Identification of Arfaptin-2 as a potential therapeutic target for Amyotrophic Lateral Sclerosis using iPSC-derived motor neurons and zebrafish as models of ALS

Anushka Bhargava

Thesis Submitted for the degree of Doctor of Philosophy (PhD)

Sheffield Institute for Translational Neuroscience

February 2023

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a rare progressive adult-onset neurodegenerative disorder that is characterized by an accelerated and irreversible loss or degeneration of upper motor neurons (UMNs) and lower motor neurons (LMNs). Motor neuron degeneration leads to muscle wasting, paralysis, and eventual death due to respiratory failure. The pathogenesis of neuronal degeneration in ALS is highly varied and multifactorial. Insoluble protein aggregates are a pathological hallmark feature of neurodegenerative disorders. The molecular makeup of protein aggregates deepens the understanding of disease pathogenesis involved in ALS and potentially provides novel targets for therapeutic interventions.

Previously, two independent studies have shown that Arfaptin-2 (*ARFIP2*) is involved in protein aggregation in Huntington's disease (HD) and its C-terminal component (HC-ARFIP2) can confer neuroprotective effects. Since neurodegenerative disorders share some common features of pathogenesis, therapeutic strategies may be transferrable. Preliminary work in our lab implied the involvement of ARFIP2 in ALS and to our knowledge, this has not been reported before. Accordingly, we sought to investigate the therapeutic potential of ARFIP2 in ALS. In particular, we explored whether HC-ARFIP2 could improve disease characteristics of ALS using *in vitro* and *in vivo* models that carry the *C9ORF72* mutation, the most common genetic cause of ALS. In this thesis, we have shown that the endogenous levels of ARFIP2 are elevated in the ALS disease background, using both human brain tissue sections and *C9ORF72*-ALS iPSC-derived motor neurons (MNs). To our knowledge, increased levels of ARFIP2 in ALS is a novel finding and this finding highlights its possible involvement in the disease. This warranted further assessment of the effects of targeting ARFIP2 on the pathogenesis of ALS.

This thesis describes the generation and validation of essential plasmids and appropriate delivery vectors used to investigate the effects of HC-ARFIP2 on *C9ORF72*-ALS disease characteristics in both *in vitro* and *in vivo* models of ALS. Validation of the various viral vectors and tagged constructs in a simple cell model (HEK293T cells) showed that transgenes were successfully expressed. The tools generated were then used in the more complex *in vitro* and *in vivo* ALS disease models

with the *C9ORF72* expansion mutation. While HC-ARFIP2 expression had no measurable effects on RNA foci both *in vitro* and *in vivo*, there was an effect of HC-ARFIP2 expression on dipeptide repeat proteins (DPRs). Results from a poly-GP electrochemiluminescence (Meso Scale Discovery (MSD)) ELISA showed that expressing HC-ARFIP2 significantly decreases the levels of the poly-GP DPR in *C9ORF72*-ALS iPSC-derived MNs, reducing the levels to background values. Furthermore, HC-ARFIP2 expression was shown to improve cell viability in *C9ORF72*-ALS iPSC-derived MNs, while the transgene expression was well tolerated over time.

With the observed therapeutic potential of HC-ARFIP2 expression in the *in vitro* model, the effect of HC-ARFIP2 was assessed in zebrafish models of ALS. Heterozygous HC-ARFIP2 expressing transgenic lines were successfully established and crossed with the heterozygous *C9ORF72*-ALS sense and antisense lines, which generated four genetically distinct experimental populations that were used to assess the effects of HC-ARFIP2 on *C9ORF72*-ALS disease characteristics in the *in vivo* model. Interestingly, preliminary data suggested that HC-ARFIP2 expression may have a similar effect on poly-GP DPRs, as seen in the *in vitro* model, in zebrafish with the *C9ORF72* mutation. However, a measurable effect was seen only in the sense genotype background. Using the *in vivo* model, the neuro-muscular integrity of the fish and the effect of HC-ARFIP2 on motor performance was assessed by conducting a swim tunnel experiment. 6-month-old adult zebrafish showed a strong phenotype of motor performance deficits. Preliminary data showed that HC-ARFIP2 expression has a measurable effect on the motor performance deficit phenotype and is able to rescue it, further highlighting the therapeutic potential of HC-ARFIP2. Furthermore, HC-ARFIP2 expression does not show any adverse effects on motor performance. The genotype, gender and weight of the fish were considered when required in these studies.

Taken together, the data presented in this thesis identifies the therapeutic potential of targeting ARFIP2 in ALS and shows that expressing HC-ARFIP2 may ameliorate the *C9ORF72*-ALS-associated disease phenotypes. Though the data presented is promising and encouraging, further work is required to completely comprehend the involvement and effects of HC-ARFIP2 in ALS. Nonetheless, it is important to highlight

that the current project shows findings of great importance that may contribute to the identification and development of a novel therapeutic target for ALS.

ACKNOWLEDGEMENTS

I am thankful to my supervisor, Dr. Ke Ning, for giving me this opportunity and providing me with unwavering support. I would like to thank Dr. Tennore Ramesh for his guidance, encouragement, and knowledge of zebrafish, which has been a large part of my Ph.D. project. I am grateful to him for his constant support and providing me with opportunities to further my career. I am also thankful to Prof. Pamela J Shaw for her valuable feedback and insights which have been critical to the success of my research.

I would like to especially extend my gratitude to Dr. Adrian Higginbottom for taking the time to provide me with invaluable feedback and constructive criticism on this thesis. His expertise, patience, and inspiration have been instrumental in shaping the direction of my work and helping me complete this thesis. It would not be possible to get through the writing period of my Ph.D. without his guidance. I would also like to thank Dr. Andrew Grierson for providing me with an empathetic ear, support and resources I needed to overcome the hurdles faced during my Ph.D.

I am grateful to my colleagues in SITraN, past and present, who have shared their knowledge and provided a supportive environment for learning and growth. Thank you to Dr. Vinay Godena, Dr. Eva Karyka, Dr. Dave Burrows, and Allan Shaw for taking the time out of their busy schedules to answer all my questions and help me with my project. Particularly, I would like to express my heartfelt gratitude to Dr. Cleide dos Santos de Souza for teaching me everything I know about cell culture and the stimulating discussions that shaped me into the researcher I am today.

I will forever cherish the friends that I have made during my time at SITraN. Their constant encouragement has been a source of strength and motivation. I am grateful for the memories, laughter, CPR trips and their unwavering support. Thank you to Emily, Matt, Emily, Toby, Shaila, Ale, Aytac and Marco for making Sheffield a home away from home.

I would also like to thank my biggest cheerleaders, my family – Ma, Papa, Didi and Jiju. They have been the driving force behind my success. Thank you for always believing in me and encouraging me to pursue my dreams. I could not have done this without their love and support. Finally, thank you to my fiancé, Shreyash, for making my dull days brighter and being the positive light in my life.

This PhD thesis is dedicated to my Ma and Papa

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

2-Mer	Positive Control Embryos That Contain 2 G4c2 Repeats
AD	Autosomal Dominant
ALS	Amyotrophic Lateral Sclerosis
AMP	Amplitude
ANG	Angiogenin
AR	Autosomal Recessive
ARFIP2	Arfaptin-2/ADP-Ribosylation Factor-Interacting Protein 2
ASO	Antisense Oligonucleotide
ATP	Adenosine Triphosphate
C9ORF72	Chromosome 9 Open Reading Frame 72
CSF	Cerebrospinal Fluid
DAO	D-Amino Acid Oxidase
DCTN	Dynactin
DEPC	Diethyl Pyrocarbonate
DMSO	Dimethyl Sulfoxide
DPR	Dipeptide Repeat Proteins
ELISA	Enzyme-Linked Immunosorbent Assay
FDA	Food And Drug Administration
FFPE	Formalin-Fixed Paraffin-Embedded
FISH	Fluorescence RNA In Situ Hybridisation
FL-ARFIP2	Full Length Arfaptin-2
FTD	Frontotemporal Dementia
FUS	Fused In Sarcoma
GFP	Green Fluorescent Protein
HC-ARFIP2	C-Terminal Component Of Arfaptin-2
HD	Huntington's Disease
HEK293T	Human Embryonic Kidney 293 Cells
HFE	Human Haemochromatosis
HRE	Hexanucleotide Repeat Expansion
HSR	Heat Shock Response
ICC	Immunocytochemistry
IF	Immunofluorescence

IHC	Immunohistochemistry
iPSC	Induced Pluripotent Stem Cells
ISO-ALS	Isogenic Control For ALS Lines
LMN	Lower Motor Neurons
LNA	Locked Nucleic Acid
LV	Lentivirus
MN	Motor Neurons
MND	Motor Neuron Disease
MOI	Multiplicity Of Infection
MSD	Meso Scale Discovery
NES	Nuclear Export Signal
NPC	Neuron Progenitor Cell
NTG	Non-Transgenic
OPTN	Optineurin
PBP	Progressive Bulbar Palsy
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEI	Polyethylenimine
PFA	Paraformaldehyde
PGRN	Granulin Precursor
PIC	Protease Inhibitor Cocktail
PLS	Primary Lateral Sclerosis
PMA	Progressive Muscular Atrophy
Poly-GA	Poly-Gly-Ala
Poly-GP	Poly-Gly-Pro
Poly-GR	Poly-Gly-Arg
Poly-PA	Poly-Pro-Ala
Poly-PR	Poly-Pro-Arg
PON	Paraoxonase 1
PRPH	Peripherin
qPCR	Quantitative Polymerase Chain Reaction
RFC	Relative Centrifugal Force
RNA	Ribonucleic Acid
ROS	Reactive Oxygen Species

RPM	Revolutions Per Minute
RT	Room Temperature
SBTB	Sheffield Brain Tissue Bank
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel
SETX	Senataxin
SMN	Survival Motor Neuron
SOD1	Superoxide Dismutase 1
SPG	Spastic Paraplegia
SQSTM1/p62	Sequestosome 1
TARDBP	TAR DNA-Binding Protein
TBS	Tris-Buffer Saline
TDP-43	TAR-DNA Binding Protein Of 43 kDa
UBQLN2	Ubiquilin-2
UMN	Upper Motor Neurons
UPS	Ubiquitin-Proteasome System
UT	Untransfected
VAPB	VAMP Associated Protein B
VCP	Valosin Containing Protein

1. INTRODUCTION

Motor neuron disease (MND) encompasses a group of neurodegenerative disorders affecting motor neurons (MNs). The premature degeneration of motor neurons, upper and/or lower, causes the loss of control over voluntary movements over time. The onset of MND is highly variable and occurs in both adults and children (spinal muscular atrophy). The cause of most MNDs is unknown, however, some may be inherited (National Institute of Neurological Disorders and Stroke (NINDS), 2014). The spectrum of diseases grouped under MND is differentiated depending on if the disease is sporadic or inherited; and if degeneration affects upper motor neurons (UMN), lower motor neurons (LMN), or both (Statland *et al.*, 2015). MND can be classified into four subclasses: progressive muscular atrophy (PMA), which predominantly affects LMNs; primary lateral sclerosis (PLS), which affects the UMNs; amyotrophic lateral sclerosis (ALS), where both UMNs and LMNs are affected; and progressive bulbar palsy (PBP), which mainly affects the bulbar musculature. Approximately 90% of MND patients have the ALS form of the disease and the rest within other subclasses at initial presentation usually eventually progress into ALS (Shaw, 2005). Consequently, the two terms, MND and ALS, are used interchangeably.

1.1 Amyotrophic Lateral Sclerosis

ALS, also known as Lou Gehrig's Disease in the US and MND in the UK, is a progressive neurodegenerative disorder that is characterized by an accelerated and irreversible loss or degeneration of UMNs and LMNs (Rowland and Shneider, 2001). These neurons are responsible for controlling voluntary muscles and their degeneration leads to muscle wasting, paralysis, and eventual death due to respiratory failure (Martin, Al Khleifat and Al-Chalabi, 2017). Though the disease mainly affects the motor neurons, ALS is now thought of as a multisystem neurodegenerative disorder due to the involvement of other types of neurons and the heterogeneity of the disease at the clinical, neuropathological, and genetic levels (Ferraiuolo *et al.*, 2011; Hardiman *et al.*, 2017; Brown and Al-Chalabi, 2017; van Es *et al.*, 2017). The onset is most common in late middle age (55-60 years) and men over the age of 50 are more susceptible than women. In some cases, juvenile-onset has also been reported. The average age of survival after the first sign of symptoms is reported at ~2-5 years, with respiratory failure as an end-stage occurrence (Ferraiuolo *et al.*, 2011; Shaw and Wood-Allum, 2010; Fávero *et al.*, 2017). Furthermore, a key defining feature of the disease is that the spread and progression of ALS are mainly contiguous. Studies on

patterns of symptom development have suggested and shown that symptoms of neurodegeneration usually develop in adjacent anatomical regions during the progression of the disease (Brooks, 1991; Walhout *et al.*, 2017; Zhenfei *et al.*, 2019).

1.1.1 Historical aspects of ALS

Jean-Martin Charcot examined the post-mortem tissues of individuals who had experienced muscular spasms throughout their lives and originally described ALS in 1869 (Gois *et al.*, 2020). In France, ALS is also known as la maladie de Charcot. Charcot observed a decrease in MNs, limb muscular atrophy and scarring, and tissue hardening in the corticospinal tracts (Meininger, 2011). Over the following years, several studies characterized ALS (Rowland and Shneider, 2001; Wijesekera and Leigh, 2009; Leblond *et al.*, 2014). However, to date, ALS remains challenging to diagnose at an early stage.

1.1.2 Epidemiology

ALS is a rare adult-onset disorder and due to diagnostic delay, short survival time in ALS and concentrated epidemiological studies in Europe and North America, it is challenging to estimate the global incidence and prevalence of ALS (Chio *et al.*, 2013; Linden-Junior *et al.*, 2013). Studies in Europe have shown that ALS has an estimated yearly incidence of approximately 1.75-3 per 100,000 and whilst significant geographical differences exist, ALS has an average prevalence of 10-12 per 100,000 (Logroscino *et al.*, 2010; Collaborators, 2018). Individuals in the 45-75 years age group have the highest risk of developing ALS and the yearly incidence of this age group is approximately 4-8 per 100,000. The mean age at the onset of symptoms is variable and dependent on the type of ALS. In sporadic ALS (sALS), the mean age at onset of symptoms is 58-63 years, and 40-60 years in familial ALS (fALS) (Logroscino *et al.*, 2010). For men and women, respectively, the estimated cumulative lifetime risk for having ALS is 1:350 and 1:400 (Johnston *et al.*, 2006; Ryan *et al.*, 2019). Similar to that in Europe, the yearly incidence of ALS in Asia ranges from 0.3-3.3 per 100,000 (Doi *et al.*, 2014). In Europe, North America and Asia, the yearly prevalence of ALS can range from 1.0-11.3 per 100,000 (Chio *et al.*, 2013; Doi *et al.*, 2014; Logroscino *et al.*, 2010; Mehta *et al.*, 2014).

1.1.3 Types of ALS

ALS can be of two types, where both are clinically and pathologically similar: sALS and fALS. These are distinguished according to the aetiology of the disease. sALS is more common in ALS patients, comprising a majority (90-95%) of ALS cases (Mulder *et al.*, 1986).

1.1.3.1 Sporadic ALS

90-95% of ALS cases are sporadic, where the patients display no family history of ALS (Mejzini *et al.*, 2019). The cause of sALS is not as well understood yet, however, studies have suggested that it has a genetic component. Twin studies conducted to estimate the heritability of sALS showed that the genetic contribution to sALS is significant and estimated at 61% (Graham, Macdonald and Hawkes, 1997; Al-Chalabi *et al.*, 2010). While identifying the genetic cause of fALS is higher, studies have been able to identify genetic causes for sALS. However, this can be masked by reduced penetrance, lack of family history/information, premature death of family members, recessive inheritance and small family size. Nonetheless, it has been identified that 11% of sALS patients have genetic mutations in genes that are associated with fALS (Renton, Chiò and Traynor, 2014). Recently, a study identified 23 novel mutations in 22 protein-coding genes associated with sALS cases (Logan *et al.*, 2022). They showed that most of the sALS patients have at least one mutation in the 22 identified genes with a 25-99% probability of developing ALS. The variability in the probability of disease development is likely to be because of reduced and incomplete penetrance and variable expressivity. The lower penetrance may be due to an individual's genetic context, i.e. there may be some protective elements in their genome. Additionally, an investigative study that assessed the disease risk to relatives of sALS patients showed that first-degree relatives were at a definite increased risk (Hanby *et al.*, 2011). Furthermore, in sALS, the causal genes identified are believed to interact with potential environmental risk factors associated with sALS (Oskarsson, Horton and Mitsumoto, 2015; Al-Chalabi and Hardiman, 2013), which could increase the probability of developing ALS. Such factors include military service (Beard and Kamel, 2015), exposure to industrial chemicals and heavy metals (Graham, Macdonald and Hawkes, 1997; Wang *et al.*, 2014; Callaghan *et al.*, 2011), and cigarette smoking (Wang *et al.*, 2011; de Jong *et al.*, 2012; Calvo *et al.*, 2016). To

date, smoking is the only environmental factor that has been accepted to be linked with ALS, though a recent study challenged this conclusion (Opie-Martin *et al.*, 2020).

1.1.3.2 Familial ALS

fALS accounts for 5-10% of ALS cases, where a genetically related family member of the patient is also affected by ALS (Mejzini *et al.*, 2019). This type of ALS is usually inherited as an autosomal dominant trait (Ferraiuolo *et al.*, 2011). Around 70% of fALS cases can currently be explained by identified genetic factors (Renton, Chiò and Traynor, 2014; Chia, Chiò and Traynor, 2018). The most common genetic mutation in fALS is in the chromosome 9 open reading frame 72 (*C9ORF72*) gene. This defect appears in ~40% of familial cases. The superoxide dismutase 1 (*SOD1*) gene was the first gene to be associated with fALS and mutations in SOD1 account for ~20% of fALS cases. Other genes involved in fALS are the fused in sarcoma (*FUS*) and TAR DNA-binding protein (*TARDBP*) genes, mutations in which explain 1-5% of familial cases. Together these genes explain more than 50% of fALS cases (van Es *et al.*, 2017).

1.1.4 Clinical features

Symptoms appear over time and ALS patients display various clinical phenotypes attributed to variable degeneration of UMNs and LMNs (Chiò *et al.*, 2011). Eventually, all ALS patients suffer from loss of voluntary control and mobility. The hallmark clinical presentation of ALS is increasing muscle weakness associated with muscle atrophy, fasciculations, muscle cramps and stiffness (Masrori and Van Damme, 2020). The region and age of the disease onset, family history, the type of motor neuron affected, the extent of additional motor involvement, and the duration of the disease can all affect the phenotypic variability of ALS (Swinnen and Robberecht, 2014).

The majority of ALS patients exhibit the spinal ALS phenotype, which includes UMN and LMN degeneration. Those with the spinal onset of ALS show symptoms of muscle weakness which can start distally or proximally in the lower or upper limbs. With disease progression, loss of LMNs leads to weakness, fasciculation, hyporeflexia, muscular atrophy and subsequent progressive paralysis. Loss of UMNs results in hypertonia, hyperreflexia, and spasticity which can affect manual dexterity, gait and

balance (Wijesekera and Leigh, 2009; Gois *et al.*, 2020). Some patients also display bulbar onset ALS. The bulbar area is affected by the bulbar phenotype, which causes tongue fasciculation, dysarthria, and dysphagia in ALS patients (Swinnen and Robberecht, 2014; Wolf *et al.*, 2014; Yunusova *et al.*, 2019). Patients with bulbar onset can also show symptoms affecting the limbs parallel with bulbar symptoms. The spinal onset of ALS occurs in approximately two-thirds of the patients, whereas bulbar onset ALS accounts for one-third of ALS patients. The PLS and PMA disorders classified under MND share clinical characteristics with typical ALS, making the diagnosis of ALS challenging due to MND heterogeneity (Gois *et al.*, 2020; Visser *et al.*, 2007).

While most ALS patients can be classed as having a typical ALS phenotype with bulbar or spinal onset, ALS is being increasingly recognized as a clinically heterogeneous disorder with distinct motor and extra-motor symptoms. The motor symptoms of ALS are quite heterogeneous, and the motor manifestations may be associated with varying extents of frontotemporal cerebral involvement. This leads to various phenotypic presentations of ALS with various disease trajectories (Masrori and Van Damme, 2020).

1.1.5 Diagnosis of ALS

Due to the complexity of symptoms which may overlap with other neurological disorders and the heterogeneous nature of ALS, an early and definite diagnosis is difficult, leading to a delay in diagnosis from the onset of symptoms. Generally, ALS is diagnosed based on clinical observations and several tests to exclude other overlapping neurodegenerative disorders and causes of UMN and LMN dysfunction (Zarei *et al.*, 2015). Standardised diagnostic criteria date back to the original El Escorial revised and Awaji criteria. They exist to rule out the differential diagnosis of ALS. Both rate the extent of diagnostic certainty by assessment of clinical features from possible to probable to definite ALS. These measures are based on the number of affected regions as well as clinical or electrophysiological findings or both (Brooks *et al.*, 2000; Costa, Swash and de Carvalho, 2012). Diagnosis using these criteria is based on how many of the four brainstem and spinal cord regions – bulbar, cervical, thoracic and lumbosacral – display a history of progressive clinical features due to the involvement of LMNs and UMNs. The El Escorial criteria describe levels of diagnosis

that are dependent on clinical evaluation of the extent of UMN and LMN injury. Though this criterion is specific to ALS, it lacks sensitivity which may lead to delay and inaccuracy of diagnosis. The Awaji criteria have enhanced diagnostic sensitivity and use electrophysiological data for the diagnosis of ALS. **Table 1.1** summarises the levels of diagnostic certainty in the two criteria (Tao and Wu, 2017; Hardiman *et al.*, 2017).

Table 1.1: Levels of diagnostic certainty for El Escorial and Awaji criteria

Criteria	Levels of Diagnostic Certainty			
	Definite	**Probable**	**Possible**	**Suspected****
El Escorial	UMN and LMN involvement in three regions	UMN and LMN involvement in at least two regions	UMN and LMN involvement in only one region Or Only UMN involvement in two or more regions	LMN involvement only
Awaji	Electrophysiological evidence of UMN and LMN involvement in the bulbar region and at least two spinal cord regions Or UMN and LMN involvement in three regions	Electrophysiological evidence of UMN and LMN involvement in at least two regions	Electrophysiological evidence of UMN and LMN involvement alone in two or more regions	–
** This has been eliminated from the El Escorial revised criteria.				

Recently, an international consensus group proposed new diagnostic criteria for ALS which recognises the heterogeneity in clinical presentation to enhance the diagnosis in the early stages of ALS when clinical manifestations are minimal (Shefner *et al.*, 2020). The Gold coast criteria define the disease by three measures – advancing progressive motor impairment which is recorded by history or repeated clinical assessment, UMN and LMN impairment in at least one region of the body or LMN impairment in at least two regions; and investigation which excludes other diseases. The proposal of the new criteria allows for an early and definitive diagnosis. It was

reported that the sensitivity of diagnosis of the Gold Coast criteria was maintained regardless of the duration of ALS and region of onset and was consistent with the sensitivity of the revised El Escorial and Awaji criteria (Hannaford *et al.*, 2021). Though the diagnostic sensitivity was similar, the Gold Coast criteria were more sensitive and specific for PMA and for distinguishing PLS as a form of MND (Hannaford *et al.*, 2021; Pugdahl *et al.*, 2021). In the El Escorial revised and Awaji criteria, PLS meets the category of possible ALS. Overall, the new Gold Coast criteria will aid in diagnosis and help patients and their families feel less uncertain, as well as improve the recruitment of patients into clinical trials.

1.1.6 Current therapeutic strategies and management

Currently, therapeutic strategies for ALS are limited and there is no known cure for ALS. However, there are various disease management options available focusing on treating symptoms and giving supportive care. These can slow disease development and improve quality of life. Previously, only two pharmacological interventions, Riluzole and Edaravone, were approved and available in some countries for prolonging survival and slowing down the disease progression, respectively (Andrews *et al.*, 2020; Witzel *et al.*, 2022). Riluzole, approved in 1995, is a widely available drug to treat ALS and the only drug that prolongs survival by approximately 3 months (van Es *et al.*, 2017). The mechanism of action of Riluzole is not completely understood, however, it is believed that the drug is an anti-glutamate treatment that reduces motor neuron injury. It decreases the levels of the excitatory neurotransmitter glutamate by reducing its release from the pre-synaptic terminal and therefore inhibiting glutamate-mediated excitotoxicity (Corcia *et al.*, 2021). However, its direct interaction with glutamate receptors has been debated. In addition to the anti-glutamatergic regulation of excitotoxic pathways, effects on mitochondrial function, alteration to fat metabolism, effects on persistent sodium channel function and potentiation of calcium-dependent potassium currents have been highlighted (Cheah *et al.*, 2010; Jaiswal, 2016; Thakor *et al.*, 2021; Chowdhury *et al.*, 2008; Thompson *et al.*, 2018; Deflorio *et al.*, 2014). It is debated whether the slowing of the disease course happens at all stages of ALS or only at advanced disease stages. A retrospective analysis of data from a dose-ranging study investigated the effects of Riluzole across different stages of the disease (Fang *et al.*, 2018). The results suggested that the disease-modifying effects of Riluzole occur at later stages of ALS. This highlights that

Riluzole might affect different pathways depending on the stage of ALS. In the study, the earliest stage of ALS was not analysed. Several open-label non-randomised trials have indicated that Riluzole is most effective at earlier stages of ALS and suggested early disease modulation (Traynor *et al.*, 2003; Zoing *et al.*, 2006; Riviere *et al.*, 1998; Geevasinga *et al.*, 2016). Given the lower likelihood that any treatment could confer considerable therapeutic effects within a depleted population of motor neurons in the advanced stage of ALS, earlier efficacy seems more probable than later effects.

