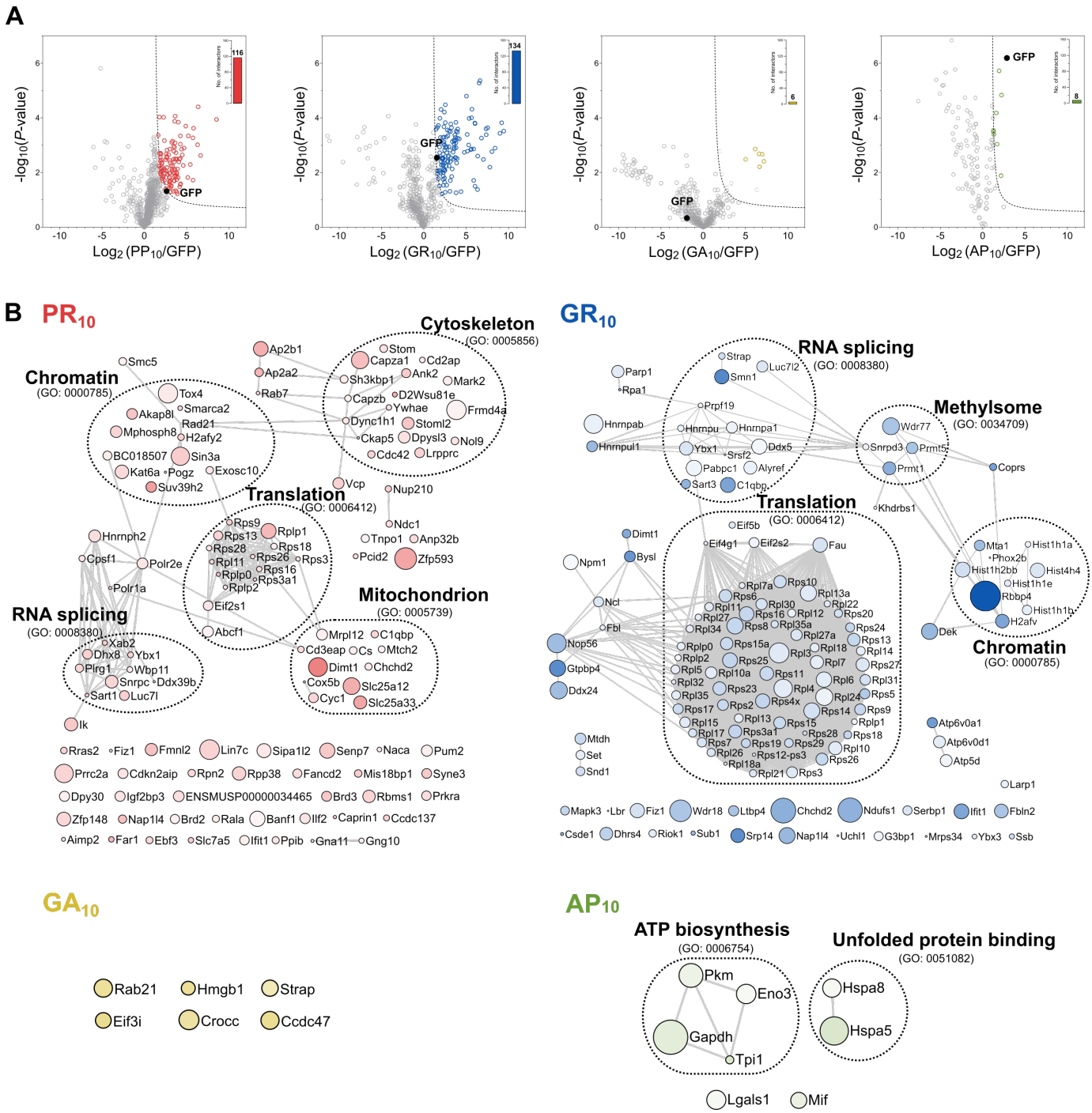

Figure 3-S4 | Interactome analysis of the DPR10 variants. Relates to Figure 3-2. A. Volcano Plots of each DPR_{10}-GFP versus GFP-only control of quantitative proteomics analysis of GFP-Trap immunoprecipitates of DPR_{10}- GFP transfected in Neuro2a cells harvested 48h after transfection. Significant binders (shown in colored circles) were classified with False Discovery Rate of ≤ 0.01 (dotted lines). The number of interactors is indicated. B. STRING (v10) interaction maps for proteins significantly enriched with confidence set at 0.9 (highest stringency). Circle sizes are proportional to –log10 (*P*-value). The color intensity is proportional to the log (fold change). Selected significantly enriched GO terms (GOCC, GOPB, and UniProt keywords) are displayed (with FDR cutoff *P* value < 0.05)

CHAPTER 4

Inclusion bodies formed by polyglutamine and poly(glycine-alanine) are enriched with distinct proteomes but converge in proteins that are risk factors for disease and involved in protein degradation

4 CHAPTER 4: Inclusion bodies formed by polyglutamine and poly(glycine-alanine) are enriched with distinct proteomes but converge in proteins that are risk factors for disease and involved in protein degradation*

4.1 Abstract

Poly(glycine-alanine) (polyGA) is one of the dipolypeptides expressed in Motor Neuron Disease caused by C9ORF72 mutations and accumulates as inclusion bodies in the brain of patients. Superficially these inclusions are similar to those formed by polyglutamine (polyQ) in Huntington's disease and both have been reported to form an amyloid-like structure suggesting they might aggregate via similar mechanisms to confer cellular dysfunction similarly. Here we investigated which endogenous proteins were enriched in these inclusions and whether aggregation-prone lengths of polyQ (Q_{97}), in context of Huntingtin exon 1, shared similar patterns to aggregation-prone lengths of polyGA (101_{GA}). When co-expressed in the same cell, $polyGA_{101}$ and $HttQ_{97}$ inclusions adopted distinct phases with no overlap suggesting different endogenous proteins would be enriched. Proteomic analyses indeed yielded distinct sets of endogenous proteins recruited into the inclusion types. The proteosome, microtubules, TriC chaperones, and translational machinery were enriched in polyGA aggregates, whereas Dnaj chaperones, nuclear envelope and RNA splicing proteins were enriched in polyQ aggregates. Both structures revealed a synergy of degradation machinery including proteins in the polyQ aggregates that are risk factors for other neurodegenerative diseases involving protein aggregation when mutated, which suggests a convergence point in the pathomechanisms of these diseases.

4.2 Introduction

The formation of protein inclusions is a hallmark of many neurodegenerative diseases. In Huntington's disease, amino-terminal fragments of mutant Huntingtin (Htt) protein aggregate into intraneuronal inclusions (DiFiglia et al., 1997). The aggregation of mutant Htt is triggered by an

* This chapter represents the following manuscript:

Radwan M, Lilley JD, Ang CS, Reid GE, Hatters DM (2020). Inclusion bodies formed by polyglutamine and poly(glycine-alanine) are enriched with distinct proteomes but converge in proteins that are risk factors for disease and involved in protein degradation. bioRxiv 2020.05.04.076547; doi: https://doi.org/10.1101/2020.05.04.076547.

abnormally expanded polyglutamine (polyQ) sequence encoded in exon 1 that arises by CAG trinucleotide repeat expansions (MacDonald et al., 1993; Scherzinger et al., 1999). Long polyglutamine sequences form cytoplasmic or nuclear inclusions in animal and mouse models and are associated with a pathological cascade of events (reviewed in (Cox et al., 2018)).

In motor neuron disease caused by C9ORF72 GGGGCC hexanucleotide repeat expansion mutations, protein inclusions arise from the aggregation of dipolydipeptides (DPRs) expressed abnormally from the expanded GGGGCC hexanucleotide repeat sequence. 5 different DPRs are expressed, namely dipeptide polymers of glycine-alanine (polyGA), proline-arginine (polyPR), glycine-arginine (polyGR), proline-alanine (polyPA), and proline-glycine (PolyPG). Of these polyPR and polyGR are profoundly toxic when expressed in cell culture and animal models targeting mechanisms in ribosome biogenesis, translation, actin cytoskeleton among others (Kwon et al., 2014; Zhang et al., 2015; Kanekura et al., 2016; Lee et al., 2016a; Freibaum and Taylor, 2017; Zhang et al., 2018; Radwan et al., 2020). PolyGA appears less toxic than the others although has been reported in some models to confer toxicity (Al-Sarraj et al., 2011; Ash et al., 2013; Bigio et al., 2013; Gendron et al., 2013; Mann et al., 2013; Mori et al., 2013a; Zu et al., 2013; May et al., 2014; Mizielinska and Isaacs, 2014). We previously reported polyGA to be mildly toxic to cultured Neuro2a cells and to induce a distinct network of proteome changes that occur compared to the Arg-rich DPRs (Radwan et al., 2020). We also noted a distinction to the other DPRs in forming large inclusions that morphologically are similar to the inclusions formed by polyQ. Furthermore, it has been reported that polyGA inclusions are, like polyQ, SDS-insoluble and amyloid-like (May et al., 2014; Chang et al., 2016). PolyGA is also more widespread in MND-patient brain tissue compared to the other DPRs (Mackenzie et al., 2015).

Here we investigated the proteinaceous composition of polyGA inclusions and compared the profile quantitatively to inclusions of polyQ using a Huntingtin exon 1 model ($Httex1Q_{97}$) in a mouse neuroblastoma cell culture model using a novel proteomics-based approach to enrich the inclusions from cells under mild lysis conditions. We find distinct recruitment patterns to each inclusion type and also some similarities in biological mechanisms pertaining to degradation and a convergent pathomechanism for neurodegenerative diseases involving inappropriate protein aggregation.

4.3 Results and discussion

Previously we found that $polyGA_{101}$ as a fusion to GFP formed cytosolic inclusions in Neuro2a cells when transiently transfected. To further investigate the rate that this occurs, we used live cell imaging to track cells from 24 h after transfection onwards. Of cells with detectable

levels of expression by 24 h, almost all of them had formed inclusions by 60 h **(Figure 4-1A)**. This was faster than comparable experiments with Httex1Q_{97} as a fusion to mCherry, which is well known to extensively aggregate in cell culture (Arrasate et al., 2004; Ormsby et al., 2013) **(Figure 4-1A)**. We noted that some lower-expressing cells showed detectable expression of polyGA only after 24 h (which we did not track) and that these were likely to form aggregates more slowly. When we co-expressed Httex1Q_{97} as a fusion to mCherry, we found the polyGA_{101} and Httex1Q_{97} formed discrete inclusions in the same cell with no apparent colocalization **(Figure 4-1B)**. This suggested that any concomitant co-aggregation patterns that arise with endogenous proteins may involve distinct proteins.

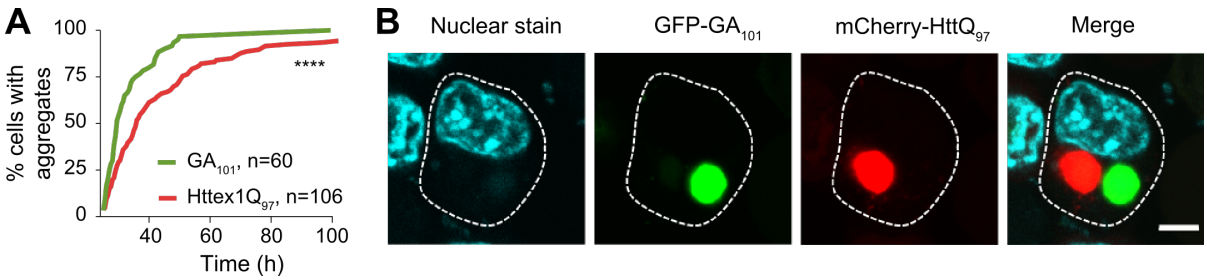

Figure 4-1 | Httex1Q_{97} and polyGA_{101} rapidly form distinct inclusions in neuro2a cells.

A. Kaplan-Meier curves of polyGA_{101}- or Httex1Q_{97} expressing cells that form visible inclusions as assessed by longitudinal imaging. Cells were tracked from 24 h post-transfection. *P*-values correspond to log-rank (Mantel-Cox) test. B. Confocal micrographs of neuro2a cells co-expressing GFP-tagged GA_{101} and mCherry-tagged Httex1Q_{97}, fixed 24 hr post-transfection and stained with Hoechst33258 (cyan) to visualize nuclei. Scale bar represents 5 µm.

To investigate these potential differences, pellets recovered from lysates of neuro2a cells expressing GFP-tagged Httex1Q_{97} or GFP-tagged polyGA_{101} were sorted to purify the aggregates using flow cytometry via monitoring the GFP fluorescence. Quantitative proteomics, by way of dimethyl isotope labelling, was used to define the proteins enriched in each aggregate class after normalization to the total mass of protein. We observed 737 proteins. Of these 69 were significantly enriched in polyGA inclusions (3 replicates, a permutation-based FDR cut-off of 5% and S0 of 0.1) and 48 were enriched in Httex1Q_{97} **(Table 4-S1, Table 4-S2 and Figure 4-2)**.

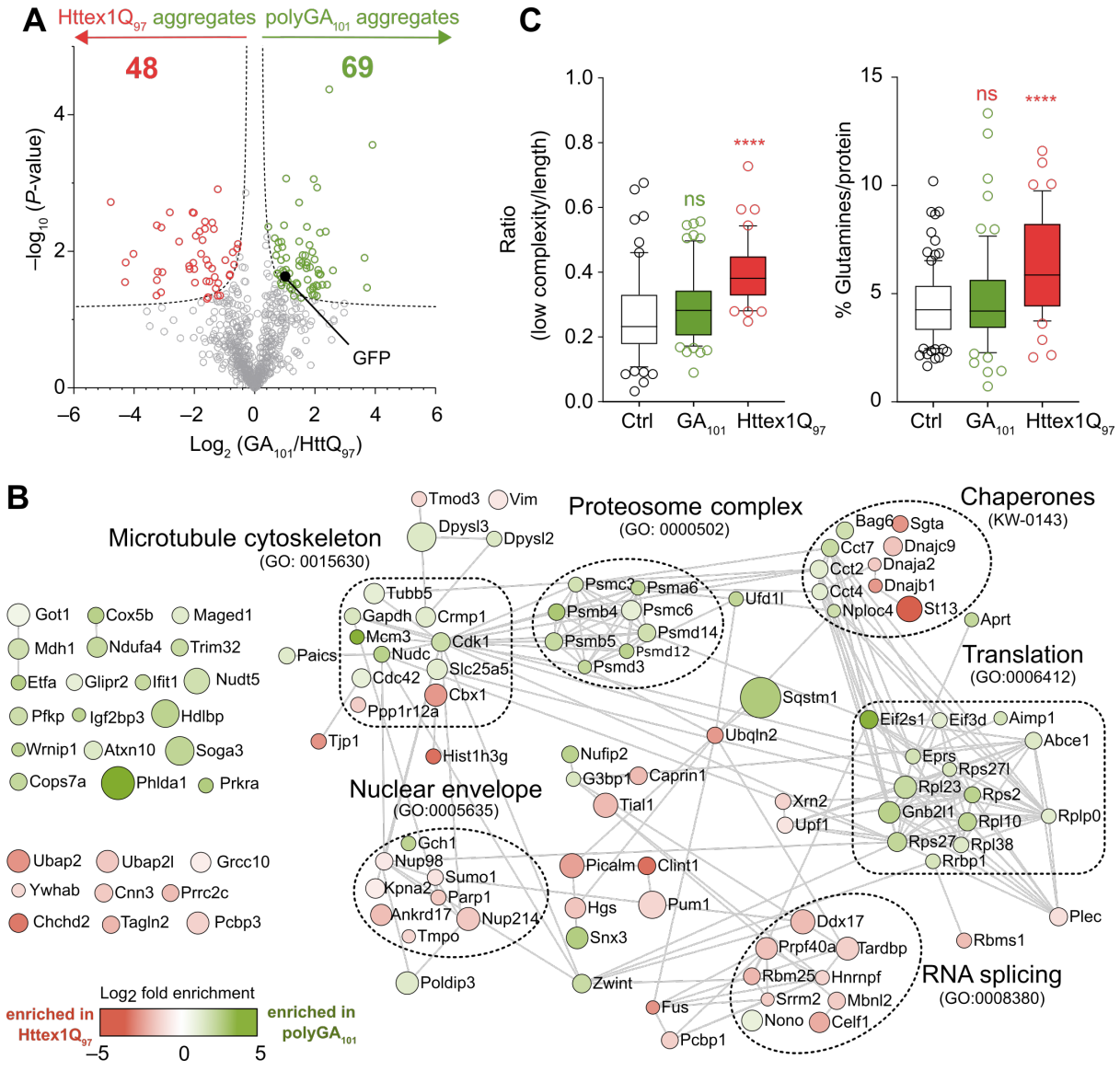

Figure 4-2 | Proteome recruitment patterns to polyGA$_{101}$ and Httex1Q$_{97}$ inclusions.

A. Volcano plot of proteins identified in the inclusions. *P*-values were calculated by a two-sided one samples t-test with null hypothesis that abundances were unchanged and the Log_2 ratio was equal to 0. Proteins meeting stringency thresholds (hyperbolic curves, FDR≤0.05, S_0=0.1) are shown as colored circles. B. STRING interaction maps (v.11) determined in Cytoscape (v3.7) for proteins significantly enriched in the inclusions (the full list of proteins are in **Table 4-S2**). The analysis was done at the highest confidence setting. Each protein was represented by a colored circle sized proportionally to -log_{10} (*P*-value). The color scale represents logarithm of fold change. Selected significantly enriched GO terms (GOCC, GOPB, and UniProt keywords) are displayed (Full terms are shown in **Table 4-S3**). C. Analysis of enriched proteomes for low-complexity regions (IUPred-L) and high glutamine content. Significance of difference was assessed against a control dataset of random mouse proteins **(Table 4-S1)** with the Mann–Whitney–Wilcoxon test. Whiskers extend from 10 to 90%.

We observed notable features in this dataset consistent with known pathological markers of polyGA inclusions. Namely in C9ORF72 mediated MND, a subset of inclusions is non-reactive to TDP-43 (Cruts et al., 2013). In most other forms of MND, TDP-43-reactive inclusions are a key pathological signature of neurons in disease (Neumann et al., 2006). These TDP-43 negative inclusions were previously found to be immunoreactive for polyGA, suggesting they form by polyGA aggregation (DeJesus-Hernandez et al., 2011; Renton et al., 2011). We observed a lack of TDP43 in the polyGA inclusions by virtue of enrichment in the Httex1Q_{97} inclusions **(Table 4-1)**. In addition, the TDP43-negative inclusions seen *in vivo* are immunoreactive to p62 (Pikkarainen et al., 2010) and lack immunoreactivity to FUS, optineurin, alpha-internexin and neurofilament (King et al., 2009; King et al., 2011). In our data p62 (also called sequestesome 1) is one of the most enriched proteins in polyGA inclusions, whereas Fus appeared excluded by virtue of its enrichment in Httex1Q_{97} inclusions, which has been observed previously in cell models of polyQ aggregation and human pathology (Doi et al., 2008; Doi et al., 2010; Wear et al., 2015; Ramdzan et al., 2017). Hence these data point to the cell model of polyGA inclusions mimicking the process of aggregation and recruitment seen *in vivo* and also providing specificity of co-recruitment relative to Httex1Q_{97}.

Analysis of the differences is shown visually in **Figure 4-2B** by a (STRING) protein-protein interaction map and annotation to functional networks. Overall both inclusions yielded an enrichment for gene ontology and KEGG networks of microtubule cytoskeleton, proteasome complex, chaperones, RNA splicing and nuclear envelope **(Figure 4-2B; Table 4-S3)**. These findings are in accordance with prior findings that protein aggregation impacts these biological processes and in particular involvement in machinery for their clearance and degradation (Ross and Poirier, 2005; Woerner et al., 2016; Boland et al., 2018; Nussbacher et al., 2019).

In addition, the data points more directly to proteins and genes implicated in MND phenotype and mechanisms. Phlda1 was one of the proteins enriched in polyGA inclusions. Previously it was found that Phlda1 was upregulated in Fus-mutant motor neurons and that this was an adaptive response to protect against apoptosis (De Santis et al., 2017). Phlda1 was also observed upregulated in sporadic MND fibroblasts treated to stress compared to controls (Raman et al., 2015). Nudt5 was also found enriched in inclusions, and expression of this gene was significantly increased in motor neurons derived from induced pluripotent cells from MND patients over controls (Egawa et al., 2012). Another protein of note enriched in the polyGA aggregates was Dpysl3. A missense mutation that has been linked to MND risk in the French population and in culture expression of the mutation leads to shorted neuronal survival (Blasco

et al., 2013). Hence it remains plausible that co-aggregation of these proteins into polyGA inclusions sequesters their activity and renders cells less resilient to stress triggers.

The Httex1Q_{97}-enriched proteome also yielded noteworthy findings. Previously it was found that polyQ can preferentially co-recruit proteins containing LCRs and proteins enriched in glutamine (Wear et al., 2015; Ramdzan et al., 2017). These patterns were also observed in our data **(Figure 4-2B)**. However, polyGA did not show these enrichment patterns, indicative of specificity for polyQ in recruiting LCRs and Q-rich proteins.

Table 4-1 | Proteins enriched in inclusions of polyGA$_{101}$ and Httex1Q$_{97}$*

Enriched in polyGA$_{101}$		
Description	**Gene id**	**Log$_2$ enrichment (Mean ± SD)**
Pleckstrin homology-like domain family a member 1	Phlda1	3.91 ±1.18
DNA replication licensing factor mcm3	Mcm3	3.73 ±10.23
Eukaryotic translation initiation factor 2 subunit 1	Eif2s1	3.65 ±5.69
Proteasome subunit beta type-4	Psmb4	2.6 ±2.39
Sequestosome-1 (p62)	**Sqstm1**	**2.47 ±0.11**
Interferon-inducible double-stranded RNA-dependent protein kinase activator a	Prkra	2.4 ±2.63
Sorting nexin-3	Snx3	2.36 ±1.05
Electron transfer flavoprotein subunit alpha, mitochondrial	Etfa	2.24 ±2.41
Cytochrome c oxidase subunit 5b, mitochondrial	Cox5b	2.19 ±2.64
Nuclear migration protein Nudc	Nudc	2.17 ±1.67
Receptor of activated protein c kinase 1	Rack1	2.17 ±0.91
40s ribosomal protein s2	Rps2	2.12 ±1.85
Nuclear fragile x mental retardation-interacting protein 2	Nufip2	2.11 ±1.41
26s proteasome non-ATPase regulatory subunit 12	Psmd12	2.09 ±2.28
Vigilin	Hdlbp	2.07 ±0.36
Insulin-like growth factor 2 mRNA-binding protein 3	Igf2bp3	2.06 ±2.16
Gtp cyclohydrolase 1	Gch1	2.04 ±2.12
60s ribosomal protein l10	Rpl10	1.99 ±1.03
ATPase wrnip1	Wrnip1	1.98 ±2.39

Protein soga3	Soga3	1.96 ±0.27
Ubiquitin fusion degradation protein 1 homolog	Ufd1l	1.95 ±1.62
Proteasome subunit alpha type-6	Psma6	1.93 ±1.54
26s proteasome non-ATPase regulatory subunit 3	Psmd3	1.91 ±2.13
40s ribosomal protein s27	Rps27	1.91 ±0.84
Adenine phosphoribosyltransferase	Aprt	1.84 ±1.62
Cop9 signalosome complex subunit 7a	Cops7a	1.83 ±1.14
Cytochrome c oxidase subunit ndufa4	Ndufa4	1.82 ±0.68
Interferon-induced protein with tetratricopeptide repeats 1	Ifit1	1.76 ±1.24
Proteasome subunit beta type-5	Psmb5	1.74 ±0.84
60s ribosomal protein l23	Rpl23	1.73 ±0.47
E3 ubiquitin-protein ligase trim32	Trim32	1.72 ±1
T-complex protein 1 subunit eta	Cct7	1.72 ±0.86
Zw10 interactor	Zwint	1.68 ±0.74
Cyclin-dependent kinase 1	Cdk1	1.59 ±0.65
Atp-dependent 6-phosphofructokinase, platelet type	Pfkp	1.57 ±0.67
Nuclear protein localization protein 4 homolog	Nploc4	1.54 ±1.38
Large proline-rich protein bag6	Bag6	1.48 ±0.66
26s proteasome non-ATPase regulatory subunit 14	Psmd14	1.48 ±0.56
Malate dehydrogenase, cytoplasmic	Mdh1	1.48 ±0.43
Adp-sugar pyrophosphatase	Nudt5	1.47 ±0.21
26s protease regulatory subunit 6a	Psmc3	1.42 ±0.87
Bifunctional glutamate/proline--trna ligase	Eprs	1.39 ±0.81
Aminoacyl tRNA synthase complex-interacting multifunctional protein 1	Aimp1	1.3 ±0.91
Ribosome-binding protein 1	Rrbp1	1.26 ±0.71
40s ribosomal protein s27-like	Rps27l	1.23 ±0.68
Ras GTPase-activating protein-binding protein 1	G3bp1	1.18 ±0.66
Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	1.15 ±0.48
Dihydropyrimidinase-related protein 2	Dpysl2	1.09 ±0.45
Ataxin-10	Atxn10	1.05 ±0.29
Dihydropyrimidinase-related protein 3	**Dpysl3**	**1.05 ±0.08**
ATP-binding cassette sub-family e member 1	Abce1	1.04 ±0.34
60s ribosomal protein l38	Rpl38	1.03 ±0.39
Green fluorescent protein	Gfp	1.01 ±0.4

Multifunctional protein ade2	Paics	0.98 ±0.35
Polymerase delta-interacting protein 3	Poldip3	0.97 ±0.15
Melanoma-associated antigen d1	Maged1	0.94 ±0.31
Dihydropyrimidinase-related protein 1	Crmp1	0.92 ±0.19
60s acidic ribosomal protein p0	Rplp0	0.89 ±0.35
Adp/atp translocase 2	Slc25a5	0.89 ±0.17
T-complex protein 1 subunit beta	Cct2	0.88 ±0.22
Eukaryotic translation initiation factor 3 subunit d	Eif3d	0.87 ±0.32
T-complex protein 1 subunit delta	Cct4	0.84 ±0.27
Tubulin beta-5 chain	Tubb5	0.79 ±0.14
26s protease regulatory subunit 10b	Psmc6	0.76 ±0.15
Golgi-associated plant pathogenesis-related protein 1	Glipr2	0.75 ±0.22
IgE-binding protein	Iap	0.75 ±0.22
Cell division control protein 42 homolog	Cdc42	0.71 ±0.15
Non-POU domain-containing octamer-binding protein	Nono	0.67 ±0.1
Aspartate aminotransferase, cytoplasmic	Got1	0.45 ±0.05
Enriched in Httex1Q$_{97}$		
Description	**Gene id**	**Log$_2$ enrichment (Mean ± SD)**
Hsc70-interacting protein	St13	4.78 ±0.36
Histone h3.1	Hist1h3a	4.3 ±1.28
Clathrin interactor 1	Clint1	4.28 ±0.91
Hsc70-interacting protein	St13	4.78 ±0.36
Histone h3.1	Hist1h3a	4.3 ±1.28
Clathrin interactor 1	Clint1	4.28 ±0.91
Coiled-coil-helix-coiled-coil-helix domain-containing protein 2	Chchd2	4.02 ±0.73
RNA-binding protein FUS	**Fus**	**3.26 ±1.23**
Tight junction protein ZO-1	Tjp1	3.25 ±0.93
Ubiquitin-associated protein 2	Ubap2	3.24 ±0.36
Small glutamine-rich tetratricopeptide repeat-containing protein alpha	Sgta	3.22 ±0.8
Dnaj homolog subfamily b member 1	Dnajb1	3.11 ±1.11