Edaravone, approved in 2015 in Japan and in 2017 in the US, is an antioxidant and free radical scavenger and has been shown to significantly slow down the motor decline in ALS (Ito *et al.*, 2008; Group and Group, 2017). Previously, the drug was available only intravenously and recently the FDA has approved an oral formulation of Edaravone called Radicava ORS (Dorst and Genge, 2022; Aschenbrenner, 2022). The approval in Japan and the U.S. was largely based on a single trial conducted in Japan involving early-stage ALS patients (Cruz, 2018; Abe *et al.*, 2014; The Writing Group, 2017). While the trials of the antioxidant Edaravone show efficacy (The Writing Group, 2017; Edaravone (MCI-186) ALS 16 Study Group, 2017), there was a lack of generalisability to the wider population of ALS patients and there are concerns regarding the associated safety and benefits. Therefore, Edaravone has not yet been approved worldwide and was declined by the European Medicines Agency.

In addition to the drugs previously approved, the FDA has recently approved a new experimental drug for ALS called Relyvrio (Brown, 2022). It is a co-formulation of sodium phenylbutyrate and taurursodiol and has been found to reduce neuronal death in ALS models and slow down the functional decline (Paganoni *et al.*, 2020). Furthermore, a long-term survival analysis of the participants from the trial has shown that Relyvrio has a long-term survival benefit with early initiation of the drug in ALS, highlighting that the drug has functional as well as survival benefits in ALS (Paganoni *et al.*, 2021). Currently, the mechanism of action of Relyvrio is unknown and is speculated from the processes that are affected by sodium phenylbutyrate and taurursodiol. Relyvrio is therefore believed to have positive effects on the function of mitochondria and the endoplasmic reticulum. Furthermore, an investigational drug called Tofersen (also known as BIIB067) is under the peer review period by the FDA. Tofersen is an antisense oligonucleotide (ASO) that targets superoxide dismutase 1

(SOD1) ALS. Administered through a lumbar puncture, the ASO can facilitate the degradation of *SOD1* messenger RNA and thereby reduces the synthesis of the SOD1 protein. The recently published Phase 3 trial results showed positive effects on biomarkers by 6 months and evidence of clinical benefits by 12 months of treatment administration (Miller *et al.*, 2022).

Though pharmacological options are available, they either have modest effects where survival is prolonged by a few months, or further clinical trials are still required post-approval of the drug. Therefore, supportive care remains the foundation of ALS management. Care is most effectively provided by multi-disciplinary teams of healthcare professionals (van Es *et al.*, 2017). Palliative care from a multidisciplinary team is likely to have a measurable effect, especially with those with a bulbar onset of ALS because the patients may get access to nutritionists and advisors for quality care of living which can lead to small positive effects. Intervening with non-invasive ventilation when respiratory muscle weakness causes neuromuscular respiratory failure has been shown to have significantly positive effects on both life expectancy and quality of life (Bourke *et al.*, 2006).

1.1.7   Pathogenesis

The pathogenesis of neuronal degeneration in ALS is highly varied and multifactorial (**Figure 1.1**). The exact mechanisms and molecular pathways causing motor neuron death are yet to be fully understood, however, it is likely that it may be due to a complex interplay between various identified pathological factors that contribute to the pathogenesis of the disease, combined with genetic factors, environmental factors and age-related dysfunction (Mejzini *et al.*, 2019; Masrori and Van Damme, 2020; Yang *et al.*, 2021). The identified factors are heavily influenced by mutations in various genes identified over the years, which makes ALS a genetically heterogeneous disorder. In addition to genetic mutations, the most common factors implicated in ALS are oxidative stress, mitochondrial dysfunction, excitotoxicity, altered RNA metabolism and protein aggregation.
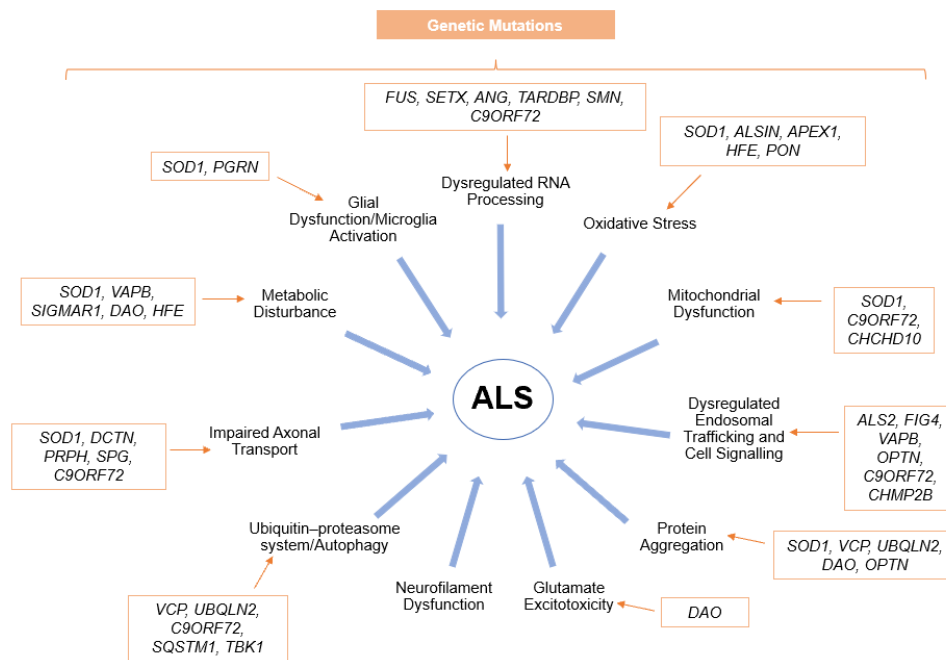
Figure 1.1: Identified factors that contribute to the pathogenesis of ALS
There are various cellular pathways and mechanisms implicated for neurodegeneration in ALS. Genetic mutations are a known cause of ALS. Mutations in genes that affect the identified pathways are highlighted. Alsin (*ALS2*), Angiogenin (*ANG*), Apurinic/Apyrimidinic Endodeoxyribonuclease 1(APEX1), D-amino acid oxidase (*DAO*), Dynactin (*DCTN*), Polyphosphoinositide phosphatase (*FIG4*), Fused in Sarcoma (*FUS*), Human haemochromatosis (*HFE*), Optineurin (OPTN), Granulin Precursor (*PGRN*), Paraoxonase 1 (*PON*), Peripherin (*PRPH*), Senataxin (*SETX*), Sigma Non-Opioid Intracellular Receptor 1 (*SIGMAR1*), Survival Motor Neuron (*SMN*), Superoxide Dismutase 1 (*SOD1*), Spastic Paraplegia (*SPG*), TAR DNA Binding Protein (*TARDBP*), Ubiquilin 2 (*UBQLN2*), VAMP Associated Protein B (*VAPB*), Valosin Containing Protein (*VCP*), Charged Multivesicular Body Protein 2B (*CHMP2B*), Chromosome 9 Open Reading Frame 72 (*C9ORF72*).

1.1.7.1 Genetic Mutations

Research conducted over many years studying family pedigrees and inheritance has allowed the identification of multiple genetic mutations that predispose individuals to ALS. This helps to provide an understanding of the underlying pathogenesis of ALS. To date, more than 20 genes have been found in association with ALS (**Table 1.2**). There is high overall heritability in ALS and in sALS patients the heritability is estimated at 30%-60% (Ryan *et al.*, 2019; Al-Chalabi *et al.*, 2010). Inheritance of mutations in these genes is mostly autosomal dominant, nevertheless, autosomal recessive and X-linked inheritance have also been described (Corcia *et al.*, 2017). The effects of

associated genetic mutations are not well understood and there are speculations that their involvement may be due to a toxic gain-of-function, loss-of-function, or through influence on essential cellular pathways. Mutations in *SOD1, TDP43, FUS* and *C9ORF72* are the most frequent genetic causes of ALS and have been extensively characterized.

Table 1.2: Genes that have been found to be associated with ALS

GENE	LOCUS	ONSET/INHERITANCE	PATHOGENIC PATHWAY *	PROTEIN AGGREGATE FORMATION **	REF.
ALS2	2q33	Juvenile/AR	Endosomal trafficking, cell signalling, oxidative stress	–	(Yang *et al.*, 2001)
ANG	14q11.2	Adult/AD	Aberrant RNA processing	TDP-43 aggregates	(Greenway *et al.*, 2006)
ATXN2	12q24.12	Adult/AD	RNA metabolism and processing	TDP-43 positive	(Elden *et al.*, 2010)
C9ORF72	9q21-q22	Adult/AD	RNA processing (mRNA splicing regulation, transcription), endosomal trafficking, autophagy	TDP-43 aggregates, p62-positive neuronal cytoplasmic inclusions, characteristic TDP-43 negative aggregates in neurons of cerebellum and hippocampus (these are SQSTM1-positive and contain UBQLN2)	(DeJesus-Hernandez *et al.*, 2011; Brettschneider *et al.*, 2012; Renton *et al.*, 2011)

CCNF	16p13.3	–/AD	Abnormal ubiquitination, protein degradation and aggregation	Ubiquitinated protein inclusions, TDP-43 positive, SCF substrate positive	(Williams *et al.*, 2016)
CHCHD10	22q11.23	Adult/AD	Mitochondrial dysfunction	–	(Bannwarth *et al.*, 2014)
CHMP2B	3p11.2	Adult/AD	Endosomal trafficking	p62 positive inclusions	(Parkinson *et al.*, 2006)
DAO	12q24	Adult/AD	Metabolic disturbance, apoptosis, protein aggregation, glutamate excitotoxicity	–	(Mitchell *et al.*, 2010)
ERBB4	2q34	Adult/AD	Disruption of the NRG1-ErbB4 pathway	–	(Takahashi *et al.*, 2013)
FIG4	6q21	Adult/AD	Apoptosis, endosomal trafficking, cell signalling	Not known	(Chow *et al.*, 2009)
FUS/TLS	16p11.2	Adult/AD	Aberrant RNA processing and metabolism	FUS aggregates, no TDP-43 aggregates	(Kwiatkowski *et al.*, 2009)
GLE1	9q34	–/AR	RNA processing and metabolism, haploinsufficiency	–	(Kaneb *et al.*, 2015)
HNRNPA1	12q13	Adult/AD	RNA metabolism – mRNA processing, splicing and transport	Cytoplasmic inclusions of hnRNPA1	(Kim *et al.*, 2013b)

MATR3	5q31	AD	RNA processing	MATR3-positive inclusions (C9ORF72 background)	(Johnson *et al.*, 2014a)
OPTN	10p13	Adult/AD, AR	Protein aggregation, endosomal trafficking, cell signalling	TDP-43 aggregation with Glu478Gly genetic mutation background, OPTN inclusions in the spinal cord	(Maruyama *et al.*, 2010)
PFN1	17p13.3	Adult/AD	Inhibits axonal growth, cytoskeletal pathway disruption	Ubiquitinated protein aggregates, some TDP43 positive	(Wu *et al.*, 2012)
SETX	9q34	Juvenile/AD	Dysfunction of helicase activity, RNA processing	–	(Chen *et al.*, 2004)
SIGMAR1	9p13.3	Adult/AD Juvenile/AR	Metabolic disturbance	TDP-43 cytoplasmic inclusions, FUS-positive inclusions (sigma-1 negative)	(Al-Saif, Al-Mohanna and Bohlega, 2011; Luty *et al.*, 2010)
SOD1	21q22.1	Adult/AD (AR)	Oxidative Stress, protein aggregation, mitochondrial and axonal dysfunction	Cytoplasmic aggregates of SOD1, absence of TDP-43 aggregates	(Rosen *et al.*, 1993; Morgan and Orrell, 2016)
SPG11	15q15-q21	Juvenile/AR	Unknown	–	(Orlacchio *et al.*, 2010)

SQSTM1/ p62	5q35.3	Adult/AD	Protein degradation, autophagy	Ubiquitin/p62 positive aggregates	(Fecto *et al.,* 2011)
TARDBP	1p36.2	Adult/AD	Aberrant RNA processing, dysregulated nuclear RNA metabolism	TDP-43 aggregates	(Sreedharan *et al.,* 2008)
TBK1	12q4.2	Adult/AD	TBK1 haploinsufficiency, autophagy and mitophagy	TDP-43 positive, p62-positive, ubiquitin-positive	(Cirulli *et al.,* 2015; Freischmidt *et al.,* 2015; Oakes, Davies and Collins, 2017)
TUBA4A	2q35	Adult/AD	Cytoskeletal defects	–	(Smith *et al.,* 2014)
UBQLN2	Xp11	Adult/X-linked	Ubiquitin/protein aggregation and degradation	Ubiquilin-2, TDP-43 positive and FUS-positive	(Deng *et al.,* 2011b)
VAPB	20q13.3	Adult/AD	Metabolic disturbance, apoptosis, endosomal trafficking, cell signalling	Possible TDP-43 aggregates	(Nishimura *et al.,* 2004)
VCP	9p13-p12	Adult/AD	Ubiquitin/protein aggregation and degradation, apoptosis, autophagy	–	(Johnson *et al.,* 2010)
***(Chen *et al.*, 2013), **(Turner *et al.*, 2013).**					

AD = Autosomal Dominant, AR = Autosomal Recessive

1.1.7.1.1 SOD1

In 1993, Rosen *et al.* identified the first causal gene, *SOD1*, that was associated with ALS. Mutations have been reported in all of the five exons of the gene, which account for ~20% of fALS cases, ~1%-2% of sALS cases and only approximately 2% of all ALS cases (Rosen *et al.*, 1993; Pramatarova *et al.*, 1995; Andersen *et al.*, 2003). The locus of the gene is on chromosome 21q22.1 and encodes for a Copper and Zinc (Cu/Zn) binding superoxide dismutase enzyme. This ubiquitously expressed 153 amino acid metalloenzyme functions as an extremely stable homodimer to scavenge the toxic intracellular superoxide radicals (charged oxygen molecules). The major function of the SOD1 enzyme is to attach to Cu/Zn molecules and catalyse the reaction where the superoxide radicals, produced as a toxic by-product of mitochondrial respiration, are converted into oxygen and hydrogen peroxide. Over the years, more than 185 different ALS-associated mutations in *SOD1* have been identified throughout the gene and most of them are missense mutations (Yamashita and Ando, 2015). Disease characteristics, progression, duration and severity differ depending on the mutations. Quick progression and shorter survival time are observed in those with the A4V (the most common mutation in the US population), H43R, L84V, G85R, N86S, and G93A mutations, whereas carriers of the G93C, D90A, or H46R mutations have relatively longer survival (Yamashita and Ando, 2015). Mutations in the *SOD1* gene are dominantly inherited, however, both dominant and recessive inheritance patterns have been reported for the D90A mutation (Andersen *et al.*, 1995; Andersen *et al.*, 1996; Pansarasa *et al.*, 2018; Robberecht *et al.*, 1996). Homozygous patients with the D90A variant show a slow disease progression, whereas D90A heterozygous patients show a faster progression (Mejzini *et al.*, 2019). Interestingly, the D90A-*SOD1* variant is more frequent in Scandinavia (2.5%) than anywhere else in the world (0.05%), yet the prevalence of ALS is not higher than average in Scandinavia. An early study using a Scandinavian cohort showed that D90A recessive families shared a common founder, while many were identified for dominant families and therefore identified separate founders (Al-Chalabi *et al.*, 1998). This suggests that new cases of the D90A mutation have shown a dominant inheritance pattern many times, however, recessive inheritance was seen only once. This shows that the D90A mutation, like all other SOD1 mutations, should show a dominant inheritance and that the recessive families

must have a protective factor linked to the variant in their gene pool which leads to reduced penetrance of the mutant allele and slower disease progression. This explains the high frequency of the allele and low prevalence of the disease in Scandinavia.

Despite being the most studied gene in ALS, the pathways through which *SOD1* mutations lead to neuronal death remain incompletely understood. However, it has been indicated that toxicity is due to a toxic gain-of-function rather than a loss of protein and therefore a loss-of-function (Cleveland *et al.*, 1995; Reaume *et al.*, 1996). Furthermore, recent gene therapy trials with the investigational drug Tofersen (Miller *et al.*, 2020; Miller *et al.*, 2022) support the idea that a toxic gain of function is the mode of action for SOD1 in ALS. Mutations can cause changes in the conformation and function of SOD1 that confer toxicity through several mechanisms. Identified pathological processes include oxidative stress via increased ROS, excitotoxicity, endoplasmic reticulum stress, mitochondrial dysfunction and protein aggregation due to prion-like properties (Hayashi, Homma and Ichijo, 2016). Misfolded protein is marked for degradation by ubiquitylation, however, misfolded mutant SOD1 escapes this and eventually forms aggregates. Mutant SOD1 may cause impairment of the proteasomal pathway and autophagy. However, whether aggregates lead to toxicity is uncertain (Shaw, 2005; Gros-Louis, Gaspar and Rouleau, 2006; Robberecht and Philips, 2013). The impairment of the protein degradation pathways may be due to biological aging and with the burden of misfolded proteins over the years, the degradation pathways are no longer able to cope. This is consistent with the late age of onset of ALS.

1.1.7.1.2 TDP43

In addition to ALS, frontotemporal dementia (FTD) is another neurodegenerative disorder that affects the frontal and temporal lobes. In FTD, progressive changes in personality, behaviour and/or speech are observed (Van Langenhove, van der Zee and Van Broeckhoven, 2012). Roughly 15% of patients with FTD display symptoms of ALS, on the other hand, around 50% of ALS patients show symptoms indicting impairment of frontal and temporal lobe function (Ng, Rademakers and Miller, 2015). It has been shown that FTD and ALS share a disease spectrum that has a common molecular basis and shows the involvement of *TDP-43* and *FUS* (Ito and Suzuki, 2011). *TDP-43* was first discovered to be involved in ALS in 2006 (Arai

et al., 2006; Neumann *et al.*, 2006). Mutations in *TDP-43* account for 4-5% of fALS cases and less than 1% of sALS cases (Sreedharan *et al.*, 2008). TDP-43 is a DNA/RNA binding protein, made up of 414 amino acids and encoded by the TAR DNA binding protein (*TARDBP*) gene. TDP-43 contains a nuclear localisation signal (NLS) and normally localises in the nucleus, however, it also contains a nuclear export signal (NES) and therefore it is constantly transported between the nucleus and cytoplasm (Ayala *et al.*, 2008). However, it has been suggested that the NES may be non-functional (Ederle *et al.*, 2018; Pinarbasi *et al.*, 2018). In the presence of ALS-associated mutations, TDP-43 is reported to show a disrupted nucleocytoplasmic distribution (Barmada *et al.*, 2010; Guerrero *et al.*, 2019). This suggests that TDP-43 dysfunction due to the mutation can lead to mislocalisation.

Barmada et al. (2010) showed that toxicity in neurons due to mutant TDP-43 expression was associated with the mislocalisation of TDP-43 from the nucleus to the cytoplasm. Furthermore, the presence of inclusion bodies was not required for toxicity and neuronal death, which may highlight that the mislocalisation of TDP-43 leads to cellular toxicity and death. However, it is difficult to determine if the cellular phenotypes are due to the mislocalisation or the genetic mutation itself. TDP43 functions as a gene expression regulator and is involved in many RNA processing processes including mRNA splicing, regulation of mRNA stability and non-coding RNAs, mRNA transport and translation (Buratti and Baralle, 2010; Ratti and Buratti, 2016). Around 48 mutations in the *TARDBP* gene have been identified to be associated with ALS (Lattante, Rouleau and Kabashi, 2013; Mejzini *et al.*, 2019) and most of them are missense mutations in the glycine-rich domain at the C-terminal. The C-terminal interacts with other ribonucleoproteins and plays a role in regulating mRNA splicing (Buratti *et al.*, 2005; Neumann *et al.*, 2006). Mutations in the C-terminal of TDP-43 can therefore hinder RNA processing.

Due to the correlation between the cytoplasmic increase of TDP-43 and the loss of nuclear TDP-43, it has been hypothesised that pathogenesis includes either a loss of normal TDP-43 functions in the nucleus, a toxic gain of function, or both. To assess whether toxicity was due to loss of function, many animal models were generated. It was shown that TDP43 was essential for embryonic development as homozygous TDP43 null mice were not viable (Sephton *et al.*, 2010; Kraemer *et al.*, 2010). *TARDBP*

deletion heterozygous mice showed deficits in motor performance but no motor neuron degeneration and reduction in TDP43 expression (Kraemer *et al.*, 2010). The gain of function hypothesis is supported by the overexpression model of TDP43. It was reported that TDP43 overexpression causes phenotypes of neurodegeneration (Kabashi *et al.*, 2010; Wils *et al.*, 2010). The loss and overexpression of TDP43 causing ALS phenotypes show the importance of fine-regulation of TDP43. An abundance of TDP43 in the cytoplasm can cause the formation of inclusions and result in cellular dysfunction. Whereas reduced TDP43 in the nucleus can cause extensive mRNA metabolism dysregulation as loss of TDP-43 has been shown to result in differential splicing and expression of targets (Highley *et al.*, 2014; Klim *et al.*, 2019). TDP-43 has been shown to function as a repressor of cryptic exon inclusion during RNA splicing (Ling *et al.*, 2015; Klim *et al.*, 2019). Recent studies in ALS have shown that loss of TDP-43 and its nuclear mislocalisation is associated with cryptic exon inclusion in genes such as stathmin-2 (*STMN2*) and unc-13 homolog A (*UNC13A*), leading to their reduced protein expression critical for cellular function (Klim *et al.*, 2019; Ma *et al.*, 2022).

1.1.7.1.3 FUS

In 2009, mutations in another RNA-binding protein gene called *FUS* (fused in sarcoma) were identified (Kwiatkowski *et al.*, 2009; Vance *et al.*, 2009). Genetic studies showed that mutations in the *FUS* gene account for 4% of fALS and 1% of sALS cases (Lagier-Tourenne, Polymenidou and Cleveland, 2010) and are linked to early-onset and juvenile cases (Gromicho *et al.*, 2017). FUS is ubiquitously expressed and consists of 526 amino acids. It is normally mainly localised in the nucleus. Like TDP43, FUS is involved in the regulation of gene expression including splicing, RNA transport, transcription and translation (Ratti and Buratti, 2016). FUS is also involved in DNA repair mechanisms (Wang *et al.*, 2013). To date, more than 50 mutations in *FUS* have been reported in ALS and the majority are missense mutations, however, insertions, deletions, splicing and nonsense mutations have been shown to occur in some cases (Lattante, Rouleau and Kabashi, 2013). Several mutations are found in the NLS and therefore cause the translocation of the FUS protein into the cytoplasm (Vance *et al.*, 2013). Mutations in FUS increase the propensity of FUS to aggregate and contribute to the pathogenesis of ALS (Nomura *et al.*, 2014).

Whether the disease is caused due to loss of function or gain of function is debated in *FUS*-ALS. A study showed that FUS knockout mice were able to survive into adulthood, however, the mice did not show phenotypes of ALS (Kino *et al.*, 2015). This highlighted that loss of function is not enough to cause the disease. On the contrary, another study using drosophila showed that FUS knockdown caused neuronal degeneration and locomotive deficits (Sasayama *et al.*, 2012). While the studies show contradicting results in support of the mode of action in *FUS*-ALS, it is possible that the differences in phenotypes observed were due to physiological differences of the models and therefore is a model effect. The gain of function in *FUS*-ALS is supported by a study that showed that overexpressing FUS in mice resulted in aggressive motor neurodegeneration and cytoplasmic accumulation of FUS (Mitchell *et al.*, 2013). However, the resultant phenotypes showing toxicity may be due to the overexpression of the protein rather than a toxic gain of function. It is debated whether toxicity is due to aggregation of FUS or an increase in cytoplasmic FUS following mislocalisation. It has been reported that FUS aggregates can cause toxic effects directly, where severe motor phenotypes were seen in mouse models (Robinson *et al.*, 2015; Shelkovnikova *et al.*, 2013). In contrast, there are studies using rodent models that show toxicity due to soluble cytoplasmic FUS, where neurodegeneration was observed in parallel with increased cytoplasmic FUS but without aggregates (Sharma *et al.*, 2016; Scekic-Zahirovic *et al.*, 2016). Aggregation of FUS may be a protective mechanism. Aggregation may be a compensatory way of protecting the cells from the toxic increase in cytoplasmic FUS. The importance of the propensity of FUS to aggregate in normal cellular function to build subnuclear paraspeckles which regulate gene expression offers support for this (Hennig *et al.*, 2015). However, it is also likely that the aggregation may overwhelm the protein degradation pathways, sequester functional proteins or disrupt cellular pathways which in turn may burden the cell and lead to toxicity. Impaired intracellular functions may also be due to the direct effects of *FUS* mutations that can lead to defects in splicing and DNA damage (Qiu *et al.*, 2014).

1.1.7.1.4 C9ORF72

First described in 2011, mutations in the *C9ORF72* gene were identified in both ALS and FTD. The mutation is a GGGGCC (G_4C_2) hexanucleotide repeat expansion in the intron 1 region (the non-coding region between exons 1 and 1b) of the gene (Renton *et al.*, 2011; DeJesus-Hernandez *et al.*, 2011). This is the most common

pathological genetic mutation in ALS and FTD (Majounie *et al.*, 2012). Aberrant expansion mutations in *C9ORF72* account for 40% of fALS and approximately 7% of sALS cases (Majounie *et al.*, 2012). In ALS, a repeat length of more than 30 units is usually considered pathogenic, whereas an expansion of hundreds to thousands of repeat units can be observed in patients with ALS (Van Mossevelde *et al.*, 2017). Healthy individuals usually have repeat lengths varying from 2 to 30 units on both alleles, with 2, 5 and 8 units being the most common (Van Mossevelde *et al.*, 2017; Fredi *et al.*, 2019). Some individuals can have an intermediate expansion ranging from 20 to 30 repeats in size and many studies suggest that 20 or 23 repeats can be used to differentiate between neutral and pathogenic expansion (Renton *et al.*, 2011; Van Mossevelde *et al.*, 2017; Byrne *et al.*, 2014). Furthermore, intermediate expansions between 24 and 30 repeats have been shown to be associated with ALS (Iacoangeli *et al.*, 2019). This further highlights that the threshold could be 20 or 23 repeats, however, the exact threshold for distinguishing between normal and pathogenic expansion has not yet been determined. The cytoplasmic C9ORF72 protein is abundantly expressed in neurons of the cerebral cortex, motor neurons and presynaptic terminals (DeJesus-Hernandez *et al.*, 2011; Renton *et al.*, 2011; Rizzu *et al.*, 2016). Its function is not well known; however, evidence suggests its involvement in RNA processing, regulating endosomal trafficking, autophagy in neurons and regulating DNA damage repair (Farg *et al.*, 2014; Nassif, Woehlbier and Manque, 2017; Webster *et al.*, 2016; He *et al.*, 2022).

The pathological mechanism by which the *C9ORF72* mutation causes ALS remains incompletely understood, however, three distinct mechanisms have been proposed (**Figure 1.2**). Firstly, the hexanucleotide repeat expansion could interfere with transcription processes leading to a reduction in the levels of the C9ORF72 protein and therefore a loss of function of the gene. This phenomenon is called haploinsufficiency (Waite *et al.*, 2014). Secondly, through bidirectional transcription in the sense and antisense directions, the repeat expansion may form secondary sense and antisense RNA structures and nuclear foci causing sequestration of functionally important RNA binding proteins (DeJesus-Hernandez *et al.*, 2011) including ALYREF, hnRNPs, nucleolin, SRSF1 and 2 (Babić Leko *et al.*, 2019). This leads to RNA toxicity and could result in the loss of function of the sequestered proteins. Thirdly, through non-ATG repeat-associated (RAN) translation, potentially aberrant dipeptide repeat

proteins (DPRs) are generated from the sense and antisense strands (Ash *et al.*, 2013; Gendron *et al.*, 2013). These mechanisms are all likely to contribute to toxicity (van Blitterswijk, DeJesus-Hernandez and Rademakers, 2012; Gendron and Petrucelli, 2018).
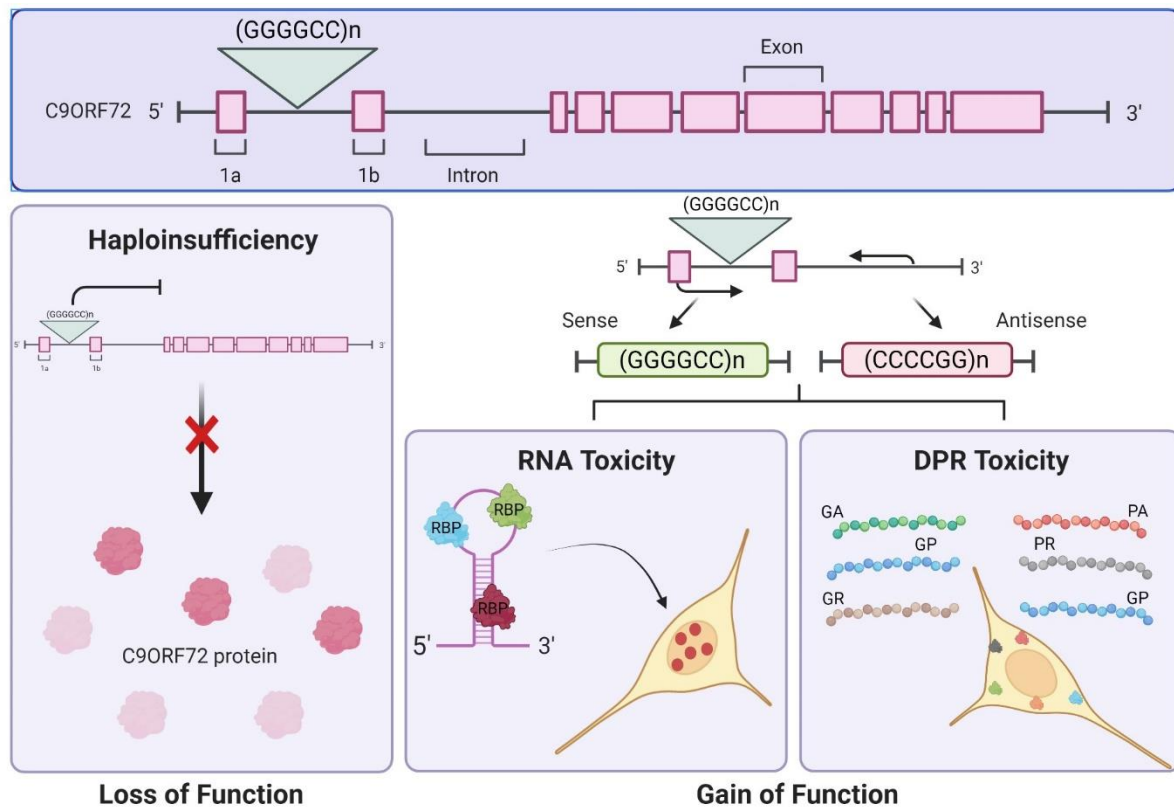
Figure 1.2: Proposed mechanisms of C9ORF72 mutation
The mechanisms by which C9ORF72 mutation causes ALS is incompletely understood, however, haploinsufficiency (loss-of-function); bidirectional transcription leads to the formation of RNA foci causing sequestration of RNA-binding proteins and the production of toxic dipeptide repeat proteins (DPRs) translated from the hexanucleotide expansion are all likely to contribute to toxicity. Image made using BioRender.

A hallmark feature of C9ORF72 pathology is the presence of aberrant cytoplasmic DPRs that are generated from the sense and antisense transcripts through RAN translation (Mori *et al.*, 2013c; Zu *et al.*, 2013; Ash *et al.*, 2013). In *C9ORF72*-ALS, five different DPR species are synthesised. Poly-Gly-Ala (poly-GA), and poly-Gly-Arg (poly-GR) are translated from the sense transcript from different open reading frames, and poly-Pro-Ala (poly-PA) and poly-Pro-Arg (poly-PR) are generated from the antisense transcript. Poly-Gly-Pro (poly-GP) is common to both and is generated from both the sense and antisense transcripts (Mori *et al.*, 2013a; Mori *et al.*, 2013c; Ash *et*

al., 2013). It has been shown that DPR toxicity is mostly due to arginine-rich DPRs, specifically poly-PR (Mizielinska *et al.*, 2014; Kwon *et al.*, 2014; Lee *et al.*, 2016). Arginine-containing DPRs can disrupt many cellular processes including transport between the nucleus and cytoplasm, RNA processing and translation dysregulation, nucleolar stress, dysregulation of the ubiquitin-proteasome system (UPS) and formation of stress granules (extensively reviewed in (Schmitz *et al.*, 2021; Babić Leko *et al.*, 2019). This can be explained by the unique positive charge and the high polarity that is due to the arginine residues which may enable more intracellular protein interactions. In addition to cytoplasmic toxicity, arginine-rich DPRs can also lead to nuclear toxicity. Due to their potential ability to resemble arginine-rich nuclear localization signal domains, poly-GR and poly-PR are also easily transported into the nucleus (Kwon *et al.*, 2014). The arginine-rich DPRs have been shown to cause nucleocytoplasmic transport defects and impairment of stress granule dynamics (Freibaum *et al.*, 2015; Zhang *et al.*, 2015; Ryan *et al.*, 2022; Zhang *et al.*, 2018; Lee *et al.*, 2016). These pathomechanisms are linked to TDP-43 pathology and emphasize the multifactorial nature of ALS. The non-arginine-rich DPR, poly-GA, has a high translation efficiency and predicted structural properties, which make poly-GA easily detectable in cytoplasmic inclusions. It has been shown that poly-GA can disrupt the UPS and lead to endoplasmic reticulum stress (Zhang *et al.*, 2014). It was reported that the poly-GA aggregates colocalise with a proteasomal subunit and therefore inhibit the proteasome (Khosravi *et al.*, 2020). The study also showed that poly-GA, via proteasome inhibition promoted TDP-43 pathology. Another study has shown that poly-GR aggregation can also induce TDP-43 pathology. This further highlights the relationship between the several genes and mechanisms associated with ALS.

1.1.7.2 Oxidative Stress

Reactive oxygen species (ROS) or free radicals, such as superoxide anions, hydrogen peroxide and hydroxyl radicals, are unstable molecules and natural by-products of oxygen metabolism. ROS has many important physiological actions including redox homeostasis and proper functioning of the cardiovascular system and immune system (Patel *et al.*, 2017). ROS also plays a key role in cell signalling, regulation of cell survival and death, cell proliferation and differentiation, gene expression and post-translational protein modification (Murphy *et al.*, 2011; Ray, Huang and Tsuji, 2012; Zhang *et al.*, 2022). Oxidative stress is caused when there is

an imbalance between the levels of ROS and the ability to detoxify the reactive species (Pizzino *et al.*, 2017). Accumulation of ROS and therefore oxidative stress results in irreversible damage to proteins, lipids, DNA and RNA. High levels of ROS and oxidative damage have been reported in cerebrospinal fluid (CSF), urine samples and serum of patients with ALS (Zarei *et al.*, 2015). The SOD1 enzyme plays a vital role to prevent damage via oxidative stress and dismutate the superoxide radicals leaked from the mitochondria during the generation of ATP. Mutations in SOD1 can interfere with protein function and lead to cytotoxicity (Bonafede and Mariotti, 2017). Studies of CSF in ALS patients and transgenic mice reported that oxidative stress can also be due to defective oxidative phosphorylation (Tohgi *et al.*, 1999; Bacman, Bradley and Moraes, 2006). This related the increase in levels of ROS with mitochondrial dysfunction, which is another pathogenetic mechanism in ALS. Furthermore, it has been reported that oxidative stress can lead to mislocalisation and an increased tendency of aggregation of TDP-43 and FUS (Vance *et al.*, 2013; Cohen *et al.*, 2015). This may lead to RNA dysregulation, protein aggregation as well as impaired proteostasis. This highlights how different proposed mechanisms in ALS may be inter-related.

1.1.7.3 Mitochondrial Dysfunction

In eukaryotic cells, mitochondria are vital organelles that function to produce ATP through oxidative phosphorylation for cellular energy requirements, cellular respiration and cytosolic calcium homeostasis. Mitochondrial damage resulting in a reduction of ATP and metabolic changes has been implicated in many neurodegenerative disorders, including ALS (Fiorito, Chiabrando and Tolosano, 2018; Sasaki, Horie and Iwata, 2007; Dupuis *et al.*, 2011). Changes in mitochondrial dynamics, morphology and localization have been reported to contribute to the pathogenesis of ALS. These have been observed in spinal motor neurons and skeletal muscles of sALS and fALS patients and in mouse models of ALS (Sasaki and Iwata, 2007; Magrané and Manfredi, 2009; Higgins, Jung and Xu, 2003). In ALS, defective mitochondrial axonal transport has been shown to cause the accumulation of aberrant mitochondria in the axons of motor neurons (Muyderman and Chen, 2014; Magrané and Manfredi, 2009; Mórotz *et al.*, 2012). Transport of mitochondria to areas with high requirements of ATP and calcium homeostasis, like synaptic terminals, is of key importance and therefore defects can lead to metabolic alterations. Mitochondria in

ALS patients have been shown to display impaired calcium homeostasis and increased production of ROS (Beal *et al.*, 1997; Beal, 2002). It has been suggested that mitochondrial dysfunction can be due to mutations or deletions of mitochondrial DNA (mtDNA) (Wiedemann *et al.*, 2002) which can be caused by ROS (and hence oxidative stress) which also damage mitochondrial proteins and lipids on the membrane (Karbowski and Neutzner, 2012). Furthermore, mitochondrial dysfunction can cause reduced calcium uptake from the cytoplasm and increase sensitivity to excitotoxicity (Bonafede and Mariotti, 2017). Additionally, mitochondrial dysfunction has been associated with biological aging and is considered one of the hallmarks of aging (López-Otín *et al.*, 2013). This age-related mitochondrial dysfunction can further lead to increased oxidative injury and is consistent with the late age of onset of ALS. This also further highlights the complex interplay of proposed mechanisms that lead to disease pathogenesis in ALS.

1.1.7.4 Excitotoxicity

One of the first proposed mechanisms for the pathogenesis of ALS was glutamate excitotoxicity (Bendotti and Carrì, 2004). Excitotoxicity is a proposed mechanism for the pathogenesis of ALS caused by the overstimulation of glutamate receptors by the glutamate neurotransmitter that leads to neuronal injury or degeneration and eventual death of the involved neuron (Mejzini *et al.*, 2019). Glutamate is an excitatory neurotransmitter that is produced in the presynaptic terminal and released into the synaptic cleft where it diffuses to activate specific postsynaptic receptors, including α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA) and N-methyl-D-aspartate (NMDA) receptors, and generates action potentials (Mejzini *et al.*, 2019). After triggering action potentials, glutamate is cleared from the synaptic cleft by glial and neuronal cell transporter proteins, the excitatory amino acid transporters (EAATs) (Sundaram, Gowtham and Nayak, 2012). To avoid excitotoxicity, glutamate concentrations in the synaptic cleft are highly regulated (Shaw and Eggett, 2000). Prolonged activation of the receptors by glutamate causes depolarization of neurons, the excessive firing of action potentials and an influx of calcium. Sustained elevation of intracellular calcium causes mitochondrial damage leading to increased ROS formation, which has been implicated in neuronal degeneration (Vucic, Rothstein and Kiernan, 2014). Therefore, rapid removal and regulation of glutamate are necessary to prevent neuronal toxicity. In ALS patients and

mutant *SOD1* transgenic mouse models, reduced levels of the astroglial EAAT2 (isoform 2) were reported in the motor cortex and spinal cord predicted to result in reduced glutamate clearance and increased concentrations of synaptic glutamate and therefore overstimulation of receptor and excitotoxic neuronal degeneration (Lin *et al.*, 1998; Trotti *et al.*, 1999; Zarei *et al.*, 2015; Boston-Howes *et al.*, 2006). Furthermore, it was shown that loss of EAAT2 is a contributor to motor neuron degeneration in ALS, but not however, a primary cause (Guo *et al.*, 2003). Other factors contributing to excitotoxicity have been extensively reviewed (King *et al.*, 2016) and include intrinsic excitability of motor neurons, alteration in regulation by interneuron populations and intracellular glutamate release from damaged neurons and glial cells including astrocytes and microglia. In addition to glutamate-mediated excitotoxicity, abnormal regulation by interneurons may also contribute to excitotoxicity and therefore neurodegeneration in ALS (Turner and Kiernan, 2012). Interneurons are a type of neuron that regulates neuronal signalling in the brain and spinal cord, majority of which are inhibitory in the cortex and use γ-Aminobutyric acid (GABA) or glycine neurotransmitters to inhibit the activity of others (Armada-Moreira *et al.*, 2020). In healthy individuals, a subthreshold stimulus of the motor cortex leads to the activation of inhibitory GABAergic interneurons, which reduces subsequent excitability in a process called short intracortical inhibition and helps to regulate neuronal activity and prevent excitotoxicity (Wagle-Shukla *et al.*, 2009). However, in ALS patients, this process is impaired and has been shown to be an adverse prognostic factor in ALS (Shibuya *et al.*, 2016). Several studies show a loss of interneurons in addition to or preceding motor neuron loss and highlight the importance of understanding and maintaining the excitation-inhibition balance (glutamate vs GABA) (Stephens *et al.*, 2006; Thielsen *et al.*, 2013; Minciacchi *et al.*, 2009; Martin *et al.*, 2007; McGown *et al.*, 2013; Foerster *et al.*, 2013). Nonetheless, the fact that Riluzole, a glutamate neurotransmission inhibitor, is the only partially-effective treatment for ALS supports that glutamate excitotoxicity plays a role in the aetiology of ALS (Bensimon, Lacomblez and Meininger, 1994; Saitoh and Takahashi, 2020). However, the impact of the treatment is modest and suggests that other mechanisms are also involved in the pathophysiology of the disease.

1.1.7.5 Altered RNA Metabolism

Identification of mutations in RNA processing genes, *TDP-43* and *FUS,* introduced RNA dysregulation as a pathological mechanism in ALS (Buratti *et al.*, 2010; Masrori and Van Damme, 2020). Since then, many RNA-binding proteins have been implicated in the pathogenesis of ALS, including angiogenin (*ANG*), senataxin (SETX), heterogeneous nuclear ribonucleoproteins A1 (*hnRNPA1*) and A2B1 (*hnRNPA2B1*), matrin-3 (*MATR3*) and ataxin-2 (*ATXN2*) (Boeynaems *et al.*, 2016; Kim *et al.*, 2013a; Johnson *et al.*, 2014b). This offers further support that altered RNA metabolism is involved in ALS and has led to an increased interest in RNA metabolism in neurodegenerative disorders. RNA binding proteins play various roles in RNA metabolism, including mRNA processing involving splicing, transcription, translation, non-coding RNA metabolism, micro RNA biogenesis and storage in stress granules (Mejzini *et al.*, 2019). Normally, these proteins localise mainly in the nucleus, and depletion in nuclear levels of these proteins due to mutations or mislocalisation to the cytoplasm due to aggregate formation can cause toxicity and induce overall transcriptome abnormalities (Masrori and Van Damme, 2020).

1.1.7.6 Protein Aggregation and impaired proteostasis

Insoluble protein aggregates or intracellular cytoplasmic inclusions are a pathological hallmark feature of neurodegenerative disorders, including Parkinson's, Alzheimer's, Huntington's disease and ALS. These inclusions characteristically comprise misfolded proteins. It has been proposed that aggregates form when the degradation capacity of a cell is exceeded by dysfunctional protein levels or due to defects in the protein degradation pathways which include the ubiquitin-proteasome system (UPS) and autophagosome-lysosome pathway/autophagy (Johnston, Ward and Kopito, 1998; Takalo *et al.*, 2013). Inclusions may also form due to the prion-like properties of proteins which make them aggregation-prone. However, the exact mechanisms that lead to aggregation are still being studied.

Different forms of inclusions have been reported in ALS, which include Bunina bodies, hyaline conglomerate inclusions (HCIs) and ubiquitinated inclusions (UBIs) (Wijesekera and Leigh, 2009; Wood, Beaujeux and Shaw, 2003). UBIs are predominantly characteristic of ALS. These are either skein-like (filamentous), Lewy body-like (compact and spherical) or a mixture of the two. Proteins that accumulate in

these inclusions are misfolded and ubiquitinated. In most neurodegenerative disorders, inclusions are tau-positive, however, ubiquitin-positive and tau-negative inclusions are seen in ALS. The protein composition of aggregates varies according to disease background in ALS; however, several common constituents of protein aggregates have been identified (**Figure 1.3**).
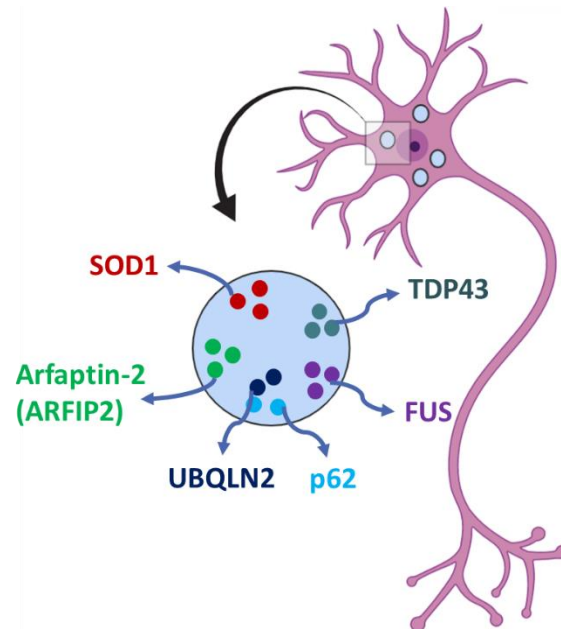
Figure 1.3: Common constituents of a protein aggregate in ALS
Protein aggregates in neurons of ALS patients are insoluble intracellular cytoplasmic inclusions which usually consist of misfolded proteins. The protein composition of aggregates varies according to disease background in ALS; however, several common constituents have been identified. These include SOD1, TDP43, FUS, p62, UBQLN2. Most recently a protein called Arfaptin-2 (ARFIP2) has been found to colocalize with proteins in aggregates of neurodegenerative diseases. SOD1 = Superoxide dismutase 1, FUS = Fused in Sarcoma, UBQLN2 = Ubiquilin-2. Image made using BioRender.

The first protein to be identified in aggregates was SOD1 (Rosen *et al.*, 1993). Aggregates comprising mutant SOD1 proteins have been found in sALS and fALS patients as well as in the mutant SOD1 mouse model (Boillée, Vande Velde and Cleveland, 2006). While misfolded mutant SOD1 aggregates are commonly found in *SOD1*-fALS, oxidised and misfolded wild-type-SOD1 inclusions have been reported in motor neurons of sALS cases that lack SOD1 mutations (Bosco *et al.*, 2010). Mutations in the SOD1 gene cause the mutant protein to misfold which escapes the degradation pathways. SOD1 is shown to be involved in the impairment of protein

degradation pathways (Bendotti *et al.*, 2012; Chen *et al.*, 2012). Mutant SOD1 has an increased aggregation propensity to oligomerize with itself or other intracellular proteins, in turn forming aggregates of high molecular weight and leading to a stress response (Shaw and Valentine, 2007). Furthermore, mutant SOD1 has been shown to induce the misfolded pathologic conformation of the protein on wild-type SOD1 in a prion-like manner, contributing to the protein aggregation pathogenesis of ALS (Maniecka and Polymenidou, 2015; Grad *et al.*, 2011; Münch, O'Brien and Bertolotti, 2011).