Chromobox protein homolog 1	Cbx1	3.1 ±0.36
Ubiquilin-2	**Ubqln2**	**3.06 ±0.77**
Phosphatidylinositol-binding clathrin assembly protein	**Picalm**	**2.81 ±0.25**
Cugbp elav-like family member 1	Celf1	2.51 ±0.37
Transgelin-2	Tagln2	2.17 ±0.49
Rna-binding protein 25	Rbm25	2.16 ±0.51
Nucleolysin tiar	Tial1	2.06 ±0.19
Caprin-1	Caprin1	2.04 ±0.43
Probable atp-dependent rna helicase ddx17	Ddx17	2.02 ±0.18
Protein prrc2c	Prrc2c	2.01 ±0.48
Rna-binding motif, single-stranded-interacting protein 1	Rbms1	1.98 ±0.13
Ankyrin repeat domain-containing protein 17	Ankrd17	1.95 ±0.26
Pre-mRNA-processing factor 40 homolog a	Prpf40a	1.82 ±0.23
Dnaj homolog subfamily c member 9	Dnajc9	1.78 ±0.32
Hepatocyte growth factor-regulated tyrosine kinase substrate	Hgs	1.73 ±0.25
Ubiquitin-associated protein 2-like	Ubap2l	1.65 ±0.2
Nuclear pore complex protein nup214	Nup214	1.65 ±0.17
Poly [adp-ribose] polymerase 1	Parp1	1.6 ±0.47
Calponin-3	Cnn3	1.6 ±0.32
Dnaj homolog subfamily a member 2	Dnaja2	1.58 ±0.63
Serine/arginine repetitive matrix protein 2	Srrm2	1.55 ±0.6
Muscleblind-like protein 2	Mbnl2	1.54 ±0.34
Protein phosphatase 1 regulatory subunit 12a	Ppp1r12a	1.5 ±0.44
Poly(rc)-binding protein 1	Pcbp1	1.41 ±0.33
Tar DNA-binding protein 43	**Tardbp**	**1.41 ±0.15**
Poly(rc)-binding protein 3	Pcbp3	1.35 ±0.16
5'-3' exoribonuclease 2	Xrn2	1.32 ±0.36
Heterogeneous nuclear ribonucleoprotein f	Hnrnpf	1.29 ±0.43
Lamina-associated polypeptide 2, isoforms alpha/zeta	Tmpo	1.22 ±0.46
Pumilio homolog 1	Pum1	1.22 ±0.07
Tropomodulin-3	Tmod3	1.21 ±0.37
14-3-3 protein beta/alpha	Ywhab	1.15 ±0.43
Plectin	Plec	0.97 ±0.2

Small ubiquitin-related modifier 1	Sumo1	0.83 ±0.22
Regulator of nonsense transcripts 1	Upf1	0.82 ±0.21
Vimentin	Vim	0.72 ±0.13
Nuclear pore complex protein nup98-nup96	Nup98	0.68 ±0.15
Importin subunit alpha-1	Kpna2	0.58 ±0.1
Protein c10	Grcc10	0.56 ±0.01

* Only proteins that meet significance cut-off (hyperbolic curves, permutation-based FDR≤0.05, S_0=0.1). Full table of proteins are shown in **Table 4-S2**; Bold are genes with known causes or risk factors for MND (or other neurodegenerative diseases in the case of Picalm); Italics are cellular proteins previously seen to become more insoluble when Httex1Q_{97} formed inclusions (Sui et al., 2020).

To assess whether the changes in polyGA inclusion formation had other effects on proteome abundance, we expressed polyGA_{101} and at 48 h after transfection sorted live cells into those with visible aggregates from those without by flow cytometry sorting method and pulse shape analysis (Ramdzan et al., 2012) **(Figure 4-3 A)**. We found cells with inclusions were more reactive to Sytox **(Figure 4-3 A inset)**, which is indicative of dying and dead cells than cells without inclusions so we excluded these cells from analysis. 35% of the remaining live cells expressing polyGA had inclusions **(Figure 4-3 A)**. This was a lower yield of aggregates than we measured by live cell imaging at 48h in **Figure 4-1A** (about 90%), which we attribute to this experiment being more inclusive of the lower-expressing cells and other differences in the experimental conditions that affect aggregation rates such as phototoxicity of live cell imaging. Nonetheless, the yield obtained made the experiment amenable for comparing cells with versus without inclusions. Out of 2420 proteins identified, we observed 56 proteins that significantly changed abundance in these sorted cell populations **(Figure 4-3 ; Table 4-S4)**. There was no overlap in the proteins seen enriched in polyGA inclusions with proteins that changed expression due to polyGA aggregation. This provides firmer confidence that the enrichment seen in the polyGA aggregates arises from co-aggregation rather than changes in gene expression. Of the genes that changed expression, protein interaction networks yielded significant enrichment in networks including nuclear speck (GO: 0016607), ribosome biogenesis (GO: 0042254), chromosome (GO:0005694), mitochondrion (GO:0005739) and Golgi-to-ER-traffic (MMU-6811442) **(Table 4-S5)**. These pathways would be anticipated to be activated by stress responses incurred by protein aggregation, however, we did not note any striking changes that pertained to novel mechanisms other than that from this data.

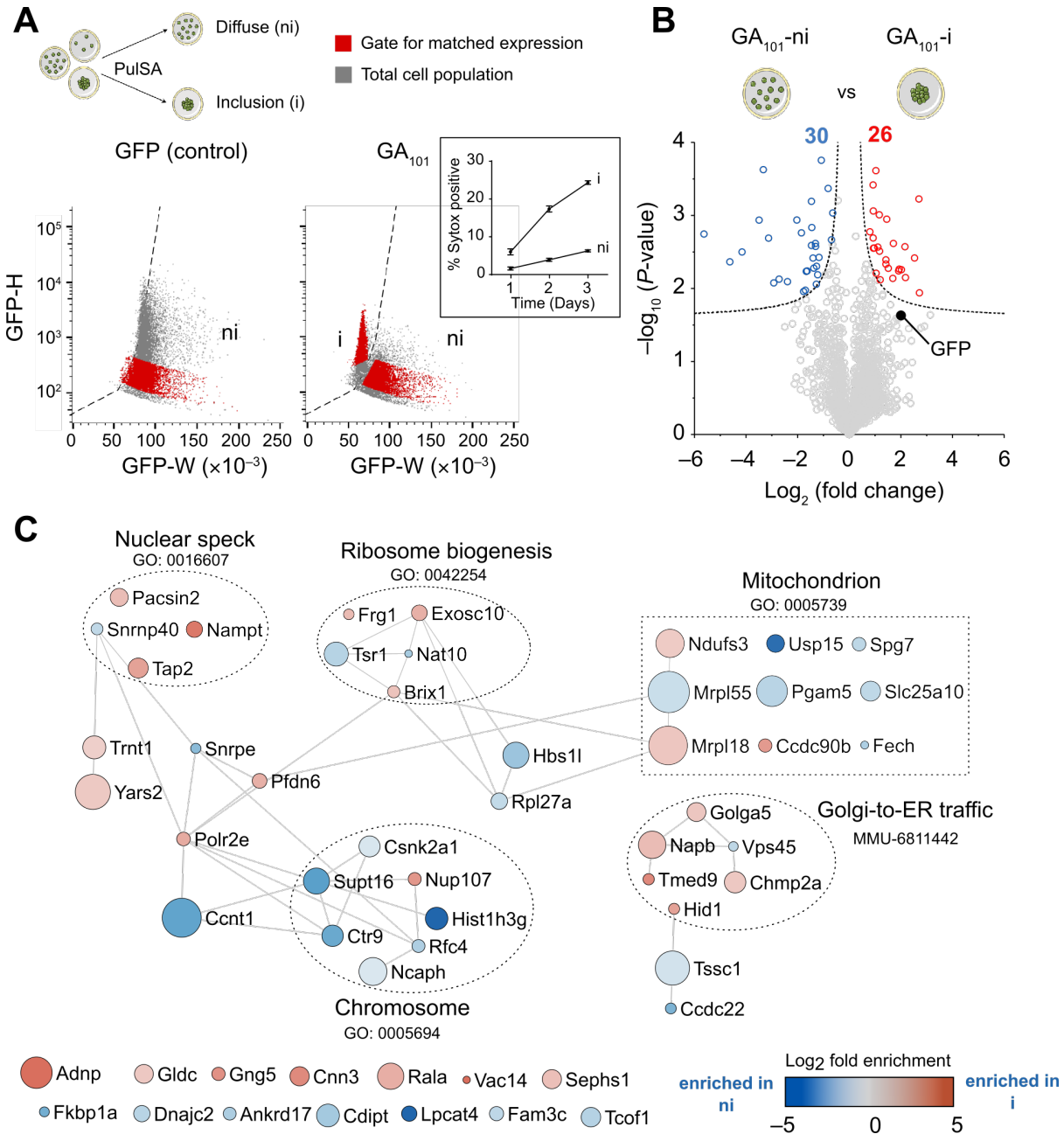

Figure 4-3 | Cellular protein abundance changes arising from polyGA$_{101}$ aggregation.

A. Schematic of flow cytometry method of pulse shape analysis (PulSA) to sort cells enriched with inclusions (i) from those without inclusions (ni). Cells with inclusions display shorter width (W) fluorescence values versus cells with soluble protein, and typically higher height values (H) arising from the condense foci of fluorescence inside the cells. Cells were sorted to exclude dead cells by Sytox reactivity. Inset shows the percentage of transfected cells reactive to Sytox by time after transfection. n=4, means ± SD shown. B. Volcano plots of proteins that changed their abundance upon polyGA aggregation. The data for all proteins are plotted as $\log_2$-fold change versus the $-\log_{10}$ of the *P*-value. The dotted line indicates significance cut-off (hyperbolic curves, FDR≤0.05, S_0=0.1) and proteins meeting stringency thresholds are shown as colored circles. C. Protein-protein interaction network (STRING v11) of proteins significantly

changed in abundance upon polyGA aggregation (the full list of proteins are in **Table 4-S4**). The analysis was done at the highest confidence setting. Each protein was represented by a colored circle sized proportionally to $-\log_{10}(P\text{-value})$. The color scale represents the logarithm of fold change. Selected significantly enriched GO terms (GOCC, GOPB, and UniProt keywords) are displayed **(Table 4-S5)**.

Lastly, we investigated the overlap of proteins enriched in Httex1Q_{97} inclusions with our previously reported changes in solubility of whole cell proteome before versus after inclusions had formed (Sui et al., 2020). In that dataset, we observed 17 proteins that significantly decreased in solubility as cells expressing Httex1Q_{97} shifted from a dispersed unaggregated state to forming inclusions (Sui et al., 2020). Of these, 5 proteins were found in our list of 48 proteins significantly enriched in Httex1Q97 inclusions (Picalm, Hgs, Clint1, Ubqln2 and Dnajb1). Four of these proteins (all except Dnajb1) form a robust protein-protein interaction network with a significant gene ontology enrichment for clathrin coat assembly (GO:0048268; FDR of 0.0031) suggestive that this mechanism is involved in polyQ aggregation. Clint1 and Ubqln2 were previously shown to colocalize to polyQ inclusions, supporting this conclusion (Wear et al., 2015; Yang et al., 2018). An interesting note with respect to mechanism is that UBQLN2 targets ubiquitinated substrates for degradation in ERAD and autophagy (Renaud et al., 2019). Furthermore, mutations in UBQLN2 cause MND, and appear to operate by impairing protein degradation of ubiquitinated proteins (Deng et al., 2011). Further supporting an important role linking protein aggregation, degradation more broadly to these neurodegenerative diseases is the enrichment of Picalm in the polyQ inclusions. Picalm is an phosphatidylinositol-binding clathrin assembly protein and has been shown via GWAS as a top ten risk for Alzheimer's disease (Harold et al., 2009; Elias-Sonnenschein et al., 2012). It has been reported to modulate intracellular APP processing and plaque pathogenesis (Xiao et al., 2012), modulate autophagy and alter tau clearance (Moreau et al., 2014).

Collectively the data here reports proteins that co-aggregate into two very different neurodegenerative disease proteinaceous deposits. The findings provide specificity of proteins to the aggregation type that provides a useful perspective to that reported by others. Moreover, the mechanisms of protein clearance mechanism appear relevant to both aggregation types and notably of a number of proteins in the Httex1Q_{97} aggregates that when mutated are modifiers of MND risk. Therefore, the findings identify a synergy of biological mechanisms involved in protein degradation that appear central to at least two different neurodegenerative diseases, and possibly more applicable to the other neurodegenerative diseases involving inappropriate protein aggregation.

4.4 Acknowledgements

We thank the Bio21 Melbourne Mass Spectrometry and Proteomics Facility. This work was funded by grants to DMH (National Health and Medical Research Council APP1161803 and Motor Neuron Disease Research Institute, Australia small grant) and to DMH and GER (Australian Research Council DP170103093). MR acknowledges support from an Australian Government Research Training Program (RTP) Scholarship and an Egyptian Ministry of Higher Education and Scientific Research PhD scholarship.

4.5 Contribution to this publication

As first author of this publication, I contributed to the study design and had major contributions to the experimental work. In detail, I performed all experiments shown in this publication. J.D.L. has participated in purifying protein aggregates using fluorescence-activated cell sorting (FACS) for subsequent aggregate proteome analysis.

4.6 Supplemental information

Table 4-S1. List of random proteins from mouse Uniprot database. Supplied separately.

Table 4-S2. Proteins enriched in inclusions of polyGA$_{101}$ and Httex1Q$_{97}$. Relates to Table 4-1 & Fig 4-2. Supplied separately.

Table 4-S3. Gene Ontology terms enriched among proteins identified in polyGA$_{101}$ or Httex1Q$_{97}$ inclusions. Relates to Fig 4-2. Supplied separately.

Table 4-S4. Cellular abundances of proteins caused by polyGA$_{101}$ aggregation. Relates to Fig 4-3. Supplied separately.

Table 4-S5. Gene Ontology terms enriched among proteins that changed abundance upon polyGA$_{101}$ aggregation. Relates to Fig 4-3. Supplied separately.

CHAPTER 5

General Discussion & Conclusions

5 General Discussion & Conclusion

5.1 Thesis summary

Approximately 88% of patients with ALS-FTD, including 40 % of fALS, 26 % of fFTD cases and <10% of sALS and sFTD, harbour a hexanucleotide repeat expansion in the first exon of C9ORF72 gene (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Gijselinck et al., 2012). The identification of this mutation has led to a surge of research in recent years that has focused on identifying the molecular mechanism responsible for causing disease. Various *in vitro* and *in vivo* models have demonstrated that dissection of the disease mechanism is not straightforward, and it is likely that neurodegeneration results from a combination of loss-of-function and gain-of-function factors. Nevertheless, to fully understand the global disease mechanisms, it is necessary to appreciate the individual mechanisms by which each independent factor may contribute towards neurodegeneration. With this in mind, the work in this dissertation focused primarily on adding to the understanding of unique dipeptide repeat proteins produced as a consequence of the repeat expansion by utilising a simple mouse neuroblastoma cell line and quantitative proteomics. This thesis describes a systematic investigation of how the expression of the different dipeptide polymers engage with the proteome and determined how these interactions relate to molecular dysfunction. Key themes emerging throughout this work highlight a possible convergence of multipronged mechanisms involving disruption of translation, cellular actin organization and methylation processes as a novel explanation of many aspects of C9FTD and C9ALS biology. In Chapter 4, I shed light into which endogenous proteins were enriched in poly-GA inclusions and whether aggregation-prone lengths of polyQ (Q97), in context of Huntingtin exon 1, shared similar patterns that may uncover novel disease pathways eliciting neurodegeneration. Despite the significant contribution of the main findings of this thesis to the ALS and FTD field, many questions remain for future study. In this section, I outline important steps to take in moving forward from this work.

5.2 Key findings and future directions

5.2.1 Varying DPR repeat length produce markedly different phenotypes

With the discovery of G4C2 repeats in C9ORF72, it is still unknown what the repeat length constituting DPR pathological aggregates in C9ORF72-mutation carriers. Work in this thesis has examined how different repeat lengths of each DPRs impact their localization and toxicity. We synthesized DNA vectors coding for short (10x) and long (101x) polymers corresponding to the four C9ORF72 RAN products derived from the sense and antisense strand and assessed their

relative toxicity when expressed in mouse neuroblastoma cell line. We observed a striking nuclear accumulation of long and short poly-PR as well as short poly-GR polymer. The longer repeat length of poly-GR was excluded from the nucleus and poly-GA has formed cytoplasmic inclusions only with 101x repeat unit lengths whereas no distinct localisation was observed following expression of poly-AP at both repeat lengths. This simple model also has shown the poly-PR and poly-GR are highly toxic at both polymer lengths; long poly-GA is associated with milder toxicity whilst poly-AP did not appear to confer toxicity. The fact that we observed differences in the localization and toxicity of the peptides between short and long repeating-units emphasizes the importance of future studies investigating different DPR lengths.

5.2.2 Arginine-rich DPRs toxicity: Does ribosome-associated protein quality control (RQC) network kick-start the toxic cascade?

Blocking the global protein translation is among the reported mechanisms that explain the arginine-rich DPRs toxicity (Kanekura et al., 2016; Hartmann et al., 2018; Moens et al., 2019). How Arg-rich DPRs repress translational is still an intriguing avenue of exploration. It has been previously predicted to be due to formation of cytoplasmic stress granules (Lee et al., 2016a) or reduced ribosome biogenesis (Kramer et al., 2018) observed in multiple poly-PR and poly-GR models. Other studies have suggested that defective translation is attributed to the direct engagement of poly-GR and poly-PR with different components of translation process. *In vitro* translation assays revealed that poly-GR and poly-PR formed insoluble complexes with mRNA and inhibited translation. This study suggested that the direct binding of the arginine-containing DPRs to mRNA restain the access of translation machinery to mRNA and thereby block transation (Kanekura et al., 2016). However, several proteomic studies demonstrated that the Arg-rich DPRs can still bind to the ribosome even in the presence of RNAse (Kanekura et al., 2016; Lee et al., 2016a; Lin et al., 2016; Lopez-Gonzalez et al., 2016; Yin et al., 2017; Hartmann et al., 2018). Indeed, several ribosomal proteins have robustly localized to the poly-GR/PR inclusions in C9ORF72 patient brains, suggeting that depletion of ribosomal proteins may drive the translation arrest (Hartmann et al., 2018; Zhang et al., 2018).

Data presented in chapter 3 have provided a novel clarification of the previously ambiguous understanding as to how the arginine peptides cause defective translation. This was the first study to notice the interaction of polyPR with a member of the ribosome-associated quality control (RQC) network, namely ABCE1. The data is consistent with a mechanism by which the GR and PR proteins stall on the ribosome during translation. We postulate that Arg-rich peptide domains may electrostatically jam the ribosome exit tunnel during synthesis. Indeed,

antimicrobial peptides enriched for PR-containing motifs have been demonstrated to bind to the ribosomal exit tunnel and inhibit bacterial protein synthesis (Gagnon et al., 2016). It is therefore readily conceivable that emergent DPRs may also re-enter and plug the exit tunnel through electrostatic interactions. Overall, these experiments imply that ribosome stalling is the main feature of problematic translation associated with long Arg-rich DPRs, but we cannot exclude additional effector mechanisms since translation is regulated at many levels. The next important step is to systemically analyse members of RQC machinery and their potential colocalization with GR/PR aggregates using proximity ligation assay or fluorescence resonance energy transfer in a more relevant disease model including patient-derived tissue, cells or induced pluripotent stem cell-derived neurons, and *in vivo* disease models. Also, it will be interesting to test whether resolving the ribosome stalling happening during Arg-rich DPRs translation either genetically or pharmacologically can rescue the poly-PR/GR-induced translation repression.

The RQC is a newly discovered pathway that monitors the process of translation for possible errors, processes damaged or aberrant RNAs and ribosomal subunits and works upstream the chaperone system to control the nascent polypeptide folding. The RQC is proposed to act in the following sequence of events. Firstly, the ribosome-recycling factors; Pelota and Hbs1 recognize the stalled ribosome and ABCE1allows its splitting. The full RQC complex is then assembled around the 60S ribosomal subunit to allow ubiquitination of nascent polypeptides by NEMF and Listerin. The TCF25 and VCP complex extract the nascent chain from the 60S subunit and target it for degradation by the proteasome. The components of the RQC and ribosome are then recycled (Brandman and Hegde, 2016).

Little research has been performed on the involvement of the RQC mechanisms in ALS or FTD. In 2009, a study reported that mutations in the E3 ubiquitin ligase Listerin cause ALS-like symptoms in mice (Chu et al., 2009). Deletion of Ltn1; the yeast analogue of the mammalian Listerin, results in accumulation of stalled protein products into inclusion bodies (Bengtson and Joazeiro, 2010). Also, mutation of a CNS-specific tRNAArg UCU gene disrupts the activity of Pelota and leads to persistent ribosome stalling and widespread neurodegeneration(Ishimura et al., 2014). In the context of ABCE1, its physiological level, at least in yeast cells, is below the threshold sufficient for complete ribosome recycling (Young et al., 2015). ABCE1 is innately metastable and prone to sequestration by disease-associated aggregates (Olzscha et al., 2011). Additionally, ABCE1 is a twin-ATPase protein with a unique N-terminal iron–sulfur (FeS) cluster domain that contains two [4Fe-4S]2+ clusters (Barthelme et al., 2007; Karcher et al., 2008). Fe-S association is essential for ABCE1 function (Alhebshi et al., 2012) and the age-associated oxidative stress impairs the activity of ABCE1 via limiting or chelating iron and inhibiting the

assembly or stability of ABCE1(Brumaghim et al., 2003). This result explains the accumulation of ribosomes-associated 3'UTR of mRNAs in aged brains (Sudmant et al., 2018). So, depletion of ABCE1 via sequestration into polyGR/PR aggregates may synergize the inhibition of ABCE1 activity induced by age-associated oxidative stress leading to RQC failure. Clearly, failure of RQC to recycle the stalled ribosomal complex in order to re-enter the cellular pool can diminish the translational machinery that is extremely limited in neurons leading to neurodegeneration (Kapur et al., 2017).

This model proposes ribosome stalling during Arg-rich DPR translation as the molecular mechanism behind the predominant ribosome binding, eventually leading to impairment of translation. Four main lines of significance are anticipated for this proposed mechanism. First, it is proposed that ribosome stalling may be the earliest step in the toxic cascade elicited by Arg-rich DPR expression. Second, ribosome stalling may give clues to the repeat length constituting DPR pathological aggregates in C9ORF72-mutation carriers. Pathogenic G4C2 repeat length range from over 30 to hundreds or thousands of repeats (DeJesus-Hernandez et al., 2011; Renton et al., 2011; Gijselinck et al., 2012). This high variability in G_4C_2 repeat size has made it challenging to develop disease models that express DPRs of a size similar to those observed in patients. This is particularly of importance since several animal and cellular models have already demonstrated that DPR length has an effect, with longer repeats having greater toxicity (Mizielinska et al., 2014a; Wen et al., 2014; Chew et al., 2015; Jovičić et al., 2015). This thesis has shed light on the limited ability of ribosomes to synthesize a long stretch af arginine-rich peptides irrespective of the repeat length that the parent mRNA encodes for. As a next step, future ribosome footprint profiling experiments are warranted to measure the ribosome occupancy at the G_4C_2 expanded RNA and map where ribosome stops. This knowledge will help define the exact DPR length observed in human patients to better undermine the physiological mechanisms that are affected by DPR presence. Third, these findings introduce ribosome stalling as a new player in determining the actual DPR length translated from G_4C_2 repeat transcript besides the repeat length in the parent mRNA. Hence, this may help explain the lack of correlation between G_4C_2 repeat length the patient harbours and disease severity reported in several studies (Beck et al., 2013; Dols-Icardo et al., 2013; van Blitterswijk et al., 2013; Nordin et al., 2015).

Finally, poly-GR and poly-PR frequently form cytoplasmic aggregates in patients (Mori et al., 2013a; Mori et al., 2013c; Schludi et al., 2015) in contrast to the predominant nucleolar localization observed in most *in vitro* systems. It is proposed that the presence of functional RQC in *in vitro* models allow the resolution of ribosome stalling during DPR expression **(Figure 5-1)**. The extracted nascent DPR polypeptide chain can then translocate to the nucleus and localize

into the nucleolus. On the other hand, RQC system seems to be defective during aging as evident by the accumulation of ribosomes-associated 3'UTR of mRNAs in aged brains (Sudmant et al., 2018). This may result in failure to resolve the stalled ribosomes with the consequent aggregation of nascent DPR chain-ribosome complex into cytoplasmic inclusions in patient brains. This notion is bolstered by interactome data from rat primary cortical neurons, in which cytoplasmic GFP-(GR)149 has been shown to bind already-assembled ribosomes. Our hypothesis may also explain the severe toxicity in some poly-PR/GR models that is hard to reconcile with the prodromal expression at least of poly-GR many years before disease onset (Vatsavayai et al., 2016). Therefore, current models likely exaggerate toxicity although it is possible that cytoplasmic poly-GR/PR inclusions trigger quite different pathways *in vivo* with milder effects.