A major constituent of ubiquitinated protein inclusions in more than 95% of ALS cases, is TDP-43 (Arai *et al.*, 2006; Neumann *et al.*, 2006) and TDP-43 proteinopathy is now considered a pathological hallmark of ALS. Wild-type TDP-43 has been reported to accumulate in cytoplasmic inclusions in spinal motor neurons, neurons in the hippocampus and frontal cortex and glial cells in sALS cases and the majority of *SOD1*-negative fALS patients, however, this was not the case in *SOD1*-ALS patients (Mackenzie *et al.*, 2007; Tan *et al.*, 2007). This protein normally localizes to the nucleus, however, in ALS, the protein can be mislocalised to the cytoplasm and may lead to cytoplasmic aggregates (Neumann *et al.*, 2006; Suk and Rousseaux, 2020). Furthermore, TDP-43 has a tendency to be prone to aggregation which depends on its C-terminal (which has a prion-like domain) and is increased by ALS-associated mutations on TDP-43 (Johnson *et al.*, 2008; Johnson *et al.*, 2009; Nonaka *et al.*, 2013). In the event of stressors, i.e., oxidative stress, TDP-43 forms stress granules in the cytoplasm. Stress granules are transient complexes of RNA and proteins that form under cellular stress. They sequester aberrant or unnecessary mRNA and facilitate cell survival by preventing the translation of non-essential mRNA and pro-apoptotic proteins under stress (Protter and Parker, 2016). Further to TDP-43 being an aggregation-prone protein (Nonaka *et al.*, 2013), it is proposed that stress granules have a pro-aggregation property and lead to the formation of inclusions specific to ALS (Dewey *et al.*, 2012). In the cell, the TDP-43 protein is cleaved producing C-terminal fragments, and the fragments as well as the full-length protein are hyperphosphorylated (Neumann *et al.*, 2006; Inukai *et al.*, 2008). It has been shown that TDP-43 C-terminal fragments are found in spinal cord aggregates in ALS and are predominantly aggregation-prone as well as showing toxicity when overexpressed (Barmada *et al.*, 2010; Igaz *et al.*, 2009; Zhang *et al.*, 2009). Phosphorylation of TDP-

43 can regulate aggregation and its toxicity. It has been reported that, in C. elegans, inhibiting phosphorylation of TDP-43 ameliorates the neurodegenerative effect of *TDP-43* mutants in ALS (Liachko, Guthrie and Kraemer, 2010). On the other hand, it has been reported that mutating phosphorylation sites increases aggregation and hyperphosphorylation decreases aggregate formation and toxicity (Li *et al.*, 2011). Consequently, uncertainties concerning *TDP-43*'s pathogenic processes in ALS remain, despite the fact that it plays a significant and fundamental role in ALS pathophysiology.

Mutations in the *FUS* gene, also known as translocated in liposarcoma (*TLS*), are associated with ALS. FUS-positive cytoplasmic inclusions have been described in ALS patients carrying *FUS* mutations and allow characterisation of *FUS*-ALS (Kwiatkowski *et al.*, 2009; Vance *et al.*, 2009) and in SOD1-negative ALS cases (Deng *et al.*, 2010). In ALS patients with *FUS* mutations, predominant degeneration of lower motor neurons with FUS-positive cytoplasmic inclusions and normally distributed TDP-43 was seen in post-mortem tissue (Vance *et al.,* 2009; Hewitt *et al.*, 2010). This observation distinguishes *FUS*-ALS cases from other ALS patients. Like TDP-43, FUS is a nuclear RNA-binding protein and plays a role in DNA repair, transcription and mRNA processing (Lagier-Tourenne, Polymenidou and Cleveland, 2010). Previously, it was reported that most of the ALS-associated mutations reported in *FUS* are localised in the NLS of the protein which prevents its transport into the nucleus, causing mislocalization and accumulation of FUS in the cytoplasm (Dormann *et al.*, 2010; Dormann and Haass, 2011) where mutant FUS forms cytoplasmic aggregates (Kwiatkowski *et al.*, 2009; Vance *et al.*, 2009). Like TDP-43, FUS is prone to aggregation, however, unlike TDP-43 this is not influenced by the ALS-associated mutation (Sun *et al.,* 2011). Furthermore, FUS also has prion-like domains and is involved in self-aggregation in the cytoplasm (Feuillette *et al.*, 2017) and mutant FUS is shown to be engaged in stress granules that lead to cytoplasmic inclusions (Dormann *et al.*, 2010).

In addition to TDP-43-positive inclusions observed in motor neurons, which are characteristic of ALS (DeJesus-Hernandez *et al.*, 2011), it has been previously reported that, pathologically, ALS patients with the hexanucleotide repeat expansion mutation in the *C9ORF72* gene also show pathognomonic inclusions that are UBQLN-

and SQSTM1/p62-positive by immunohistochemistry but negative for TDP-43 (Al-Sarraj *et al.*, 2011; Troakes *et al.*, 2012). A recent study further supports this and implies that these inclusions may be due to an impairment of TDP-43 inclusion body formation in *C9ORF72*-ALS (Lee *et al.*, 2019). It is possible that this may be due to the pathological features associated with the *C9ORF72* repeat expansion mutation. Cytoplasmic inclusions also contain DPRs that are a result of non-ATG-initiated translation from the mutant intronic repeat expansion transcripts (Mori *et al.*, 2013c; Ash *et al.*, 2013). It may be possible that the DPRs are competing with TDP-43 to become components of the inclusions, or the DPRs and RNA foci could interfere with TDP-43 inclusion formation. The additional TDP-43 negative inclusions highlighted in previous studies may help distinguish the *C9ORF72* expanded repeat carriers from other ALS cases, however, studies have shown the presence of TDP-43 inclusions in *C9ORF72*-ALS (Chew *et al.*, 2019; Chew *et al.*, 2015a; Cook *et al.*, 2020; Liu *et al.*, 2016). Recently, Cook et al. (2020) showed that the poly-GR DPR sequesters endogenous TDP-43 and promotes the aggregation of TDP-43. Therefore, instead of interfering with TDP-43 inclusion formation, DPRs may induce TDP-43 proteinopathy. Aggregation of mutant RNAs is common in repeat expansion disorders. In C9ORF72-ALS cases, aggregates of RNA molecules/transcripts have been shown to sequester several RNA binding proteins and contribute to the protein aggregation pathology (Mori *et al.*, 2013b; Lee *et al.*, 2013; Cooper-Knock *et al.*, 2014). Of the several RNA binding proteins found sequestered to RNA molecule aggregates, hnRNPA3 was also found as a constituent of the p62-positive/TDP43-negative cytoplasmic inclusions in the hippocampus of *C9ORF72*-ALS patients.

Disruption of protein clearance pathways (UPS and autophagy) is suggested in ALS. In UPS, the proteins are marked by ubiquitination for degradation before they are recognised and degraded by the proteasome (Finley, 2009). Therefore, ubiquitin-positive inclusions suggest proteostasis impairment or overwhelmed protein degradation pathways. Autophagy is an intracellular protein clearance pathway that degrades and recycles long-lived proteins and organelles in the cytoplasm, maintaining homeostasis (Kelekar, 2005). ALS-associated genetic mutations have been found and reported in several genes involved in protein degradation pathways. These include mutations in *UBQLN2, OPTN, SQSTM1, C9ORF72* and *VCP.* Mutations in these genes give an insight into the aggregate formation and highlight

that dysregulation of protein homeostasis plays a pathogenic role in ALS. Proteasome-associated proteins encoded by these genes are all reported within cytoplasmic inclusions and further implicate the role of impaired proteostasis in ALS. Furthermore, it has been reported that with motor neuron-specific disruption of the proteasome, inclusions positive for TDP-43, FUS, UBQLN2 and OPTN form which shows that dysfunction of the proteasome is sufficient to cause aggregation of ALS-linked proteins and degeneration of motor neurons (Tashiro *et al.*, 2012). Studies have identified the SQSTM1/p62 protein as a common constituent present in nearly all pathological inclusions in protein aggregation diseases, including ALS (Teyssou *et al.*, 2013; Mizuno *et al.*, 2006). OPTN was initially detected in aggregates of only non-SOD1-ALS patients, however, some conflicting studies show the presence of OPTN in some *SOD1*-ALS cases (Maruyama *et al.*, 2010; Deng *et al.*, 2011a). Additionally, UBQLN2-positive inclusions have been reported in ALS with *SOD1, FUS, TARDBP* and *C9ORF72* mutations (Deng *et al.*, 2011b).

The exact mechanisms for the formation of cytoplasmic protein aggregations in ALS are still under study. Thus far, the literature has suggested a contribution through high propensity to the aggregation of proteins with prion-like domains, alterations in the formation of RNA stress granules and dysfunction of protein degradation pathways. Though the presence of cytoplasmic protein aggregates is a hallmark feature of ALS, their role in disease pathogenesis is still unclear. It is debated whether these inclusions are toxic and lead to neuronal dysfunction and eventual death; harmless by-products resulting from other toxic events; or function to protect cells by sequestration of misfolded toxic proteins (Baloh, 2011).

Proteins found in aggregates tell us a lot about the pathogenesis of ALS and enable researchers to identify therapeutic targets for the disease. Recently, studies have shown that Arfaptin-2 (ARFIP2) colocalizes with proteins in aggregates of neurodegenerative diseases and exerts neuroprotective effects (Peters *et al.*, 2002; Rangone *et al.*, 2005).

1.2   ADP-ribosylation factor-interacting protein 2 (Arfaptin-2)

Arfaptin-2, encoded by the ARFIP2 gene on chromosome 11p15.4, consists of 341 amino acids and has a calculated molecular mass of 38.6 kDa. It is highly

conserved between species including humans, mice and zebrafish, with 332 orthologs. In humans, ARFIP2 is ubiquitously expressed in many tissues and cell types, with high expression seen in the brain, including the cerebral cortex, hippocampus and cerebellum, as well as in several tissues including the heart, liver, kidney, lung and skeletal muscles (Kanoh, Williger and Exton, 1997; *Human Protein Atlas*). Subcellularly, ARFIP2 is diffusely located throughout the cytoplasm, localised to the perinuclear region and is found prominently around the microtubule-organising centre and Golgi apparatus. Additionally, its distribution is seen on membrane ruffles. When exogenously expressed, ARFIP2 was seen in the perinuclear region and dense aggregates within the nucleus. However, this pattern may be due to its overexpression (Peters *et al.*, 2002; Rangone *et al.*, 2005). The precise function of ARFIP2 is currently unknown, nevertheless, its structure and protein-protein interactions indicate possible functions.

1.2.1 Structure

As indicated through primary structural analysis (**Figure 1.4**), ARFIP2 is a hydrophilic protein with some hydrophobic regions. It has various phosphorylation sites and a leucine zipper motif. This motif confers hydrophobic regions to proteins to form dimers. Zipper proteins, most often transcription factors, generally have a positively charged domain adjacent to the N-terminus of the zipper that allows DNA binding. *ARFIP2*, however, lacks such a domain and therefore is unlikely to bind to DNA, but it can still form dimers (Kanoh, Williger and Exton, 1997). A Bin/Amphiphysin/Rvs (BAR) domain in the C-terminus of ARFIP2 allows dimerization and indicates that it may be involved in membrane binding, sensing membrane curvature and inducing membrane tubulation (Peter *et al.*, 2004; Ambroggio *et al.*, 2013). Preceding the BAR domain, ARFIP2 has an amphipathic helix (AH) which allows it to localise and bind at the trans-Golgi network (TGN). It has been shown that here ARFIP2 may play a vital role in the biogenesis of secretory storage granules (Cruz-Garcia *et al.*, 2013).
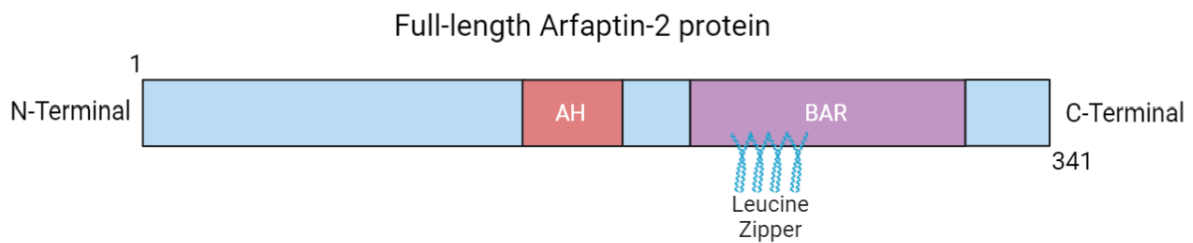
Figure 1.4: Schematic representation of the structure of Arfaptin-2
Known structural domains in ARFIP2 are shown – Amphipathic Helix (AH), Bin/Amphiphysin/Rvs (BAR) and leucine Zipper. Adapted from (Mohammedeid, Lukashchuk and Ning, 2014). Image made using BioRender.

1.2.2   Functions

ARFIP2 is also known as a partner of Rac1 (*POR1*) and interacts with small GTPases which include ADP-ribosylation factor (Arf) proteins and Rac1. These GTPases, belonging to the Ras superfamily, are enzymes that function as 'molecular switches' and regulate a plethora of cellular processes. GTP-dependent interaction of ARFIP2 with Arfs suggests its involvement in regulating intracellular vesicular transport, including between the endoplasmic reticulum and Golgi (cis), the formation of coated vesicles, endosomal trafficking and actin cytoskeleton organisation and rearrangement (Kanoh, Williger and Exton, 1997; D'Souza-Schorey *et al.*, 1997). Hinted by its name POR1, ARFIP2 also binds to Rac1 in a GTP-dependent manner which suggests a role in Rac1-mediated pathways including induction of membrane ruffling and lamellipodium assembly (Van Aelst, Joneson and Bar-Sagi, 1996). ARFIP2 also binds to Arl1 via its BAR domain. Arl1 recruits ARFIP2 to trans-Golgi membranes and regulates this association. Here, it is shown that ARFIP2 may be involved in inducing vesicular and tubular structures stemming from the trans-Golgi membranes (Man *et al.*, 2011). Furthermore, Recently, ARFIP2 has been shown to be involved in autophagy by regulating the amino acid starvation-dependent distribution of ATG9A vesicles which then deliver the PI4-kinase, PI4KIIIβ, to the autophagosome initiation site (Judith *et al.*, 2019). Previously, studies into ARFIP2 in neurodegenerative disorders have added further to its proposed functions. It has been

shown to be involved in regulating protein aggregation and exerting neuroprotective effects.

1.2.3 Arfaptin-2 in neurodegenerative diseases

Huntington's disease (HD) is a progressive neurodegenerative disorder caused by an aberrant expanded polyglutamine (polyQ) tract due to expanded CAG repeats at the N-terminus of the huntingtin (HTT) protein. The presence of nuclear and cytoplasmic aggregates of mutant-huntingtin and ubiquitinated proteins is characteristic of HD. Impairment of the proteasomal degradation pathway has been shown to be involved in the pathogenesis of the disease (Jana *et al.*, 2001). Previous literature has shown that the expression of ARFIP2 was significantly increased in the brain of HD mouse models and patients, and it has been implicated in the modulation of protein aggregation in HD (Peters *et al.*, 2002; Rangone *et al.*, 2005). **Figure 1.5** below summarises the key findings of the previous work showing the involvement of ARFIP2 in neurodegenerative diseases.
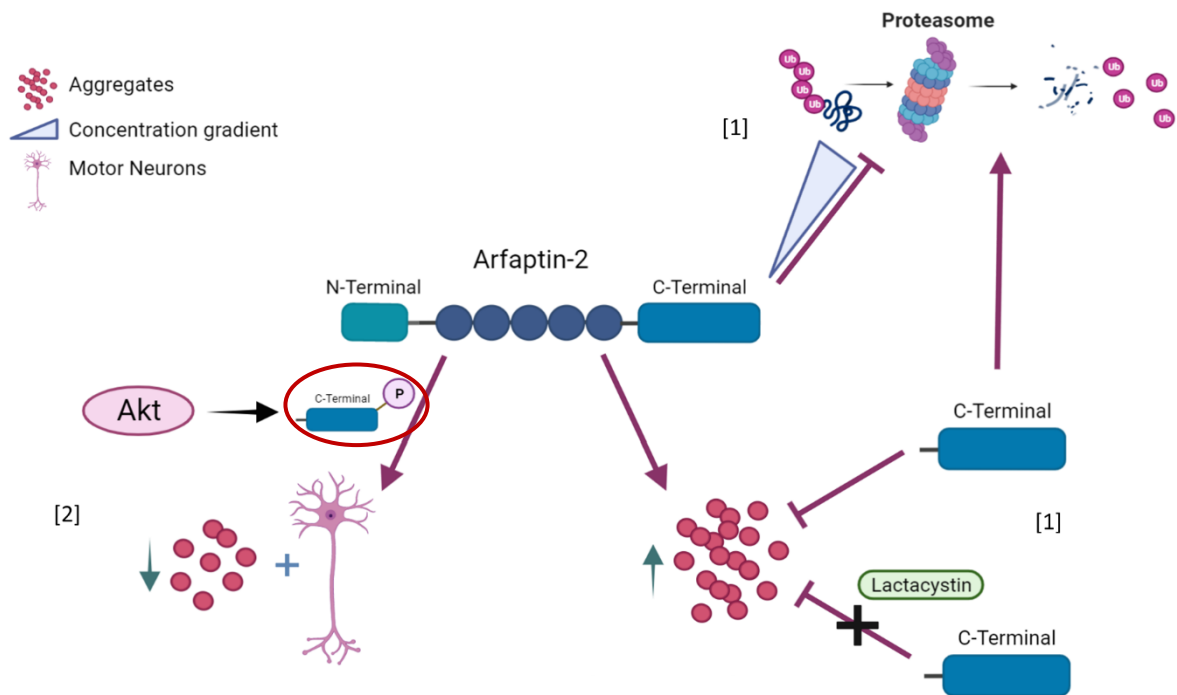
Figure 1.5: Schematic diagram summarising previous publications of ARFIP2 involvement in neurodegenerative diseases
Peters et al. (2002) showed that the C-terminal of ARFIP2 (HC-ARFIP2) dominantly inhibits aggregation of mutant Huntingtin, but not in the presence of proteasome inhibitors, and rescues proteasome activity [1]. Rangone et al. (2005) showed that full-length and phosphorylated ARFIP2 exerts a protective effect in HD that is mediated via phosphorylation of Ser260 by Akt [2]. It was noteworthy to see that the phosphorylation site (Ser260) is on the C-terminus of ARFIP2. Image made using BioRender.

Peters et al. (2002) showed that ARFIP2 localises within aggregates in HD cell models which may indicate a role in aggregate formation. Furthermore, they reported that it induces protein aggregate formation, like those seen in HD, in cultured cells and regulates protein aggregation. It was shown that, while full-length and N-terminal ARFIP2 promoted protein aggregation, expressing only the C-terminal of Arfaptin-2 (HC-ARFIP2) which comprises the last 172 amino acids, inhibited aggregation but not in the presence of a proteasome inhibitor called Lactacystin. This hinted that ARFIP2 may be working by affecting the proteasome activity. Here it was also shown that an increase in the concentration of ARFIP2, decreased proteasome activity and that expressing HC-ARFIP2 dominantly rescues proteasomal activity which is reported to be impaired in neurodegenerative disorders including HD and ALS (Waelter *et al.*,

2001; Mejzini *et al.*, 2019). Thus it was demonstrated that HC-ARFIP2 exerts neuroprotective effects in HD.

Support for the role of ARFIP2 in regulating protein aggregation is offered by Rangone et al. (2005). The study identified that the neuroprotective effects of full-length ARFIP2 were facilitated through its phosphorylation state at serine residue 260 (Ser^{260}). Phosphorylation of the protein was shown to decrease aggregation and promote survival in neuronal models of HD. The study gives further support by showing that ARFIP2 phosphorylation inhibits the proteasome blockade and rescues proteasome impairment. This phosphorylation was shown to be mediated by Akt, a pro-survival kinase, and it was shown to be effective in a huntingtin phosphorylation-independent manner. In contrast to Peters et al. (2002), it was proposed here that wild-type ARFIP2 does not affect aggregation which may be due to experimental differences or because the wild-type ARFIP2 was phosphorylated. Furthermore, the neuroprotective effects of full-length ARFIP2 were reported.

Though these studies show that ARFIP2 regulates protein aggregation and supports the retention of proteasome function thereby increasing neuron survival, they propose that ARFIP2 does so through different approaches. While one study suggests that full-length ARFIP2 promotes aggregation by inhibiting the proteasome degradation pathway and the truncated HC-ARFIP2 inhibits aggregation by rescuing the proteasome activity, the other study suggested that ARFIP2 is neuroprotective when it is phosphorylated (Rangone *et al.*, 2005; Peters *et al.*, 2002). It is noteworthy that the phosphorylation site (Ser^{260}) is on the C-terminal of the protein. Furthermore, C-terminal deletion mutants that no longer had the phosphorylation site, enhanced protein aggregation (Peters *et al.*, 2002). Therefore, although the two studies described different approaches, the pathway by which ARFIP2 is neuroprotective may be the same. Whether the neuroprotective effects of ARFIP2 are specific to HD or may be involved in other neurodegenerative diseases is yet to be determined.

1.3   Hypothesis

Since protein aggregation is a common feature in neurodegenerative disorders, therapeutic strategies may be transferrable. Previously, our lab looked into ARFIP2 in ALS and implied its involvement (Mohammedeid, 2015). Preliminary data showed that

HC-ARFIP2 significantly enhanced motor neuron survival in primary motor neuron cultures.

Accordingly, based on these findings and previous studies that highlight that HC-ARFIP2 can have neuroprotective effects in a neurodegenerative disorder background, we hypothesize that ARFIP2 may be a novel potential therapeutic target for ALS, targeting which may ameliorate the disease phenotype. In particular, we aim to assess the effects of expressing the dominant negative mutant of ARFIP2 (HC-ARFIP2) on disease characteristics in ALS using *in vitro* and *in vivo* models carrying the C9ORF72 mutation.

1.3.1 Aims

To assess the effects of HC-ARFIP2 on *C9ORF72*-ALS disease characteristics, we aimed to:

1. Express HC-ARFIP2 in the *in vitro* (iPSC-derived motor neurons) and *in vivo* (zebrafish) models of ALS used in this project. In order to achieve this, plasmids and appropriate delivery vectors were made and validated.

2. Investigate the expression and distribution of endogenous ARFIP2 in control and ALS biosamples, which include human brain tissue sections and *C9ORF72*-ALS iPSC-derived MNs.

3. Assess whether the expression of HC-ARFIP2 affects disease characteristics in *C9ORF72*-ALS iPSC-derived motor neurons.

4. Finally, with a therapeutic potential of HC-ARFIP2 seen in the *in vitro* model, we aimed to assess whether the expression of HC-ARFIP2 affects disease pathogenesis in a zebrafish model of *C9ORF72*-ALS.

2. MATERIALS AND METHODS

2.1 Molecular Biology Experiments

2.1.1 Materials

2.1.1.1 LV-based plasmids

The cDNA vector constructs encoding human wild-type Arfaptin-2, and Arfaptin-2 with C-terminus alone (HC-ARFIP2) were generous gifts from Dr. Crislyn D'Souza-Schorey (University of Notre Dame, Notre Dame, Indiana, USA). The LV-based plasmid stocks were made and maintained by Dr. Aida Mohammedeid during her Ph.D. Full-length Arfaptin-2 (FL-*ARFIP2*) and the C-terminal of Arfaptin-2 from base 510-1026 (HC-*ARFIP2*) were expressed in individual pLV_SIN-W-PGK-cPPT plasmids (LV based plasmids – pLV-FL and pLV-HC, respectively). Both gene inserts contain a Kozak sequence and an HA-tag epitope. pLV-FL and pLV-HC carrying the genes of interest were the starting constructs for further cloning. They used to generate other constructs used in the project. The individual pLV_SIN-W-PGK-cPPT plasmid expressing the green fluorescent protein (GFP) gene (pLV-*GFP*) was kindly provided by Prof. Mimoun Azzouz's lab. The plasmids maps of the constructs can be found in chapter 3 (section 3.3.1).

2.1.2 Methods

2.1.2.1 Subcloning of FL-ARFIP2 and HC-ARFIP2 into pcDNA3.1

pLV-FL and pLV-HC were digested (section 2.1.2.6) and sequenced (section 2.1.2.15) to check the presence and sequence of FL-*ARFIP2* and HC-*ARFIP2* genes. These genes from the LV plasmid were subcloned into the engineered pCDNA3.1 backbone to produce the final pCDNA3.1 plasmid with the genes of interest. A plasmid with the engineered pCDNA3.1 backbone was provided by Dr. Tennore Ramesh. The principle of the subcloning is summarised in **Figure 2.1**.
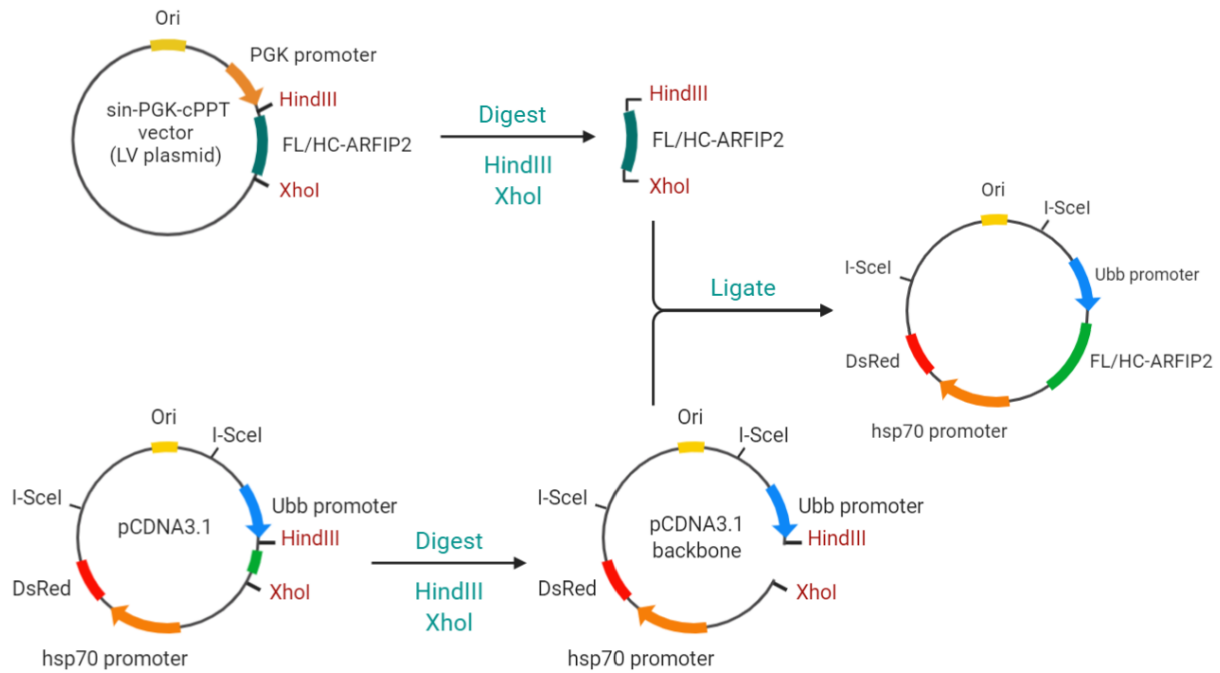
Figure 2.1: Schematic presentation of the subcloning method
The LV plasmid was digested with HindIII and XhoI restriction enzymes to isolate the gene of interest, ARFIP2 (FL/HC). The pCDNA3.1 recipient plasmid was also digested using the same restriction enzymes to linearize the plasmid. The isolated gene of interest was then ligated into the pCDNA3.1 plasmid backbone using DNA ligase.

2.1.2.1.1 Amplification of FL-ARFIP2 and HC-ARFIP2

The genes of interest were amplified using PCR (section 2.1.2.10), with some changes. The DNA templates used, and primer sequences are listed in **Table 2.1**. The primers were designed to incorporate the restriction enzyme sites into the PCR products for cloning into pcDNA3.1. A 50 µl PCR reaction mix was prepared (**Table 2.2**) in a tube. 3% DMSO was added to the PCR reaction mix for FL-*ARFIP2* only to resolve any secondary structures. The PCR was performed using the programme outlined in **Table 2.3**.

Table 2.1: DNA templates and primer sequences used for amplifying ARFIP2

DNA Template	Gene of Interest	Forward Primer (Fv)	Reverse Primer (Rv)
Sin-PGK-cPPT-FL-ARFIP2 (pLV-FL)	FL-*ARFIP2*	AT**AAGCTT**GCCACCATGACGGACGGGATC	CT**CTCGAG**TCAAGCGTAATCTGGAACATCGT

Sin-PGK-cPPT-HC-ARFIP2 (pLV-HC)	HC-*ARFIP2*	AT**AAGCTT**GCCACCATGCTCAGCCAGAAG	
The sequence in bold corresponds to HindIII. The sequence for <u>XhoI</u> is bold and underlined.			

Table 2.2: Reaction mix for PCR of FL-ARFIP2 and HC-ARFIP2

Reagent	Volume (µl)
DNA template (2 ng/µl)	1
Forward primer (10 µM)	2
Reverse Primer (10 µM)	2
DNTPs	1
Phusion Polymerase	1
5X Phusion Buffer	10
DMSO (3%)	1.5
NF-H2O	made up to 50 µl

Table 2.3: Standard PCR programme

Step	Temperature (°C)	Time (min:sec)	Cycles
Initial denaturation	94	2:00	
PCR Denaturation Annealing Elongation	 94 55 68	 0:15 0:30 3:00	 10X
PCR Denaturation Annealing Elongation	 94 58 68	 0:15 0:30 3:00	 20X
Final elongation	68	7:00	
Hold	10	Infinite	

2.1.2.1.2 Cloning PCR DNA fragments into the vector backbone

After the PCR was run, 1X DNA loading buffer blue (Bioline, bio37045) was added to 3 µl of PCR products and made up to 20 µl with NF-H2O. The 20 µl samples were loaded onto a 1% (w/v) agarose gel and the DNA fragments were separated using gel electrophoresis (section 2.1.2.11) to validate the PCR.

The remaining volume of PCR products was purified to strip off the ingredients of PCR, i.e., DNTPs, polymerase, primers, and buffer. The PCR products were purified using the QIAquick Gel Extraction Kit (QIAGEN, 28704), following the manufacturer's gel extraction protocol. Once the purified PCR product was eluted, it was quantified (section 2.1.2.14).

The PCR products were then digested with HindIII and XhoI restriction enzymes (section 2.1.2.6) and run on a 1% agarose gel (section 2.1.2.11) to separate the DNA fragments. The desired bands corresponding to the purified and digested PCR products (gene of interest inserts) were then excised and purified using the QIAquick Gel Extraction Kit (section 2.1.2.12).

Following the gel extraction and purification of the inserts, the FL-*ARFIP2* and HC-*ARFIP2* inserts and the already digested pcDNA3.1 vector backbone (with HindIII and XhoI) were ligated (section 2.1.2.13). The ligated constructs were transformed into DH5alpha competent cells (section 2.1.2.3), expanded (section 2.1.2.4), and isolated (section 2.1.2.5). The plasmids were validated by restriction enzyme digests (section 2.1.2.6) and sequencing (section 2.1.2.15).

The entire plasmid for pCDNA3.1-FL-*ARFIP2* and pCDNA3.1-HC-*ARFIP2* was sequenced. 12 primer pairs (forward and reverse) were designed using the ApE – A plasmid Editor software and NCBI primer blast. Primers that were 18-24 bp in length were chosen with a 40-60% GC content. The primers that were chosen either began or ended with at least one G/C pair. Primers had a primer melting temperature (Tm) of 50-60°C. The primer pairs had a Tm within 5°C of each other and did not have complementary regions. Primers at every ~1 Kb were chosen. Since good sequencing will produce 300-900 bp of useable sequence (sanger sequencing maximum read length), overlapping primer pairs were designed. **Table 2.4** below lists the primers used for both, pCDNA3.1-FL-ARFIP2 and pCDNA3.1-HC-ARFIP2.

Table 2.4: Primers used for sequencing pCDNA3.1-FL-ARFIP2 and pCDNA3.1-HC-ARFIP2

Forward Primer	Sequence	Reverse Primer	Sequence
Fv1	GCACTCTCAGTACAATCTGCTC	Rv1	GAGTGAACAGCAGCACAAAG
Fv2	GGGAATTTGGAACAGGTTTGTC	Rv2	ATGCAAGTTACTGGTCCTCC
Fv3	CAGTAGGACTCTCGACCATC	Rv3	TCGACCATATTAGTGGCACAG
Fv4	TCCTAGTTCAGTTTAGCCAACC	Rv4	CTCCATTGTGGCTGCCTTCC
Fv5	AAGGGTGACCAAATGCAAAG	Rv5	GGAACATCGTATGGGTACTGC
Fv6	GGGACAAGTATGAGAAGCTGC	Rv6	TGCGGTTGTGTGGTGTGTTG
Fv7	CGTTAAACGAATCTGACCAAGC	Rv7	ACAGGATGTCCCAGGCGAAG
Fv8	CTACTCCGAGAAAGTCATCACC	Rv8	AATCGGAACCCTAAAGGGAG
Fv9	CTGAGGCGGAAAGAACCAGC	Rv9	ATGGATACTTTCTCGGCAGG
Fv10	ACAATCGGCTGCTCTGATGC	Rv10	AATACGCAAACCGCCTCTCC
Fv11	AGCTAGAGCTTGGCGTAATC	Rv11	ACGAAATAGACAGATCGCTGAG
Fv12	GTCTGACGCTCAGTGGAACG	Rv12	AATTGTCGGTCAAGCCTTGC
All primers are the same for FL and HC, **except for Rv4 for HC**. The Rv4 for HC is**: GGACTTCTGGCTGAGCATGG**			

2.1.2.2 Production of ARFIP2-shRNA constructs

The steps involved in constructing and validating the shRNA plasmids are summarised below in **Figure 2.2.**
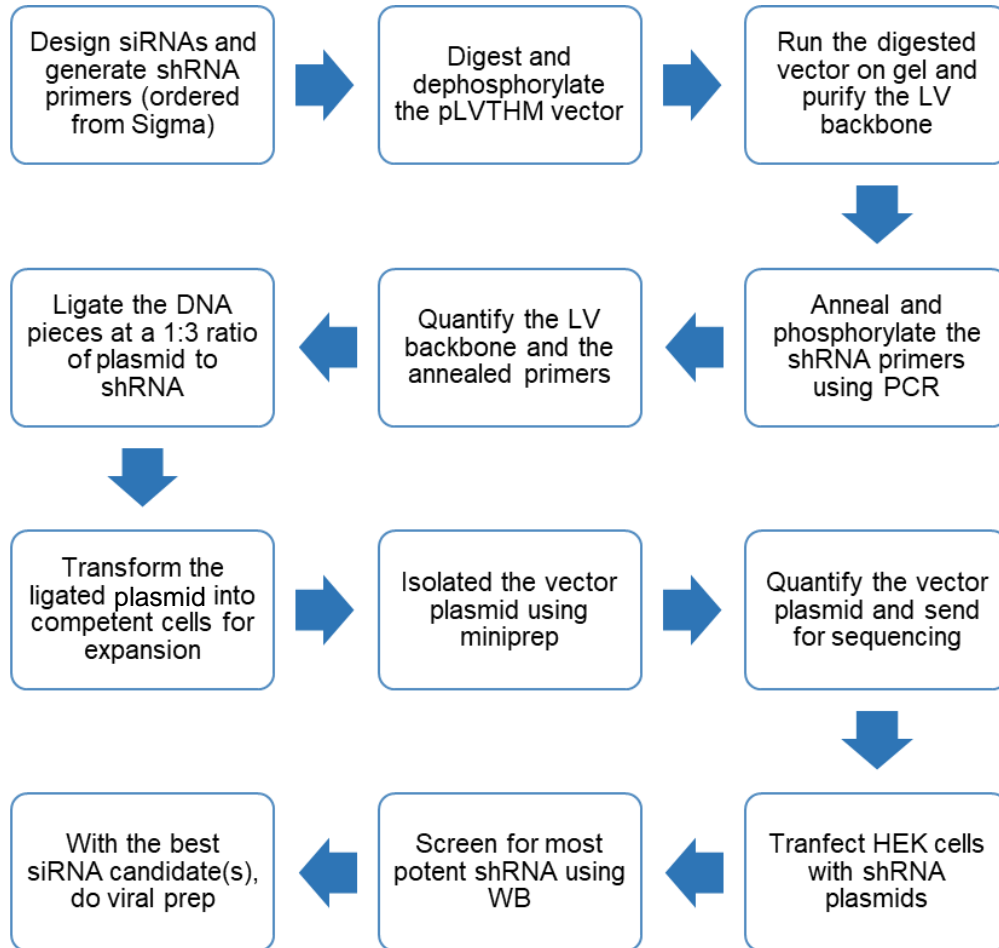
Figure 2.2: Workflow chart for production and validation of shRNA constructs
The chart above briefly describes the steps involved in the production of shRNA plasmids.

2.1.2.2.1 shRNA primer design

Short hairpin RNA or small hairpin RNA (shRNA) oligonucleotides (now referred to as oligos) were designed to be cloned into a vector for specific protein depletion guided by Dr. Evangelia Karyka. The oligos were designed according to the following template:

Forward –

CGCGTCCCC	(SENSE N19)	TTCAAGAGA	(ANTISENSE N19)	TTTTTGGAAAT

Reverse –

| CGATTTCCAAAAA | (SENSE N19) | ==TCTCTTGAA== | (ANTISENSE N19) | GGGGA |

The oligos sequence above allows the generation of a duplex with ends that would be compatible with those generated by restriction enzyme digestion of the vector backbone (MluI and ClaI restriction enzymes were used, section 2.1.2.6). The resulting construct would enable the duplex to be translated into shRNA oligos. The sequence highlighted in yellow shows the loop of the shRNA. To generate and order our shRNA oligos, the 19-nucleotide sense sequence (sense N19) was designed using the Stealth RNAi siRNAs target designer option on the Block-IT RNAi Designer online software (Invitrogen). The human ARFIP2 nucleotide sequence was used to search and select the target siRNA sequences which would be the sense N19 sequences in the template. Open reading frames (ORF) were selected as regions for target design. The three most appropriate open reading frame (ORF) siRNAs with a 45-55% G/C percentage were chosen (471, 480, and 749) and inputted into the template. The antisense 19-nucleotide sequence was generated using the ApE – A plasmid Editor software. The oligos generated that were ordered are listed in **Table 2.5** below.

Table 2.5: shRNA primers designed and ordered that target ARFIP2

	Sense strand	shRNA Loop	Anti-sense strand	
shRNA 1	siRNA_471			
Forward				
CGCGTCCCC	GGAGCTAGAGCTGCAGATT	==TTCAAGAGA==	AATCTGCAGCTCTAGCTCC	**TTTTTGGAAAT**
Reverse				
CGATTTCCAAAAA	GGAGCTAGAGCTGCAGATT	==TCTCTTGAA==	AATCTGCAGCTCTAGCTCC	**GGGGA**
shRNA 2	siRNA_480			
Forward				
CGCGTCCCC	GCTGCAGATTGAGTTGCTG	==TTCAAGAGA==	CAGCAACTCAATCTGCAGC	**TTTTTGGAAAT**
Reverse				
CGATTTCCAAAAA	GCTGCAGATTGAGTTGCTG	==TCTCTTGAA==	CAGCAACTCAATCTGCAGC	**GGGGA**
shRNA 3	siRNA_749			
Forward				
CGCGTCCCC	CCATGGAAGACACGCTCAT	==TTCAAGAGA==	ATGAGCGTGTCTTCCATGG	**TTTTTGGAAAT**
Reverse				
CGATTTCCAAAAA	CCATGGAAGACACGCTCAT	==TCTCTTGAA==	ATGAGCGTGTCTTCCATGG	**GGGGA**

2.1.2.2.2 Preparation of pLVTHM vector backbone

The pLVTHM vector was used for direct cloning of shRNA oligos into the vector. The vector was digested with MluI-HF and ClaI restriction enzymes (following the methodology described in section 2.1.2.6), with minor changes. The digestion reaction mix consisted of 1 µg of the pLVTHM vector, 1 µl of each enzyme (10 Units/µl), 1X CutSmart buffer, 1% BSA, and was made up to 50 µl with NF-H2O. The digestion mix was incubated for 1 hour at 37°C. Post digestion, the vector backbone was de-phosphorylated (section 2.1.2.7). The digested and dephosphorylated vector was electrophoresed on an agarose gel (section 2.1.2.11). The band corresponding to the linear LV-vector backbone was extracted and purified (section 2.1.2.12) and quantified (section 2.1.2.14).

2.1.2.2.3 Preparation of the shRNA oligos

The forward and reverse oligos needed to be prepared by phosphorylation and annealing for ligation into the vector. The oligos were phosphorylated using T4 polynucleotide kinase (section 2.1.2.8) and spun down before annealing using a thermal cycler (section 2.1.2.9). Once annealed, the samples were quantified (section 2.1.2.14).

2.1.2.2.4 Construction of the recombinant plasmid

The dephosphorylated vector backbone and the phosphorylated and annealed shRNA oligos were then ligated following the methodology described (section 2.1.2.13). The ligated shRNA construct was transformed into NEB stable cells (section 2.1.2.3) and the plasmids were isolated and purified from colonies using the QIAprep spin miniprep kit (QIAGEN) (section 2.1.2.5). Successful ligation was checked by digestion with restriction enzymes (section 2.1.2.6) and confirmed by sequencing (section 2.1.2.15). The primers used for sequencing were:

H1 forward: 5' TCGCTATGTGTTCTGGGAAA 3'

R2 reverse: 5' CGGCATCAGAGCAGATTGTA 3'

2.1.2.3 Transformation of competent cells

Competent bacterial cells were used for bacterial transformations. The DH5alpha competent cells were ordered from Invitrogen (ThermoFisher Scientific, 18265017) and the NEB Stable Competent *E. coli* and NEB 10-β Competent *E. coli*

cells were ordered from New England BioLabs (NEB) (C3040H and C3019I, respectively). Plasmid constructs or ligated DNA constructs were transformed into respective competent cells. The cells were thawed on ice for around 20 mins. In a 1.5 ml Eppendorf, 1- 5 µl (usually 10 pg - 100 ng) (1 µl for intact plasmid and 5 µl for ligated plasmid) of the plasmid of interest was added to 25 µl of competent cells. The DNA and cells were mixed gently by flicking the tube 4-5 times and incubated on ice for 30 mins. The cells were then heat shocked at 42°C for 45 seconds and then immediately cooled on ice for 2 mins. Room temperature (RT) SOC media (200 µl) was added to the tube and the cells were incubated at 37°C for an hour with vigorous shaking at 250 rpm (NB-205 mini shaking incubator, HandyLAB systems). 10 cm LB agar selection plates containing carbenicillin (LBa/Carb selection plates) (50 µg/ml), a substitute for ampicillin, were warmed to 37°C and 200 µl of transformed cells were added and spread on the selection plates using silica beads. The plates were incubated overnight at 37°C. The next day, individual colonies were picked using a sterile p200 pipette tip.

2.1.2.4 Plasmid expansion by bacterial culture

All steps were conducted under aseptic conditions. When a bacterial culture larger than 5 ml was needed, a starter culture was initiated. The volume of bacterial cultures was dependent on the amount of DNA required and the DNA isolation kit used (discussed further in section 2.1.2.5). Individual colonies were picked from the transformed bacterial selection plates using a sterile p200 pipette tip and inoculated into a starter culture with 5 ml of LB/Carb. The starter culture was incubated at 37°C for approximately 8 hours with shaking at 250 rpm as above. This starter culture was then used to inoculate the overnight culture of the volume needed as per the isolation kit used. The starter culture was diluted by 1:500 in a flask containing the appropriate volume of LB/Carb. The culture was incubated overnight at 37°C while shaking at 250 rpm. For bacterial cultures of 5 ml, individual picked colonies from the selection plates were directly placed into falcons or flasks containing Lysogeny broth (LB) supplemented with 50 µg/ml of carbenicillin. The bacterial cultures were incubated overnight at 37°C with vigorous shaking at 250 rpm (NB-205 mini shaking incubator, HandyLAB systems) for growth.

2.1.2.5 Plasmid isolation and purification

The plasmid purification kits were purchased from QIAGEN. Plasmids were isolated and purified from positively transformed bacterial cultures using the QIAGEN plasmid prep kits (Miniprep/Midiprep/Megaprep). The size of the prep was dependent on the amount of DNA desired. The Miniprep kit was used to isolate less than 20 µg of DNA using a 5 ml bacterial culture while the Midiprep and Megaprep kits were used to isolate up to 250 µg or 2.5 mg of DNA, respectively, using a 50 ml or 500 ml bacterial culture, respectively. The isolation and purification of plasmids were carried out following the manufacturer's instructions provided with the kits. The concentration of DNA was quantified (section 2.1.2.14) and constructs were stored at -20°C for future experiments.

2.1.2.6 Restriction enzyme digest

Plasmids were digested using specific restriction enzymes to check for inserts, prepare the vector backbone, and check for an intact vector backbone. To determine the choice of restriction enzymes to digest a specific plasmid, the SnapGene Viewer version 6.0.3 and A Plasmid Editor software (ApE) sequence analysis programmes were used. The reaction buffer was determined according to the enzymes chosen. In an Eppendorf, a digest reaction mix was prepared by mixing 1 µg of DNA, 1µl of each chosen restriction enzyme (10 units/µl), 1X of the appropriate restriction enzyme buffer, and made up to the desired volume using nuclease-free water (NF-H$_2$O) (Ambion/ThermoFisher Scientific, AM9938). The overall reaction volume ranges between 10-50 µl, depending on the application, and was mostly governed by the amount of DNA to be cut. The reaction mix was mixed gently by pipetting and incubated at 37°C for 1 hour. Restriction enzymes and their optimal buffers were purchased from NEB.

2.1.2.7 De-phosphorylation of vector backbone

Calf-intestinal alkaline phosphatase (CIP) was used to remove 5'-phosphate groups from the vector backbone to prevent self-ligation of the digested vector. This reduces the background in cloning. Post a restriction enzyme digestion of the vector, 1µl (1 unit/1 pmol of DNA) of CIP (NEB M0290) was added to the digest and incubated for 1 hour at 37°C. The CIP was then inactivated at 65°C for 30 mins.

2.1.2.8 Phosphorylation of oligonucleotides

If the vector backbone has been dephosphorylated, the inserts must be phosphorylated because ligation requires the presence of a 5' phosphate. Oligos were phosphorylated using T4 polynucleotide kinase (NEB M0201S). A 50 µl reaction mix was prepared which contained 3 µl of each oligo (100 µM stock concentration), 1µl of T4 polynucleotide kinase (10,000 units/ml stock concentration), 1X T4 ligase buffer, and NF-H$_2$O to make up the desired volume. The samples were spun down briefly and incubated in a thermal cycler at 37°C for 30 min first and then at 65°C for 20 min (heat inactivation).

2.1.2.9 Annealing polynucleotides

Single-stranded oligonucleotides with complementary sequences were annealed in order to generate a double-stranded molecule that can be used as an insert for ligation. The oligos were annealed using a thermal cycler. Equal volumes of the 100 µM oligos were mixed in a tube and were incubated at 95°C for 4 min first and then at 70°C for 10 min. The samples were then placed on the benchtop to cool down.

2.1.2.10   Polymerase chain reaction (PCR)

The genes of interest were amplified using PCR. A PCR reaction mix with the required reagents or 1X master mix, 1 µl of each 10 µM forward and reverse primers and 1 ug DNA template was loaded into each well of a 96-well PCR plate or a 0.2 ml thin-walled PCR tube and made up to the required volume with NF-H2O (reaction mixes for specific applications have been outlined in corresponding sections). The reaction mix was pipetted up and down to ensure the reagents were mixed properly and briefly centrifuged. PCR was conducted using a GS2 thermal cycler (G-Storm) with specific PCR programmes (outlined in relevant sections for an application). All PCR primers were ordered from Sigma. Primer designs are discussed in relevant sections (sections 2.1.2.1.1 and 2.3.2.5).

2.1.2.11   Agarose gel electrophoresis

PCR products or DNA digest reactions were run on either a 1% or 2% (w/v) agarose gel (measured amount of agarose (Melford, MB1200) in 1X TAE buffer (40 mM Tris base, 20mM glacial acetic acid, and 1mM EDTA) with 0.5 µg/ml ethidium bromide, EtBr (Sigma-Aldrich, E1510)) alongside a DNA molecular weight marker. The

concentration of the gel was dependent on the expected sizes of bands that needed to be separated. When the expected DNA fragments were <500bp, a 2% gel was used to resolve smaller bands from each other. Larger bands were separated using a 1% gel. Before loading the samples into the wells of the gel, a loading buffer dye was added to the samples. PCR reactions containing FIREPol already have a loading buffer in the samples and therefore the loading dye was added to samples that did not contain a dye. A DNA loading dye (DNA loading buffer blue (Bioline, bio37045)) was added to these samples. DNA was loaded into the wells of the agarose gel and run at 100-120V for 30 mins. The gel was imaged using the G-Box (Syngene).

2.1.2.12   Agarose gel extraction and purification

The gel extraction kits were purchased from QIAGEN. Post agarose gel electrophoresis (section 2.1.2.11), the DNA bands were visualised using the UV trans-illuminator. Bands at the desired heights were cut out from the gel using a blade and placed into a 1.5 ml Eppendorf. The DNA was extracted and purified by following the manufacturer's gel extraction protocol using the QIAquick Gel Extraction Kit (QIAGEN, 28704). This was done to strip off ingredients from any steps before DNA extraction. Once the purified DNA was eluted, it was quantified (section 2.1.2.14).

2.1.2.13   Ligation

DNA inserts and the previously digested vector backbone (with appropriate restriction enzymes) were ligated using T4 DNA ligase (NEB, M02025) according to the manufacturer's protocol. Before setting up the ligation reaction, the amount of digested insert and vector to be used for the reaction was determined (section 2.1.2.14). 50 ng of the vector was used and the vector: insert molar ratio was 1:3. To determine the required amount of insert, the following equation was used:

$$Weight\ of\ insert\ (ng) = weight\ of\ vector\ (ng)\ x\ \frac{size\ of\ insert\ (Kbp)}{size\ of\ vector\ (Kbp)}\ x\ molar\ ratio\ \frac{insert}{vector}$$

A ligation reaction mix was prepared to contain 50 ng of the digested vector backbone, the required amount of insert (ng), 2µl of 10X T4 ligase buffer, and 1µl of T4 DNA 5 U/µl ligase, and the volume was made up to 20µl with NF-H$_2$O. The mix was briefly

centrifuged and then incubated overnight at 16°C in a thermal cycler (G-Storm GS1 PCR/Thermal Cycler, Cambridge Scientific).