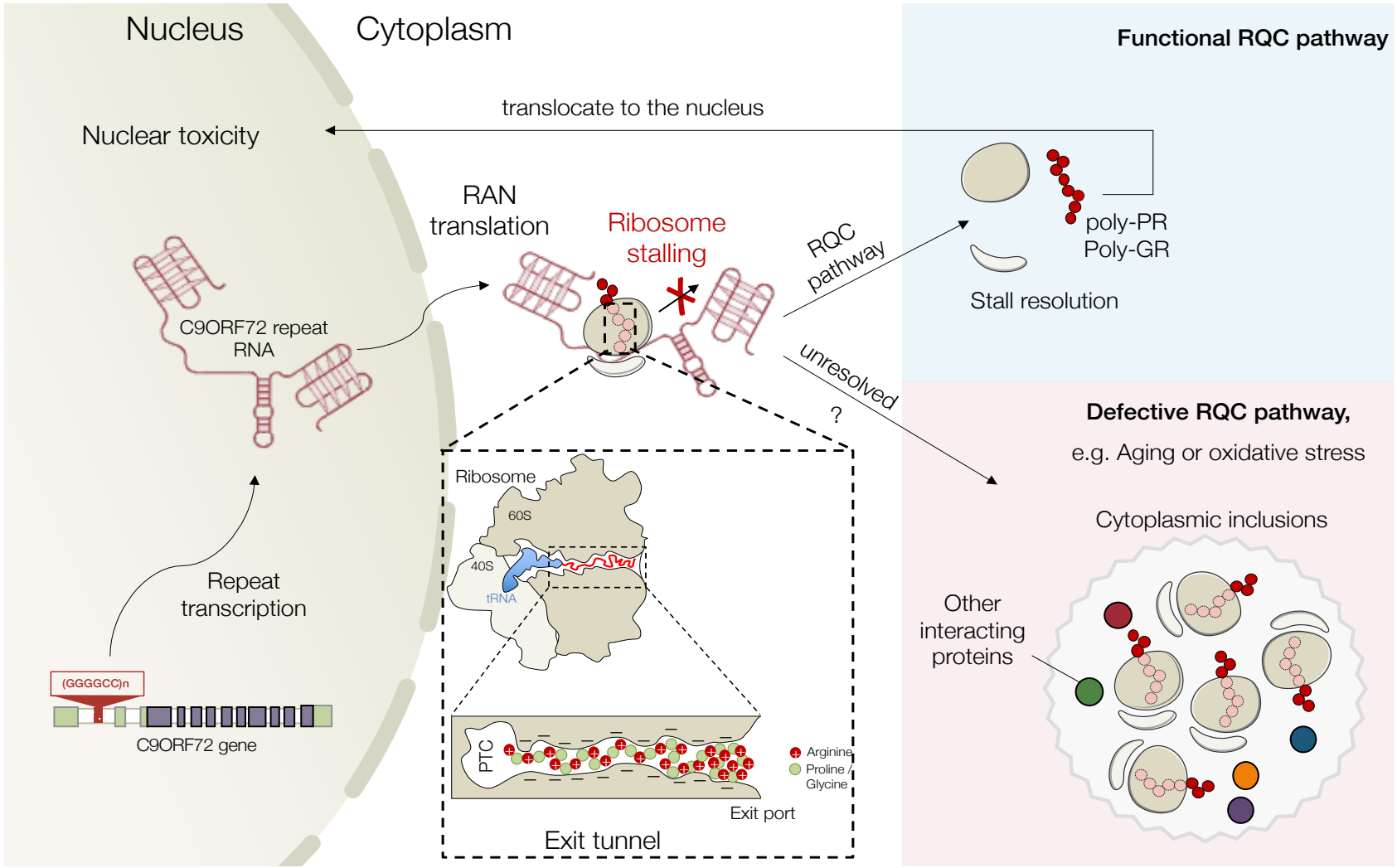

Figure 5-1 | Proposed model of ribosome stalling induced by Arg-rich DPRs translation and its predicted consequences.

5.2.3 Arginine-rich DPRs alter the arginine methylation patterns

The methylation of arginine residues within a protein is a common post-translational modification that can have a significant impact on the localisation and behaviour of proteins (Bulau et al., 2006; Uhlmann et al., 2012; Larsen et al., 2016). Given that the poly-GR and poly-PR proteins contain a significant amount of arginine residues and these DPR proteins have been consistently shown to be the most toxic DPR protein species in a number of model systems (Kwon et al., 2014; Mizielinska et al., 2014a; Wen et al., 2014; Jovičić et al., 2015; Tao et al., 2015; Lee et al., 2016a; Rudich et al., 2017; Xu and Xu, 2018), it is possible that the methylation of arginine residues within these proteins could play a significant role in their behaviour and influence toxicity.

Recent studies have detected dimethylated poly-GR inclusions in C9ORF72-FTLD and C9ORF72-ALS patient post-mortem tissue and demonstrated that poly-GR aggregates may contain a mixture of unmethylated, symmetric dimethylated (SDMA) and asymmetric (ADMA) forms following immunohistochemical detection of DPR protein inclusions using antibodies designed to detect different arginine methylation states (Mori et al., 2013a; Boeynaems et al., 2016; Chitiprolu et al., 2018; Sakae et al., 2018; Gittings et al., 2020), suggesting that arginine methylation of DPR protein occurs *in vivo*. Gittings et al. (2020) have detected a consistent positive correlation between SDMA modification of poly-GR and disease duration and age at death suggesting that this posttranslational modification may have a significant effect on disease outcome. These authors went on to hypothesize this positive correlation to be due to the symmetric dimethylation weakening both electrostatic and pi-pi interactions between arginine residues of poly-GR molecules. In doing so, poly-GR become less liable to phase separate. Since phase separation of poly-GR contributes to its toxicity (Lee et al., 2016a), these modifications were found to reduce poly-GR toxicity in primary neuronal cultures (Gittings et al., 2020). An earlier study has come to the same conclusion when knockdown of four out of ten arginine methyltransferases enhanced the toxicity of poly-PR expression in Drosophila. One of these enzymes, PRMT1, was subsequently found to co-localise with both poly-PR and poly-GR in HeLa cells, suggesting that this enzyme may be involved in the methylation of these proteins (Boeynaems et al., 2016). Another study demonstrated that treating HEK 293 cells with adenosine dialdehyde (AdOx), a global methyltransferase inhibitor, decreased the formation of ADMA cytoplasmic aggregates of GFP tagged poly-GR consisting of 100 repeats in HEK 293 cells, suggesting that poly-GR aggregation may be modulated by this arginine methylation. These studies signify the role that arginine methylation may play in C9ORF72 disease pathogenesis.

Thus far, all previous studies have focused on studying the biophysical and functional effect of arginine methylation on the Arg-rich DPRs molecules themselves. In chapter 3, the whole proteome analyses revealed that GR_{101} expression leads to hypomethylation of endogenous proteins, positing that the Arg-rich dipeptide polymers are mimetics of endogenous arginine methylase substrates and recruit these methylase proteins. A consequence is hypomethylation of the proteome, including key proteins known to be involved in ALS mechanisms **(Figure 5-2)**. An important next step will be to examine the global arginine methylation patterns in a more relevant disease model or C9ORF72 patient brain tissues.

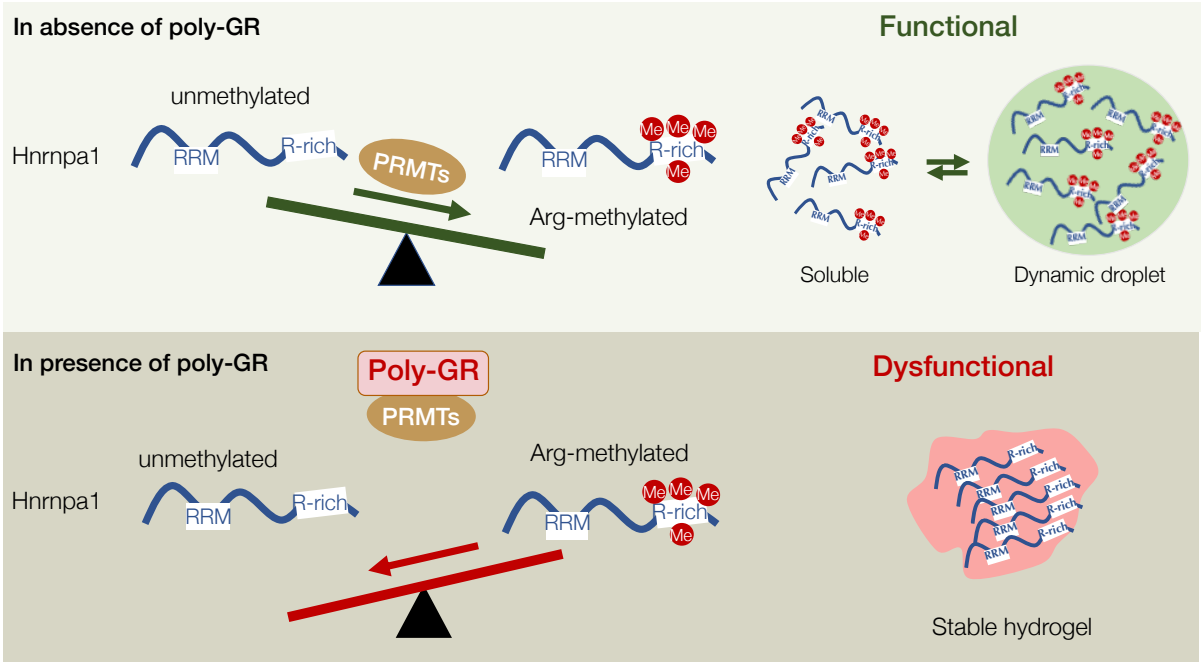

Figure 5-2 | Proposed model of poly-GR induced hypomethylation of Hnrnp family proteins.

How perturbation of arginine methylation may be linked to the pathogenesis of FTD and ALS is yet to be explored. However, like other posttranslational modifications, defects in arginine methylation is expected to have many downstream effects. For instance, several proteins associated with ALS and FTD, which are capable of liquid-liquid phase transition, contain RGG/RG-rich domains and are known targets of arginine methylation (Ong et al., 2004; Hung et al., 2009; Altmeyer et al., 2015; Molliex et al., 2015), suggesting that perhaps the regulation of phase transition by arginine methylation may be a mechanism common to these proteins. Potential candidate proteins include some members of the hnRNP A family, which aggregate in rare familial forms of ALS (Kim et al., 2013a; Mori et al., 2013b), and EWS and TAF15, both of which accumulate in pathological inclusions in FTLD-FUS (Neumann et al., 2011).

Also, arginine methylation of GAR (Gly-Arg rich) motifs in the human MRE11, a key enzyme in DNA double-strand break repair and stability of genomic DNA is essential for targeting MRE11 to DNA damage foci and colocalization with the phosphorylated H2A histone family member X (γ-H2AX) (Déry et al., 2008). Markers of DNA damage response (DDR) including γ-H2AX were up-regulated in C9ORF72 ALS patient spinal cord motor neurons (Farg et al., 2017). A following study has confirmed that Arg-rich DPRs in addition to the C9ORF72 RNA is capable of inducing DNA damage (Pottier et al., 2015). Whether the observed DNA damage is a downstream effect of disruption in arginine methylation require investigation.

Another important class of methylase substrates is arginine-rich histones including H3 and H4 histones (Thimmegowda and Yeldur, 2016). PRMT1 is responsible for the asymmetric di-methylation of Arginine 3 on Histone H4 (H4R3me2asym), which promotes histone acetylation and gene transcription (Litt et al., 2009). In a FUS mouse model of ALS, loss of PRMT1 function due to recruitment into FUS aggregates has led to reduction of asymmetric di-methylation of Arginine 3 on Histone H4 (H4R3me2asym) causing a drop in the acetylation of Histone H3 on Lysine 9 and 14, ultimately leading to transcriptional silencing (Tibshirani et al., 2015). Overexpression of PRMT1 was found to rescue neurite growth in the same model (Jun et al., 2017). Also, the chromatin modelling enzyme POLR2A undergoes symmetrical di-methylation by PRMT5 and this modification recruits the Tudor domain of survival of motor neuron (SMN) (Yanling Zhao et al., 2016). SMN interacts with senataxin, which is sometimes mutated in ALS (Hirano et al., 2011). Perturbations in this pathway lead to the formation of R-loops, genomic instability and pre-mature transcription termination causing neurodegeneration (Yanling Zhao et al., 2016). In the context of C9ORF72, aberrations of histone methylation have previously been reported in a mouse model of PR_{50} (Zhang et al., 2019). Further studies are necessary to examine the possible correlation between arginine hypomethylation and histone anomalies linked to the pathogenesis of FTD and ALS.

From the above, it is tempting to view perturbations in global arginine methylation as a unifying mechanism for the different defects reported in C9ORF72 ALS/FTD including DNA damage and histone anomalies. In addition, the defect in this posttranslational modification can be extended to other forms of ALS and FTD via promoting phase separation and aggregation of disease-associated RNA-binding proteins.

5.2.4 Arginine-rich DPRs impede the assembly of the actin cytoskeleton

The results from this thesis demonstrated a potentially significant relationship between C9ORF72-associated DPRs and the actin cytoskeleton. Within neurons, actin plays several roles

including support of membranes, regulating axonal transport, motility of the developing growth cone and pre- and post-synapse formation, and is the substrate for myosin motor proteins (e.g. myosin Va) which link cytoskeletal elements. Disruption to the actin cytoskeleton can perturb synapse formation and neuronal integrity (Eira et al., 2016) and has been reported in several ALS models(Hensel and Claus, 2018). Mutation in profilin-1 protein, which mediates the conversion of G-actin to F-actin, is linked to ALS (Wu et al., 2012). Other cytoskeletal genes have also been linked to ALS including TUBA4A and DCTN1 and KIF5A (Taylor et al., 2016a; Nicolas et al., 2018). Data presented in chapter 3 have shown that Arg-rich DPRs destabilize the actin cytoskeleton. Lack of direct binding with actin suggested that poly-PR and poly-GR may interfere with machinery implicated in actin filament assembly. This is the first study to establish a direct link between the actin cytoskeleton defects and Arg-rich DPRs.

A previous immunocytochemical study has detected the colocalization of C9ORF72 and actin (Farg et al., 2014). In a follow-up study, shRNA knockdown of C9ORF72 in cultured motor neurons reduced actin filament production, and truncated axons and growth cones which is also observed in iPSC-derived motor neurons and post mortem tissue from patients with C9ORF72 expansion (Sivadasan et al., 2016). Hence, it is expected that Arg-rich DPRs can synergize the negative impact of C9ORF72 haploinsufficiency on the neuronal actin cytoskeleton. Another intriguing possibility is that the defects in actin cytoskeleton may be the mechanism underlying the observed nucleocytoplasmic transport pathology in C9ORF72-ALS/FTD (Freibaum et al., 2015; Jovičić et al., 2015; Zhang et al., 2015; Boeynaems et al., 2016; Shi et al., 2017a). This postulation is supported by a group of evidences. First, a study has observed that poly-PR or poly-GR impair the nucleocytoplasmic transport, albeit indirectly (Vanneste et al., 2019). A following study has found that alteration to the actin cytoskeleton disrupts nucleocytoplasmic transport in ALS caused by mutant Profilin1 (PFN1) and by C9ORF72 repeat expansion (Giampetruzzi et al., 2019b). The authors have also reported that promoting the actin filament assembly can alleviate defects in nuclear-cytoplasmic transport defects which supports the conclusion that destabilization of the actin cytoskeleton is pertinent to disease pathomechanisms.

To extend the key findings of Chapter 3, a correlation between the R-rich DPRs and cellular toxicity via this mechanism could be investigated in a more relevant disease model or C9ORF72 patient brain tissues. Additionally, it is crucial to study in more depth the exact mechanism by which each Arg-rich DPR species interfere with the actin cytoskeleton integrity. Our proteomic data showed the specific interaction of polyPR with the actin capping proteins (capza1, capza2 and capzb), whereas polyGR has been shown to interact with the actin-binding

proteins (e.g. Coro1a, Nexn and Fscn1), thus pointing toward divergent interference mechanisms. Another point that warrant investigation is whether the actin cytoskeleton defect is a consequence of other cellular effects, such as hypomethylation or ribosome stalling. Interestingly it has been reported that arginine methylation of an arginine methylase (PRMT2) regulates the activity of actin nucleator protein Cobl, which suggests a possible role for arginine methylation defects being an upstream mediator of effects on the actin cytoskeleton (Hou et al., 2018).

5.2.5 polyGA inclusion are enriched with distinct proteins involved in protein degradation

PolyGA inclusions are more widespread in FTD-ALS patient brain tissue compared to the other PDRs (Mackenzie et al., 2015). In chapter 3, we reported polyGA$_{101}$ to be mildly toxic to cultured Neuro2a cells and to induce a distinct network of proteome changes that occur compared to the arg-rich PDRs. We also noted a distinction to the other PDRs in forming large inclusions that morphologically are similar to the inclusions formed by polyQ. Both of polyGA and polyQ inclusions are SDS-insoluble and assemble by an amyloid-like mechanism (May et al., 2014; Chang et al., 2016).

In chapter 4, the use of FACS to purify aggregates, in combination with quantitative MS, enabled us to perform the first in-depth characterization of polyGA aggregates compared to HttEx1 polyQ aggregates. This generic approach provided an unbiased tool to identify and quantify their constituent proteins. A notable connections to ALS biology were observed in the ensemble of proteins found in the polyGA inclusions suggesting that the cell model of polyGA inclusions mimicks the process of aggregation and recruitment seen *in vivo*. The screen also yielded proteins implicated in MND phenotype and mechanisms including Bag6, Phlda1, Nudt5 and Dpysl3 (Egawa et al., 2012; Blasco et al., 2013; Raman et al., 2015). We propose that co-aggregation of these proteins into polyGA inclusions sequesters their activity and renders cells less resilient to stress triggers. However, additional experiments are needed to elucidate if these proteins can also be found in patient aggregates and how their loss of function through sequestration might contribute to neurotoxicity. We also noted the GFP moiety (which was identified from both Httex1 and polyGA fusions) was enriched in the polyGA aggregates suggesting that the inclusions formed by polyGA contain less mass of associated proteins than those formed by Httex1. This could arise by more non-specific interactions to the Httex1 inclusions or by polyGA forming a greater proportion (by mass) of the inclusion composition than for Httex1 inclusions.

In the case of HttEx1 aggregates, we observed that polyQ more selectively co-aggregates with other proteins enriched with glutamine and LCRs suggesting that inclusion assembly may be directed by cellular quality control mechanisms for proteins with some degree of unfoldedness and similar physicochemical properties. One such candidate is CREB Binding Protein, which is dysfunctional in Huntington's Disease. CREB Binding protein contains a glutamine repeat and is co-aggregated into inclusions formed by mutant Htt (Luthi-Carter et al., 2000; Wyttenbach et al., 2001; Mantamadiotis et al., 2002; Moily et al., 2017). Hence the co-aggregation of CREB binding protein may result in its loss of activity as one contribution to pathogenesis. Lastly, we observed an overlap of proteins enriched in Httex1Q97 inclusions with proteins identified in Httex1 inclusions from other study (Mitsui et al., 2002).

Chapter 4 has also focused on examining how the aggregation status of poly-GA impact the whole proteome using a combination of flow cytometry and PulSA with MS analysis. We observed that the stress responses incurred by protein aggregation have resulted in significant changes in abundance of proteins belonged to nuclear speck, ribosome biogenesis, chromosome, mitochondrion and Golgi-to-ER-traffic.

Altogether, the proposed experiments outlined in this section (5.2) will drive further investigation and promise an interesting study that will be useful for identifying critical targets for effective therapies.

5.3 General conclusion

The discovery of the C9ORF72 repeat expansion mutation has revolutionised the ALS and FTD research field in a relatively short space of time. It has led to the identification of C9ORF72-specific pathological features, the development of several model systems and has resulted in the formation of many new and exciting hypothesised mechanism of disease. The work in this thesis has primarily focused on toxicity induced by unique DPR proteins produced by RAN translation in C9ORF72-linked disease. Using quantitative proteomics in a Neuro2a cell model, this thesis demonstrated that the ATG-driven arginine-containing DPRs (polyPR, polyGR), that have been recodonised to lack a repetitive RNA intermediate, are profoundly toxic and stall the ribosome likely by electrostatic interaction with the ribosome exit tunnel. Further poly-GR also bound arginine methylases, and thereby inducing hypomethylation of endogenous proteins and also destabilized the actin cytoskeleton **(Figure 5-3)**. Poly-GA, on the other hand, recruited into its cytoplasmic inclusions proteins that are involved in the degradation machinery and are risk factors for neurodegenerative diseases. In this view, this thesis offered novel explanations for phenotypes seen in patient tissue from C9ORF72-forms of ALS and FTD. Although many

questions remain to be answered, it is hoped that the discovered mechanisms will make significant advancements in our understanding of the mechanisms of FTD and ALS that will ultimately aid in the development of successful disease-modifying treatments for these devastating neurodegenerative diseases.

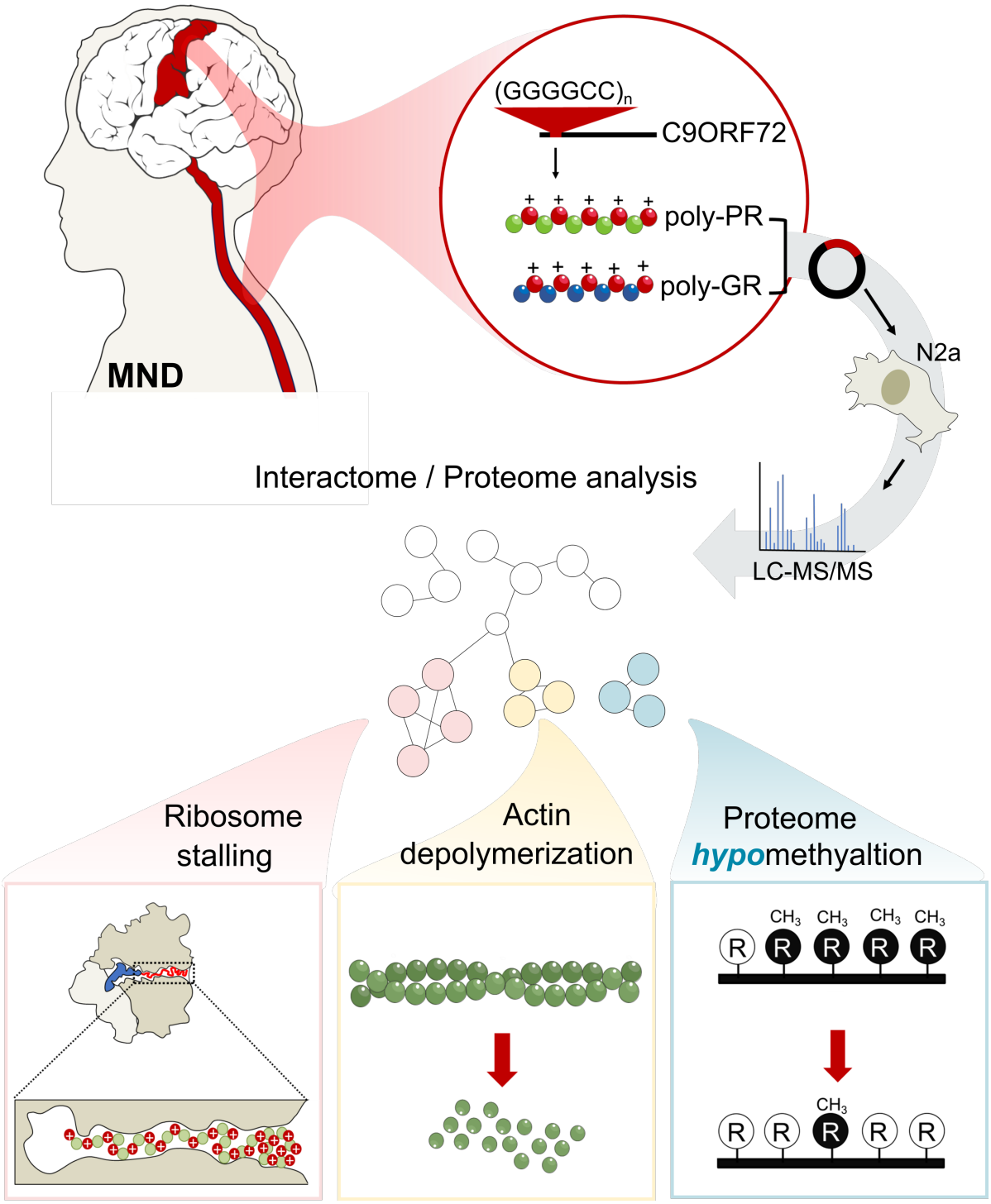

Figure 5-3 | A schematic summary of main findings of this thesis depicting three novel mechanisms linked to the Arg-rich DPRs.

REFERENCES

6 References

ABE, K., AOKI, M., TSUJI, S., ITOYAMA, Y., SOBUE, G., TOGO, M., HAMADA, C., TANAKA, M., AKIMOTO, M., NAKAMURA, K., TAKAHASHI, F., KONDO, K., YOSHINO, H., ABE, K., AOKI, M., TSUJI, S., ITOYAMA, Y., SOBUE, G., TOGO, M., HAMADA, C., SASAKI, H., YABE, I., DOI, S., WARITA, H., IMAI, T., ITO, H., FUKUCHI, M., OSUMI, E., WADA, M., NAKANO, I., MORITA, M., OGATA, K., MARUKI, Y., ITO, K., KANO, O., YAMAZAKI, M., TAKAHASHI, Y., ISHIURA, H., OGINO, M., KOIKE, R., ISHIDA, C., UCHIYAMA, T., MIZOGUCHI, K., OBI, T., WATANABE, H., ATSUTA, N., AIBA, I., TANIGUCHI, A., SAWADA, H., HAZAMA, T., FUJIMURA, H., KUSAKA, H., KUNIEDA, T., KIKUCHI, H., MATSUO, H., UEYAMA, H., UEKAWA, K., TANAKA, M., AKIMOTO, M., UEDA, M., MURAKAMI, A., SUMII, R., KUDOU, T., NAKAMURA, K., MORIMOTO, K., YONEOKA, T., HIRAI, M., SASAKI, K., TERAI, H., NATORI, T., MATSUI, H., KOTANI, K., YOSHIDA, K., IWASAKI, T., TAKAHASHI, F., KONDO, K., and YOSHINO, H. (2017). Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *The Lancet Neurology* 16**,** 505-512.

ABRAMOFF, M.D., MAGALHAES, P.J., and RAM, S.J. (2004). Image Processing with ImageJ. *Biophotonics International* 11**,** 36-42.

AGUZZI, A., and ALTMEYER, M. (2016). Phase Separation: Linking Cellular Compartmentalization to Disease. *Trends Cell Biol* 26**,** 547-558.

AL-SARRAJ, S., KING, A., TROAKES, C., SMITH, B., MAEKAWA, S., BODI, I., ROGELJ, B., AL-CHALABI, A., HORTOBAGYI, T., and SHAW, C.E. (2011). p62 positive, TDP-43 negative, neuronal cytoplasmic and intranuclear inclusions in the cerebellum and hippocampus define the pathology of C9orf72-linked FTLD and MND/ALS. *Acta Neuropathol* 122**,** 691-702.

ALHEBSHI, A., SIDERI, T.C., HOLLAND, S.L., and AVERY, S.V. (2012). The essential iron-sulfur protein Rli1 is an important target accounting for inhibition of cell growth by reactive oxygen species. *Molecular biology of the cell* 23**,** 3582-3590.

ALTMEYER, M., NEELSEN, K.J., TELONI, F., POZDNYAKOVA, I., PELLEGRINO, S., GRØFTE, M., RASK, M.-B.D., STREICHER, W., JUNGMICHEL, S., NIELSEN, M.L., and LUKAS, J. (2015). Liquid demixing of intrinsically disordered proteins is seeded by poly(ADP-ribose). *Nature Communications* 6**,** 8088.