2.1.2.14   Quantification of DNA

All DNA was quantified using the Nanodrop spectrophotometer (Labtech) with the ND-1000 v3.2.1 software. Before quantifying any samples, the spectrometer was 'blanked' using the solution the DNA was eluted in to measure the background because of the machine or solvent. After 'blanking', 1 µl of the DNA was pipetted onto the pedestal of the spectrometer. The lid was closed, and the measurement was taken. The concentration of the DNA was recorded.

2.1.2.15   Sequencing of plasmids

Plasmids were sequenced by Sanger sequencing. This was performed by the DNA sequencing service at the University of Sheffield Core Genomics Facility. 10 µl of purified PCR products at 50 ng/µl or 10 µl of purified plasmid DNA in water at 100 ng/µl were provided per reaction with 10 µl of primers at 1 pmol/µl (µM) per reaction. All primers were ordered and purchased from Sigma. Primers used for sequencing specific plasmids are mentioned in relevant sections (sections 2.1.2.1.2 and 2.1.2.2.4)

2.1.2.16   Lentivirus production

The lentivirus (LV) production was conducted in a class II safety cabinet and appropriate cleaning protocols were followed. The packaging plasmids were kindly provided by Professor Mimoun Azzouz's group. HEK293T cells were used. FL-ARFIP2, HC-ARFIP2, and GFP were expressed in a lentiviral self-inactivating transfer vector (SIN) containing a woodchuck hepatitis virus post-regulatory element (W) to express FL-ARFIP2, HC-ARFIP2 or GFP under a PGK promoter (pLV-FL, pLV-HC or pLV-GFP). All plasmids were validated by sequencing (section 2.1.2.15). The LV-based plasmids were used to produce the corresponding virus (LV-FL, LV-HC, or LV-GFP). **Figure 2.3** shows a summary schematic representation of the lentiviral production method used.
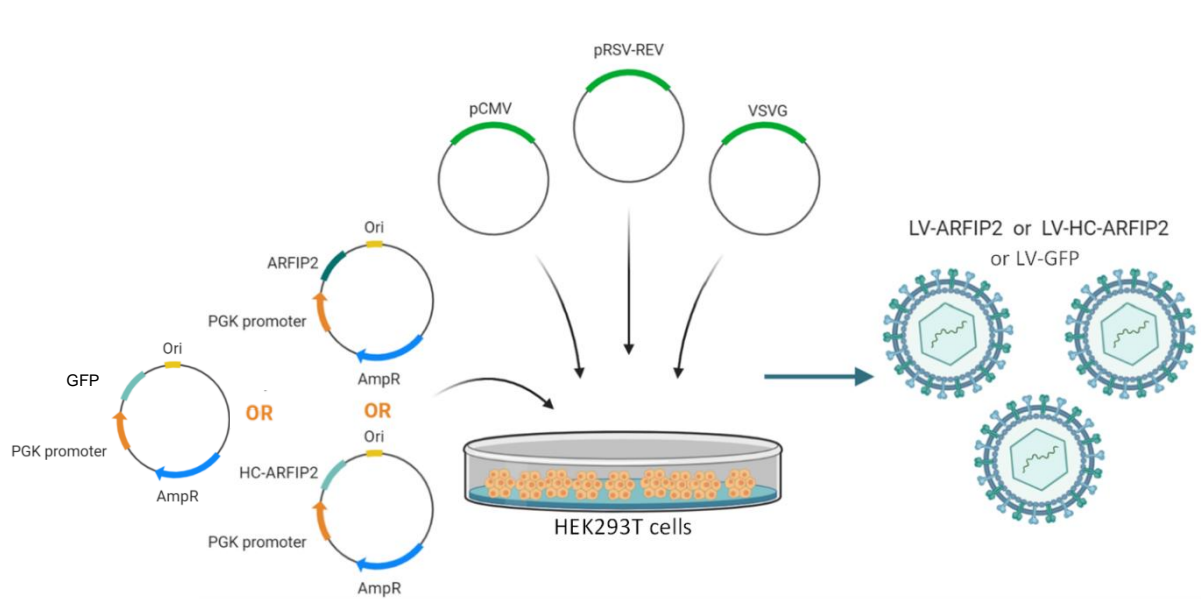
Figure 2.3: Four plasmid transfection method used for lentiviral production
The schematic summarises the lentiviral vector production method. HEK293T cells were transfected with LV based plasmids carrying FL-ARFIP2 (pLV-FL), HC-ARFIP2 (pLV-HC) or GFP (pLV-GFP) were transfected alongside 3 packaging plasmids. The cells were incubated and 3 days post transfection, the virus was harvested from the cell media and purified.

2.1.2.16.1   HEK293T cells plating for LV production

Cells were maintained in a T175 flask and were used at passage 21. At the beginning of the week of virus production, the cells were lifted from the T175 flasks using trypsin (section 2.2.2.1.3) and collected in a 50 ml falcon. 10 µl of the cell suspension was added onto a haemocytometer and the cells were counted (section 2.2.2.3). Twenty 10 cm tissue culture dishes were plated at a cell density of 3×10^6 cells per dish in 10 ml of DMEM media with 10% FBS and 1% P/S. The dishes were incubated at 37°C with 5% CO_2 overnight in an incubator.

2.1.2.16.2   Calcium phosphate transfection for LV production

The day after plating 3×10^6 HEK293T cells in each of the twenty 10 cm dishes (section 2.1.2.16.1), the twenty dishes were transfected with an LV-insert plasmid and three packaging plasmids using the calcium phosphate precipitation method (section 2.2.2.5). Previously optimised conditions for LV production were used. For one 10 cm dish, 1 ml of transfection mix was prepared which consisted of 13 µg of the

pCMVΔR9.92 packaging plasmid, 3 μg of the pRSV-REV plasmid encoding the HIV-1 Rev protein essential for viral replication, 3.75 μg of the pMD.G plasmid encoding the VSVG (vesicular stomatitis virus G) envelope protein, 13 μg of the LV-insert plasmid, 250 μl of 0.5 M calcium chloride ($CaCl_2$) and 500 μl of 2X HEPES-buffered saline containing sodium phosphate (HBS – 100 mM HEPES buffer, 280 mM NaCl, 1.5 mM Na2HPO4, adjusted pH to 7.1 with NaOH. Sterilization was achieved by passing through a 0.22 μm filter). The DNA was combined first and made up to a volume of 250 μl with sterile dH2O. $CaCl_2$ was then added dropwise to the DNA and then the DNA/$CaCl_2$ mix was added dropwise to bubbled 2X HBS. As described in section 2.2.2.5, the transfection mix was then mixed, precipitated, and added dropwise to the 10 ml of media already in each 10 cm dish, covering the whole plate. The plates were incubated at 37°C with 5% CO_2 and the complete media was changed after 6 hours of incubation. The plates were incubated for 48-72 hours before harvesting the virus (section 2.1.2.16.3).

2.1.2.16.3   LV harvest and purification

The Beckman Coulter – Optima L-100K ultracentrifuge (Beckman Coulter, 393253) was used for all centrifugation steps with the SW28 rotor. Before use, all pots and lids to be used for the rotor were disinfected with Distel (1:100) and 70% industrial methylated spirit (IMS) and then washed with dH_2O for 1 min each. All the media containing the virus from the transfected 10 cm dishes was collected and combined into one sterile flask. From this flask, the media was filtered into a new sterile flask using a 0.45 μm filter. The media was equally divided into Ultra-clear centrifuge tubes (Beckman, 344058), where each tube contained media from approximately 3.3 plates. The tubes were placed into SW28 pots and the pairs of pots containing the tubes were carefully balanced for centrifugation. The tubes were centrifuged using the SW28 rotor at 65,000xg in the Beckman Ultracentrifuge for 90 min at 4°C. Once spun down, the supernatant was decanted into virkon without disturbing the viral pellet and the tubes with the virus pellet were placed on ice. 1% BSA in PBS was prepared and filtered using a 0.22 μm filter. For a standard viral titre ranging between $1x10^8$ – $1x10^9$ TU/ml, 70 μl of the 1% BSA/PBS was added per transfected plate to each of the tubes containing a pellet (pellet was from approximately 3.3 plates). The tubes were incubated on ice for 1 hour. Post incubation, the pellets were resuspended in the 1% BSA/PBS, and all the suspensions were pooled and mixed thoroughly to ensure the

even distribution of the virus. The viral suspension was aliquoted to 0.5 ml tubes with 50 µl in each. All aliquots were stored at -80°C.

2.1.2.16.4   Genomic extraction of transduced cells

An estimated titre of each virus produced was determined by using an LV-GFP virus with a known biological titre (known through Fluorescence-activated cell sorting, FACS) for reference. A 12-well plate was seeded with 1×10^5 HEK293T cells/well with 1ml of DMEM media (10% FBS, 1% P/S) in each well. The next day, the cells were transduced (section 2.2.2.6.2) with LV-GFP (of known functional titre), and LV-insert (LV-HC or LV-FL) viruses serially diluted to 1:100, 1:1000, and 1:10,000 in a total volume of 500 µl/well. 72 hours post-transduction, genomic DNA from the transduced cells was isolated using the GenElute^{TM} Mammalian Genomic DNA miniprep kit (Sigma-Aldrich, G1N350) according to the manufacturer's instructions. The cultured and transduced cells were harvested by first washing the cells twice with ice-cold PBS and then scraping the cells in 500µl PBS using a sterile cell scraper. The cells were then pelleted down for 5 min at 300 x g and the supernatant was completely removed and discarded. Each pellet was re-suspended thoroughly in 200µl of re-suspension solution. The cells were then lysed by the addition of 20µl of the proteinase K solution, followed by 200 µl of lysis solution C. The mix was vortexed thoroughly and incubated for 10 min at 70°C. DNA binding columns were prepared by adding 500µl of the column preparation solution to each of the pre-assembled binding columns, centrifuging at 12,000xg for 1 min, and discarding the flow-through. To prepare the DNA for binding, 200 µl of 95-100% ethanol was added to the lysate and mixed thoroughly by vortexing before transferring the lysate mix into the treated binding column. The columns were centrifuged at ≥6500xg for 1 min. The column was washed twice with 500 µl of wash solution and centrifuged for 1 min at ≥6500xg and 3 min at maximum speed, respectively. Finally, the DNA was eluted by pipetting 200 µl of elution solution into the centre of the column and centrifuging for 1 min at ≥6500xg. The concentration of the genomic DNA from cultured and transduced cells was quantified (section 2.1.2.14)

2.1.2.16.5   Determining the functional viral titres using qPCR

Viral titres were determined through qPCR comparison against a virus of known titre, using the woodchuck hepatitis virus post-regulatory element (WPRE) primers: CCCGTACGGCTTTCGTTTTC (Fwd) and CAAACACAGAGCACACCACG (Rvs) as a

standard to assess the copy number of stably integrated lentiviruses. The qPCR reaction mix (10 µl) for LV titration is listed in **Table 2.6.** qPCR was run using the parameters listed in **Table 2.7**. The relative quantification of WPRE and amounts were achieved by using the ΔΔCT method. The subsequent values were used to calculate the approximate titre of each lentiviral vector relative to the known functional LV-GFP titre through FACS titration.

Table 2.6: Reaction mix for determining LV titre using qPCR

Reagent	Volume (µl)
SYBR green master mix	5
5 µM Forward primer	1
5 µM Reverse primer	1
NF-H2O	2
10 ng/µl DNA	1

Table 2.7: qPCR programme for determining LV titre

Step	Temperature	Time	Cycles
PCR initial activation	95	5 min	
Denaturation	95	10 secs	
Annealing	60	30 secs	40X
Melt curve analysis	65	5 secs	
Melt curve analysis	Ramp	(+ 0.5°C/5 secs)	
Melt curve analysis	95	-	
Store	10	Infinite	

2.1.2.17 Protein extraction from cells

Cells were harvested after 3 days post-transduction where the total protein was extracted. The media was aspirated, and the cells were washed once with RT PBS. Ice-cold PBS was added to each well. Using a sterile cell scraper, the cells (in PBS) were scraped and resuspended in PBS using a p1000 pipette until the cells in each well were detached. The cell/PBS mix was added into a 1.5 ml Eppendorf and centrifuged at 500xg for 5 min at 4°C. After the spin, the supernatant was aspirated and discarded, and the cell pellet was lysed by adding 200 µl of a lysis solution made

up of 1X RIPA lysis buffer (Merck Millipore 20-188), 1X protease inhibitor cocktail (100x) (PIC) (Thermo Scientific™, 78429) and made up to the desired volume with NF-H$_2$O. The cell pellet was resuspended in the lysis solution and incubated on ice for 30 min. The cells were then sonicated 3 times for 10 seconds each at an amplitude of 25% (Soniprep 150, MSE and Sonics, Vibra-cell) while on ice. This was followed by centrifuging the cells at full speed (17,000xg using the Fresco™ 17 Microcentrifuge, Thermo Scientific) for 10 min at 4°C. Following the spin, the pellet was discarded, and the supernatant was collected and stored at -20°C.

2.1.2.18 Protein extraction from embryos for SDS-PAGE

Batch protein extraction of zebrafish embryos at 5 dpf was done for western blot analysis. Zebrafish embryos obtained from crossing lines were kept in 10 cm petri dishes in E3 medium at 28°C. To the petri dish, 1 ml of pure Tricane was added to immobilise embryos. Around 50 embryos were transferred into 1.5 ml microcentrifuge tubes and all the E3 medium was removed. Embryos were washed once with PBS containing Tricane (1:20). The embryos were then washed with 1ml deionised water to remove excess salt and 1 ml of dH$_2$O was added to allow embryos to sediment to then remove the supernatant. 2x Laemlli buffer (125 mM Tris/HCl pH 6.8, 20% (v/v) glycerol, 10% 2-mercaptoethanol added fresh immediately before use, 4% SDS and 0.02% bromophenol blue) was prepared and 8 µl per embryo was added. The samples were then heated for 5 min at 95°C and pipetted up and down to disrupt embryos. The samples were sonicated for 10 seconds at 30% AMP (Soniprep 150, MSE and Sonics, Vibra-cell) followed by heating for another 5 min at 95°C. The samples were then placed on ice for 5 min and centrifuged at 17,000xg (Fresco™ 17 Microcentrifuge, Thermo Scientific) for 5 min using a benchtop microcentrifuge to remove any insoluble particles (*e.g.* pigment). The supernatant was pipetted into a new labelled tube while the pellet was discarded. The samples were stored at -20°C for further experiments. The samples were heated for 10 min and re-centrifuged for 5 min at 17,000xg (Fresco™ 17 Microcentrifuge, Thermo Scientific) prior to loading for western blot (20 µl/well).

2.1.2.19 Bicinchoninic acid (BCA) protein quantification assay for cell lysates

To determine the concentration of total protein in an extracted sample, a Pierce™ BCA protein assay kit (ThermoFisher Scientific, 23225) was used following

the manufacturer's protocol. On a 96-well flat-bottom plate, standards (100 µl/well) were prepared from a 2 mg/ml bovine serum albumin (BSA) stock. Different concentrations for the standards were achieved by doing a serial dilution. The first well of the standard was prepared with a concentration of 1mg/ml of BSA by diluting the stock in NF-H$_2$O. Following this, the standards were serially diluted using a 1:2 ratio in NF-H$_2$O. All standards were run in duplicate. To determine the concentration of samples, dilutions for the samples were prepared for a final volume of 100 µl/well. A small volume of lysate from the protein extraction samples was removed and diluted to 1:50 in NF-H$_2$O. This was then serially diluted using a 1:2 ratio to achieve a 1:100 and 1:200 dilution of the samples. A blank well was prepared that contained only NF-H$_2$O. The BCA working reagent (WR) was made by mixing reagent A and reagent B (50:1, respectively). 100 µl of the WR was added to all wells of the standards and samples. The plate was then incubated at 37°C for 30 min. Using the PHERAstar FSX (BMG Labtech) plate reader, the optical density (OD) was measured at a 562nm wavelength. Using Microsoft Excel, a standard curve was plotted, and the concentrations of the samples were determined.

2.1.2.20   Western Blot

The materials used for western blotting were purchased from BIO-RAD. **Table 2.8** summarises the buffers and solution used for western blotting and the reagents used to make them up

Table 2.8: Buffers and solutions for western blotting

Buffer/Solution	Reagents
10X Running buffer	25 mM tris-base, 250 mM glycine, and 0.1% SDS made up the desired volume in dH2O
4X Laemmli sample buffer (10ml stock)	2 ml 1 M Tris-HCL pH 6.8 (stacking gel buffer), 4 ml 1 M β-ME, 0.8 g SDS, 40 mg bromophenol blue, 3.2 ml glycerol, and adjust the final volume to 10 ml with Milli-Q water
5X Transfer Buffer	125 mM tris-base, 960 mM glycine, pH 8.3, and made up to the desired volume with dH2O
Resolving gel Buffer (1.5 M)	375 mM Tris base and 0.1% SDS in deionized water (dH2O), pH 8.8

Stacking gel buffer (1.0 M)	125 mM Tris base and 0.1% SDS in dH2O, pH 6.8
1X Tris-buffered saline, 0.1% Tween-20 (TBST)	137 mM NaCl, 26 mM tris-base, 0.1% tween-20, pH 7.6
Western blocking buffer (5% milk)	2.5 g of Non-fat dry milk powder in 50 ml of 1X TBST

2.1.2.20.1 Sample prep

The total protein was extracted from the samples (section 2.1.2.17 and 2.1.2.18) and the protein concentration for each cell lysate sample was determined (section 2.1.2.19). The samples were prepared to load onto the sodium dodecyl sulphate polyacrylamide gel (SDS-PAGE). For each cell lysate sample to contain 20 µg of total protein, a calculated volume of lysate was diluted in NF-H$_2$O and 4X Laemmeli sample buffer was added to each sample (final concentration of 1X). To reduce and denature the samples, the samples were boiled at 95°C for 5 min.

2.1.2.20.2 SDS-PAGE gel electrophoresis

A 10% or 15% separating gel was made by adding 10% or 15% ProtoGel 30% (w/v) acrylamide mix (National Diagnostics, Geneflow, A2-0072), 1.5M separating gel buffer (Tris-HCl pH8.8), 0.1% SDS, 0.1% ammonium persulfate (APS) and 0.1% TEMED [N, N, N', N'-Tetramethylethylenediamine] (Melford, T18000-0.1) to distilled water (dH$_2$O). This mix was loaded onto a 1.5mm thick SDS-PAGE plate and left to polymerise. Isopropanol was added to remove bubbles when necessary. Once the gel was set, a 4% stacking gel was made by adding 4% acrylamide mix, 1M stacking gel buffer (Tris-HCl pH 6.8), 0.1% SDS, 0.0.6% APS and 0.3% TEMED made up to a total volume of 4mls with dH$_2$O. The isopropanol was removed, the stacking gel mix was dispensed on top of the separating gel and a 1.5mm comb was placed into the stacking gel part of the gel to form wells.

2.1.2.20.3 Loading and running the gel

The set SDS-PAGE gel was placed into a tank with a clamp and filled with 1X running buffer. The comb was taken out of the gel and samples of 20 µg total protein

were loaded into wells of the gel along with 5 µl of the protein ladder (BioRad, ~245 kDa). The gel was electrophoresed at 70V until the samples ran through the stacking gel. Upon reaching the resolving gel, the voltage was increased to 170V, and the gel run was stopped when the dye front was at the bottom of the gel or ran off the bottom.

2.1.2.20.4   Electrophoretic transfer

While the gel was running, 1X transfer buffer was prepared with 20% methanol made up to the desired volume with dH_2O. Once the gel was stopped, the stacking gel with the wells was cut off and the protein gel was blotted onto a polyvinylidene difluoride (PVDF) membrane. Before the transfer, the membrane was activated by soaking it in methanol for 1 min. A transfer sandwich was set up and electrophoresed in the prepared 1X transfer buffer. The transfer was run at 270mA for 90 mins. To check if the protein bands had been transferred onto the membrane, Ponceau S staining dye was poured onto the membrane to visualise the bands and subsequently washed off with pure water.

2.1.2.20.5   Immunostaining

Before probing with antibodies, the membrane was blocked with 5% milk in TBST (blocking buffer) for 1 hour at RT. The membrane was then probed with diluted specific primary antibodies (**Table 2.9**) made up in blocking buffer and incubated overnight at 4°C. The next day, the primary antibody was removed, and the membrane was washed three times for 5 min each with TBST. Secondary antibodies (**Table 2.10**) specific to the primary antibody species were diluted and made up in a blocking buffer. The secondary antibodies were added to the membranes and incubated for an hour at RT. Post incubation, the secondary antibody was removed, and the membrane was washed three times for 5 min each with TBST.

Table 2.9: Primary antibodies used for western blotting

Antibody	Dilution Factor	Host	Manufacturer
Anti-Arfaptin2 (polyclonal)	1:500	Rabbit	Invitrogen (40-2400)
Anti-α-tubulin (monoclonal)	1:10,000	Mouse	Sigma-Aldrich (T9026)
Anti-HA-tag (monoclonal) (C29F4)	1:500	Rabbit	Cell Signalling Technologies, 3724

Table 2.10: Secondary antibodies used for western blotting

Antibody	Dilution Factor	Host	Manufacturer
Goat anti-rabbit IgG antibody, HRP-conjugate	1:5000	Goat	Sigma-Aldrich (12-348)
Goat anti-mouse IgG (H+L)-HRP conjugate	1:10,000	Goat	Bio-Rad (1706516)
Goat anti-Mouse IgG (H+L)-HRP	1:10,000	Goat	Invitrogen (31430)

2.1.2.20.6   Signal development and Imaging

After washing, the membrane was developed using the Pierce™ Enhanced Chemiluminescence (ECL) Western Blotting Substrate (Thermo Scientific™, 32106). Equal amounts of the two ECL reagents (1:1) were added onto the membrane and incubated for approximately one minute. The membrane was then visualised and imaged using the G-Box system (Syngene).

The ImageJ-win64 software (ImageJ) or GeneTools analysis software (Syngene) were used to analyse the images of the membrane and determine and compare the densities of the protein bands. Equal-sized rectangles were drawn around the bands to select each band. For each image, the rectangle size was carefully selected to avoid mixing between the bands in the adjacent lanes and was taller than the band itself to include the white background which allowed the software to compare the dark protein bands against the background as well as adjust for background. Over-exposure of the blot was avoided, and suitable exposures were used. The protein band densities were measured using the various tools of the software and the densities were represented as arbitrary units. The densities of the sample bands were normalised relative to the densities of the loading control bands of α-tubulin. Tubulin was chosen as the loading control as its highly conserved, has a high expression level and shows stability under many experimental conditions. However, in the future a total protein stain would offer as a better loading control as it directly measures the total amount of sample protein in each lane and increases the accuracy of normalisation. Expression levels were determined by comparing the normalised levels of the samples to those of the untransfected or control samples.

2.1.2.21   MSD ELISA

A poly (GP) Meso Scale Discovery (MSD) enzyme-linked immunosorbent assay (ELISA) was established in-house using a custom-made Rb-anti-GP antibody (Eurogentec) and adapted from previously described methods (Simone *et al.*, 2018).

2.1.2.21.1   Cell sample prep

Cells were harvested as described before in section 2.1.2.17. Briefly, iPSC-derived MNs were lysed in 1X RIPA buffer (Merck Millipore, 20-188) with a protease inhibitor cocktail (PIC) (Thermo Scientific™, 78429) and sonicated at 25% amplitude (Soniprep 150, MSE and Sonics, Vibra-cell) for 10 seconds on and off 3 times. The lysates were centrifuged at 17,000xg (Fresco™ 17 Microcentrifuge, Thermo Scientific) for 10 mins at 4°C and stored at -20°C till further use.

2.1.2.21.2   Zebrafish sample prep

Immobilised zebrafish embryos were added into 1.5ml screw-cap microcentrifuge tubes (alpha laboratories, CP5911) and all the E3 medium was removed. 150 µl of lysis solution (1X RIPA, 1X PIC, 3M Urea, 2% SDS, and water) was added to each tube along with a scoop of 1.4 mm ceramic (zirconium oxide) beads and placed on ice. The fish were then homogenised using the Precellys Evolution homogenizer (Bertin Instruments). The lysates were centrifuged for 15 min at 17,000xg (Fresco™ 17 Microcentrifuge, Thermo Scientific) and the supernatant was stored at -20°C. The MSD ELISA was run as previously described in section 2.1.2.21.3.

2.1.2.21.3   Assay

For the MSD ELISA, 30 µl of poly (GP) antibody (1 µg ml-1) diluted in 1X tris-buffer saline (TBS) was added per well of a 96-well SECTOR plate (MSD) and incubated overnight at 4°C. Plates were washed three times in 1X TBS-Tween®-20 (0.1%) and blocked for 1 h shaking in 50 µl 5% milk-TBS-T solution (RT, 700rpm). Plates were washed again in TBS-T and 50 µl $GP_{(7)}$ standard or zebrafish or cell lysate (1 mg/ml) was added per well and incubated for 2 h shaking (RT, 700 rpm). Plates were washed again in TBS-T and 50 µl MSD® SULFO-TAG labelled streptavidin (1 µg ml-1) and biotinylated poly (GP) antibody (1 µg ml-1, Eurogentec) were added per well diluted in blocking solution. The wells were washed again, 150 µl MSD® read

buffer-T was added and plates were imaged using the MESO SECTOR 2400 imager. Using the known concentration of the standards and the corresponding response, a standard curve was generated using GraphPad Prism version 9.3.1 and used to interpolate the amount of poly-GP in samples.

2.1.2.22 Histology for human brain tissue

The human brain tissue sections were requested from the Sheffield Brain Tissue Bank (SBTB). Tissue sections were requested from 6 healthy cases, 6 sporadic ALS patients, and 6 ALS patients carrying the C9ORF72 mutation. The ethical approval form can be found in **Appendix A**. Immunohistochemistry (IHC) was used to detect the expression and distribution of endogenous ARFIP2 (using the rabbit Arfaptin-2 Polyclonal antibody, Bioss, bs-1442r) in Formalin-Fixed Paraffin-Embedded (FFPE) cerebellum and neocortex (frontal lobe sections) from human brain tissue.

2.1.2.22.1 Antigen retrieval

To deparaffinise the FFPE slides, the slides were immersed in xylene twice for 5 min each and then rehydrated by sequential immersion into 100% ethanol twice for 5 min each, followed by placing in 95% and then 70% (Fisher Scientific, UK) for 5 min each. The section slides were then briefly washed in tap water for 5 min. After the wash, the slides were blocked for endogenous peroxidase by incubation in methanol with 3% hydrogen peroxidase (H_2O_2) for 20 min. The slides were washed in tap water for 5 min and enzymatic antigen retrieval was performed with 10 mM TSC buffer (3g of trisodium citrate powder in 1L distilled water with the pH adjusted to 6.0 using HCl) at pH 6.0 using the microwave method. The slides were placed in a TSC buffer at pH 6.0 in a container and microwaved for 15 min. The slides were then cooled gradually by placing them under running tap water for 1 min. The slides were then washed in TRIS-buffered saline (TBS) and incubated with TBS for 5 min.

2.1.2.22.2 IHC

Slides were immunostained with the Vectastain Elite ABC-HRP kits (Peroxidase, Rabbit IgG) as per the manufacturer's instructions, specific to the host species in which the primary antibody was raised. TBS was removed and the sections were incubated with the blocking solution (kit recipe) for 30 min followed by incubation in primary antibody overnight at 4°C in a humidity chamber. The secondary-only

control was incubated in a blocking solution to ensure consistent conditions. The next day, the slides were rinsed with TBS twice for 5 min each. TBS was removed and species-specific biotinylated secondary antibody (supplied with the kit, goat-anti-rabbit) was made up and added to the slides. Slides were incubated for 30 min at RT and then washed twice with TBS for 5 min each. The signal was amplified by incubation with the ABC reagent mix (supplied with kit) for 30 min at RT followed by two 5 min washes with TBS. Antibody staining was visualised using a 3,3'-diaminobenzidine (DAB) kit (Vector Laboratories) for ~ 6 mins and the reaction was quenched by rinsing the sections with dH_2O followed by washing in tap water for 5 min. The nuclei on the slides were counterstained with Harry's haematoxylin for 2 min and then washed in running tap water for 1 min. The haematoxylin was differentiated by dipping the slides three times in acid alcohol and then rinsing them with tap water. The slides were then incubated in Scott's tap water for 10 seconds to allow the haematoxylin to turn blue from purple. The slides were washed in tap water for 1 min and then dehydrated by immersion in progressively concentrated ethanol solutions, 70% and 90% for 1 min each and then twice in 100% ethanol for 1 min each. The slides were incubated twice in xylene for 10 min each followed by coverslipping (Fisher Scientific, UK) in DPX mounting media (Leica, UK). To set the DPX, the slides were dried overnight at 37°C in an oven.

2.1.2.22.3   Imaging and analysis of the FFPE sections

The stained slides were imaged on the NanoZoomer -XR Digital Slide Scanner, S210 (Hamamatsu). The slides were scanned by Daniel J Fillingham. The slides were analysed using the NDP.view2 Viewing software (Hamamatsu). Analysis of the staining intensity was done by Daniel J Fillingham.

2.2   In Vitro Experiments

Lentiviral vector-based gene transfer to transduced terminally differentiated cells and neuronal cells was conducted under the lentivirus licence number GMO2006_07B.

2.2.1 Materials

2.2.1.1 DEPC treatment of solutions

Diethyl Pyrocarbonate (DEPC) is a strong nuclear inhibitor and inactivates RNase in solution, thereby protecting RNA from being degraded. DEPC (Sigma Aldrich, D5758) was added to PBS or water at a final concentration of 0.1% (v/v). The solution was stirred at RT and the DEPC-mixed solutions were incubated overnight at 37°C. The next day, the solutions were autoclaved. DEPC-treated PBS (DEPC-PBS) was used to make 4% PFA when needed.

2.2.1.2 Lentiviral vectors

Lentiviruses produced (section 2.1.2.16) expressing genes of interest (LV-FL, LV-HC, and LV-GFP) were used to transduce iPSC-derived MNs and HEK293Tcells (section 2.2.2.6.1 and 2.2.2.6.2). LV-HC was transduced at MOI 40 and LV-HC was transduced at MOI 10.

2.2.2 Methods

2.2.2.1 Culturing HEK293T cells

2.2.2.1.1 Storage and cell maintenance

Vials of the human embryonic kidney cell line that expresses a mutant SV40 large T antigen (HEK293T) were stored in liquid nitrogen in 1 ml of 10% dimethyl sulfoxide (DMSO) and 90% foetal bovine serum (FBS). The cell line was kindly provided by Prof. Mimoun Azzouz's lab. The cells were cultured in Dulbecco's Modified Eagle's Medium (DMEM) with L-Glutamine and 4.5 g/L glucose without sodium pyruvate (Lonza, BE12-741F), 10% heat-inactivated foetal bovine serum (FBS) (Life Science Production 5-001A-H1-BR) and 1% penicillin/streptomycin (P/S) (Lonza DE17-603E).

2.2.2.1.2 Thawing and culturing HEK293T cells

All cell culture flasks and plates were purchased from CELLSTAR®. To plate the stored frozen cells, a vial of cells was thawed for approximately 1 min in a 37°C water bath, quickly suspended in 9ml of culture media, and cultured in a T25 flask. The flasks were incubated at 37°C and 5% CO_2. The following day, cell media was changed to ensure the removal of DMSO. Once 80-100% confluency was reached,

the cells were passaged in a 1:2 split ratio into T75 flasks to scale up the cell line. Eventually, the cells were maintained in a T175 flask.

2.2.2.1.3 Harvesting and passaging HEK293T cells

Cells were passaged using trypsin (Lonza BE02-007E) once 80-100% confluency was reached. To passage and lift the cells, cell media was removed, and the cells were washed once with 10ml PBS. 3ml of 1X trypsin was added to the T175 flask to detach the cells (1ml for T25 and 2ml for T75). The flasks were then incubated for 2-3 min (depending on cell adherence) at 37°C and 5% CO_2 to allow cells to detach. Following incubation, a volume of culture media was added to make up to 10 ml and the cells were suspended. The cell suspension was collected in a falcon. To maintain cell growth, the cells were passaged twice every week using a 1:10 split ratio into a fresh T175 flask. If cells were to be seeded into desired plates post passaging, the cells were counted using a haemocytometer (Section 2.2.2.3).

2.2.2.2 Culturing iPSC-derived motor neurons

iPSC cell culture was conducted under a cell culture GM licence (2001_14aGMAG (Neurodegeneration Cell Culture models Feb 2018 update)). Human-induced pluripotent stem cells (iPSCs) were differentiated into mature functional motor neurons (MNs). The origins of induced pluripotent stem cells (iPSCs) are listed in **Table 2.11**. All iPSCs were cultured and maintained by Dr. Cleide Dos Santos Souza until they reached the neuron progenitor cell (NPC) stage, at which stage they were provided. From the NPC stage onwards, the cells were maintained and differentiated using the protocols described below. All factors needed at various stages of the differentiation were added to basal media (**Table 2.12**). **Figure 2.4** briefly illustrates the established motor neuron differentiation protocol that was used.

Table 2.11: iPSC lines used and their origins

iPSC control Lines	Short name	Source Tissue	Clinical remarks	Mutation	Race	Gender	Age at Sampling	Supplier
CS14iCTR-nxx	Control CS14	Fibroblast	Clinically normal and healthy volunteer	N/A	CAUC	Female	30-35	cedars-sinai https://www.cedars-sinai.edu/Research/Research-Cores/Induced-Pluripotent-Stem-Cell-Core-/Stem-Cell-Lines.aspx#Amyotrophic%20Lateral%20Sclerosis%20(ALS)
MIFF1	Control MIFF1	Foetal foreskin fibroblasts	Clinically normal and healthy volunteer	N/A	Unknown	Male	Unknown	TUoS https://www.ncbi.nlm.nih.gov/pubmed/26810087
GM23338	Control Coriel/ PGP	Fibroblast	Clinically normal and healthy volunteer	N/A	CAUC	Male	55	Coriel https://www.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM23338
ALS C9orf72 lines	**Short name**	**Source Tissue**	**Clinical remarks**	**Mutation**	**Race**	**Gender**	**Age at Sampling**	**Supplier**
CS28iALS-C9nxx	ALS-28	Fibroblast	Age of onset: 46; Site of onset: left upper extremity.	C9orf72 hexanucleotide repeat expansion (6-8kb expanded allele)	CAUC	Male	47	cedars-sinai https://www.cedars-sinai.edu/Research/Research-Cores/Induced-Pluripotent-Stem-Cell-Core-/Stem-Cell-Lines.aspx#Amyotrophic%20Lateral%20Sclerosis%20(ALS)
CS29iALS-C9nxx	ALS-29	Fibroblast	Age of onset: 46; Site of onset: left upper extremity.	C9ORF72 hexanucleotide repeat expansion (6-8kb expanded allele)	CAUC	Male	47	cedars-sinai https://www.cedars-sinai.edu/Research/Research-Cores/Induced-Pluripotent-Stem-Cell-Core-/Stem-Cell-Lines.aspx#Amyotrophic%20Lateral%20Sclerosis%20(ALS)

CS52iALS-C9nxx	ALS-52	Fibroblast	ALS, Age of onset: 57; disease duration, 48 months; Site of onset: Left upper extremity.	C9orf72 hexanucleotide repeat expansion (6-8 kb)	Unknown	Male	49	cedars-sinai https://www.cedars-sinai.edu/Research/Research-Cores/Induced-Pluripotent-Stem-Cell-Core-/Stem-Cell-Lines.aspx#Amyotrophic%20Lateral%20Sclerosis%20(ALS)
ALS Isogenic Control Lines	**Short name**	**Source Tissue**	**Clinical remarks**	**Mutation**	**Race**	**Gender**	**Age at Sampling**	**Supplier**
CS29iALS-C9n1.ISOxx	ISO 29	Fibroblast	Age of onset: 46; Site of onset: left upper extremity.	Isogenic control line of CS29iALS-nxx	CAUC	Male	47	cedars-sinai https://www.cedars-sinai.edu/Research/Research-Cores/Induced-Pluripotent-Stem-Cell-Core-/Stem-Cell-Lines.aspx#Amyotrophic%20Lateral%20Sclerosis%20(ALS)
CS52iALS-C9n6.ISOxx	ISO 52	Fibroblast	ALS, Age of onset: 57; disease duration, 48 months; Site of onset: Left upper extremity.	Isogenic control line of CS52iALS-nxx	Unknown	Male	49	cedars-sinai https://www.cedars-sinai.edu/Research/Research-Cores/Induced-Pluripotent-Stem-Cell-Core-/Stem-Cell-Lines.aspx#Amyotrophic%20Lateral%20Sclerosis%20(ALS)

Table 2.12: Basal media composition

Component	Concentration	Reasoning
KnockOut DMEM/F12 media	48% (v/v)	–
Neurobasal media	48% (v/v)	–

B-27 supplement	1% (v/v)	Added to serum-free media for the long-term survival of neurons and to promote growth and proliferation of cells without differentiation.
N-2 supplement	0.5% (v/v)	Helps in the initial commitment and differentiation, survival, and expression of post-mitotic neurons in culture. Also favours the expression of neural-specific genes and inhibits the growth of non-neuronal cells in culture.
GlutaMAX™	1% (v/v)	Direct substitute for L-glutamine in media as it is more stable in aqueous solutions. L-glutamine is often added to media as it is an essential amino acid additive and also serves as a secondary energy source when cells are rapidly dividing, as well as a source of nitrogen for the synthesis of proteins and nucleic acids. L-glutamine is also the most potent anti-oxidant.
Penicillin/Streptomycin	1% (v/v)	Prevents bacterial contamination and maintains sterile conditions
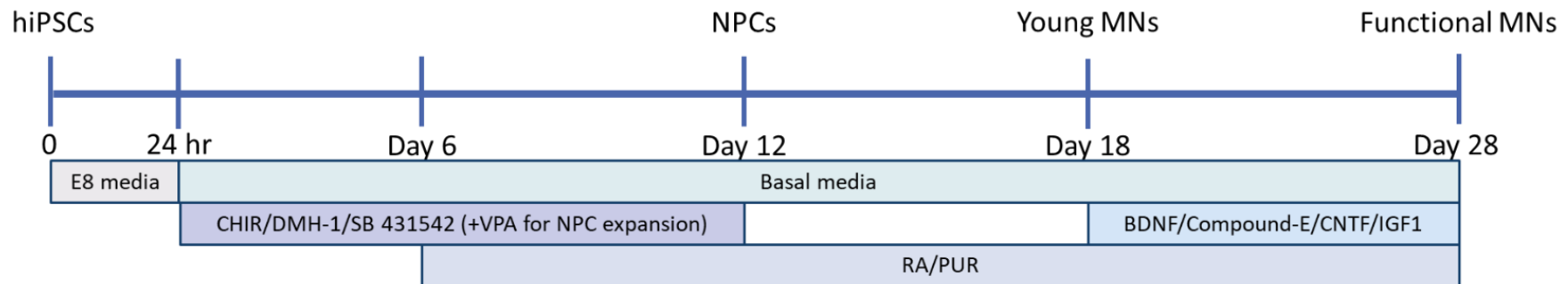
Figure 2.4: Established differentiation protocol for iPSC directed differentiation into functional motor neurons
The timeline here shows the established differentiation protocol that was used. Human induced pluripotent stem cells (hiPSCs) go through neural induction and caudalization until day 12, at which stage they are neural progenitor cells (NPCs). From here, with neuronal induction and maturation, they are differentiated into functional motor neurons (MNs). Scale bar = 10 µm. CHIR = CHIR99021, DMH-1 = dorsomorphin homolog 1, VPA = Valproic acid, RA = Retinoic acid, PUR = Purmorphamine, BDNF = Brain-derived neurotrophic factor, CNTF = Ciliary neurotrophic factor, IGF-1 = insulin-like growth factor 1.

2.2.2.2.1 Preparation of plates for culture

96-well plates, 12-well plates, and 6-well plates were used in culture along with 24-well plates with 13 mm sterile coverslips. The coverslips were immersed in 70% ethanol for 15 min and then washed with sterile water three times before placing one coverslip in each well of a 24-well plate. All plates were prepared for culture by coating with poly-ornithine (10 µg/ml) and Matrigel (Corning, 356230) matrix (0.5 mg in 6 ml Knockout DMEM media, ThermoFisher 10829018). The plates were coated with poly-ornithine and incubated overnight in the hood at RT. The plates were then sealed and stored at 4°C until needed for use. On the day of plating cells, poly-ornithine was aspirated, and the plates were washed three times with sterile PBS. Matrigel was prepared and added to the plates. The Matrigel-coated plates were sealed and incubated in the hoot for at least 1 hour before use. Matrigel plates were stored for a maximum of a week at 4°C.

2.2.2.2.2 Differentiating iPSCs to NPCs

iPSCs were cultured and maintained by Dr. Cleide Dos Santos Souza in mTeSR™1 media (STEMCELL Technologies, 85850). iPSCs were differentiated into NPCs by first culturing iPSCs in E8 media (STEMCELL Technologies, 05940) for 24 hours at 37°C with 5% CO_2. After this, the cells were cultured in iPSC-NPC day 1-6 differentiation media (**Table 2.13**) to lead the iPSCs in the direction of a neural pathway. The cells were kept at 37°C with 5% CO_2 and the media was changed every 24 hours. At this stage, an even neuroepithelial sheet formed. On day 7, cells were cultured in iPSC-NPC day 7-12 differentiation media (**Table 2.14**). The media was changed every 24 hours and the cells were expanded. The cells no longer displayed the rounded iPSC morphology, and instead displayed the morphology of neural rosettes. On day 12, the cells were confirmed and validated to show that they were NPCs.

Table 2.13: iPSC-NPC differentiation media (day 1-6)

Component	Concentration	Reasoning
Basal media (Table 2.3)	Media base	–
CHIR 99021	3 µM	GSK3 inhibitor/WNT activator and maintains neural stem cell; further increases neuralisation efficiency

| DMH-1 | 2 µM | Induces differentiation of human iPSCs to SOX1 and PAX6 expressing neural precursor cells |
| SB 431542 | 2 µM | Enhances differentiation of NPCs from human iPSCs |

Table 2.14: iPSC-NPC differentiation media (day 7-12)

Component	Concentration	Reasoning
Basal media (Table 2.3)	Media base	–
CHIR 99021	1 µM	As above
DMH-1	2 µM	
SB 431542	2 µM	
Retinoic acid (RA)	0.1 µM	Can promote the growth of dendrites and neuronal differentiation of neural stem cells
Purmorphamine (PUR)	0.5 µM	Activates the Hedgehog pathway and promotes differentiation of MNs

2.2.2.2.3 Expanding and passaging of NPCs

NPCs were maintained and cultured in NPC expansion media (**Table 2.15**). The media was changed every 2 days. When the cells were at 80-100% confluency, the cells were passaged. On the day of passaging, 10 µM ROCK inhibitor was supplemented into the media in each well and the plate was incubated for 1 hour at 37°C with 5% CO_2. Post incubation, the media in the wells was aspirated and the cells were washed once with PBS. The cells were then lifted from the plate surface by adding accutase (Merck Millipore A6964-100ML), enough to cover the cell surface per well. The plates were incubated for 7 mins. The cells in accutase were then pipetted up and down to detach and dislodge the cells. The cells were collected in 15 ml falcons and 1 ml of NPC expansion media (**Table 2.15**) was added to reduce the stress on the cells. The cells were centrifuged at 200 RFC for 4 mins at RT and the supernatant was removed. Each cell pellet was resuspended in 1 ml of NPC expansion media. The cells were either plated onto Matrigel prepared plated (section 2.2.2.2.1) using a 1:2 or 1:3 split ratio (dependent on the confluency of cells) or counted (section 2.2.2.3) to be plated at a specific density for experiments.

Table 2.15: NPC expansion media

Component	Concentration	Reasoning
Basal media (Table 2.3)	Media base	–
CHIR 99021	3 µM	As above
DMH-1	2 µM	
SB 431542	2 µM	
Retinoic acid (RA)	0.1 µM	
Purmorphamine (PUR)	0.5 µM	
Valproic Acid (VPA)	0.5 µM	Promotes neuronal differentiation and neurite outgrowth of neural stem cells

2.2.2.2.4 Freezing and storage of NPCs

For long-term use, cells were cryopreserved at the NPC stage of the differentiation protocol. To ensure recovery of NPCs, the cells were treated with 10 µM ROCK inhibitor, lifted with accutase, and pelleted down (as described in section 2.2.2.2.3). Each cell pellet was then resuspended in 2 ml of NPC expansion media supplemented with 10% DMSO. The cell suspension was then aliquoted in pre-chilled cryo-vials (1ml per vial) and the cryo-vials were immediately put into a pre-chilled Mister Frosty. The cells were at -80°C. For long-term storage, the vials were moved to liquid nitrogen.

2.2.2.2.5 Thawing of NPCs

To thaw and plate frozen NPCs, the frozen NPC cryo-vial was rapidly thawed in a pre-heated 37°C water bath for approximately 30 seconds. The contents of the vial were gently transferred to a single well of a Matrigel-coated 6-well plate. In each well, 5 ml of NPC expansion media (**Table 2.15**), supplemented with 10 µM ROCK inhibitor was added. The plate was then incubated for 24 hours at 37°C with 5% CO_2. The next day, the medium was replaced with 3 ml of fresh NPC expansion media per well without ROCK inhibitor. The cells were then regularly monitored, and the media was changed every 2 days until 80-100% confluency was reached at which point the cells were passaged (section 2.2.2.2.3).

2.2.2.2.6 Differentiating NPCS to mature MNs

The NPCs were maintained until 90-100% confluency was reached for motor neuron differentiation. Henceforth, the NPCs were differentiated into mature MNs by following an established protocol (Du *et al.*, 2015). Once 90-100% confluent, the media was switched to MN differentiation media for days 13-18 (**Table 2.16**) and then for days 19-28 (**Table 2.17**). The cells were usually passaged (section 2.2.2.2.3) around day 20 of the differentiation protocol to be plated at specific cell densities for experiments. The cells were then maintained as per protocol until day 28 at which point, the cells were differentiated into mature MNs. Markers were used to label mature MNs for validation. For experiment purposes, the cells were maintained in 19–28-day media until day 40 to ensure the development of disease phenotypes.

Table 2.16: MN differentiation media (day 13-18)

Component	Concentration	Reasoning
Basal media (Table 2.3)	Media base	–
Retinoic acid (RA)	0.5 µM	As above
Purmorphamine (PUR)	0.1 µM	

Table 2.17: MN differentiation media (day 19-28)

Component	Concentration	Reasoning
Basal media (Table 2.3)	Media base	–
Retinoic acid (RA)	0.5 µM	As above
Purmorphamine (PUR)	0.1 µM	
Compound E	0.1 µM	γ-secretase inhibitor and notch signalling pathway inhibitor; accelerates neuralisation
BDNF	10 ng/ml	Promotes differentiation and is a crucial regulator of neuronal survival
CNTF	10 ng/ml	Plays a significant role in neurogenesis and promotes cholinergic differentiation
IGF-1	10 ng/ml	Growth factor for neuroprotection and neurogenesis; promotes maturation

2.2.2.3 Cell counting using a Haemocytometer

Cells were counted using the Neubauer Improved haemocytometer for plating cells at specific cell densities. If the cells in suspension were too concentrated (i.e. 2.5×10^6 cells/well), the cells were diluted before to obtain a final concentration of around 1×10^6 cells/well. It was ensured that a cell count of at least 100 cells was achieved. To count cells, a glass coverslip cleaned with 70% ethanol was used to cover the clean chamber area. The coverslip was pressed on the chamber until newton rings were observed. This was to ensure that a 10 µl volume was consistently loaded. 10 µl of the cell suspension was introduced carefully into the Neubauer chamber, avoiding bubbles and moving the coverslip. The chamber was visualised under a microscope and the cells in the 4 quadrants were counted. The number of cells counted was divided by the number of quadrants to get an average (average number of cells counted). Where a dilution was done, the dilution factor was considered in the calculation. The specific volume needed to plate cells at the desired cell density was calculated using the equation below

$$\textit{The volume of cell suspension needed (µl/well)} =$$

$$\frac{\textit{Number of cells needed per well}}{\textit{average number of cells counted x Dilution factor x } 10^4} \; x \; 1000$$

In general, iPSC-derived MNs were plated at a cell density of 20,000 cells/well in a 96-well plate, 100,000 cells/well on coverslips in a 24-well plate, 1×10^6 cells/well in a 12-well plate and 2×10^6 cells/well in a 6-well plate.

2.2.2.4 Transfection of HEK293T cells using PEI

When the cells were 80-100% confluent, the cells were passaged (section 2.2.2.1.3), counted (section 2.2.2.3), and seeded onto a 6-well plate at a cell density of 300,000 cells per well. The cells were incubated at 37°C with 5% CO_2 overnight. The next day, plasmids were introduced into the cells by transfection with polyethylenimine (PEI), a stable cationic polymer. For a 6-well plate, the optimised PEI:DNA ratio is 3:1 (w/w). Therefore, 6 µg of PEI and 2 µg of DNA were used per well of a 6-well plate. A transfection mix was prepared by diluting the stated amounts of PEI and DNA in 300 µl of serum-free media (SFM) (DMEM) for one well. The transfection mix was vortexed for 10 seconds and then incubated at room temperature

(RT) for 15 mins. 300 µl of the transfection mix was added to the 3 ml of media in each well with a cell density of 300,000 cells/well. Cells were incubated at 37°C with 5% CO_2 for 3 days before harvesting for total protein extraction (section 2.1.2.17).