AMBIKAIRAJAH, A., DEVENNEY, E., FLANAGAN, E., YEW, B., MIOSHI, E., KIERNAN, M.C., HODGES, J.R., and HORNBERGER, M. (2014). A visual MRI atrophy rating scale for the amyotrophic lateral sclerosis-frontotemporal dementia continuum. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 15**,** 226-234.

AOKI, Y., MANZANO, R., LEE, Y., DAFINCA, R., AOKI, M., DOUGLAS, A.G.L., VARELA, M.A., SATHYAPRAKASH, C., SCABER, J., BARBAGALLO, P., VADER, P., MAGER, I., EZZAT, K., TURNER, M.R., ITO, N., GASCO, S., OHBAYASHI, N., EL ANDALOUSSI, S., TAKEDA, S., FUKUDA, M., TALBOT, K., and WOOD, M.J.A. (2017). C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain* 140**,** 887-897.

ARRASATE, M., MITRA, S., SCHWEITZER, E.S., SEGAL, M.R., and FINKBEINER, S. (2004). Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431**,** 805.

ASH, P.E., BIENIEK, K.F., GENDRON, T.F., CAULFIELD, T., LIN, W.L., DEJESUS-HERNANDEZ, M., VAN BLITTERSWIJK, M.M., JANSEN-WEST, K., PAUL, J.W., 3RD, RADEMAKERS, R., BOYLAN, K.B., DICKSON, D.W., and PETRUCELLI, L. (2013). Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* 77**,** 639-646.

ATANASIO, A., DECMAN, V., WHITE, D., RAMOS, M., IKIZ, B., LEE, H.C., SIAO, C.J., BRYDGES, S., LAROSA, E., BAI, Y., FURY, W., BURFEIND, P., ZAMFIROVA, R., WARSHAW, G., ORENGO, J., OYEJIDE, A., FRALISH, M., AUERBACH, W., POUEYMIROU, W., FREUDENBERG, J., GONG, G., ZAMBROWICZ, B., VALENZUELA, D., YANCOPOULOS, G., MURPHY, A., THURSTON, G., and LAI, K.M. (2016). C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci Rep* 6**,** 23204.

BAKER, M., MACKENZIE, I.R., PICKERING-BROWN, S.M., GASS, J., RADEMAKERS, R., LINDHOLM, C., SNOWDEN, J., ADAMSON, J., SADOVNICK, A.D., ROLLINSON, S., CANNON, A., DWOSH, E., NEARY, D., MELQUIST, S., RICHARDSON, A., DICKSON, D., BERGER, Z., ERIKSEN, J., ROBINSON, T., ZEHR, C., DICKEY, C.A., CROOK, R., MCGOWAN, E., MANN, D., BOEVE, B., FELDMAN, H., and HUTTON, M. (2006). Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature* 442**,** 916-919.

BAÑEZ-CORONEL, M., AYHAN, F., TARABOCHIA, A.D., ZU, T., PEREZ, B.A., TUSI, S.K., PLETNIKOVA, O., BORCHELT, D.R., ROSS, C.A., MARGOLIS, R.L., YACHNIS, A.T., TRONCOSO, J.C., and RANUM, L.P. (2015). RAN Translation in Huntington Disease. *Neuron* 88**,** 667-677.

BARTHELME, D., SCHEELE, U., DINKELAKER, S., JANOSCHKA, A., MACMILLAN, F., ALBERS, S.-V., DRIESSEN, A.J.M., STAGNI, M.S., BILL, E., MEYER-KLAUCKE, W., SCHÜNEMANN, V., and TAMPÉ, R. (2007). Structural Organization of Essential Iron-Sulfur Clusters in the Evolutionarily Highly Conserved ATP-binding Cassette Protein ABCE1. *Journal of Biological Chemistry* 282**,** 14598-14607.

BECK, J., POULTER, M., HENSMAN, D., ROHRER, J.D., MAHONEY, C.J., ADAMSON, G., CAMPBELL, T., UPHILL, J., BORG, A., FRATTA, P., ORRELL, R.W., MALASPINA, A., ROWE, J., BROWN, J., HODGES, J., SIDLE, K., POLKE, J.M., HOULDEN, H., SCHOTT, J.M., FOX, N.C., ROSSOR, M.N., TABRIZI, S.J., ISAACS, A.M., HARDY, J., WARREN, J.D., COLLINGE, J., and MEAD, S. (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**,** 345-353.

BELZIL, V.V., BAUER, P.O., PRUDENCIO, M., GENDRON, T.F., STETLER, C.T., YAN, I.K., PREGENT, L., DAUGHRITY, L., BAKER, M.C., RADEMAKERS, R., BOYLAN, K., PATEL, T.C., DICKSON, D.W., and PETRUCELLI, L. (2013). Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol* 126**,** 895-905.

BENAJIBA, L., LE BER, I., CAMUZAT, A., LACOSTE, M., THOMAS-ANTERION, C., COURATIER, P., LEGALLIC, S., SALACHAS, F., HANNEQUIN, D., DECOUSUS, M., LACOMBLEZ, L., GUEDJ, E., GOLFIER, V., CAMU, W., DUBOIS, B., CAMPION, D., MEININGER, V., BRICE, A., FRENCH, C., and GENETIC RESEARCH NETWORK ON FRONTOTEMPORAL LOBAR DEGENERATION/FRONTOTEMPORAL LOBAR DEGENERATION WITH MOTONEURON, D. (2009). TARDBP mutations in motoneuron disease with frontotemporal lobar degeneration. *Annals of Neurology* 65**,** 470-473.

BENGTSON, M.H., and JOAZEIRO, C.A.P. (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467**,** 470-473.

BENJAMINI, Y., and HOCHBERG, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 57**,** 289-300.

BIGIO, E.H., WEINTRAUB, S., RADEMAKERS, R., BAKER, M., AHMADIAN, S.S., RADEMAKER, A., WEITNER, B.B., MAO, Q., LEE, K.H., MISHRA, M., GANTI, R.A., and MESULAM, M.M. (2013). Frontotemporal lobar degeneration with TDP-43

proteinopathy and chromosome 9p repeat expansion in C9ORF72: clinicopathologic correlation. *Neuropathology* 33**,** 122-133.

BLANC, R.S., and RICHARD, S. (2017). Arginine Methylation: The Coming of Age. *Mol Cell* 65**,** 8-24.

BLASCO, H., BERNARD-MARISSAL, N., VOURC'H, P., GUETTARD, Y., SUNYACH, C., AUGEREAU, O., KHEDERCHAH, J., MOUZAT, K., ANTAR, C., GORDON, P., VEYRAT-DUREBEX, C., BESSON, G., ANDERSEN, P., SALACHAS, F., MEININGER, V., CAMU, W., PETTMANN, B., ANDRES, C., and CORCIA, P. (2013). A rare motor neuron deleterious missense mutation in the DPYSL3 (CRMP4) gene is associated with ALS. *Human mutation* 34.

BOEYNAEMS, S., BOGAERT, E., KOVACS, D., KONIJNENBERG, A., TIMMERMAN, E., VOLKOV, A., GUHAROY, M., DE DECKER, M., JASPERS, T., RYAN, V.H., JANKE, A.M., BAATSEN, P., VERCRUYSSE, T., KOLAITIS, R.-M., DAELEMANS, D., TAYLOR, J.P., KEDERSHA, N., ANDERSON, P., IMPENS, F., SOBOTT, F., SCHYMKOWITZ, J., ROUSSEAU, F., FAWZI, N.L., ROBBERECHT, W., VAN DAMME, P., TOMPA, P., and VAN DEN BOSCH, L. (2017). Phase Separation of C9orf72 Dipeptide Repeats Perturbs Stress Granule Dynamics. *Molecular Cell* 65**,** 1044-1055.e1045.

BOEYNAEMS, S., BOGAERT, E., MICHIELS, E., GIJSELINCK, I., SIEBEN, A., JOVICIC, A., DE BAETS, G., SCHEVENEELS, W., STEYAERT, J., CUIJT, I., VERSTREPEN, K.J., CALLAERTS, P., ROUSSEAU, F., SCHYMKOWITZ, J., CRUTS, M., VAN BROECKHOVEN, C., VAN DAMME, P., GITLER, A.D., ROBBERECHT, W., and VAN DEN BOSCH, L. (2016). Drosophila screen connects nuclear transport genes to DPR pathology in c9ALS/FTD. *Sci Rep* 6**,** 20877.

BOISVERT, F.-M., VAN KONINGSBRUGGEN, S., NAVASCUÉS, J., and LAMOND, A.I. (2007). The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology* 8**,** 574-585.

BOLAND, B., YU, W.H., CORTI, O., MOLLEREAU, B., HENRIQUES, A., BEZARD, E., PASTORES, G.M., RUBINSZTEIN, D.C., NIXON, R.A., DUCHEN, M.R., MALLUCCI, G.R., KROEMER, G., LEVINE, B., ESKELINEN, E.L., MOCHEL, F., SPEDDING, M., LOUIS, C., MARTIN, O.R., and MILLAN, M.J. (2018). Promoting the clearance of neurotoxic proteins in neurodegenerative disorders of ageing. *Nat Rev Drug Discov* 17**,** 660-688.

BORRONI, B., BONVICINI, C., ALBERICI, A., BURATTI, E., AGOSTI, C., ARCHETTI, S., PAPETTI, A., STUANI, C., DI LUCA, M., GENNARELLI, M., and PADOVANI, A. (2009). Mutation within TARDBP leads to Frontotemporal Dementia without motor neuron disease. *Human Mutation* 30**,** E974-E983.

BOXER, A.L., MACKENZIE, I.R., BOEVE, B.F., BAKER, M., SEELEY, W.W., CROOK, R., FELDMAN, H., HSIUNG, G.-Y.R., RUTHERFORD, N., LALUZ, V., WHITWELL, J., FOTI, D., MCDADE, E., MOLANO, J., KARYDAS, A., WOJTAS, A., GOLDMAN, J., MIRSKY, J., SENGDY, P., DEARMOND, S., MILLER, B.L., and RADEMAKERS, R. (2011). Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *Journal of neurology, neurosurgery, and psychiatry* 82**,** 196-203.

BRANDMAN, O., and HEGDE, R.S. (2016). Ribosome-associated protein quality control. *Nat Struct Mol Biol* 23**,** 7-15.

BRANGWYNNE, CLIFFORD P., TOMPA, P., and PAPPU, ROHIT V. (2015). Polymer physics of intracellular phase transitions. *Nature Physics* 11**,** 899-904.

BRETTSCHNEIDER, J., TOLEDO, J.B., VAN DEERLIN, V.M., ELMAN, L., MCCLUSKEY, L., LEE, V.M.Y., and TROJANOWSKI, J.Q. (2012). Microglial activation correlates with disease progression and upper motor neuron clinical symptoms in amyotrophic lateral sclerosis. *PloS one* 7**,** e39216-e39216.

BRUMAGHIM, J.L., LI, Y., HENLE, E., and LINN, S. (2003). Effects of Hydrogen Peroxide upon Nicotinamide Nucleotide Metabolism in Escherichia coli: CHANGES IN ENZYME LEVELS AND NICOTINAMIDE NUCLEOTIDE POOLS AND STUDIES OF

THE OXIDATION OF NAD(P)H BY Fe(III). *Journal of Biological Chemistry* 278**,** 42495-42504.

BULAU, P., ZAKRZEWICZ, D., KITOWSKA, K., WARDEGA, B., KREUDER, J., and EICKELBERG, O. (2006). Quantitative assessment of arginine methylation in free versus protein-incorporated amino acids in vitro and in vivo using protein hydrolysis and high-performance liquid chromatography. *BioTechniques* 40**,** 305-310.

BURBERRY, A., SUZUKI, N., WANG, J.Y., MOCCIA, R., MORDES, D.A., STEWART, M.H., SUZUKI-UEMATSU, S., GHOSH, S., SINGH, A., MERKLE, F.T., KOSZKA, K., LI, Q.Z., ZON, L., ROSSI, D.J., TROWBRIDGE, J.J., NOTARANGELO, L.D., and EGGAN, K. (2016). Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci Transl Med* 8**,** 347ra393.

BURRELL, J.R., HALLIDAY, G.M., KRIL, J.J., ITTNER, L.M., GÖTZ, J., KIERNAN, M.C., and HODGES, J.R. (2016). The frontotemporal dementia-motor neuron disease continuum. *The Lancet* 388**,** 919-931.

CHANG, F.N., NAVICKAS, I.J., CHANG, C.N., and DANCIS, B.M. (1976). Methylation of ribosomal proteins in HeLa cells. *Arch Biochem Biophys* 172**,** 627-633.

CHANG, Y.-J., JENG, U.S., CHIANG, Y.-L., HWANG, I.-S., and CHEN, Y.-R. (2016). The Glycine-Alanine Dipeptide Repeat from C9orf72 Hexanucleotide Expansions Forms Toxic Amyloids Possessing Cell-to-Cell Transmission Properties. *The Journal of biological chemistry* 291**,** 4903-4911.

CHARCOT, J.M., and JOFFORY, A. (1869). Deux cas d'atrophie musculaire progressive avec lesions de la substance grise et des faisceaux antero-lateraux de la moelle epiniere. *Arch. Physiol. Neurol. Pathol* 2**,** 744-754.

CHEN, Y., LIN, Z., CHEN, X., CAO, B., WEI, Q., OU, R., ZHAO, B., SONG, W., WU, Y., and SHANG, H.F. (2016). Large C9orf72 repeat expansions are seen in Chinese patients with sporadic amyotrophic lateral sclerosis. *Neurobiol Aging* 38**,** 217.e215-217.e222.

CHENG, W., WANG, S., MESTRE, A.A., FU, C., MAKAREM, A., XIAN, F., HAYES, L.R., LOPEZ-GONZALEZ, R., DRENNER, K., JIANG, J., CLEVELAND, D.W., and SUN, S. (2018). C9ORF72 GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2α phosphorylation. *Nat Commun* 9**,** 51.

CHEW, J., GENDRON, T.F., PRUDENCIO, M., SASAGURI, H., ZHANG, Y.-J., CASTANEDES-CASEY, M., LEE, C.W., JANSEN-WEST, K., KURTI, A., MURRAY, M.E., BIENIEK, K.F., BAUER, P.O., WHITELAW, E.C., ROUSSEAU, L., STANKOWSKI, J.N., STETLER, C., DAUGHRITY, L.M., PERKERSON, E.A., DESARO, P., JOHNSTON, A., OVERSTREET, K., EDBAUER, D., RADEMAKERS, R., BOYLAN, K.B., DICKSON, D.W., FRYER, J.D., and PETRUCELLI, L. (2015). Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science (New York, N.Y.)* 348**,** 1151-1154.

CHIÒ, A., LOGROSCINO, G., TRAYNOR, B.J., COLLINS, J., SIMEONE, J.C., GOLDSTEIN, L.A., and WHITE, L.A. (2013). Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature. *Neuroepidemiology* 41**,** 118-130.

CHITIPROLU, M., JAGOW, C., TREMBLAY, V., BONDY-CHORNEY, E., PARIS, G., SAVARD, A., PALIDWOR, G., BARRY, F.A., ZINMAN, L., KEITH, J., ROGAEVA, E., ROBERTSON, J., LAVALLEE-ADAM, M., WOULFE, J., COUTURE, J.F., COTE, J., and GIBBINGS, D. (2018). A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. *Nat Commun* 9**,** 2794.

CHU, J., HONG, N.A., MASUDA, C.A., JENKINS, B.V., NELMS, K.A., GOODNOW, C.C., GLYNNE, R.J., WU, H., MASLIAH, E., JOAZEIRO, C.A.P., and KAY, S.A. (2009). A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America* 106**,** 2097-2103.

CIURA, S., LATTANTE, S., LE BER, I., LATOUCHE, M., TOSTIVINT, H., BRICE, A., and KABASHI, E. (2013). Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann Neurol* 74**,** 180-187.

CLEARY, J.D., and RANUM, L.P. (2014). Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. *Curr Opin Genet Dev* 26**,** 6-15.

CONLON, E.G., LU, L., SHARMA, A., YAMAZAKI, T., TANG, T., SHNEIDER, N.A., and MANLEY, J.L. (2016). The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *Elife* 5.

COOPER-KNOCK, J., BURY, J.J., HEATH, P.R., WYLES, M., HIGGINBOTTOM, A., GELSTHORPE, C., HIGHLEY, J.R., HAUTBERGUE, G., RATTRAY, M., KIRBY, J., and SHAW, P.J. (2015). C9ORF72 GGGGCC Expanded Repeats Produce Splicing Dysregulation which Correlates with Disease Severity in Amyotrophic Lateral Sclerosis. *PLOS ONE* 10**,** e0127376.

COOPER-KNOCK, J., HEWITT, C., HIGHLEY, J.R., BROCKINGTON, A., MILANO, A., MAN, S., MARTINDALE, J., HARTLEY, J., WALSH, T., GELSTHORPE, C., BAXTER, L., FORSTER, G., FOX, M., BURY, J., MOK, K., MCDERMOTT, C.J., TRAYNOR, B.J., KIRBY, J., WHARTON, S.B., INCE, P.G., HARDY, J., and SHAW, P.J. (2012). Clinico-pathological features in amyotrophic lateral sclerosis with expansions in C9ORF72. *Brain* 135**,** 751-764.

COOPER-KNOCK, J., WALSH, M.J., HIGGINBOTTOM, A., ROBIN HIGHLEY, J., DICKMAN, M.J., EDBAUER, D., INCE, P.G., WHARTON, S.B., WILSON, S.A., KIRBY, J., HAUTBERGUE, G.M., and SHAW, P.J. (2014). Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* 137**,** 2040-2051.

COX, D., RAEBURN, C., SUI, X., and HATTERS, D.M. (2018). Protein aggregation in cell biology: An aggregomics perspective of health and disease. *Semin Cell Dev Biol*.

CRISTOFANI, R., CRIPPA, V., RUSMINI, P., CICARDI, M.E., MERONI, M., LICATA, N.V., SALA, G., GIORGETTI, E., GRUNSEICH, C., GALBIATI, M., PICCOLELLA, M., MESSI, E., FERRARESE, C., CARRA, S., and POLETTI, A. (2017). Inhibition of retrograde transport modulates misfolded protein accumulation and clearance in motoneuron diseases. *Autophagy* 13**,** 1280-1303.

CRUTS, M., GIJSELINCK, I., VAN LANGENHOVE, T., VAN DER ZEE, J., and VAN BROECKHOVEN, C. (2013). Current insights into the C9orf72 repeat expansion diseases of the FTLD/ALS spectrum. *Trends Neurosci* 36**,** 450-459.

DE SANTIS, R., SANTINI, L., COLANTONI, A., PERUZZI, G., DE TURRIS, V., ALFANO, V., BOZZONI, I., and ROSA, A. (2017). FUS Mutant Human Motoneurons Display Altered Transcriptome and microRNA Pathways with Implications for ALS Pathogenesis. *Stem Cell Reports* 9**,** 1450-1462.

DEJESUS-HERNANDEZ, M., MACKENZIE, I.R., BOEVE, B.F., BOXER, A.L., BAKER, M., RUTHERFORD, N.J., NICHOLSON, A.M., FINCH, N.A., FLYNN, H., ADAMSON, J., KOURI, N., WOJTAS, A., SENGDY, P., HSIUNG, G.Y., KARYDAS, A., SEELEY, W.W., JOSEPHS, K.A., COPPOLA, G., GESCHWIND, D.H., WSZOLEK, Z.K., FELDMAN, H., KNOPMAN, D.S., PETERSEN, R.C., MILLER, B.L., DICKSON, D.W., BOYLAN, K.B., GRAFF-RADFORD, N.R., and RADEMAKERS, R. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72**,** 245-256.

DENG, H.-X., CHEN, W., HONG, S.-T., BOYCOTT, K.M., GORRIE, G.H., SIDDIQUE, N., YANG, Y., FECTO, F., SHI, Y., ZHAI, H., JIANG, H., HIRANO, M., RAMPERSAUD, E., JANSEN, G.H., DONKERVOORT, S., BIGIO, E.H., BROOKS, B.R., AJROUD, K., SUFIT, R.L., HAINES, J.L., MUGNAINI, E., PERICAK-VANCE, M.A., and SIDDIQUE, T. (2011). Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 477**,** 211-215.

DÉRY, U., COULOMBE, Y., RODRIGUE, A., STASIAK, A., RICHARD, S., and MASSON, J.-Y. (2008). A glycine-arginine domain in control of the human MRE11 DNA repair protein. *Molecular and cellular biology* 28**,** 3058-3069.

DIFIGLIA, M., SAPP, E., CHASE, K.O., DAVIES, S.W., BATES, G.P., VONSATTEL, J.P., and ARONIN, N. (1997). Aggregation of huntingtin in neuronal intranuclear inclusions and dystrophic neurites in brain. *Science* 277**,** 1990-1993.

DOI, H., KOYANO, S., SUZUKI, Y., NUKINA, N., and KUROIWA, Y. (2010). The RNA-binding protein FUS/TLS is a common aggregate-interacting protein in polyglutamine diseases. *Neuroscience Research* 66**,** 131-133.

DOI, H., OKAMURA, K., BAUER, P.O., FURUKAWA, Y., SHIMIZU, H., KUROSAWA, M., MACHIDA, Y., MIYAZAKI, H., MITSUI, K., KUROIWA, Y., and NUKINA, N. (2008). RNA-binding Protein TLS Is a Major Nuclear Aggregate-interacting Protein in Huntingtin Exon 1 with Expanded Polyglutamine-expressing Cells. *Journal of Biological Chemistry* 283**,** 6489-6500.

DOLS-ICARDO, O., GARCÍA-REDONDO, A., ROJAS-GARCÍA, R., SÁNCHEZ-VALLE, R., NOGUERA, A., GÓMEZ-TORTOSA, E., PASTOR, P., HERNÁNDEZ, I., ESTEBAN-PÉREZ, J., SUÁREZ-CALVET, M., ANTÓN-AGUIRRE, S., AMER, G., ORTEGA-CUBERO, S., BLESA, R., FORTEA, J., ALCOLEA, D., CAPDEVILA, A., ANTONELL, A., LLADÓ, A., MUÑOZ-BLANCO, J.L., MORA, J.S., GALÁN-DÁVILA, L., RODRÍGUEZ DE RIVERA, F.J., LLEÓ, A., and CLARIMÓN, J. (2013). Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. *Human Molecular Genetics* 23**,** 749-754.

DONNELLY, C.J., ZHANG, P.W., PHAM, J.T., HAEUSLER, A.R., MISTRY, N.A., VIDENSKY, S., DALEY, E.L., POTH, E.M., HOOVER, B., FINES, D.M., MARAGAKIS, N., TIENARI, P.J., PETRUCELLI, L., TRAYNOR, B.J., WANG, J., RIGO, F., BENNETT, C.F., BLACKSHAW, S., SATTLER, R., and ROTHSTEIN, J.D. (2013). RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* 80**,** 415-428.

DORMANN, D., MADL, T., VALORI, C.F., BENTMANN, E., TAHIROVIC, S., ABOU-AJRAM, C., KREMMER, E., ANSORGE, O., MACKENZIE, I.R., NEUMANN, M., and HAASS, C. (2012). Arginine methylation next to the PY-NLS modulates Transportin binding and nuclear import of FUS. *EMBO J* 31**,** 4258-4275.

DOSZTÁNYI, Z., CSIZMOK, V., TOMPA, P., and SIMON, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* 21**,** 3433-3434.

DOTY, K.R., GUILLOT-SESTIER, M.V., and TOWN, T. (2015). The role of the immune system in neurodegenerative disorders: Adaptive or maladaptive? *Brain Res* 1617**,** 155-173.

DREYFUSS, G., MATUNIS, M.J., PINOL-ROMA, S., and BURD, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. *Annu Rev Biochem* 62**,** 289-321.

DUYAO, M., AMBROSE, C., MYERS, R., NOVELLETTO, A., PERSICHETTI, F., FRONTALI, M., FOLSTEIN, S., ROSS, C., FRANZ, M., ABBOTT, M., and ET AL. (1993). Trinucleotide repeat length instability and age of onset in Huntington's disease. *Nat Genet* 4**,** 387-392.

EGAWA, N., KITAOKA, S., TSUKITA, K., NAITOH, M., TAKAHASHI, K., YAMAMOTO, T., ADACHI, F., KONDO, T., OKITA, K., ASAKA, I., AOI, T., WATANABE, A., YAMADA, Y., MORIZANE, A., TAKAHASHI, J., AYAKI, T., ITO, H., YOSHIKAWA, K., YAMAWAKI, S., SUZUKI, S., WATANABE, D., HIOKI, H., KANEKO, T., MAKIOKA, K., OKAMOTO, K., TAKUMA, H., TAMAOKA, A., HASEGAWA, K., NONAKA, T., HASEGAWA, M., KAWATA, A., YOSHIDA, M., NAKAHATA, T., TAKAHASHI, R., MARCHETTO, M.C., GAGE, F.H., YAMANAKA, S., and INOUE, H. (2012). Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci Transl Med* 4**,** 145ra104.

EIRA, J., SILVA, C.S., SOUSA, M.M., and LIZ, M.A. (2016). The cytoskeleton as a novel therapeutic target for old neurodegenerative disorders. *Progress in Neurobiology* 141**,** 61-82.

ELIAS-SONNENSCHEIN, L.S., BERTRAM, L., and VISSER, P.J. (2012). Relationship between genetic risk factors and markers for Alzheimer's disease pathology. *Biomark Med* 6**,** 477-495.

FARG, M.A., KONOPKA, A., SOO, K.Y., ITO, D., and ATKIN, J.D. (2017). The DNA damage response (DDR) is induced by the C9orf72 repeat expansion in amyotrophic lateral sclerosis. *Hum Mol Genet* 26**,** 2882-2896.

FARG, M.A., SUNDARAMOORTHY, V., SULTANA, J.M., YANG, S., ATKINSON, R.A.K., LEVINA, V., HALLORAN, M.A., GLEESON, P.A., BLAIR, I.P., SOO, K.Y., KING, A.E., and ATKIN, J.D. (2014). C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Human molecular genetics* 23**,** 3579-3595.

FAY, M.M., ANDERSON, P.J., and IVANOV, P. (2017). ALS/FTD-Associated C9ORF72 Repeat RNA Promotes Phase Transitions In Vitro and in Cells. *Cell Rep* 21**,** 3573-3584.

FLORES, B.N., DULCHAVSKY, M.E., KRANS, A., SAWAYA, M.R., PAULSON, H.L., TODD, P.K., BARMADA, S.J., and IVANOVA, M.I. (2016). Distinct C9orf72-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS One* 11**,** e0165084.

FORMAN, M.S., MACKENZIE, I.R., CAIRNS, N.J., SWANSON, E., BOYER, P.J., DRACHMAN, D.A., JHAVERI, B.S., KARLAWISH, J.H., PESTRONK, A., SMITH, T.W., TU, P.-H., WATTS, G.D.J., MARKESBERY, W.R., SMITH, C.D., and KIMONIS, V.E. (2006). Novel Ubiquitin Neuropathology in Frontotemporal Dementia With Valosin-Containing Protein Gene Mutations. *Journal of Neuropathology & Experimental Neurology* 65**,** 571-581.

FRATTA, P., POULTER, M., LASHLEY, T., ROHRER, J.D., POLKE, J.M., BECK, J., RYAN, N., HENSMAN, D., MIZIELINSKA, S., WAITE, A.J., LAI, M.C., GENDRON, T.F., PETRUCELLI, L., FISHER, E.M., REVESZ, T., WARREN, J.D., COLLINGE, J., ISAACS, A.M., and MEAD, S. (2013). Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. *Acta Neuropathol* 126**,** 401-409.

FREIBAUM, B.D., LU, Y., LOPEZ-GONZALEZ, R., KIM, N.C., ALMEIDA, S., LEE, K.-H., BADDERS, N., VALENTINE, M., MILLER, B.L., WONG, P.C., PETRUCELLI, L., KIM, H.J., GAO, F.-B., and TAYLOR, J.P. (2015). GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* 525**,** 129-133.

FREIBAUM, B.D., and TAYLOR, J.P. (2017). The Role of Dipeptide Repeats in C9ORF72-Related ALS-FTD. *Front Mol Neurosci* 10**,** 35.

FREISCHMIDT, A., WIELAND, T., RICHTER, B., RUF, W., SCHAEFFER, V., MÜLLER, K., MARROQUIN, N., NORDIN, F., HÜBERS, A., WEYDT, P., PINTO, S., PRESS, R., MILLECAMPS, S., MOLKO, N., BERNARD, E., DESNUELLE, C., SORIANI, M.-H., DORST, J., GRAF, E., NORDSTRÖM, U., FEILER, M.S., PUTZ, S., BOECKERS, T.M., MEYER, T., WINKLER, A.S., WINKELMAN, J., DE CARVALHO, M., THAL, D.R., OTTO, M., BRÄNNSTRÖM, T., VOLK, A.E., KURSULA, P., DANZER, K.M., LICHTNER, P., DIKIC, I., MEITINGER, T., LUDOLPH, A.C., STROM, T.M., ANDERSEN, P.M., and WEISHAUPT, J.H. (2015). Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nature Neuroscience* 18**,** 631-636.

GAGNON, M.G., ROY, R.N., LOMAKIN, I.B., FLORIN, T., MANKIN, A.S., and STEITZ, T.A. (2016). Structures of proline-rich peptides bound to the ribosome reveal a common mechanism of protein synthesis inhibition. *Nucleic Acids Res* 44**,** 2439-2450.

GAMI, P., MURRAY, C., SCHOTTLAENDER, L., BETTENCOURT, C., DE PABLO FERNANDEZ, E., MUDANOHWO, E., MIZIELINSKA, S., POLKE, J.M., HOLTON, J.L., ISAACS, A.M., HOULDEN, H., REVESZ, T., and LASHLEY, T. (2015). A 30-unit

hexanucleotide repeat expansion in C9orf72 induces pathological lesions with dipeptide-repeat proteins and RNA foci, but not TDP-43 inclusions and clinical disease. *Acta Neuropathologica* 130**,** 599-601.

GARCIA-REDONDO, A., DOLS-ICARDO, O., ROJAS-GARCIA, R., ESTEBAN-PEREZ, J., CORDERO-VAZQUEZ, P., MUNOZ-BLANCO, J.L., CATALINA, I., GONZALEZ-MUNOZ, M., VARONA, L., SARASOLA, E., POVEDANO, M., SEVILLA, T., GUERRERO, A., PARDO, J., LOPEZ DE MUNAIN, A., MARQUEZ-INFANTE, C., DE RIVERA, F.J., PASTOR, P., JERICO, I., DE ARCAYA, A.A., MORA, J.S., CLARIMON, J., GONZALO-MARTINEZ, J.F., JUAREZ-RUFIAN, A., ATENCIA, G., JIMENEZ-BAUTISTA, R., MORAN, Y., MASCIAS, J., HERNANDEZ-BARRAL, M., KAPETANOVIC, S., GARCIA-BARCINA, M., ALCALA, C., VELA, A., RAMIREZ-RAMOS, C., GALAN, L., PEREZ-TUR, J., QUINTANS, B., SOBRIDO, M.J., FERNANDEZ-TORRON, R., POZA, J.J., GOROSTIDI, A., PARADAS, C., VILLOSLADA, P., LARRODE, P., CAPABLO, J.L., PASCUAL-CALVET, J., GONI, M., MORGADO, Y., GUITART, M., MORENO-LAGUNA, S., RUEDA, A., MARTIN-ESTEFANIA, C., CEMILLAN, C., BLESA, R., and LLEO, A. (2013). Analysis of the C9orf72 gene in patients with amyotrophic lateral sclerosis in Spain and different populations worldwide. *Hum Mutat* 34**,** 79-82.

GENDRON, T.F., BIENIEK, K.F., ZHANG, Y.J., JANSEN-WEST, K., ASH, P.E., CAULFIELD, T., DAUGHRITY, L., DUNMORE, J.H., CASTANEDES-CASEY, M., CHEW, J., COSIO, D.M., VAN BLITTERSWIJK, M., LEE, W.C., RADEMAKERS, R., BOYLAN, K.B., DICKSON, D.W., and PETRUCELLI, L. (2013). Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol* 126**,** 829-844.

GENDRON, T.F., CHEW, J., STANKOWSKI, J.N., HAYES, L.R., ZHANG, Y.J., PRUDENCIO, M., CARLOMAGNO, Y., DAUGHRITY, L.M., JANSEN-WEST, K., PERKERSON, E.A., O'RAW, A., COOK, C., PREGENT, L., BELZIL, V., VAN BLITTERSWIJK, M., TABASSIAN, L.J., LEE, C.W., YUE, M., TONG, J., SONG, Y., CASTANEDES-CASEY, M., ROUSSEAU, L., PHILLIPS, V., DICKSON, D.W., RADEMAKERS, R., FRYER, J.D., RUSH, B.K., PEDRAZA, O., CAPUTO, A.M., DESARO, P., PALMUCCI, C., ROBERTSON, A., HECKMAN, M.G., DIEHL, N.N., WIGGS, E., TIERNEY, M., BRAUN, L., FARREN, J., LACOMIS, D., LADHA, S., FOURNIER, C.N., MCCLUSKEY, L.F., ELMAN, L.B., TOLEDO, J.B., MCBRIDE, J.D., TILOCA, C., MORELLI, C., POLETTI, B., SOLCA, F., PRELLE, A., WUU, J., JOCKEL-BALSAROTTI, J., RIGO, F., AMBROSE, C., DATTA, A., YANG, W., RAITCHEVA, D., ANTOGNETTI, G., MCCAMPBELL, A., VAN SWIETEN, J.C., MILLER, B.L., BOXER, A.L., BROWN, R.H., BOWSER, R., MILLER, T.M., TROJANOWSKI, J.Q., GROSSMAN, M., BERRY, J.D., HU, W.T., RATTI, A., TRAYNOR, B.J., DISNEY, M.D., BENATAR, M., SILANI, V., GLASS, J.D., FLOETER, M.K., ROTHSTEIN, J.D., BOYLAN, K.B., and PETRUCELLI, L. (2017). Poly(GP) proteins are a useful pharmacodynamic marker for C9ORF72-associated amyotrophic lateral sclerosis. *Sci Transl Med* 9.

GENDRON, T.F., VAN BLITTERSWIJK, M., BIENIEK, K.F., DAUGHRITY, L.M., JIANG, J., RUSH, B.K., PEDRAZA, O., LUCAS, J.A., MURRAY, M.E., DESARO, P., ROBERTSON, A., OVERSTREET, K., THOMAS, C.S., CROOK, J.E., CASTANEDES-CASEY, M., ROUSSEAU, L., JOSEPHS, K.A., PARISI, J.E., KNOPMAN, D.S., PETERSEN, R.C., BOEVE, B.F., GRAFF-RADFORD, N.R., RADEMAKERS, R., LAGIER-TOURENNE, C., EDBAUER, D., CLEVELAND, D.W., DICKSON, D.W., PETRUCELLI, L., and BOYLAN, K.B. (2015). Cerebellar c9RAN proteins associate with clinical and neuropathological characteristics of C9ORF72 repeat expansion carriers. *Acta Neuropathol* 130**,** 559-573.

GESER, F., LEE, V.M.Y., and TROJANOWSKI, J.Q. (2010). Amyotrophic lateral sclerosis and frontotemporal lobar degeneration: a spectrum of TDP-43 proteinopathies.

Neuropathology : official journal of the Japanese Society of Neuropathology 30**,** 103-112.

GHOSH, S., and LIPPA, C.F. (2013). Clinical Subtypes of Frontotemporal Dementia. *American Journal of Alzheimer's Disease & Other Dementiasr* 30**,** 653-661.

GIAMPETRUZZI, A., DANIELSON, E.W., GUMINA, V., JEON, M., BOOPATHY, S., BROWN, R.H., RATTI, A., LANDERS, J.E., and FALLINI, C. (2019a). Modulation of actin polymerization affects nucleocytoplasmic transport in multiple forms of amyotrophic lateral sclerosis. *Nat Commun* 10**,** 3827.

GIAMPETRUZZI, A., DANIELSON, E.W., GUMINA, V., JEON, M., BOOPATHY, S., BROWN, R.H., RATTI, A., LANDERS, J.E., and FALLINI, C. (2019b). Modulation of actin polymerization affects nucleocytoplasmic transport in multiple forms of amyotrophic lateral sclerosis. *Nature Communications* 10**,** 3827.

GIJSELINCK, I., CRUTS, M., and VAN BROECKHOVEN, C. (2018). The Genetics of C9orf72 Expansions. *Cold Spring Harbor perspectives in medicine* 8**,** a026757.

GIJSELINCK, I., ENGELBORGHS, S., MAES, G., CUIJT, I., PEETERS, K., MATTHEIJSSENS, M., JORIS, G., CRAS, P., MARTIN, J.-J., DE DEYN, P.P., KUMAR-SINGH, S., VAN BROECKHOVEN, C., and CRUTS, M. (2010). Identification of 2 Loci at Chromosomes 9 and 14 in a Multiplex Family With Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Archives of Neurology* 67**,** 606-616.

GIJSELINCK, I., VAN LANGENHOVE, T., VAN DER ZEE, J., SLEEGERS, K., PHILTJENS, S., KLEINBERGER, G., JANSSENS, J., BETTENS, K., VAN CAUWENBERGHE, C., PERESON, S., ENGELBORGHS, S., SIEBEN, A., DE JONGHE, P., VANDENBERGHE, R., SANTENS, P., DE BLEECKER, J., MAES, G., BÄUMER, V., DILLEN, L., JORIS, G., CUIJT, I., CORSMIT, E., ELINCK, E., VAN DONGEN, J., VERMEULEN, S., VAN DEN BROECK, M., VAERENBERG, C., MATTHEIJSSENS, M., PEETERS, K., ROBBERECHT, W., CRAS, P., MARTIN, J.-J., DE DEYN, P.P., CRUTS, M., and VAN BROECKHOVEN, C. (2012). A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *The Lancet Neurology* 11**,** 54-65.

GITLER, A.D., and TSUIJI, H. (2016). There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS. *Brain Res* 1647**,** 19-29.

GITTINGS, L.M., BOEYNAEMS, S., LIGHTWOOD, D., CLARGO, A., TOPIA, S., NAKAYAMA, L., TROAKES, C., MANN, D.M.A., GITLER, A.D., LASHLEY, T., and ISAACS, A.M. (2020). Symmetric dimethylation of poly-GR correlates with disease duration in C9orf72 FTLD and ALS and reduces poly-GR phase separation and toxicity. *Acta Neuropathologica* 139**,** 407-410.

GOMEZ-TORTOSA, E., GALLEGO, J., GUERRERO-LOPEZ, R., MARCOS, A., GIL-NECIGA, E., SAINZ, M.J., DIAZ, A., FRANCO-MACIAS, E., TRUJILLO-TIEBAS, M.J., AYUSO, C., and PEREZ-PEREZ, J. (2013). C9ORF72 hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. *Neurology* 80**,** 366-370.

GREEN, K.M., GLINEBURG, M.R., KEARSE, M.G., FLORES, B.N., LINSALATA, A.E., FEDAK, S.J., GOLDSTROHM, A.C., BARMADA, S.J., and TODD, P.K. (2017). RAN translation at C9orf72-associated repeat expansions is selectively enhanced by the integrated stress response. *Nat Commun* 8**,** 2005.

GROSSE, R., COPELAND, J.W., NEWSOME, T.P., WAY, M., and TREISMAN, R. (2003). A role for VASP in RhoA-Diaphanous signalling to actin dynamics and SRF activity. *EMBO J* 22**,** 3050-3061.

HAEUSLER, A.R., DONNELLY, C.J., PERIZ, G., SIMKO, E.A., SHAW, P.G., KIM, M.S., MARAGAKIS, N.J., TRONCOSO, J.C., PANDEY, A., SATTLER, R., ROTHSTEIN, J.D., and WANG, J. (2014). C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* 507**,** 195-200.

HAEUSLER, A.R., DONNELLY, C.J., and ROTHSTEIN, J.D. (2016). The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. *Nature Reviews Neuroscience* 17**,** 383-395.

HARDIMAN, O., AL-CHALABI, A., CHIO, A., CORR, E.M., LOGROSCINO, G., ROBBERECHT, W., SHAW, P.J., SIMMONS, Z., and VAN DEN BERG, L.H. (2017). Amyotrophic lateral sclerosis. *Nature Reviews Disease Primers* 3**,** 17071.

HARDIMAN, O., and VAN DEN BERG, L.H. (2017). Edaravone: a new treatment for ALS on the horizon? *The Lancet Neurology* 16**,** 490-491.

HARMS, M.B., CADY, J., ZAIDMAN, C., COOPER, P., BALI, T., ALLRED, P., CRUCHAGA, C., BAUGHN, M., LIBBY, R.T., PESTRONK, A., GOATE, A., RAVITS, J., and BALOH, R.H. (2013). Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. *Neurobiology of aging* 34**,** 2234.e2213-2234.e2232.2234E2219.

HAROLD, D., ABRAHAM, R., HOLLINGWORTH, P., SIMS, R., GERRISH, A., HAMSHERE, M.L., PAHWA, J.S., MOSKVINA, V., DOWZELL, K., WILLIAMS, A., JONES, N., THOMAS, C., STRETTON, A., MORGAN, A.R., LOVESTONE, S., POWELL, J., PROITSI, P., LUPTON, M.K., BRAYNE, C., RUBINSZTEIN, D.C., GILL, M., LAWLOR, B., LYNCH, A., MORGAN, K., BROWN, K.S., PASSMORE, P.A., CRAIG, D., MCGUINNESS, B., TODD, S., HOLMES, C., MANN, D., SMITH, A.D., LOVE, S., KEHOE, P.G., HARDY, J., MEAD, S., FOX, N., ROSSOR, M., COLLINGE, J., MAIER, W., JESSEN, F., SCHURMANN, B., HEUN, R., VAN DEN BUSSCHE, H., HEUSER, I., KORNHUBER, J., WILTFANG, J., DICHGANS, M., FROLICH, L., HAMPEL, H., HULL, M., RUJESCU, D., GOATE, A.M., KAUWE, J.S., CRUCHAGA, C., NOWOTNY, P., MORRIS, J.C., MAYO, K., SLEEGERS, K., BETTENS, K., ENGELBORGHS, S., DE DEYN, P.P., VAN BROECKHOVEN, C., LIVINGSTON, G., BASS, N.J., GURLING, H., MCQUILLIN, A., GWILLIAM, R., DELOUKAS, P., AL-CHALABI, A., SHAW, C.E., TSOLAKI, M., SINGLETON, A.B., GUERREIRO, R., MUHLEISEN, T.W., NOTHEN, M.M., MOEBUS, S., JOCKEL, K.H., KLOPP, N., WICHMANN, H.E., CARRASQUILLO, M.M., PANKRATZ, V.S., YOUNKIN, S.G., HOLMANS, P.A., O'DONOVAN, M., OWEN, M.J., and WILLIAMS, J. (2009). Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet* 41**,** 1088-1093.

HARTMANN, H., HORNBURG, D., CZUPPA, M., BADER, J., MICHAELSEN, M., FARNY, D., ARZBERGER, T., MANN, M., MEISSNER, F., and EDBAUER, D. (2018). Proteomics and C9orf72 neuropathology identify ribosomes as poly-GR/PR interactors driving toxicity. *Life Science Alliance* 1.

HENSEL, N., and CLAUS, P. (2018). The Actin Cytoskeleton in SMA and ALS: How Does It Contribute to Motoneuron Degeneration? *Neuroscientist* 24**,** 54-72.

HIPP, M.S., PARK, S.H., and HARTL, F.U. (2014). Proteostasis impairment in protein-misfolding and -aggregation diseases. *Trends Cell Biol* 24**,** 506-514.

HIRANO, M., QUINZII, C.M., MITSUMOTO, H., HAYS, A.P., ROBERTS, J.K., RICHARD, P., and ROWLAND, L.P. (2011). Senataxin mutations and amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis* 12**,** 223-227.

HOGAN, D.B., JETTÉ, N., FIEST, K.M., ROBERTS, J.I., PEARSON, D., SMITH, E.E., ROACH, P., KIRK, A., PRINGSHEIM, T., and MAXWELL, C.J. (2016). The Prevalence and Incidence of Frontotemporal Dementia: a Systematic Review. *Canadian Journal of Neurological Sciences / Journal Canadien des Sciences Neurologiques* 43**,** S96-S109.

HOU, W., NEMITZ, S., SCHOPPER, S., NIELSEN, M.L., KESSELS, M.M., and QUALMANN, B. (2018). Arginine Methylation by PRMT2 Controls the Functions of the Actin Nucleator Cobl. *Dev Cell* 45**,** 262-275 e268.

HUNG, C.-J., LEE, Y.-J., CHEN, D.-H., and LI, C. (2009). Proteomic Analysis of Methylarginine-Containing Proteins in HeLa Cells by Two-Dimensional Gel Electrophoresis and Immunoblotting with a Methylarginine-Specific Antibody. *The Protein Journal* 28**,** 139-147.

HUTTON, M., LENDON, C.L., RIZZU, P., BAKER, M., FROELICH, S., HOULDEN, H., PICKERING-BROWN, S., CHAKRAVERTY, S., ISAACS, A., GROVER, A., HACKETT, J., ADAMSON, J., LINCOLN, S., DICKSON, D., DAVIES, P., PETERSEN, R.C., STEVENS, M., DE GRAAFF, E., WAUTERS, E., VAN BAREN, J., HILLEBRAND, M., JOOSSE, M., KWON, J.M., NOWOTNY, P., CHE, L.K., NORTON, J., MORRIS, J.C., REED, L.A., TROJANOWSKI, J., BASUN, H., LANNFELT, L., NEYSTAT, M., FAHN, S., DARK, F., TANNENBERG, T., DODD, P.R., HAYWARD, N., KWOK, J.B.J., SCHOFIELD, P.R., ANDREADIS, A., SNOWDEN, J., CRAUFURD, D., NEARY, D., OWEN, F., OOSTRA, B.A., HARDY, J., GOATE, A., VAN SWIETEN, J., MANN, D., LYNCH, T., and HEUTINK, P. (1998). Association of missense and 5′-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature* 393**,** 702-705.

ISHIMURA, R., NAGY, G., DOTU, I., ZHOU, H., YANG, X.-L., SCHIMMEL, P., SENJU, S., NISHIMURA, Y., CHUANG, J.H., and ACKERMAN, S.L. (2014). RNA function. Ribosome stalling induced by mutation of a CNS-specific tRNA causes neurodegeneration. *Science (New York, N.Y.)* 345**,** 455-459.

ITO-HARASHIMA, S., KUROHA, K., TATEMATSU, T., and INADA, T. (2007). Translation of the poly(A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes Dev* 21**,** 519-524.

JAIN, S., WHEELER, J.R., WALTERS, R.W., AGRAWAL, A., BARSIC, A., and PARKER, R. (2016). ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* 164**,** 487-498.

JIANG, J., ZHU, Q., GENDRON, T.F., SABERI, S., MCALONIS-DOWNES, M., SEELMAN, A., STAUFFER, J.E., JAFAR-NEJAD, P., DRENNER, K., SCHULTE, D., CHUN, S., SUN, S., LING, S.C., MYERS, B., ENGELHARDT, J., KATZ, M., BAUGHN, M., PLATOSHYN, O., MARSALA, M., WATT, A., HEYSER, C.J., ARD, M.C., DE MUYNCK, L., DAUGHRITY, L.M., SWING, D.A., TESSAROLLO, L., JUNG, C.J., DELPOUX, A., UTZSCHNEIDER, D.T., HEDRICK, S.M., DE JONG, P.J., EDBAUER, D., VAN DAMME, P., PETRUCELLI, L., SHAW, C.E., BENNETT, C.F., DA CRUZ, S., RAVITS, J., RIGO, F., CLEVELAND, D.W., and LAGIER-TOURENNE, C. (2016). Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron* 90**,** 535-550.

JOHNSON, J.O., MANDRIOLI, J., BENATAR, M., ABRAMZON, Y., VAN DEERLIN, V.M., TROJANOWSKI, J.Q., GIBBS, J.R., BRUNETTI, M., GRONKA, S., WUU, J., DING, J., MCCLUSKEY, L., MARTINEZ-LAGE, M., FALCONE, D., HERNANDEZ, D.G., AREPALLI, S., CHONG, S., SCHYMICK, J.C., ROTHSTEIN, J., LANDI, F., WANG, Y.-D., CALVO, A., MORA, G., SABATELLI, M., MONSURRÒ, M.R., BATTISTINI, S., SALVI, F., SPATARO, R., SOLA, P., BORGHERO, G., CONSORTIUM, I., GALASSI, G., SCHOLZ, S.W., TAYLOR, J.P., RESTAGNO, G., CHIÒ, A., and TRAYNOR, B.J. (2010). Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* 68**,** 857-864.

JOVIČIĆ, A., MERTENS, J., BOEYNAEMS, S., BOGAERT, E., CHAI, N., YAMADA, S.B., PAUL, J.W., SUN, S., HERDY, J.R., BIERI, G., KRAMER, N.J., GAGE, F.H., DEN BOSCH, L.V., ROBBERECHT, W., and GITLER, A.D. (2015). Modifiers of C9orf72 DPR toxicity implicate nucleocytoplasmic transport impairments in c9FTD/ALS. *Nature neuroscience* 18**,** 1226-1229.

JUN, M.-H., RYU, H.-H., JUN, Y.-W., LIU, T., LI, Y., LIM, C.-S., LEE, Y.-S., KAANG, B.-K., JANG, D.-J., and LEE, J.-A. (2017). Sequestration of PRMT1 and Nd1-L mRNA into ALS-linked FUS mutant R521C-positive aggregates contributes to neurite degeneration upon oxidative stress. *Scientific Reports* 7**,** 40474.

JUSZKIEWICZ, S., and HEGDE, R.S. (2017). Initiation of Quality Control during Poly(A) Translation Requires Site-Specific Ribosome Ubiquitination. *Mol Cell* 65**,** 743-750 e744.

KANEKURA, K., YAGI, T., CAMMACK, A.J., MAHADEVAN, J., KURODA, M., HARMS, M.B., MILLER, T.M., and URANO, F. (2016). Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. *Hum Mol Genet* 25**,** 1803-1813.

KANSAL, K., MAREDDY, M., SLOANE, K.L., MINC, A.A., RABINS, P.V., MCGREADY, J.B., and ONYIKE, C.U. (2016). Survival in Frontotemporal Dementia Phenotypes: A Meta-Analysis. *Dementia and Geriatric Cognitive Disorders* 41**,** 109-122.

KAPUR, M., MONAGHAN, C.E., and ACKERMAN, S.L. (2017). Regulation of mRNA Translation in Neurons-A Matter of Life and Death. *Neuron* 96**,** 616-637.

KARCHER, A., SCHELE, A., and HOPFNER, K.-P. (2008). X-ray Structure of the Complete ABC Enzyme ABCE1 from Pyrococcus abyssi. *Journal of Biological Chemistry* 283**,** 7962-7971.

KEARSE, M.G., GREEN, K.M., KRANS, A., RODRIGUEZ, C.M., LINSALATA, A.E., GOLDSTROHM, A.C., and TODD, P.K. (2016). CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Mol Cell* 62**,** 314-322.

KHOSRAVI, B., HARTMANN, H., MAY, S., MÖHL, C., EDERLE, H., MICHAELSEN, M., SCHLUDI, M.H., DORMANN, D., and EDBAUER, D. (2017). Cytoplasmic poly-GA aggregates impair nuclear import of TDP-43 in C9orf72 ALS/FTLD. *Human molecular genetics* 26**,** 790-800.

KIERNAN, M.C., VUCIC, S., CHEAH, B.C., TURNER, M.R., EISEN, A., HARDIMAN, O., BURRELL, J.R., and ZOING, M.C. (2011). Amyotrophic lateral sclerosis. *The Lancet* 377**,** 942-955.

KIM, H.J., KIM, N.C., WANG, Y.-D., SCARBOROUGH, E.A., MOORE, J., DIAZ, Z., MACLEA, K.S., FREIBAUM, B., LI, S., MOLLIEX, A., KANAGARAJ, A.P., CARTER, R., BOYLAN, K.B., WOJTAS, A.M., RADEMAKERS, R., PINKUS, J.L., GREENBERG, S.A., TROJANOWSKI, J.Q., TRAYNOR, B.J., SMITH, B.N., TOPP, S., GKAZI, A.-S., MILLER, J., SHAW, C.E., KOTTLORS, M., KIRSCHNER, J., PESTRONK, A., LI, Y.R., FORD, A.F., GITLER, A.D., BENATAR, M., KING, O.D., KIMONIS, V.E., ROSS, E.D., WEIHL, C.C., SHORTER, J., and TAYLOR, J.P. (2013a). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495**,** 467-473.

KIM, H.J., KIM, N.C., WANG, Y.D., SCARBOROUGH, E.A., MOORE, J., DIAZ, Z., MACLEA, K.S., FREIBAUM, B., LI, S., MOLLIEX, A., KANAGARAJ, A.P., CARTER, R., BOYLAN, K.B., WOJTAS, A.M., RADEMAKERS, R., PINKUS, J.L., GREENBERG, S.A., TROJANOWSKI, J.Q., TRAYNOR, B.J., SMITH, B.N., TOPP, S., GKAZI, A.S., MILLER, J., SHAW, C.E., KOTTLORS, M., KIRSCHNER, J., PESTRONK, A., LI, Y.R., FORD, A.F., GITLER, A.D., BENATAR, M., KING, O.D., KIMONIS, V.E., ROSS, E.D., WEIHL, C.C., SHORTER, J., and TAYLOR, J.P. (2013b). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495**,** 467-473.

KING, A., AL-SARRAJ, S., and SHAW, C. (2009). Frontotemporal lobar degeneration with ubiquitinated tau-negative inclusions and additional alpha-synuclein pathology but also unusual cerebellar ubiquitinated p62-positive, TDP-43-negative inclusions. *Neuropathology* 29**,** 466-471.

KING, A., MAEKAWA, S., BODI, I., TROAKES, C., and AL-SARRAJ, S. (2011). Ubiquitinated, p62 immunopositive cerebellar cortical neuronal inclusions are evident across the spectrum of TDP-43 proteinopathies but are only rarely additionally immunopositive for phosphorylation-dependent TDP-43. *Neuropathology* 31**,** 239-249.

KRAMER, N.J., HANEY, M.S., MORGENS, D.W., JOVIČIĆ, A., COUTHOUIS, J., LI, A., OUSEY, J., MA, R., BIERI, G., TSUI, C.K., SHI, Y., HERTZ, N.T., TESSIER-LAVIGNE, M., ICHIDA, J.K., BASSIK, M.C., and GITLER, A.D. (2018). CRISPR–Cas9 screens in human cells and primary neurons identify modifiers of C9ORF72 dipeptide-repeat-protein toxicity. *Nature Genetics* 50**,** 603-612.

KUMAR, M., THAKUR, V., and RAGHAVA, G.P.S. (2008). COPid: Composition Based Protein Identification. *In Silico Biology* 8**,** 121-128.

KWIATKOWSKI, T.J., BOSCO, D., LECLERC, A.L., TAMRAZIAN, E., VANDERBURG, C., RUSS, C., DAVIS, A., GILCHRIST, J., KASARSKIS, E., MUNSAT, T., VALDMANIS, P., ROULEAU, G., HOSLER, B.A., CORTELLI, P., DE JONG, P., YOSHINAGA, Y., HAINES, J., PERICAK-VANCE, M., YAN, J., and BROWN, R.H. (2009). Mutations in the FUS/TLS Gene on Chromosome 16 Cause Familial Amyotrophic Lateral Sclerosis. *Science (New York, N.Y.)* 323**,** 1205-1208.

KWON, I., XIANG, S., KATO, M., WU, L., THEODOROPOULOS, P., WANG, T., KIM, J., YUN, J., XIE, Y., and MCKNIGHT, S.L. (2014). Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* 345**,** 1139-1145.

LAAKSOVIRTA, H., PEURALINNA, T., SCHYMICK, J.C., SCHOLZ, S.W., LAI, S.-L., MYLLYKANGAS, L., SULKAVA, R., JANSSON, L., HERNANDEZ, D.G., GIBBS, J.R., NALLS, M.A., HECKERMAN, D., TIENARI, P.J., and TRAYNOR, B.J. (2010). Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *The Lancet. Neurology* 9**,** 978-985.

LACOMBLEZ, L., BENSIMON, G., LEIGH, P.N., DEBOVE, C., BEJUIT, R., TRUFFINET, P., and MEININGER, V. (2002). Long-term safety of riluzole in amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders* 3**,** 23-29.

LAGIER-TOURENNE, C., BAUGHN, M., RIGO, F., SUN, S., LIU, P., LI, H.R., JIANG, J., WATT, A.T., CHUN, S., KATZ, M., QIU, J., SUN, Y., LING, S.C., ZHU, Q., POLYMENIDOU, M., DRENNER, K., ARTATES, J.W., MCALONIS-DOWNES, M., MARKMILLER, S., HUTT, K.R., PIZZO, D.P., CADY, J., HARMS, M.B., BALOH, R.H., VANDENBERG, S.R., YEO, G.W., FU, X.D., BENNETT, C.F., CLEVELAND, D.W., and RAVITS, J. (2013). Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci U S A* 110**,** E4530-4539.

LARSEN, S.C., SYLVESTERSEN, K.B., MUND, A., LYON, D., MULLARI, M., MADSEN, M.V., DANIEL, J.A., JENSEN, L.J., and NIELSEN, M.L. (2016). Proteome-wide analysis of arginine monomethylation reveals widespread occurrence in human cells. *Science Signaling* 9**,** rs9.

LASHLEY, T., ROHRER, J.D., MEAD, S., and REVESZ, T. (2015). Review: An update on clinical, genetic and pathological aspects of frontotemporal lobar degenerations. *Neuropathology and Applied Neurobiology* 41**,** 858-881.

LATTANTE, S., CIURA, S., ROULEAU, G., and KABASHI, E. (2015). Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends in Genetics* 31.

LE BER, I., CAMUZAT, A., BERGER, E., HANNEQUIN, D., LAQUERRIÈRE, A., GOLFIER, V., SEILHEAN, D., VIENNET, G., COURATIER, P., VERPILLAT, P., HEATH, S., CAMU, W., MARTINAUD, O., LACOMBLEZ, L., VERCELLETTO, M., SALACHAS, F., SELLAL, F., DIDIC, M., THOMAS-ANTERION, C., PUEL, M., MICHEL, B.F., BESSE, C., DUYCKAERTS, C., MEININGER, V., CAMPION, D., DUBOIS, B., and BRICE, A. (2009). Chromosome 9p-linked families with frontotemporal dementia associated with motor neuron disease. *Neurology* 72**,** 1669.

LEE, K.-H., ZHANG, P., KIM, H.J., MITREA, D.M., SARKAR, M., FREIBAUM, B.D., CIKA, J., COUGHLIN, M., MESSING, J., MOLLIEX, A., MAXWELL, B.A., KIM, N.C., TEMIROV, J., MOORE, J., KOLAITIS, R.-M., SHAW, T.I., BAI, B., PENG, J., KRIWACKI, R.W., and TAYLOR, J.P. (2016a). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* 167**,** 774-788.e717.

LEE, K.H., ZHANG, P., KIM, H.J., MITREA, D.M., SARKAR, M., FREIBAUM, B.D., CIKA, J., COUGHLIN, M., MESSING, J., MOLLIEX, A., MAXWELL, B.A., KIM, N.C., TEMIROV, J., MOORE, J., KOLAITIS, R.M., SHAW, T.I., BAI, B., PENG, J., KRIWACKI, R.W., and TAYLOR, J.P. (2016b). C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles. *Cell* 167**,** 774-788 e717.

LEE, Y.-B., BASKARAN, P., GOMEZ-DEZA, J., CHEN, H.-J., NISHIMURA, A.L., SMITH, B.N., TROAKES, C., ADACHI, Y., STEPTO, A., PETRUCELLI, L., GALLO, J.-M., HIRTH, F., ROGELJ, B., GUTHRIE, S., and SHAW, C.E. (2017). C9orf72 poly GA RAN-translated protein plays a key role in amyotrophic lateral sclerosis via aggregation and toxicity. *Human molecular genetics* 26**,** 4765-4777.

LEE, Y.B., CHEN, H.J., PERES, J.N., GOMEZ-DEZA, J., ATTIG, J., STALEKAR, M., TROAKES, C., NISHIMURA, A.L., SCOTTER, E.L., VANCE, C., ADACHI, Y., SARDONE, V., MILLER, J.W., SMITH, B.N., GALLO, J.M., ULE, J., HIRTH, F., ROGELJ, B., HOUART, C., and SHAW, C.E. (2013). Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep* 5**,** 1178-1186.

LEHMER, C., OECKL, P., WEISHAUPT, J.H., VOLK, A.E., DIEHL-SCHMID, J., SCHROETER, M.L., LAUER, M., KORNHUBER, J., LEVIN, J., FASSBENDER, K., LANDWEHRMEYER, B., SCHLUDI, M.H., ARZBERGER, T., KREMMER, E., FLATLEY, A., FEEDERLE, R., STEINACKER, P., WEYDT, P., LUDOLPH, A.C., EDBAUER, D., and OTTO, M. (2017). Poly-GP in cerebrospinal fluid links C9orf72-associated dipeptide repeat expression to the asymptomatic phase of ALS/FTD. *EMBO Mol Med* 9**,** 859-868.

LEVINE, T.P., DANIELS, R.D., GATTA, A.T., WONG, L.H., and HAYES, M.J. (2013). The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics (Oxford, England)* 29**,** 499-503.

LIN, C.L., BRISTOL, L.A., JIN, L., DYKES-HOBERG, M., CRAWFORD, T., CLAWSON, L., and ROTHSTEIN, J.D. (1998). Aberrant RNA processing in a neurodegenerative disease: the cause for absent EAAT2, a glutamate transporter, in amyotrophic lateral sclerosis. *Neuron* 20**,** 589-602.

LIN, Y., MORI, E., KATO, M., XIANG, S., WU, L., KWON, I., and MCKNIGHT, S.L. (2016). Toxic PR Poly-Dipeptides Encoded by the C9orf72 Repeat Expansion Target LC Domain Polymers. *Cell* 167**,** 789-802 e712.

LING, S.-C., POLYMENIDOU, M., and CLEVELAND, D.W. (2013). Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* 79**,** 416-438.

LITT, M., QIU, Y., and HUANG, S. (2009). Histone arginine methylations: their roles in chromatin dynamics and transcriptional regulation. *Bioscience reports* 29**,** 131-141.

LIU, F., LIU, Q., LU, C.X., CUI, B., GUO, X.N., WANG, R.R., LIU, M.S., LI, X.G., CUI, L.Y., and ZHANG, X. (2016). Identification of a novel loss-of-function C9orf72 splice site mutation in a patient with amyotrophic lateral sclerosis. *Neurobiol Aging* 47**,** 219.e211-219.e215.

LOPEZ-GONZALEZ, R., LU, Y., GENDRON, T.F., KARYDAS, A., TRAN, H., YANG, D., PETRUCELLI, L., MILLER, B.L., ALMEIDA, S., and GAO, F.-B. (2016). Poly(GR) in C9ORF72-Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. *Neuron* 92**,** 383-391.

LOWE, J. (1994). New pathological findings in amyotrophic lateral sclerosis. *Journal of the Neurological Sciences* 124**,** 38-51.

LU, J., and DEUTSCH, C. (2008). Electrostatics in the ribosomal tunnel modulate chain elongation rates. *J Mol Biol* 384**,** 73-86.

LUTHI-CARTER, R., STRAND, A., PETERS, N.L., SOLANO, S.M., HOLLINGSWORTH, Z.R., MENON, A.S., FREY, A.S., SPEKTOR, B.S., PENNEY, E.B., SCHILLING, G., ROSS, C.A., BORCHELT, D.R., TAPSCOTT, S.J., YOUNG, A.B., CHA, J.H., and OLSON, J.M. (2000). Decreased expression of striatal signaling genes in a mouse model of Huntington's disease. *Hum Mol Genet* 9**,** 1259-1271.

LUTY, A.A., KWOK, J.B.J., THOMPSON, E.M., BLUMBERGS, P., BROOKS, W.S., LOY, C.T., DOBSON-STONE, C., PANEGYRES, P.K., HECKER, J., NICHOLSON, G.A., HALLIDAY, G.M., and SCHOFIELD, P.R. (2008). Pedigree with frontotemporal lobar

degeneration--motor neuron disease and Tar DNA binding protein-43 positive neuropathology: genetic linkage to chromosome 9. *BMC neurology* 8**,** 32-32.

MACDONALD, M.E., AMBROSE, C.M., DUYAO, M.P., MYERS, R.H., LIN, C., SRINIDHI, L., BARNES, G., TAYLOR, S.A., JAMES, M., GROOT, N., MACFARLANE, H., JENKINS, B., ANDERSON, M.A., WEXLER, N.S., GUSELLA, J.F., BATES, G.P., BAXENDALE, S., HUMMERICH, H., KIRBY, S., NORTH, M., YOUNGMAN, S., MOTT, R., ZEHETNER, G., SEDLACEK, Z., POUSTKA, A., FRISCHAUF, A.-M., LEHRACH, H., BUCKLER, A.J., CHURCH, D., DOUCETTE-STAMM, L., O'DONOVAN, M.C., RIBA-RAMIREZ, L., SHAH, M., STANTON, V.P., STROBEL, S.A., DRATHS, K.M., WALES, J.L., DERVAN, P., HOUSMAN, D.E., ALTHERR, M., SHIANG, R., THOMPSON, L., FIELDER, T., WASMUTH, J.J., TAGLE, D., VALDES, J., ELMER, L., ALLARD, M., CASTILLA, L., SWAROOP, M., BLANCHARD, K., COLLINS, F.S., SNELL, R., HOLLOWAY, T., GILLESPIE, K., DATSON, N., SHAW, D., and HARPER, P.S. (1993). A Novel Gene Containing a Trinucleotide Repeat that is Expanded and Unstable on Huntington's Disease Chromosomes. *Cell* 72**,** 971-983.

MACKENZIE, I.R., FRICK, P., GRASSER, F.A., GENDRON, T.F., PETRUCELLI, L., CASHMAN, N.R., EDBAUER, D., KREMMER, E., PRUDLO, J., TROOST, D., and NEUMANN, M. (2015). Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta Neuropathol* 130**,** 845-861.

MAJOUNIE, E., RENTON, A.E., MOK, K., DOPPER, E.G., WAITE, A., ROLLINSON, S., CHIO, A., RESTAGNO, G., NICOLAOU, N., SIMON-SANCHEZ, J., VAN SWIETEN, J.C., ABRAMZON, Y., JOHNSON, J.O., SENDTNER, M., PAMPHLETT, R., ORRELL, R.W., MEAD, S., SIDLE, K.C., HOULDEN, H., ROHRER, J.D., MORRISON, K.E., PALL, H., TALBOT, K., ANSORGE, O., CHROMOSOME, A.L.S.F.T.D.C., FRENCH RESEARCH NETWORK ON, F.F.A., CONSORTIUM, I., HERNANDEZ, D.G., AREPALLI, S., SABATELLI, M., MORA, G., CORBO, M., GIANNINI, F., CALVO, A., ENGLUND, E., BORGHERO, G., FLORIS, G.L., REMES, A.M., LAAKSOVIRTA, H., MCCLUSKEY, L., TROJANOWSKI, J.Q., VAN DEERLIN, V.M., SCHELLENBERG, G.D., NALLS, M.A., DRORY, V.E., LU, C.S., YEH, T.H., ISHIURA, H., TAKAHASHI, Y., TSUJI, S., LE BER, I., BRICE, A., DREPPER, C., WILLIAMS, N., KIRBY, J., SHAW, P., HARDY, J., TIENARI, P.J., HEUTINK, P., MORRIS, H.R., PICKERING-BROWN, S., and TRAYNOR, B.J. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol* 11**,** 323-330.

MANN, D.M., ROLLINSON, S., ROBINSON, A., BENNION CALLISTER, J., THOMPSON, J.C., SNOWDEN, J.S., GENDRON, T., PETRUCELLI, L., MASUDA-SUZUKAKE, M., HASEGAWA, M., DAVIDSON, Y., and PICKERING-BROWN, S. (2013). Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun* 1**,** 68.

MANTAMADIOTIS, T., LEMBERGER, T., BLECKMANN, S.C., KERN, H., KRETZ, O., MARTIN VILLALBA, A., TRONCHE, F., KELLENDONK, C., GAU, D., KAPFHAMMER, J., OTTO, C., SCHMID, W., and SCHUTZ, G. (2002). Disruption of CREB function in brain leads to neurodegeneration. *Nat Genet* 31**,** 47-54.

MAY, S., HORNBURG, D., SCHLUDI, M.H., ARZBERGER, T., RENTZSCH, K., SCHWENK, B.M., GRÄSSER, F.A., MORI, K., KREMMER, E., BANZHAF-STRATHMANN, J., MANN, M., MEISSNER, F., and EDBAUER, D. (2014). C9orf72 FTLD/ALS-associated Gly-Ala dipeptide repeat proteins cause neuronal toxicity and Unc119 sequestration. *Acta Neuropathologica* 128**,** 485-503.

MCGOLDRICK, P., ZHANG, M., VAN BLITTERSWIJK, M., SATO, C., MORENO, D., XIAO, S., ZHANG, A.B., MCKEEVER, P.M., WEICHERT, A., SCHNEIDER, R., KEITH, J., PETRUCELLI, L., RADEMAKERS, R., ZINMAN, L., ROBERTSON, J., and ROGAEVA,

E. (2018). Unaffected mosaic C9orf72 case. *Neurology* 90**,** e323.

MEETER, L.H., KAAT, L.D., ROHRER, J.D., and VAN SWIETEN, J.C. (2017). Imaging and fluid biomarkers in frontotemporal dementia. *Nature Reviews Neurology* 13**,** 406-419.

MEJZINI, R., FLYNN, L.L., PITOUT, I.L., FLETCHER, S., WILTON, S.D., and AKKARI, P.A. (2019). ALS Genetics, Mechanisms, and Therapeutics: Where Are We Now? *Frontiers in Neuroscience* 13.

MILLECAMPS, S., BOILLEE, S., LE BER, I., SEILHEAN, D., TEYSSOU, E., GIRAUDEAU, M., MOIGNEU, C., VANDENBERGHE, N., DANEL-BRUNAUD, V., CORCIA, P., PRADAT, P.F., LE FORESTIER, N., LACOMBLEZ, L., BRUNETEAU, G., CAMU, W., BRICE, A., CAZENEUVE, C., LEGUERN, E., MEININGER, V., and SALACHAS, F. (2012). Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. *J Med Genet* 49**,** 258-263.

MILLECAMPS, S., DE SEPTENVILLE, A., TEYSSOU, E., DANIAU, M., CAMUZAT, A., ALBERT, M., LEGUERN, E., GALIMBERTI, D., BRICE, A., MARIE, Y., and LE BER, I. (2014). Genetic analysis of matrin 3 gene in French amyotrophic lateral sclerosis patients and frontotemporal lobar degeneration with amyotrophic lateral sclerosis patients. *Neurobiology of Aging* 35**,** 2882.e2813-2882.e2815.

MITREA, D.M., CIKA, J.A., GUY, C.S., BAN, D., BANERJEE, P.R., STANLEY, C.B., NOURSE, A., DENIZ, A.A., and KRIWACKI, R.W. (2016). Nucleophosmin integrates within the nucleolus via multi-modal interactions with proteins displaying R-rich linear motifs and rRNA. *eLife* 5**,** e13571.

MITSUI, K., NAKAYAMA, H., AKAGI, T., NEKOOKI, M., OHTAWA, K., TAKIO, K., HASHIKAWA, T., and NUKINA, N. (2002). Purification of polyglutamine aggregates and identification of elongation factor-1alpha and heat shock protein 84 as aggregate-interacting proteins. *J Neurosci* 22**,** 9267-9277.

MIZIELINSKA, S., GRONKE, S., NICCOLI, T., RIDLER, C.E., CLAYTON, E.L., DEVOY, A., MOENS, T., NORONA, F.E., WOOLLACOTT, I.O.C., PIETRZYK, J., CLEVERLEY, K., NICOLL, A.J., PICKERING-BROWN, S., DOLS, J., CABECINHA, M., HENDRICH, O., FRATTA, P., FISHER, E.M.C., PARTRIDGE, L., and ISAACS, A.M. (2014a). C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science* 345**,** 1192-1194.

MIZIELINSKA, S., GRÖNKE, S., NICCOLI, T., RIDLER, C.E., CLAYTON, E.L., DEVOY, A., MOENS, T., NORONA, F.E., WOOLLACOTT, I.O.C., PIETRZYK, J., CLEVERLEY, K., NICOLL, A.J., PICKERING-BROWN, S., DOLS, J., CABECINHA, M., HENDRICH, O., FRATTA, P., FISHER, E.M.C., PARTRIDGE, L., and ISAACS, A.M. (2014b). C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science (New York, N.Y.)* 345**,** 1192-1194.

MIZIELINSKA, S., and ISAACS, A.M. (2014). C9orf72 amyotrophic lateral sclerosis and frontotemporal dementia: gain or loss of function? *Current opinion in neurology* 27**,** 515-523.

MIZIELINSKA, S., LASHLEY, T., NORONA, F.E., CLAYTON, E.L., RIDLER, C.E., FRATTA, P., and ISAACS, A.M. (2013). C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol* 126**,** 845-857.

MIZIELINSKA, S., RIDLER, C.E., BALENDRA, R., THOENG, A., WOODLING, N.S., GRÄSSER, F.A., PLAGNOL, V., LASHLEY, T., PARTRIDGE, L., and ISAACS, A.M. (2017). Bidirectional nucleolar dysfunction in C9orf72 frontotemporal lobar degeneration. *Acta Neuropathologica Communications* 5**,** 29.

MND-AUSTRALIA (2015). "Economic analysis of motor neurone disease in Australia.". (Gladesville, NSW).

MOENS, T.G., NICCOLI, T., WILSON, K.M., ATILANO, M.L., BIRSA, N., GITTINGS, L.M., HOLBLING, B.V., DYSON, M.C., THOENG, A., NEEVES, J., GLARIA, I., YU, L., BUSSMANN, J., STORKEBAUM, E., PARDO, M., CHOUDHARY, J.S., FRATTA, P.,

PARTRIDGE, L., and ISAACS, A.M. (2019). C9orf72 arginine-rich dipeptide proteins interact with ribosomal proteins in vivo to induce a toxic translational arrest that is rescued by eIF1A. *Acta Neuropathol* 137**,** 487-500.

MOILY, N.S., ORMSBY, A.R., STOJILOVIC, A., RAMDZAN, Y.M., DIESCH, J., HANNAN, R.D., ZAJAC, M.S., HANNAN, A.J., OSHLACK, A., and HATTERS, D.M. (2017). Transcriptional profiles for distinct aggregation states of mutant Huntingtin exon 1 protein unmask new Huntington's disease pathways. *Mol Cell Neurosci* 83**,** 103-112.

MOLLIEX, A., TEMIROV, J., LEE, J., COUGHLIN, M., KANAGARAJ, A.P., KIM, H.J., MITTAG, T., and TAYLOR, J.P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* 163**,** 123-133.

MOREAU, K., FLEMING, A., IMARISIO, S., LOPEZ RAMIREZ, A., MERCER, J.L., JIMENEZ-SANCHEZ, M., BENTO, C.F., PURI, C., ZAVODSZKY, E., SIDDIQI, F., LAVAU, C.P., BETTON, M., O'KANE, C.J., WECHSLER, D.S., and RUBINSZTEIN, D.C. (2014). PICALM modulates autophagy activity and tau accumulation. *Nat Commun* 5**,** 4998.

MORI, K., ARZBERGER, T., GRÄSSER, F.A., GIJSELINCK, I., MAY, S., RENTZSCH, K., WENG, S.-M., SCHLUDI, M.H., VAN DER ZEE, J., CRUTS, M., VAN BROECKHOVEN, C., KREMMER, E., KRETZSCHMAR, H.A., HAASS, C., and EDBAUER, D. (2013a). Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathologica* 126**,** 881-893.

MORI, K., LAMMICH, S., MACKENZIE, I.R., FORNE, I., ZILOW, S., KRETZSCHMAR, H., EDBAUER, D., JANSSENS, J., KLEINBERGER, G., CRUTS, M., HERMS, J., NEUMANN, M., VAN BROECKHOVEN, C., ARZBERGER, T., and HAASS, C. (2013b). hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol* 125**,** 413-423.

MORI, K., WENG, S.-M., ARZBERGER, T., MAY, S., RENTZSCH, K., KREMMER, E., SCHMID, B., KRETZSCHMAR, H.A., CRUTS, M., VAN BROECKHOVEN, C., HAASS, C., and EDBAUER, D. (2013c). The C9orf72 GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTLD/ALS. *Science* 339**,** 1335-1338.

MORITA, M., AL-CHALABI, A., ANDERSEN, P.M., HOSLER, B., SAPP, P., ENGLUND, E., MITCHELL, J.E., HABGOOD, J.J., DE BELLEROCHE, J., XI, J., JONGJAROENPRASERT, W., HORVITZ, H.R., GUNNARSSON, L.G., and BROWN, R.H. (2006). A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* 66**,** 839.

NEUMANN, M., BENTMANN, E., DORMANN, D., JAWAID, A., DEJESUS-HERNANDEZ, M., ANSORGE, O., ROEBER, S., KRETZSCHMAR, H.A., MUNOZ, D.G., KUSAKA, H., YOKOTA, O., ANG, L.-C., BILBAO, J., RADEMAKERS, R., HAASS, C., and MACKENZIE, I.R.A. (2011). FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. *Brain : a journal of neurology* 134**,** 2595-2609.

NEUMANN, M., RADEMAKERS, R., ROEBER, S., BAKER, M., KRETZSCHMAR, H.A., and MACKENZIE, I.R.A. (2009). A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain : a journal of neurology* 132**,** 2922-2931.

NEUMANN, M., SAMPATHU, D.M., KWONG, L.K., TRUAX, A.C., MICSENYI, M.C., CHOU, T.T., BRUCE, J., SCHUCK, T., GROSSMAN, M., CLARK, C.M., MCCLUSKEY, L.F., MILLER, B.L., MASLIAH, E., MACKENZIE, I.R., FELDMAN, H., FEIDEN, W., KRETZSCHMAR, H.A., TROJANOWSKI, J.Q., and LEE, V.M.Y. (2006). Ubiquitinated TDP-43 in Frontotemporal Lobar Degeneration and Amyotrophic Lateral Sclerosis. *Science* 314**,** 130.

NICHOLS, L., MARTINDALE-ADAMS, J., BURNS, R., GRANEY, M., and ZUBER, J. (2011). Translation of a Dementia Caregiver Support Program in a Health Care System-REACH VA. *Archives of internal medicine* 171**,** 353-359.

NICOLAS, A., KENNA, K.P., RENTON, A.E., TICOZZI, N., FAGHRI, F., CHIA, R., DOMINOV, J.A., KENNA, B.J., NALLS, M.A., KEAGLE, P., RIVERA, A.M., VAN RHEENEN, W., MURPHY, N.A., VAN VUGT, J.J.F.A., GEIGER, J.T., VAN DER SPEK, R.A., PLINER, H.A., SHANKARACHARYA, SMITH, B.N., MARANGI, G., TOPP, S.D., ABRAMZON, Y., GKAZI, A.S., EICHER, J.D., KENNA, A., LOGULLO, F.O., SIMONE, I., LOGROSCINO, G., SALVI, F., BARTOLOMEI, I., BORGHERO, G., MURRU, M.R., COSTANTINO, E., PANI, C., PUDDU, R., CAREDDA, C., PIRAS, V., TRANQUILLI, S., CUCCU, S., CORONGIU, D., MELIS, M., MILIA, A., MARROSU, F., MARROSU, M.G., FLORIS, G., CANNAS, A., TRANQUILLI, S., CAPASSO, M., CAPONNETTO, C., MANCARDI, G., ORIGONE, P., MANDICH, P., CONFORTI, F.L., CAVALLARO, S., MORA, G., MARINOU, K., SIDERI, R., PENCO, S., MOSCA, L., LUNETTA, C., PINTER, G.L., CORBO, M., RIVA, N., CARRERA, P., VOLANTI, P., MANDRIOLI, J., FINI, N., FASANO, A., TREMOLIZZO, L., AROSIO, A., FERRARESE, C., TROJSI, F., TEDESCHI, G., MONSURRÒ, M.R., PICCIRILLO, G., FEMIANO, C., TICCA, A., ORTU, E., LA BELLA, V., SPATARO, R., COLLETTI, T., SABATELLI, M., ZOLLINO, M., CONTE, A., LUIGETTI, M., LATTANTE, S., MARANGI, G., SANTARELLI, M., PETRUCCI, A., PUGLIATTI, M., PIRISI, A., PARISH, L.D., OCCHINERI, P., GIANNINI, F., BATTISTINI, S., RICCI, C., BENIGNI, M., CAU, T.B., LOI, D., CALVO, A., et al. (2018). Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron* 97**,** 1268-1283.e1266.

NISHIHIRA, Y., TAN, C.-F., ONODERA, O., TOYOSHIMA, Y., YAMADA, M., MORITA, T., NISHIZAWA, M., KAKITA, A., and TAKAHASHI, H. (2008). Sporadic amyotrophic lateral sclerosis: two pathological patterns shown by analysis of distribution of TDP-43-immunoreactive neuronal and glial cytoplasmic inclusions. *Acta Neuropathologica* 116**,** 169.

NORDIN, A., AKIMOTO, C., WUOLIKAINEN, A., ALSTERMARK, H., JONSSON, P., BIRVE, A., MARKLUND, S.L., GRAFFMO, K.S., FORSBERG, K., BRÄNNSTRÖM, T., and ANDERSEN, P.M. (2015). Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Human Molecular Genetics* 24**,** 3133-3142.

NUSSBACHER, J.K., TABET, R., YEO, G.W., and LAGIER-TOURENNE, C. (2019). Disruption of RNA Metabolism in Neurological Diseases and Emerging Therapeutic Interventions. *Neuron* 102**,** 294-320.

O'ROURKE, J.G., BOGDANIK, L., YANEZ, A., LALL, D., WOLF, A.J., MUHAMMAD, A.K., HO, R., CARMONA, S., VIT, J.P., ZARROW, J., KIM, K.J., BELL, S., HARMS, M.B., MILLER, T.M., DANGLER, C.A., UNDERHILL, D.M., GOODRIDGE, H.S., LUTZ, C.M., and BALOH, R.H. (2016). C9orf72 is required for proper macrophage and microglial function in mice. *Science* 351**,** 1324-1329.

OHKI, Y., WENNINGER-WEINZIERL, A., HRUSCHA, A., ASAKAWA, K., KAWAKAMI, K., HAASS, C., EDBAUER, D., and SCHMID, B. (2017). Glycine-alanine dipeptide repeat protein contributes to toxicity in a zebrafish model of C9orf72 associated neurodegeneration. *Molecular Neurodegeneration* 12**,** 6.

OLZSCHA, H., SCHERMANN, S.M., WOERNER, A.C., PINKERT, S., HECHT, M.H., TARTAGLIA, G.G., VENDRUSCOLO, M., HAYER-HARTL, M., HARTL, F.U., and VABULAS, R.M. (2011). Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* 144**,** 67-78.

ONG, S.-E., MITTLER, G., and MANN, M. (2004). Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nature Methods* 1**,** 119-126.

ONYIKE, C.U., and HUEY, E.D. (2013). Frontotemporal dementia and psychiatry. *International review of psychiatry (Abingdon, England)* 25**,** 127-129.

ORMSBY, A.R., RAMDZAN, Y.M., MOK, Y.-F., JOVANOSKI, K.D., and HATTERS, D.M. (2013). A platform to view huntingtin exon 1 aggregation flux in the cell reveals

divergent influences from chaperones hsp40 and hsp70. *The Journal of biological chemistry* 288**,** 37192-37203.

PEARSON, J.P., WILLIAMS, N.M., MAJOUNIE, E., WAITE, A., STOTT, J., NEWSWAY, V., MURRAY, A., HERNANDEZ, D., GUERREIRO, R., SINGLETON, A.B., NEAL, J., and MORRIS, H.R. (2011). Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *Journal of neurology* 258**,** 647-655.

PIKKARAINEN, M., HARTIKAINEN, P., and ALAFUZOFF, I. (2010). Ubiquitinated p62-positive, TDP-43-negative inclusions in cerebellum in frontotemporal lobar degeneration with TAR DNA binding protein 43. *Neuropathology* 30**,** 197-199.

PISAREV, A.V., SKABKIN, M.A., PISAREVA, V.P., SKABKINA, O.V., RAKOTONDRAFARA, A.M., HENTZE, M.W., HELLEN, C.U., and PESTOVA, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Mol Cell* 37**,** 196-210.

POTTIER, C., BIENIEK, K.F., FINCH, N., VAN DE VORST, M., BAKER, M., PERKERSEN, R., BROWN, P., RAVENSCROFT, T., VAN BLITTERSWIJK, M., NICHOLSON, A.M., DETURE, M., KNOPMAN, D.S., JOSEPHS, K.A., PARISI, J.E., PETERSEN, R.C., BOYLAN, K.B., BOEVE, B.F., GRAFF-RADFORD, N.R., VELTMAN, J.A., GILISSEN, C., MURRAY, M.E., DICKSON, D.W., and RADEMAKERS, R. (2015). Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathologica* 130**,** 77-92.

RABINOVICI, G.D., and MILLER, B.L. (2010). Frontotemporal lobar degeneration: epidemiology, pathophysiology, diagnosis and management. *CNS drugs* 24**,** 375-398.

RADEMAKERS, R., NEUMANN, M., and MACKENZIE, I.R. (2012). Advances in understanding the molecular basis of frontotemporal dementia. *Nature reviews. Neurology* 8**,** 423-434.

RADWAN, M., ANG, C.S., ORMSBY, A.R., COX, D., DALY, J.C., REID, G.E., and HATTERS, D.M. (2020). Arginine in C9ORF72 dipolypeptides mediates promiscuous proteome binding and multiple modes of toxicity. *Mol Cell Proteomics*.

RAMAN, R., ALLEN, S.P., GOODALL, E.F., KRAMER, S., PONGER, L.L., HEATH, P.R., MILO, M., HOLLINGER, H.C., WALSH, T., HIGHLEY, J.R., OLPIN, S., MCDERMOTT, C.J., SHAW, P.J., and KIRBY, J. (2015). Gene expression signatures in motor neurone disease fibroblasts reveal dysregulation of metabolism, hypoxia-response and RNA processing functions. *Neuropathol Appl Neurobiol* 41**,** 201-226.

RAMDZAN, Y.M., POLLING, S., CHIA, C.P.Z., NG, I.H.W., ORMSBY, A.R., CROFT, N.P., PURCELL, A.W., BOGOYEVITCH, M.A., NG, D.C.H., GLEESON, P.A., and HATTERS, D.M. (2012). Tracking protein aggregation and mislocalization in cells with flow cytometry. *Nature Methods* 9**,** 467.

RAMDZAN, Y.M., TRUBETSKOV, M.M., ORMSBY, A.R., NEWCOMBE, E.A., SUI, X., TOBIN, M.J., BONGIOVANNI, M.N., GRAS, S.L., DEWSON, G., MILLER, J.M.L., FINKBEINER, S., MOILY, N.S., NICLIS, J., PARISH, C.L., PURCELL, A.W., BAKER, M.J., WILCE, J.A., WARIS, S., STOJANOVSKI, D., BÖCKING, T., ANG, C.-S., ASCHER, D.B., REID, G.E., and HATTERS, D.M. (2017). Huntingtin Inclusions Trigger Cellular Quiescence, Deactivate Apoptosis, and Lead to Delayed Necrosis. *Cell Reports* 19**,** 919-927.

REDDY, K., ZAMIRI, B., STANLEY, S.Y., MACGREGOR, R.B., JR., and PEARSON, C.E. (2013). The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J Biol Chem* 288**,** 9860-9866.

RENAUD, L., PICHER-MARTEL, V., CODRON, P., and JULIEN, J.P. (2019). Key role of UBQLN2 in pathogenesis of amyotrophic lateral sclerosis and frontotemporal dementia. *Acta Neuropathol Commun* 7**,** 103.

RENTON, A.E., MAJOUNIE, E., WAITE, A., SIMON-SANCHEZ, J., ROLLINSON, S., GIBBS, J.R., SCHYMICK, J.C., LAAKSOVIRTA, H., VAN SWIETEN, J.C., MYLLYKANGAS, L., KALIMO, H., PAETAU, A., ABRAMZON, Y., REMES, A.M., KAGANOVICH, A., SCHOLZ, S.W., DUCKWORTH, J., DING, J., HARMER, D.W., HERNANDEZ, D.G., JOHNSON, J.O., MOK, K., RYTEN, M., TRABZUNI, D., GUERREIRO, R.J., ORRELL, R.W., NEAL, J., MURRAY, A., PEARSON, J., JANSEN, I.E., SONDERVAN, D., SEELAAR, H., BLAKE, D., YOUNG, K., HALLIWELL, N., CALLISTER, J.B., TOULSON, G., RICHARDSON, A., GERHARD, A., SNOWDEN, J., MANN, D., NEARY, D., NALLS, M.A., PEURALINNA, T., JANSSON, L., ISOVIITA, V.M., KAIVORINNE, A.L., HOLTTA-VUORI, M., IKONEN, E., SULKAVA, R., BENATAR, M., WUU, J., CHIO, A., RESTAGNO, G., BORGHERO, G., SABATELLI, M., CONSORTIUM, I., HECKERMAN, D., ROGAEVA, E., ZINMAN, L., ROTHSTEIN, J.D., SENDTNER, M., DREPPER, C., EICHLER, E.E., ALKAN, C., ABDULLAEV, Z., PACK, S.D., DUTRA, A., PAK, E., HARDY, J., SINGLETON, A., WILLIAMS, N.M., HEUTINK, P., PICKERING-BROWN, S., MORRIS, H.R., TIENARI, P.J., and TRAYNOR, B.J. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72**,** 257-268.

RINGHOLZ, G., APPEL, S., BRADSHAW, M., COOKE, N.A., MOSNIK, D., and SCHULZ, P. (2005). Prevalence and pattern of cognitive impairment in sporadic ALS. *Neurology* 65**,** 586-590.

RIPAUD, L., CHUMAKOVA, V., ANTONIN, M., HASTIE, A.R., PINKERT, S., KÖRNER, R., RUFF, K.M., PAPPU, R.V., HORNBURG, D., MANN, M., HARTL, F.U., and HIPP, M.S. (2014). Overexpression of Q-rich prion-like proteins suppresses polyQ cytotoxicity and alters the polyQ interactome. *Proc Natl Acad Sci U S A* 111**,** 18219-18224.

RIVALS, I., PERSONNAZ, L., TAING, L., and POTIER, M.-C. (2006). Enrichment or depletion of a GO category within a class of genes: which test? *Bioinformatics* 23**,** 401-407.

ROHRER, J.D., GUERREIRO, R., VANDROVCOVA, J., UPHILL, J., REIMAN, D., BECK, J., ISAACS, A.M., AUTHIER, A., FERRARI, R., FOX, N.C., MACKENZIE, I.R.A., WARREN, J.D., DE SILVA, R., HOLTON, J., REVESZ, T., HARDY, J., MEAD, S., and ROSSOR, M.N. (2009). The heritability and genetics of frontotemporal lobar degeneration. *Neurology* 73**,** 1451-1456.

ROHRER, J.D., ISAACS, A.M., MIZIELINSKA, S., MEAD, S., LASHLEY, T., WRAY, S., SIDLE, K., FRATTA, P., ORRELL, R.W., HARDY, J., HOLTON, J., REVESZ, T., ROSSOR, M.N., and WARREN, J.D. (2015). C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *Lancet Neurol* 14**,** 291-301.

ROSEN, D.R., SIDDIQUE, T., PATTERSON, D., FIGLEWICZ, D.A., SAPP, P., HENTATI, A., DONALDSON, D., GOTO, J., O'REGAN, J.P., DENG, H.-X., RAHMANI, Z., KRIZUS, A., MCKENNA-YASEK, D., CAYABYAB, A., GASTON, S.M., BERGER, R., TANZI, R.E., HALPERIN, J.J., HERZFELDT, B., VAN DEN BERGH, R., HUNG, W.-Y., BIRD, T., DENG, G., MULDER, D.W., SMYTH, C., LAING, N.G., SORIANO, E., PERICAK–VANCE, M.A., HAINES, J., ROULEAU, G.A., GUSELLA, J.S., HORVITZ, H.R., and BROWN, R.H. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 362**,** 59-62.

ROSS, C.A., and POIRIER, M.A. (2005). What is the role of protein aggregation in neurodegeneration? *Nat Rev Mol Cell Biol* 6**,** 891-898.

ROSSI, S., SERRANO, A., GERBINO, V., GIORGI, A., DI FRANCESCO, L., NENCINI, M., BOZZO, F., SCHININÀ, M.E., BAGNI, C., CESTRA, G., CARRÌ, M.T., ACHSEL, T., and COZZOLINO, M. (2015). Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *Journal of Cell Science* 128**,** 1787.

RUDICH, P., SNOZNIK, C., WATKINS, S.C., MONAGHAN, J., PANDEY, U.B., and LAMITINA, S.T. (2017). Nuclear localized C9orf72-associated arginine-containing dipeptides

exhibit age-dependent toxicity in C. elegans. *Human molecular genetics* 26**,** 4916-4928.

RUSS, J., LIU, E.Y., WU, K., NEAL, D., SUH, E., IRWIN, D.J., MCMILLAN, C.T., HARMS, M.B., CAIRNS, N.J., WOOD, E.M., XIE, S.X., ELMAN, L., MCCLUSKEY, L., GROSSMAN, M., VAN DEERLIN, V.M., and LEE, E.B. (2015). Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta Neuropathol* 129**,** 39-52.

RUSTEN, T.E., and STENMARK, H. (2010). p62, an autophagy hero or culprit? *Nat Cell Biol* 12**,** 207-209.

RYAN, V.H., DIGNON, G.L., ZERZE, G.H., CHABATA, C.V., SILVA, R., CONICELLA, A.E., AMAYA, J., BURKE, K.A., MITTAL, J., and FAWZI, N.L. (2018). Mechanistic View of hnRNPA2 Low-Complexity Domain Structure, Interactions, and Phase Separation Altered by Mutation and Arginine Methylation. *Mol Cell* 69**,** 465-479 e467.

SABERI, S., STAUFFER, J.E., JIANG, J., GARCIA, S.D., TAYLOR, A.E., SCHULTE, D., OHKUBO, T., SCHLOFFMAN, C.L., MALDONADO, M., BAUGHN, M., RODRIGUEZ, M.J., PIZZO, D., CLEVELAND, D., and RAVITS, J. (2018). Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. *Acta Neuropathol* 135**,** 459-474.

SAKAE, N., BIENIEK, K.F., ZHANG, Y.-J., ROSS, K., GENDRON, T.F., MURRAY, M.E., RADEMAKERS, R., PETRUCELLI, L., and DICKSON, D.W. (2018). Poly-GR dipeptide repeat polymers correlate with neurodegeneration and Clinicopathological subtypes in C9ORF72-related brain disease. *Acta neuropathologica communications* 6**,** 63-63.

SAREEN, D., O'ROURKE, J.G., MEERA, P., MUHAMMAD, A.K., GRANT, S., SIMPKINSON, M., BELL, S., CARMONA, S., ORNELAS, L., SAHABIAN, A., GENDRON, T., PETRUCELLI, L., BAUGHN, M., RAVITS, J., HARMS, M.B., RIGO, F., BENNETT, C.F., OTIS, T.S., SVENDSEN, C.N., and BALOH, R.H. (2013). Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med* 5**,** 208ra149.

SCARAMUZZINO, C., MONAGHAN, J., MILIOTO, C., LANSON, N.A., JR., MALTARE, A., AGGARWAL, T., CASCI, I., FACKELMAYER, F.O., PENNUTO, M., and PANDEY, U.B. (2013). Protein arginine methyltransferase 1 and 8 interact with FUS to modify its sub-cellular distribution and toxicity in vitro and in vivo. *PLoS One* 8**,** e61576.

SCHERZINGER, E., SITTLER, A., SCHWEIGER, K., HEISER, V., LURZ, R., HASENBANK, R., BATES, G.P., LEHRACH, H., and WANKER, E.E. (1999). Self-assembly of polyglutamine-containing huntingtin fragments into amyloid-like fibrils: implications for Huntington's disease pathology. *Proc Natl Acad Sci U S A* 96**,** 4604-4609.

SCHLUDI, M.H., BECKER, L., GARRETT, L., GENDRON, T.F., ZHOU, Q., SCHREIBER, F., POPPER, B., DIMOU, L., STROM, T.M., WINKELMANN, J., VON THADEN, A., RENTZSCH, K., MAY, S., MICHAELSEN, M., SCHWENK, B.M., TAN, J., SCHOSER, B., DIETERICH, M., PETRUCELLI, L., HÖLTER, S.M., WURST, W., FUCHS, H., GAILUS-DURNER, V., DE ANGELIS, M.H., KLOPSTOCK, T., ARZBERGER, T., and EDBAUER, D. (2017). Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. *Acta neuropathologica* 134**,** 241-254.

SCHLUDI, M.H., MAY, S., GRÄSSER, F.A., RENTZSCH, K., KREMMER, E., KÜPPER, C., KLOPSTOCK, T., GERMAN CONSORTIUM FOR FRONTOTEMPORAL LOBAR, D., BAVARIAN BRAIN BANKING, A., ARZBERGER, T., and EDBAUER, D. (2015). Distribution of dipeptide repeat proteins in cellular models and C9orf72 mutation cases suggests link to transcriptional silencing. *Acta Neuropathologica* 130**,** 537-555.

SELLIER, C., BUIJSEN, R.A.M., HE, F., NATLA, S., JUNG, L., TROPEL, P., GAUCHEROT, A., JACOBS, H., MEZIANE, H., VINCENT, A., CHAMPY, M.F., SORG, T., PAVLOVIC, G., WATTENHOFER-DONZE, M., BIRLING, M.C., OULAD-ABDELGHANI, M.,

EBERLING, P., RUFFENACH, F., JOINT, M., ANHEIM, M., MARTINEZ-CERDENO, V., TASSONE, F., WILLEMSEN, R., HUKEMA, R.K., VIVILLE, S., MARTINAT, C., TODD, P.K., and CHARLET-BERGUERAND, N. (2017). Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. *Neuron* 93**,** 331-347.

SELLIER, C., CAMPANARI, M.-L., JULIE CORBIER, C., GAUCHEROT, A., KOLB-CHEYNEL, I., OULAD-ABDELGHANI, M., RUFFENACH, F., PAGE, A., CIURA, S., KABASHI, E., and CHARLET-BERGUERAND, N. (2016). Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *The EMBO journal* 35**,** 1276-1297.

SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N.S., WANG, J.T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B., and IDEKER, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13**,** 2498-2504.

SHAO, S., MURRAY, J., BROWN, A., TAUNTON, J., RAMAKRISHNAN, V., and HEGDE, R.S. (2016). Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. *Cell* 167**,** 1229-1240.e1215.

SHATUNOV, A., MOK, K., NEWHOUSE, S., WEALE, M.E., SMITH, B., VANCE, C., JOHNSON, L., VELDINK, J.H., VAN ES, M.A., VAN DEN BERG, L.H., ROBBERECHT, W., VAN DAMME, P., HARDIMAN, O., FARMER, A.E., LEWIS, C.M., BUTLER, A.W., ABEL, O., ANDERSEN, P.M., FOGH, I., SILANI, V., CHIÒ, A., TRAYNOR, B.J., MELKI, J., MEININGER, V., LANDERS, J.E., MCGUFFIN, P., GLASS, J.D., PALL, H., LEIGH, P.N., HARDY, J., BROWN, R.H., JR., POWELL, J.F., ORRELL, R.W., MORRISON, K.E., SHAW, P.J., SHAW, C.E., and AL-CHALABI, A. (2010). Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *The Lancet. Neurology* 9**,** 986-994.

SHI, K.Y., MORI, E., NIZAMI, Z.F., LIN, Y., KATO, M., XIANG, S., WU, L.C., DING, M., YU, Y., GALL, J.G., and MCKNIGHT, S.L. (2017a). Toxic PRn poly-dipeptides encoded by the C9orf72 repeat expansion block nuclear import and export. *PNAS* 114**,** E1111-E1117.

SHI, X., TIAN, B., MA, C., LIU, L., ZHANG, N., NA, Y., LI, J., LU, J., and QIAO, Y. (2017b). GSK3β activity is essential for senescence-associated heterochromatin foci (SAHF) formation induced by HMGA2 in WI38 cells. *American journal of translational research* 9**,** 167-174.

SHI, Y., LIN, S., STAATS, K.A., LI, Y., CHANG, W.H., HUNG, S.T., HENDRICKS, E., LINARES, G.R., WANG, Y., SON, E.Y., WEN, X., KISLER, K., WILKINSON, B., MENENDEZ, L., SUGAWARA, T., WOOLWINE, P., HUANG, M., COWAN, M.J., GE, B., KOUTSODENDRIS, N., SANDOR, K.P., KOMBERG, J., VANGOOR, V.R., SENTHILKUMAR, K., HENNES, V., SEAH, C., NELSON, A.R., CHENG, T.Y., LEE, S.J., AUGUST, P.R., CHEN, J.A., WISNIEWSKI, N., HANSON-SMITH, V., BELGARD, T.G., ZHANG, A., COBA, M., GRUNSEICH, C., WARD, M.E., VAN DEN BERG, L.H., PASTERKAMP, R.J., TROTTI, D., ZLOKOVIC, B.V., and ICHIDA, J.K. (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med* 24**,** 313-325.

SIMON-SANCHEZ, J., DOPPER, E.G., COHN-HOKKE, P.E., HUKEMA, R.K., NICOLAOU, N., SEELAAR, H., DE GRAAF, J.R., DE KONING, I., VAN SCHOOR, N.M., DEEG, D.J., SMITS, M., RAAPHORST, J., VAN DEN BERG, L.H., SCHELHAAS, H.J., DE DIE-SMULDERS, C.E., MAJOOR-KRAKAUER, D., ROZEMULLER, A.J., WILLEMSEN, R., PIJNENBURG, Y.A., HEUTINK, P., and VAN SWIETEN, J.C. (2012). The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. *Brain* 135**,** 723-735.

SIMONE, R., BALENDRA, R., MOENS, T.G., PREZA, E., WILSON, K.M., HESLEGRAVE, A., WOODLING, N.S., NICCOLI, T., GILBERT-JARAMILLO, J., ABDELKARIM, S., CLAYTON, E.L., CLARKE, M., KONRAD, M.T., NICOLL, A.J., MITCHELL, J.S., CALVO, A., CHIO, A., HOULDEN, H., POLKE, J.M., ISMAIL, M.A., STEPHENS, C.E.,

VO, T., FARAHAT, A.A., WILSON, W.D., BOYKIN, D.W., ZETTERBERG, H., PARTRIDGE, L., WRAY, S., PARKINSON, G., NEIDLE, S., PATANI, R., FRATTA, P., and ISAACS, A.M. (2018). G-quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *EMBO Mol Med* 10**,** 22-31.

SIVADASAN, R., HORNBURG, D., DREPPER, C., FRANK, N., JABLONKA, S., HANSEL, A., LOJEWSKI, X., STERNECKERT, J., HERMANN, A., SHAW, P.J., INCE, P.G., MANN, M., MEISSNER, F., and SENDTNER, M. (2016). C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nature Neuroscience* 19**,** 1610-1618.

SMITH, B.N., NEWHOUSE, S., SHATUNOV, A., VANCE, C., TOPP, S., JOHNSON, L., MILLER, J., LEE, Y., TROAKES, C., SCOTT, K.M., JONES, A., GRAY, I., WRIGHT, J., HORTOBAGYI, T., AL-SARRAJ, S., ROGELJ, B., POWELL, J., LUPTON, M., LOVESTONE, S., SAPP, P.C., WEBER, M., NESTOR, P.J., SCHELHAAS, H.J., ASBROEK, A.A., SILANI, V., GELLERA, C., TARONI, F., TICOZZI, N., VAN DEN BERG, L., VELDINK, J., VAN DAMME, P., ROBBERECHT, W., SHAW, P.J., KIRBY, J., PALL, H., MORRISON, K.E., MORRIS, A., DE BELLEROCHE, J., VIANNEY DE JONG, J.M., BAAS, F., ANDERSEN, P.M., LANDERS, J., BROWN, R.H., JR., WEALE, M.E., AL-CHALABI, A., and SHAW, C.E. (2013). The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *Eur J Hum Genet* 21**,** 102-108.

SREEDHARAN, J., BLAIR, I.P., TRIPATHI, V.B., HU, X., VANCE, C., ROGELJ, B., ACKERLEY, S., DURNALL, J.C., WILLIAMS, K.L., BURATTI, E., BARALLE, F., DE BELLEROCHE, J., MITCHELL, J.D., LEIGH, P.N., AL-CHALABI, A., MILLER, C.C., NICHOLSON, G., and SHAW, C.E. (2008). TDP-43 Mutations in Familial and Sporadic Amyotrophic Lateral Sclerosis. *Science* 319**,** 1668.

STEPTO, A., GALLO, J.-M., SHAW, C.E., and HIRTH, F. (2014). Modelling C9ORF72 hexanucleotide repeat expansion in amyotrophic lateral sclerosis and frontotemporal dementia. *Acta Neuropathologica* 127**,** 377-389.

STREBEL, A., HARR, T., BACHMANN, F., WERNLI, M., and ERB, P. (2001). Green fluorescent protein as a novel tool to measure apoptosis and necrosis. *Cytometry* 43**,** 126-133.

SU, Z., ZHANG, Y., GENDRON, T.F., BAUER, P.O., CHEW, J., YANG, W.Y., FOSTVEDT, E., JANSEN-WEST, K., BELZIL, V.V., DESARO, P., JOHNSTON, A., OVERSTREET, K., OH, S.Y., TODD, P.K., BERRY, J.D., CUDKOWICZ, M.E., BOEVE, B.F., DICKSON, D., FLOETER, M.K., TRAYNOR, B.J., MORELLI, C., RATTI, A., SILANI, V., RADEMAKERS, R., BROWN, R.H., ROTHSTEIN, J.D., BOYLAN, K.B., PETRUCELLI, L., and DISNEY, M.D. (2014). Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* 83**,** 1043-1050.

SUDMANT, P.H., LEE, H., DOMINGUEZ, D., HEIMAN, M., and BURGE, C.B. (2018). Widespread Accumulation of Ribosome-Associated Isolated 3′ UTRs in Neuronal Cell Populations of the Aging Brain. *Cell Reports* 25**,** 2447-2456.e2444.

SUDRIA-LOPEZ, E., KOPPERS, M., DE WIT, M., VAN DER MEER, C., WESTENENG, H.-J., ZUNDEL, C.A.C., YOUSSEF, S.A., HARKEMA, L., DE BRUIN, A., VELDINK, J.H., VAN DEN BERG, L.H., and PASTERKAMP, R.J. (2016). Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta neuropathologica* 132**,** 145-147.

SUH, E., LEE, E.B., NEAL, D., WOOD, E.M., TOLEDO, J.B., RENNERT, L., IRWIN, D.J., MCMILLAN, C.T., KROCK, B., ELMAN, L.B., MCCLUSKEY, L.F., GROSSMAN, M., XIE, S.X., TROJANOWSKI, J.Q., and VAN DEERLIN, V.M. (2015). Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. *Acta Neuropathol* 130**,** 363-372.

SUI, X., PIRES, D.E.V., ORMSBY, A.R., COX, D., NIE, S., VECCHI, G., VENDRUSCOLO, M., ASCHER, D.B., REID, G.E., and HATTERS, D.M. (2020). Widespread remodeling of

proteome solubility in response to different protein homeostasis stresses. *Proceedings of the National Academy of Sciences* 117**,** 2422.

SULLIVAN, P.M., ZHOU, X., ROBINS, A.M., PAUSHTER, D.H., KIM, D., SMOLKA, M.B., and HU, F. (2016). The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta neuropathologica communications* 4**,** 51-51.

SUZUKI, H., SHIBAGAKI, Y., HATTORI, S., and MATSUOKA, M. (2018). The proline–arginine repeat protein linked to C9-ALS/FTD causes neuronal toxicity by inhibiting the DEAD-box RNA helicase-mediated ribosome biogenesis. *Cell Death & Disease* 9**,** 975.

SWINNEN, B., BENTO-ABREU, A., GENDRON, T.F., BOEYNAEMS, S., BOGAERT, E., NUYTS, R., TIMMERS, M., SCHEVENEELS, W., HERSMUS, N., WANG, J., MIZIELINSKA, S., ISAACS, A.M., PETRUCELLI, L., LEMMENS, R., VAN DAMME, P., VAN DEN BOSCH, L., and ROBBERECHT, W. (2018). A zebrafish model for C9orf72 ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol* 135**,** 427-443.

SWINNEN, B., and ROBBERECHT, W. (2014). The phenotypic variability of amyotrophic lateral sclerosis. *Nature Reviews Neurology* 10**,** 661-670.

SYNOFZIK, M., MAETZLER, W., GREHL, T., PRUDLO, J., VOM HAGEN, J.M., HAACK, T., REBASSOO, P., MUNZ, M., SCHÖLS, L., and BISKUP, S. (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**,** 2949.e2913-2949.e2917.

SZKLARCZYK, D., MORRIS, J.H., COOK, H., KUHN, M., WYDER, S., SIMONOVIC, M., SANTOS, A., DONCHEVA, N.T., ROTH, A., BORK, P., JENSEN, L.J., and VON MERING, C. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 45**,** D362-D368.

TABET, R., SCHAEFFER, L., FREYERMUTH, F., JAMBEAU, M., WORKMAN, M., LEE, C.Z., LIN, C.C., JIANG, J., JANSEN-WEST, K., ABOU-HAMDAN, H., DÉSAUBRY, L., GENDRON, T., PETRUCELLI, L., MARTIN, F., and LAGIER-TOURENNE, C. (2018). CUG initiation and frameshifting enable production of dipeptide repeat proteins from ALS/FTD C9ORF72 transcripts. *Nat Commun* 9**,** 152.

TANG, J., FRANKEL, A., COOK, R.J., KIM, S., PAIK, W.K., WILLIAMS, K.R., CLARKE, S., and HERSCHMAN, H.R. (2000). PRMT1 is the predominant type I protein arginine methyltransferase in mammalian cells. *J Biol Chem* 275**,** 7723-7730.

TAO, Z., WANG, H., XIA, Q., LI, K., LI, K., JIANG, X., XU, G., WANG, G., and YING, Z. (2015). Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity. *Hum Mol Genet* 24**,** 2426-2441.

TAYLOR, J.P., BROWN, R.H., JR., and CLEVELAND, D.W. (2016a). Decoding ALS: from genes to mechanism. *Nature* 539**,** 197-206.

TAYLOR, N., ELBAUM-GARFINKLE, S., VAIDYA, N., ZHANG, H., STONE, H.A., and BRANGWYNNE, C.P. (2016b). Biophysical characterization of organelle-based RNA/protein liquid phases using microfluidics. *Soft Matter* 12**,** 9142-9150.

THERRIEN, M., ROULEAU, G.A., DION, P.A., and PARKER, J.A. (2013). Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in C. elegans. *PLOS ONE* 8**,** e83450.

THIMMEGOWDA, C., and YELDUR, P.V. (2016). Occurrence, Functions and Biological Significance of Arginine-Rich Proteins. *Current Protein & Peptide Science* 17**,** 507-516.

THOMPSON, J.C., STOPFORD, C.L., SNOWDEN, J.S., and NEARY, D. (2005). Qualitative neuropsychological performance characteristics in frontotemporal dementia and Alzheimer's disease. *Journal of neurology, neurosurgery, and psychiatry* 76**,** 920-927.

TIBSHIRANI, M., TRADEWELL, M.L., MATTINA, K.R., MINOTTI, S., YANG, W., ZHOU, H., STRONG, M.J., HAYWARD, L.J., and DURHAM, H.D. (2015). Cytoplasmic

sequestration of FUS/TLS associated with ALS alters histone marks through loss of nuclear protein arginine methyltransferase 1. *Human molecular genetics* 24**,** 773-786.

TODD, P., OH, S., KRANS, A., HE, F., SELLIER, C., FRAZER, M., RENOUX, A., CHEN, K.-C., SCAGLIONE, K.M., BASRUR, V., ELENITOBA-JOHNSON, K., VONSATTEL, J., LOUIS, E., SUTTON, M., TAYLOR, J.P., MILLS, R., CHARLET-BERGUERAND, N., and PAULSON, H. (2013). CGG Repeat-Associated Translation Mediates Neurodegeneration in Fragile X Tremor Ataxia Syndrome. *Neuron* 78.

TRAN, H., ALMEIDA, S., MOORE, J., GENDRON, T.F., CHALASANI, U., LU, Y., DU, X., NICKERSON, J.A., PETRUCELLI, L., WENG, Z., and GAO, F.B. (2015). Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron* 87**,** 1207-1214.

TSVETKOV, A.S., MILLER, J., ARRASATE, M., WONG, J.S., PLEISS, M.A., and FINKBEINER, S. (2010). A small-molecule scaffold induces autophagy in primary neurons and protects against toxicity in a Huntington disease model. *Proceedings of the National Academy of Sciences* 107**,** 16982.

TURNER, M.R., AL-CHALABI, A., CHIO, A., HARDIMAN, O., KIERNAN, M.C., ROHRER, J.D., ROWE, J., SEELEY, W., and TALBOT, K. (2017). Genetic screening in sporadic ALS and FTD. *Journal of Neurology, Neurosurgery & Psychiatry* 88**,** 1042.

UGOLINO, J., JI, Y.J., CONCHINA, K., CHU, J., NIRUJOGI, R.S., PANDEY, A., BRADY, N.R., HAMACHER-BRADY, A., and WANG, J. (2016). Loss of C9orf72 Enhances Autophagic Activity via Deregulated mTOR and TFEB Signaling. *PLoS Genet* 12**,** e1006443.

UHLMANN, T., GEOGHEGAN, V.L., THOMAS, B., RIDLOVA, G., TRUDGIAN, D.C., and ACUTO, O. (2012). A method for large-scale identification of protein arginine methylation. *Molecular & cellular proteomics : MCP* 11**,** 1489-1499.

VAN BLITTERSWIJK, M., DEJESUS-HERNANDEZ, M., NIEMANTSVERDRIET, E., MURRAY, M.E., HECKMAN, M.G., DIEHL, N.N., BROWN, P.H., BAKER, M.C., FINCH, N.A., BAUER, P.O., SERRANO, G., BEACH, T.G., JOSEPHS, K.A., KNOPMAN, D.S., PETERSEN, R.C., BOEVE, B.F., GRAFF-RADFORD, N.R., BOYLAN, K.B., PETRUCELLI, L., DICKSON, D.W., and RADEMAKERS, R. (2013). Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *The Lancet. Neurology* 12**,** 978-988.

VAN DER ZEE, J., GIJSELINCK, I., DILLEN, L., VAN LANGENHOVE, T., THEUNS, J., ENGELBORGHS, S., PHILTJENS, S., VANDENBULCKE, M., SLEEGERS, K., SIEBEN, A., BAUMER, V., MAES, G., CORSMIT, E., BORRONI, B., PADOVANI, A., ARCHETTI, S., PERNECZKY, R., DIEHL-SCHMID, J., DE MENDONCA, A., MILTENBERGER-MILTENYI, G., PEREIRA, S., PIMENTEL, J., NACMIAS, B., BAGNOLI, S., SORBI, S., GRAFF, C., CHIANG, H.H., WESTERLUND, M., SANCHEZ-VALLE, R., LLADO, A., GELPI, E., SANTANA, I., ALMEIDA, M.R., SANTIAGO, B., FRISONI, G., ZANETTI, O., BONVICINI, C., SYNOFZIK, M., MAETZLER, W., VOM HAGEN, J.M., SCHOLS, L., HENEKA, M.T., JESSEN, F., MATEJ, R., PAROBKOVA, E., KOVACS, G.G., STROBEL, T., SARAFOV, S., TOURNEV, I., JORDANOVA, A., DANEK, A., ARZBERGER, T., FABRIZI, G.M., TESTI, S., SALMON, E., SANTENS, P., MARTIN, J.J., CRAS, P., VANDENBERGHE, R., DE DEYN, P.P., CRUTS, M., VAN BROECKHOVEN, C., VAN DER ZEE, J., GIJSELINCK, I., DILLEN, L., VAN LANGENHOVE, T., THEUNS, J., PHILTJENS, S., SLEEGERS, K., BAUMER, V., MAES, G., CORSMIT, E., CRUTS, M., VAN BROECKHOVEN, C., VAN DER ZEE, J., GIJSELINCK, I., DILLEN, L., VAN LANGENHOVE, T., PHILTJENS, S., THEUNS, J., SLEEGERS, K., BAUMER, V., MAES, G., CRUTS, M., VAN BROECKHOVEN, C., ENGELBORGHS, S., DE DEYN, P.P., CRAS, P., ENGELBORGHS, S., DE DEYN, P.P., VANDENBULCKE, M., VANDENBULCKE, M., BORRONI, B., PADOVANI, A., ARCHETTI, S., PERNECZKY, R., DIEHL-SCHMID, J., SYNOFZIK, M., MAETZLER, W., MULLER VOM HAGEN, J.,

et al. (2013). A pan-European study of the C9orf72 repeat associated with FTLD: geographic prevalence, genomic instability, and intermediate repeats. *Hum Mutat* 34**,** 363-373.

VAN ES, M.A., DIEKSTRA, F.P., VELDINK, J.H., BAAS, F., BOURQUE, P.R., SCHELHAAS, H.J., STRENGMAN, E., HENNEKAM, E.A.M., LINDHOUT, D., OPHOFF, R.A., and VAN DEN BERG, L.H. (2009). A CASE OF ALS-FTD IN A LARGE FALS PEDIGREE WITH A K17I ANG MUTATION. *Neurology* 72**,** 287.

VAN ES, M.A., HARDIMAN, O., CHIO, A., AL-CHALABI, A., PASTERKAMP, R.J., VELDINK, J.H., and VAN DEN BERG, L.H. (2017). Amyotrophic lateral sclerosis. *The Lancet* 390**,** 2084-2098.

VAN MOSSEVELDE, S., VAN DER ZEE, J., GIJSELINCK, I., SLEEGERS, K., DE BLEECKER, J., SIEBEN, A., VANDENBERGHE, R., VAN LANGENHOVE, T., BAETS, J., DERYCK, O., SANTENS, P., IVANOIU, A., WILLEMS, C., BAUMER, V., VAN DEN BROECK, M., PEETERS, K., MATTHEIJSSENS, M., DE JONGHE, P., CRAS, P., MARTIN, J.J., CRUTS, M., DE DEYN, P.P., ENGELBORGHS, S., and VAN BROECKHOVEN, C. (2017). Clinical Evidence of Disease Anticipation in Families Segregating a C9orf72 Repeat Expansion. *JAMA Neurol* 74**,** 445-452.

VANCE, C., AL-CHALABI, A., RUDDY, D., SMITH, B.N., HU, X., SREEDHARAN, J., SIDDIQUE, T., SCHELHAAS, H.J., KUSTERS, B., TROOST, D., BAAS, F., DE JONG, V., and SHAW, C.E. (2006). Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2,Äì21.3. *Brain* 129**,** 868-876.

VANNESTE, J., VERCRUYSSE, T., BOEYNAEMS, S., SICART, A., VAN DAMME, P., DAELEMANS, D., and VAN DEN BOSCH, L. (2019). C9orf72-generated poly-GR and poly-PR do not directly interfere with nucleocytoplasmic transport. *Scientific Reports* 9**,** 15728.

VATSAVAYAI, S.C., YOON, S.J., GARDNER, R.C., GENDRON, T.F., VARGAS, J.N.S., TRUJILLO, A., PRIBADI, M., PHILLIPS, J.J., GAUS, S.E., HIXSON, J.D., GARCIA, P.A., RABINOVICI, G.D., COPPOLA, G., GESCHWIND, D.H., PETRUCELLI, L., MILLER, B.L., and SEELEY, W.W. (2016). Timing and significance of pathological features in C9orf72 expansion-associated frontotemporal dementia. *Brain* 139**,** 3202-3216.

VIZCAINO, J.A., CSORDAS, A., DEL-TORO, N., DIANES, J.A., GRISS, J., LAVIDAS, I., MAYER, G., PEREZ-RIVEROL, Y., REISINGER, F., TERNENT, T., XU, Q.W., WANG, R., and HERMJAKOB, H. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44**,** D447-456.

WAITE, A.J., BAUMER, D., EAST, S., NEAL, J., MORRIS, H.R., ANSORGE, O., and BLAKE, D.J. (2014). Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol Aging* 35**,** 1779.e1775-1779.e1713.

WARREN, J.D., ROHRER, J.D., and ROSSOR, M.N. (2013). Clinical review. Frontotemporal dementia. *BMJ (Clinical research ed.)* 347**,** f4827-f4827.

WEAR, M.P., KRYNDUSHKIN, D., O'MEALLY, R., SONNENBERG, J.L., COLE, R.N., and SHEWMAKER, F.P. (2015). Proteins with Intrinsically Disordered Domains Are Preferentially Recruited to Polyglutamine Aggregates. *PLoS One* 10**,** e0136362.

WEBSTER, C.P., SMITH, E.F., BAUER, C.S., MOLLER, A., HAUTBERGUE, G.M., FERRAIUOLO, L., MYSZCZYNSKA, M.A., HIGGINBOTTOM, A., WALSH, M.J., WHITWORTH, A.J., KASPAR, B.K., MEYER, K., SHAW, P.J., GRIERSON, A.J., and DE VOS, K.J. (2016). The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *Embo j* 35**,** 1656-1676.

WEN, X., TAN, W., WESTERGARD, T., KRISHNAMURTHY, K., MARKANDAIAH, S.S., SHI, Y., LIN, S., SHNEIDER, N.A., MONAGHAN, J., PANDEY, U.B., PASINELLI, P., ICHIDA, J.K., and TROTTI, D. (2014). Antisense proline-arginine RAN dipeptides linked to

C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron* 84**,** 1213-1225.

WHEATON, M.W., SALAMONE, A.R., MOSNIK, D.M., MCDONALD, R.O., APPEL, S.H., SCHMOLCK, H.I., RINGHOLZ, G.M., and SCHULZ, P.E. (2007). Cognitive impairment in familial ALS. *Neurology* 69**,** 1411.

WOERNER, A.C., FROTTIN, F., HORNBURG, D., FENG, L.R., MEISSNER, F., PATRA, M., TATZELT, J., MANN, M., WINKLHOFER, K.F., HARTL, F.U., and HIPP, M.S. (2016). Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. *Science* 351**,** 173-176.

WOJCIECHOWSKA, M., OLEJNICZAK, M., GALKA-MARCINIAK, P., JAZUREK, M., and KRZYZOSIAK, W. (2014). RAN translation and frameshifting as translational challenges at simple repeats of human neurodegenerative disorders. *Nucleic acids research* 42.

WOOLLACOTT, I.O.C., and ROHRER, J.D. (2016). The clinical spectrum of sporadic and familial forms of frontotemporal dementia. *Journal of Neurochemistry* 138**,** 6-31.

WU, C.-H., FALLINI, C., TICOZZI, N., KEAGLE, P.J., SAPP, P.C., PIOTROWSKA, K., LOWE, P., KOPPERS, M., MCKENNA-YASEK, D., BARON, D.M., KOST, J.E., GONZALEZ-PEREZ, P., FOX, A.D., ADAMS, J., TARONI, F., TILOCA, C., LECLERC, A.L., CHAFE, S.C., MANGROO, D., MOORE, M.J., ZITZEWITZ, J.A., XU, Z.-S., VAN DEN BERG, L.H., GLASS, J.D., SICILIANO, G., CIRULLI, E.T., GOLDSTEIN, D.B., SALACHAS, F., MEININGER, V., ROSSOLL, W., RATTI, A., GELLERA, C., BOSCO, D.A., BASSELL, G.J., SILANI, V., DRORY, V.E., BROWN JR, R.H., and LANDERS, J.E. (2012). Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* 488**,** 499-503.

WYTTENBACH, A., SWARTZ, J., KITA, H., THYKJAER, T., CARMICHAEL, J., BRADLEY, J., BROWN, R., MAXWELL, M., SCHAPIRA, A., ORNTOFT, T.F., KATO, K., and RUBINSZTEIN, D.C. (2001). Polyglutamine expansions cause decreased CRE-mediated transcription and early gene expression changes prior to cell death in an inducible cell model of Huntington's disease. *Hum Mol Genet* 10**,** 1829-1845.

XI, Z., ZINMAN, L., MORENO, D., SCHYMICK, J., LIANG, Y., SATO, C., ZHENG, Y., GHANI, M., DIB, S., KEITH, J., ROBERTSON, J., and ROGAEVA, E. (2013). Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am J Hum Genet* 92**,** 981-989.

XIAO, Q., GIL, S.C., YAN, P., WANG, Y., HAN, S., GONZALES, E., PEREZ, R., CIRRITO, J.R., and LEE, J.M. (2012). Role of phosphatidylinositol clathrin assembly lymphoid-myeloid leukemia (PICALM) in intracellular amyloid precursor protein (APP) processing and amyloid plaque pathogenesis. *J Biol Chem* 287**,** 21279-21289.

XIAO, S., MACNAIR, L., MCGOLDRICK, P., MCKEEVER, P.M., MCLEAN, J.R., ZHANG, M., KEITH, J., ZINMAN, L., ROGAEVA, E., and ROBERTSON, J. (2015). Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Annals of Neurology* 78**,** 568-583.

XIAO, S., MACNAIR, L., MCLEAN, J., MCGOLDRICK, P., MCKEEVER, P., SOLEIMANI, S., KEITH, J., ZINMAN, L., ROGAEVA, E., and ROBERTSON, J. (2016). C9orf72 isoforms in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration. *Brain Research* 1647**,** 43-49.

XU, W., and XU, J. (2018). C9orf72 Dipeptide Repeats Cause Selective Neurodegeneration and Cell-Autonomous Excitotoxicity in Drosophila Glutamatergic Neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 38**,** 7741-7752.

XU, Z., POIDEVIN, M., LI, X., LI, Y., SHU, L., NELSON, D.L., LI, H., HALES, C.M., GEARING, M., WINGO, T.S., and JIN, P. (2013). Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc Natl Acad Sci U S A* 110**,** 7778-7783.

YAMAKAWA, M., ITO, D., HONDA, T., KUBO, K.-I., NODA, M., NAKAJIMA, K., and SUZUKI, N. (2014). *Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS.*

YANG, H., YUE, H.W., HE, W.T., HONG, J.Y., JIANG, L.L., and HU, H.Y. (2018). PolyQ-expanded huntingtin and ataxin-3 sequester ubiquitin adaptors hHR23B and UBQLN2 into aggregates via conjugated ubiquitin. *FASEB J* 32**,** 2923-2933.

YANG, M., LIANG, C., SWAMINATHAN, K., HERRLINGER, S., LAI, F., SHIEKHATTAR, R., and CHEN, J.-F. (2016). A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Science advances* 2**,** e1601167-e1601167.

YANG, Y., and BEDFORD, M.T. (2013). Protein arginine methyltransferases and cancer. *Nat Rev Cancer* 13**,** 37-50.

YANLING ZHAO, D., GISH, G., BRAUNSCHWEIG, U., LI, Y., NI, Z., SCHMITGES, F.W., ZHONG, G., LIU, K., LI, W., MOFFAT, J., VEDADI, M., MIN, J., PAWSON, T.J., BLENCOWE, B.J., and GREENBLATT, J.F. (2016). SMN and symmetric arginine dimethylation of RNA polymerase II C-terminal domain control termination. *Nature* 529**,** 48-53.

YIN, S., LOPEZ-GONZALEZ, R., KUNZ, R.C., GANGOPADHYAY, J., BORUFKA, C., GYGI, S.P., GAO, F.-B., and REED, R. (2017). Evidence that C9ORF72 Dipeptide Repeat Proteins Associate with U2 snRNP to Cause Mis-splicing in ALS/FTD Patients. *Cell Reports* 19**,** 2244-2256.

YOUNG, D.J., GUYDOSH, N.R., ZHANG, F., HINNEBUSCH, A.G., and GREEN, R. (2015). Rli1/ABCE1 Recycles Terminating Ribosomes and Controls Translation Reinitiation in 3'UTRs In Vivo. *Cell* 162**,** 872-884.

YU, C.H., TEULADE-FICHOU, M.P., and OLSTHOORN, R.C. (2014). Stimulation of ribosomal frameshifting by RNA G-quadruplex structures. *Nucleic Acids Res* 42**,** 1887-1892.

ZHANG, D., IYER, L.M., HE, F., and ARAVIND, L. (2012). Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Frontiers in genetics* 3**,** 283-283.

ZHANG, K., DONNELLY, C.J., HAEUSLER, A.R., GRIMA, J.C., MACHAMER, J.B., STEINWALD, P., DALEY, E.L., MILLER, S.J., CUNNINGHAM, K.M., VIDENSKY, S., GUPTA, S., THOMAS, M.A., HONG, I., CHIU, S.-L., HUGANIR, R.L., OSTROW, L.W., MATUNIS, M.J., WANG, J., SATTLER, R., LLOYD, T.E., and ROTHSTEIN, J.D. (2015). The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* 525**,** 56-61.

ZHANG, Y.-J., GENDRON, T.F., EBBERT, M.T.W., O'RAW, A.D., YUE, M., JANSEN-WEST, K., ZHANG, X., PRUDENCIO, M., CHEW, J., COOK, C.N., DAUGHRITY, L.M., TONG, J., SONG, Y., PICKLES, S.R., CASTANEDES-CASEY, M., KURTI, A., RADEMAKERS, R., OSKARSSON, B., DICKSON, D.W., HU, W., GITLER, A.D., FRYER, J.D., and PETRUCELLI, L. (2018). Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nature Medicine* 24**,** 1136-1142.

ZHANG, Y.-J., GENDRON, T.F., GRIMA, J.C., SASAGURI, H., JANSEN-WEST, K., XU, Y.-F., KATZMAN, R.B., GASS, J., MURRAY, M.E., SHINOHARA, M., LIN, W.-L., GARRETT, A., STANKOWSKI, J.N., DAUGHRITY, L., TONG, J., PERKERSON, E.A., YUE, M., CHEW, J., CASTANEDES-CASEY, M., KURTI, A., WANG, Z.S., LIESINGER, A.M., BAKER, J.D., JIANG, J., LAGIER-TOURENNE, C., EDBAUER, D., CLEVELAND, D.W., RADEMAKERS, R., BOYLAN, K.B., BU, G., LINK, C.D., DICKEY, C.A., ROTHSTEIN, J.D., DICKSON, D.W., FRYER, J.D., and PETRUCELLI, L. (2016). C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat Neurosci* 19**,** 668-677.

ZHANG, Y.-J., GUO, L., GONZALES, P.K., GENDRON, T.F., WU, Y., JANSEN-WEST, K., O'RAW, A.D., PICKLES, S.R., PRUDENCIO, M., CARLOMAGNO, Y., GACHECHILADZE, M.A., LUDWIG, C., TIAN, R., CHEW, J., DETURE, M., LIN, W.-L.,

TONG, J., DAUGHRITY, L.M., YUE, M., SONG, Y., ANDERSEN, J.W., CASTANEDES-CASEY, M., KURTI, A., DATTA, A., ANTOGNETTI, G., MCCAMPBELL, A., RADEMAKERS, R., OSKARSSON, B., DICKSON, D.W., KAMPMANN, M., WARD, M.E., FRYER, J.D., LINK, C.D., SHORTER, J., and PETRUCELLI, L. (2019). Heterochromatin anomalies and double-stranded RNA accumulation underlie C9orf72 poly(PR) toxicity. *Science* 363**,** eaav2606.

ZHANG, Y.-J., JANSEN-WEST, K., XU, Y.-F., GENDRON, T.F., BIENIEK, K.F., LIN, W.-L., SASAGURI, H., CAULFIELD, T., HUBBARD, J., DAUGHRITY, L., CHEW, J., BELZIL, V.V., PRUDENCIO, M., STANKOWSKI, J.N., CASTANEDES-CASEY, M., WHITELAW, E., ASH, P.E.A., DETURE, M., RADEMAKERS, R., BOYLAN, K.B., DICKSON, D.W., and PETRUCELLI, L. (2014). Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathologica* 128**,** 505-524.

ZU, T., GIBBENS, B., DOTY, N.S., GOMES-PEREIRA, M., HUGUET, A., STONE, M.D., MARGOLIS, J., PETERSON, M., MARKOWSKI, T.W., INGRAM, M.A., NAN, Z., FORSTER, C., LOW, W.C., SCHOSER, B., SOMIA, N.V., CLARK, H.B., SCHMECHEL, S., BITTERMAN, P.B., GOURDON, G., SWANSON, M.S., MOSELEY, M., and RANUM, L.P. (2011). Non-ATG-initiated translation directed by microsatellite expansions. *Proc Natl Acad Sci U S A* 108**,** 260-265.

ZU, T., LIU, Y., BANEZ-CORONEL, M., REID, T., PLETNIKOVA, O., LEWIS, J., MILLER, T.M., HARMS, M.B., FALCHOOK, A.E., SUBRAMONY, S.H., OSTROW, L.W., ROTHSTEIN, J.D., TRONCOSO, J.C., and RANUM, L.P. (2013). RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proc Natl Acad Sci U S A* 110**,** E4968-4977.