2.2.2.5 Transfection of HEK293T cells using calcium phosphate

The calcium phosphate precipitation method was used to transfect $3x10^6$ cells in a 10 cm dish. Transfection was done the day after plating cells. This method is based on the slow mixing of HBS with a solution of $CaCl_2$ containing the DNA. A DNA calcium phosphate co-precipitate develops which attaches to the cell surface and is taken up by the cell via endocytosis. On the day of transfection, the cells were prepared by aspirating old media and adding 10 ml of fresh media to a 10 cm dish. Two hours later, 1 ml transfection mix was prepared. In a sterile tube, 10-20 µg of DNA was made up to a volume of 250 µl with sterile dH2O. The amount of DNA varied according to the application (the amount of DNA for LV production is discussed in section 2.1.2.16.2). 250 µl of 0.5 M $CaCl_2$ was then added dropwise to the DNA. In a separate tube, 500 µl of 2X HBS was bubbled using a pipetboy attached to a 1 ml serological pipette fitted with a 200 µl pipette tip. While bubbling the HBS, the 500 µl DNA/$CaCl_2$ mix was added dropwise to 500 µl of 2X HBS. The transfection mix was gently mixed and left to precipitate by incubating at RT for 10 min until the mixture became cloudy or translucent. Post-incubation of the transfection mix, 1ml of the transfection mix was added dropwise to the 10 ml of media already in each 10 cm dish, covering the whole plate. The dishes were gently swirled to mix. The plates were incubated at 37°C with 5% CO_2. After incubation for 6 hours, the media of all plates was removed by aspiration, and 10 ml of fresh DMEM media was added. The plates were incubated at 37°C with 5% CO_2 for 48-72 hours.

2.2.2.6 Transduction with lentivirus
2.2.2.6.1 Transduction of iPSC-derived MNs

iPSC-derived MNs were transduced with LV-HC-ARFIP2 and LV-GFP on day 40 of the differentiation protocol. Cells were seeded at a density of $1x10^6$ cells/well (12-well plates) or 100,000 cells/well (24-well plates) or 20,000 cells/well (96-well plate). On the day of the transduction, conditioned media from the cells was collected in a sterile falcon and the plates were put back in the incubator with enough media left to cover the cell surface. To the conditioned media, an equal volume of fresh media

was added (now referred to as 1:1 media). The volume of virus needed per well was calculated using the formula below:

$$volume\ of\ virus\ (\mu l/well) = \frac{MOI\ x\ number\ of\ cells}{viral\ titre}$$

The calculated volume of each virus (µl/well) was added to a volume of 1:1 media needed per well (viral sample), enough to coat the cell surface per well (500 µl/well for a 12-well plate, 250 µl/well for a 24-well plate and 50 µl/well for a 96-well plate). From the plate to be transduced, the remaining media was discarded from each well and the viral sample was added to each well. 6 hours post-transduction, 3 parts of the 1:1 medium was added to 1 part of the viral sample. The cells were fed every 2 days with 50% fresh medium and were incubated at 37°C with 5% CO_2.

2.2.2.6.2 Transduction of HEK293T cells

HEK293T cells were transduced with LV-FL and LV-HC. The cells were seeded at a density of $0.15x10^6$ cells/well in a 12-well plate. An extra well was seeded at the same time at the same density. The next day, the cells in the extra well were harvested using trypsin (section 2.2.2.1.3) and counted (section 2.2.2.3). This was done to determine the approximate number of cells/well on the day of transduction and subsequently, the volume of virus (µl/well) needed using the formula in section 2.2.2.6.1. The volume of the virus (µl/well) was made up to 500 µl with culture media. From the plates, the media was aspirated and 500 µl of the viral sample was added onto the cells/well. The cells were incubated at 37°C with 5% CO_2 for 6 hours. Following this incubation, the media was completely changed with fresh culture media, and the cells were incubated at 37°C with 5% CO_2.

2.2.2.7 Fixing cells with 4% paraformaldehyde (PFA)

After 3 days post-transduction, the MNs were fixed using 4% PFA (4% (w/v) (Sigma, 158127) diluted in PBS). Each well was washed twice with RT PBS. Following this, 4% PFA was added to each well (enough to cover the layer of cells in a well) and the plate was incubated in the hood for 15 mins. The PFA was removed, and the wells were washed three times with 100-500 µl of PBS. After the last wash, fresh PBS was

added to each well and the plate was sealed with parafilm and kept at 4°C for future experiments.

2.2.2.8 Indirect Immunofluorescence (IF) – Immunocytochemistry (ICC)

The cells were fixed as previously described in section 2.2.2.7 and stored at 4°C. A stock of 0.5% Triton X-100-PBS was prepared for use as a permeabilization buffer and stored at 4°C for future use. 3% BSA/PBS was prepared and filtered to be used as a blocking buffer. From the plate with fixed MNs, PBS was removed, and 0.5% Triton X-100-PBS was added to each well. The plate was incubated for 10 mins at RT for permeabilization to allow penetration of antibodies. The wells were washed three times for 5 mins each with PBS and 3% BSA/PBS (filtered) was added to each well to block nonspecific antigen recognition. The plate was then incubated for 30 mins at RT. During this incubation, primary antibodies were prepared in 3% BSA/PBS according to selected dilutions and concentrations. Primary antibodies used at selected dilutions are listed in **Table 2.18**. 50 µl of primary antibody solution was either pipetted onto the parafilm and the coverslips with cells were placed face down on the parafilm or 50 µl of the primary antibody solution was added into each well of the 96-well plate. The cells were incubated overnight at 4°C in a humidity chamber. The next day, the primary antibody was removed, and the cells were washed three times for 5 mins each with PBS. Secondary antibodies were prepared in 3% BSA/PBS according to selected dilutions (**Table 2.19**) with the addition of a nuclear stain. The nucleus was stained with Hoechst (Sigma, 861405) at a working concentration of 1 µg/ml in the secondary antibody solution. Like the primary antibody solution, the secondary antibody solution (50 µl per coverslip or per well of a 96-well plate) was added to the cells and the cells were incubated in the dark at RT for 1 hour. The antibody was removed, and the cells were washed three times for 5 mins each with PBS. For mounting coverslips, microscope glass slides (Academy; 76x26mm, thickness 1.0-1.2 mm) were labelled appropriately. A drop of the VectaMount™ AQ aqueous mounting medium (Vector Laboratories, Z0214) was placed on the slides, and coverslips were mounted onto the glass slides. The slides were left in the dark overnight at RT and imaged using the confocal microscope (Leica) (section 2.2.2.14.1). The slides were stored at 4°C. The 96-well plates were kept in PBS with 0.02% sodium azide, parafilmed, and kept at 4°C until imaging on the Opera Phenix high-content screening system (section 2.2.2.14.2).

Chapter 2: Materials and Methods

Table 2.18: Primary antibodies used for IF

Antibody	Dilution Factor	Host	Manufacturer
Anti-human Arfaptin2 (polyclonal)	1:500	Rabbit	Invitrogen (40-2400)
Anti-HA-tag antibody (C29F4) (monoclonal)	1:500	Rabbit	Cell Signaling Technologies (3724)
Anti-human Arfaptin2 (polyclonal)	1:500	Mouse	(Novus Biologicals, H00023647-B01P)
Anti-human TDP43 (AP-2, polyclonal)	1:400	Rabbit	Proteintech (10782-2-AP)
Anti-TDP-43 antibody (polyclonal)	1:200	Rabbit	Proteintech (18280-1-AP)
Anti-MAP2 antibody (polyclonal)	1:1000	Guinea Pig	Synaptic Systems (188 004)
Anti-PAX6 antibody (polyclonal)	1:200	Rabbit	Abcam (ab5790)
Anti-Nestin antibody [2C1.3A11] (monoclonal)	1:200	Mouse	Abcam (ab18102)
Anti-Islet 1 antibody [EP4182] (monoclonal)	1:250	Rabbit	Abcam (ab109517)
Anti-Tubulin β 3 (TUBB3) Antibody (TUJ1) monoclonal	1:200	Mouse	Bio Legend (MMS-435P-50)
Anti-HA tag antibody [HA-7] (monoclonal)	1:500	Mouse	Abcam (ab49969)
Anti-Choline Acetyltransferase (ChAT) (polyclonal)	1:100	Goat	Sigma-Aldrich (AB144P)

Table 2.19: Secondary antibodies used for IF

Antibody	Dilution Factor	Host	Manufacturer
Anti-Mouse IgG (H+L) Highly	1:1000	Donkey	Invitrogen (A10037)

78 | P a g e

Cross-Adsorbed Secondary Antibody, Alexa Fluor 568			
Anti-Mouse IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	1:400	Goat	Invitrogen (A-21202)
Anti-Rabbit IgG (H+L) Highly Cross-Adsorbed Secondary Antibody, Alexa Fluor 488	1:1000	Donkey	Invitrogen (A-21206)
Anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 568	1:400	Goat	Invitrogen (A-11011)
Anti-Guinea pig IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 647	1:400	Goat	Invitrogen (A-21450)
Anti-Goat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 568	1:1000	Donkey	Invitrogen (A-11057)

2.2.2.9 Fluorescence RNA in situ hybridisation (FISH)

Cells plated on 96-well plates were used for this technique. Locked nucleic acid (LNA) probes were used. The 5'TYE-563-labelled LNA (16-mer fluorescent)-incorporated DNA probes were used against the sense (339500 LCD0156944-BKN/5TYE563/CCCCGGCCCCGGCCCC) and antisense (339500 LCD0147436-BKN/5TYE563/GGGGCCGGGGCCGGGG) RNA hexanucleotide repeats (Qiagen). All solutions (4% PFA, PBS, and dH$_2$O) used were treated with diethyl pyrocarbonate (DEPC) to inactive RNAse enzymes.

2.2.2.9.1 Sample prep

Cells were fixed as previously described (section 2.2.2.7) with some minor changes. The cells were fixed in RNAse-free 4% PFA/DEPC-PBS for 15 min at RT and washed three times for 5 min each with DEPC-PBS. The cells were kept at 4°C in DEPC-PBS until used.

2.2.2.9.2 In situ hybridisation

The cells on a 96-well plate were permeabilised with 4% PFA/DEPC-PBS with 0.1% Triton-X100 for 15 min at room temperature. For RNAse-treatment wells, 10 µl/ml RNAse-A/DEPC-PBS was added to wells and incubated for 30 min at 37°C while the non-treatment wells were kept in DEPC-PBS. Cells were blocked with 30 µl/well of hybridisation solution (**Table 2.20**) for 1 hour at 68°C in the Hybridisation Oven (Thermo Fisher Hybrid Oven) in a humidity chamber to ensure solutions did not dry out. A measured volume for a 40 nM final concentration of the LNA probes was incubated at 80°C for 75 s, then snap-cooled on ice for 5mins to denature any secondary structures. Following this, the cells were incubated with 40 nM of denatured RNA probes in the hybridisation solution overnight at 68°C in the oven (in a humidity chamber). The following day, the probe was removed and washed once with wash 1 (**Table 2.20)** for 10 min at RT. This is the high salt (HS) solution. The cells were then washed three times in wash 2 (**Table 2.20**) for 10 min each at 68°C. This is the low-salt solution (LS). All the solution from the wells was aspirated to ensure that the wells were as dry as possible. The cells were cross-linked using the crosslinker on the TL-2000 Ultraviolet Translinker (Ultra-violet Products) at an energy of 3000 x 100 µJ/cm^2 to provide final sterilisation. DEPC-PBS with 0.02% sodium azide was added to each well of the plate and stored at 4°C until imaging. Cells were imaged on the Opera Phenix™ high-content screening system (Perkin Elmer) and analysed using the Columbus software 2.8.2 (section 2.2.2.14.2). Immunocytochemistry after FISH (section 2.2.2.10) was then performed if needed.

Table 2.20: Solutions used in fluorescence in situ hybridisation for cells

Buffer/Solution	Reagents
Hybridisation solutions	50% formamide, 2X saline sodium citrate (SSC), 100 mg/ml dextran sulphate, 50 mM sodium phosphate pH 7.0 made up to desired volume with DEPC-H2O

Wash 1 (2X SSC)	2X SSC with 0.1% Triton X-100 and 1 µg/ml of Hoechst in DEPC-H$_2$O
Wash 2 (0.1X SSC)	0.1% SSC in DEPC-H2O

2.2.2.10   Immunocytochemistry after FISH

Before conducting ICC after FISH, the cells were imaged on the Opera Phenix™ for baseline acquisition. A DEPC-blocking buffer was prepared (5% horse serum in DEPC-PBS). DEPC-PBS was removed from the wells and 70 µl of permeabilization buffer (0.1% Triton-100 in DEPC-blocking buffer) was added per well. The cells were incubated at RT for 1 hour in the dark. Required primary antibodies at specific dilutions (**Table 2.18**) were prepared in the DEPC-blocking buffer and 70 µl of the primary antibody solution was added per well. The cells were incubated overnight at 4°C in the dark. The next day, the cells were washed with 0.1% Tween-2- in DPEC-PBS for 5 min, followed by two more washes with DEPC-PBS. Secondary antibodies at specific dilutions (**Table 2.19**) were prepared in the DEPC-blocking buffer. Alexa Fluor 568 was not used as this would clash with the RNA foci (Cy3). The cells were incubated at RT for 1 hour in the dark. The cells were washed with 0.1% Tween-2- in DEPC-PBS for 5 mins followed by two more washes with DEPC-PBS. The cells were then imaged on the Opera Phenix™ (section 2.2.2.14.2)

2.2.2.11   Proteasome activity assay – fluorometric assay

Proteasome activity was measured using Abcam's Proteasome Activity Assay Kit (fluorometric) (Ab107921) and followed according to the manufacturer's instructions. An AMC-tagged peptide substrate that releases free, highly fluorescent AMC in the presence of proteolytic activity, was used at 0.01 mM for preparing a series of dilutions for generating a standard curve (0 pmol/well to 10 pmol/well). Each dilution was run in duplicate. Cells were cultured at a density of 2x10^6 cells/well in a 6-well plate until day 40 of the differentiation protocol at which point, they were transduced with lentivirus and harvested. The cells were washed with cold PBS and resuspended in 0.5% NP-40 in PBS. The cells were then homogenised quickly by pipetting up and down and centrifuged at 4°C at 17,000xg (Fresco™ 17 Microcentrifuge, Thermo Scientific) using a cold benchtop microcentrifuge to remove insoluble material. The

supernatant was collected, and the pellet was discarded. Protease inhibitors were not used during cell lysate preparation. The dilutions of standards were loaded onto a 96-well plate in duplicate along with cell lysate samples and positive controls. All samples and the positive control were assayed with and without the proteasome inhibitor. Proteasome substrate was added to all samples and positive control wells, protected from light. The wells were mixed well and incubated at 37°C protected from light in the CLARIOstar® Plus multi-mode plate reader. The output was measured at T_1 (20 mins into incubation) at Ex/Em = 350/440 nm using the multi-mode plate reader at 37°C. The plate was incubated for a further 30 min at 37°C protected from light and the output was measured at T_2 at Ex/Em = 350/440 nm. Upon analysis, a standard curve was generated. The relative fluorescence units (RFU) generated entirely by proteasome activity for each sample were calculated using the formula:

$$\Delta RFU = (RFU_2 - iRFU_2) - (RFU_1 - iRFU_1)$$

Where, RFU_1 = output measurement in wells at T_1 without proteasome inhibitor, RFU_2 = output measurement in wells at T_2 without proteasome inhibitor, $iRFU_1$ = output measurement in wells at T_1 with a proteasome inhibitor, and $iRFU_2$ = output measurement in wells at T_2 with a proteasome inhibitor. Measurement of the wells which do not contain the proteasome Inhibitor will show total proteolytic activity RFU and the wells containing the proteasome inhibitor will show non-proteasome activity iRFU. ΔRFU shows proteasome activity RFU. The ΔRFU was applied to the standard curve to read out the amount of AMC in the sample well (B) and the activity of the proteasome in the test sample was calculated using the formula:

$$\text{Proteasome activity} = \left(\frac{B}{(T_2 - T_1) \times V} \right) x D$$

Where, B = amount of AMC in the sample wells (pmol), V= sample volume added into the reaction well (µl), T_1 = time (min) of the first reading, T_2 = time (min) of the second reading and D = sample dilution factor.

2.2.2.12   Resazurin Reduction Assay

The resazurin reduction assay was used to measure and monitor the cell viability of iPSC-derived MNs. Cells were cultured at a density of 100,000 cells/well on coverslips in a 24-well plate until day 40 of the differentiation protocol at which point, they were transduced with lentivirus (section 2.2.2.6.1). Prior to performing the assay, a resazurin stock solution was prepared at 1mM in PBS (Resazurin sodium salt - Sigma-Aldrich Cat.# R7017-1G) and sterilised using a 0.2 μm filter (stored at 4°C, protected from light). On the day of the assay (4 days post-transduction), a 10% (v/v) resazurin working solution was made using the resazurin stock solution with culture medium. The media was aspirated from the sample wells and washed once with PBS. 100 μl of the resazurin working solution was added directly to cells in each well and the cells were incubated for 2 hours at 37°C and 5% CO_2, protected from light. Viable cells with active metabolism reduced resazurin to resorufin. The resazurin solution was aspirated from the cells in the 24-well plate and added into an empty 96-well plate. Fresh cell culture media was added onto the cells upon complete removal of the resazurin working solution. The fluorescence of the 96-well plate was measured using the PHERAstar FSX (BMG Labtech) plate reader and the optical density (OD) was measured at a 540nm wavelength. The sample wells were assayed at regular intervals to measure the cell viability over time.

2.2.2.13   Live/Dead Assay

A live/dead assay was used to stain the nuclei of live cells (green) or dead cells (red) in a sample. Cells were cultured at a density of 100,000 cells/well on coverslips in a 24-well plate until day 40 of the differentiation protocol at which point, they were transduced with lentivirus (section 2.2.2.6.1). On the day of the assay, an assay working solution was prepared, consisting of 0.02% (v/v) SYTO-9 (Invitrogen, S34854) and 0.04% (v/v) Propidium Iodide (Invitrogen, P1304MP) in cell culture medium. The media was aspirated from the cells and 500 μl of the assay working solution was added to each well. The cells were incubated for 1 hour at 37°C and 5% CO_2, protected from light. The assay working solution was removed and the cells were fixed (section 2.2.2.7) and stored at 4°C until imaging using the confocal microscope (section 2.2.2.14.1).

2.2.2.14 Microscopy

2.2.2.14.1 Imaging using the Confocal microscope

Cells immunostained on coverslips were mounted on slides and imaged using the Leica SP5 confocal microscope. A 63x 1.4 oil objective lens was used. Image analysis was done using ImageJ (NIH).

2.2.2.14.2 Imaging using the Opera Phenix™ high-content screening system

Cells immunostained in 96-well plates were imaged on the Opera Phenix™ high-content screening system (Perkin Elmer). A 40x water immersion objection lens was used. The data were analysed using image analysis on the Columbus software (Perkin Elmer). All analysis workflows were generated according to experimental outcomes.

2.3 In Vivo Experiments

Zebrafish work was conducted under the GMO2012-13 and GMO2012-15 zebrafish GM licences.

2.3.1 Materials

2.3.1.1 Plasmid constructs

The plasmids used for the generation of stable transgenic zebrafish lines were made by subcloning the ARFIP2 genes from the sin-PGK-cPPT plasmids and into the engineered pCDNA3.1 backbone (section 2.1.2.1). The pCDNA3.1 construct had I-SceI meganuclease sites which increased the rate of transgenesis (Grabher, Joly and Wittbrodt, 2004; Liu, Jenkins and Copeland, 2003) along with the zebrafish heat shock protein 20 (hsp70) promoter followed by the DsRed (Discosoma sp. red fluorescent protein) marker gene. The final constructs are referred to as pCDNA3.1-FL-*ARFIP2* and pCDNA3.1-HC-*ARFIP2*. The constructs were validated by restriction enzyme digest (section 2.1.2.6) and sequencing (section 2.1.2.15).

2.3.1.1.1 Linearisation of zebrafish constructs

pCDNA3.1-HC-ARFIP2 was digested (section 2.1.2.6) with I-SceI meganuclease (New England Biolabs) to linearise the plasmids, with minor changes. A 50 µl reaction mix was made with 5 µg of DNA, 3 µl of I-SceI, 1X of the CutSmart buffer, and made up to 50 µl with NF-H2O. The reaction was incubated overnight at 37°C in a thermal cycler. The digest was loaded onto a 1% gel for electrophoresis

(section 2.1.2.11). The linearised fragment of interest was extracted from the gel and purified (section 2.1.2.12). The linearised DNA was quantified (section 2.1.2.14) and injected (section 2.3.2.3.1) into zebrafish embryos at the one-cell stage for the generation of transgenics.

2.3.2 Methods

2.3.2.1 Zebrafish maintenance

Zebrafish were housed and maintained in the Bateson Centre Aquarium at the University of Sheffield. The zebrafish were maintained at the optimal temperature of 28°C in a 14-hour light and 10-hour dark cycle. All experiments were conducted in line with the Home Office guidelines for animal research, in accordance with the Animal Scientific Procedures Act (ASPA) 1986 under the authority of project licence P26AA1EF5 and personal licence PIL I65179578. Maintenance of zebrafish was done using established practices (Westerfield, 2000).

2.3.2.2 Zebrafish breeding

Adult zebrafish were bred through the individual pair mating technique. A single male and a single female were paired in a tank with a divider separating each fish, the evening before spawning. At the beginning of the light cycle in the morning, the dividers were pulled, and the adults were allowed to mate. In the early afternoon or approximately 4 hours post fertilisation (hpf), on the day of spawning, the embryos were collected and sorted into groups of 50-60 embryos per 10 cm dish with aquarium system water. Unfertilised or dead embryos were removed, and the embryos were then transferred and maintained in E3 medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4, 0.0001% methylene blue) at 28°C. Zebrafish that were intended for raising to adulthood were transferred into tanks at 5 days post fertilisation (dpf) and maintained as described above.

2.3.2.3 Generation of HC-ARFIP2 transgenic zebrafish

The adult AB wild-type zebrafish strains were used to generate transgenics.

2.3.2.3.1 Microinjections

The microinjections of linear DNA into zebrafish embryos at the one-cell stage were done by Dr. Tennore Ramesh. The pCDNA3.1-HC-*ARFIP2* was linearised

(section 2.3.1.1.1) using the I-SceI restriction enzyme and diluted in a buffer containing 1X enzyme buffer, 1 U/ml of I-SceI, Phenol Red, and 1X injection buffer (10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 100 mM NaCl, 30 μM spermine and 70 μM spermidine) to a final DNA concentration of 40 ng/nl, and 1 nl of DNA solution was injected into early one-cell stage zebrafish embryos.

2.3.2.3.2 Transient expression assay

In the transient expression assay, the DNA-injected embryos are examined to see whether the promoter in the construct is able to drive the transgene expression. At 24 hpf, the zebrafish embryos were heat-shocked (section 2.3.2.4) for 3 hours. At 72 hpf, the embryos were screened under the fluorescence microscope to sort for zebrafish that showed strong DsRed expression levels with a high number of cells expressing the marker gene. A few of the selected embryos were used for validating the expression of the transgene, while the rest of the selected embryos were raised to adulthood to then screen for positive founders (F0) (section 2.3.2.3.3).

2.3.2.3.3 Screening and identification of F0

At certain probabilities, the injected DNA integrates into the genome of the germ cell precursors in embryos. These adults would be F0. The injected adults were crossed with adult AB wild-type zebrafish strains (section 2.3.2.2). Through the cross, if the injected fish is F0, there would be a 50% chance of offspring carrying the transgene and displaying the DsRed phenotype. At 24 hpf, the zebrafish embryos were heat-shocked (section 2.3.2.4) and at 72 hpf, the embryos were screened (as above) for DsRed expression. Those that produced viable DsRed-positive embryos were identified as F0 and chosen to generate and establish transgenic lines.

2.3.2.3.4 Establishing stable lines

Identified F0 fish were crossed to AB wild-type fish (section 2.3.2.2) and the offspring were heat-shocked and screened for DsRed expression (section 2.3.2.4). DsRed-positive embryos were raised to adulthood and maintained as founder transgenic fish (F1). With the transgenes integrated into the germ line, the DsRed and transgene were ubiquitously expressed. F1s were bred to generate individual transgenic lines. Two F0s gave rise to transgenic F1 lines (line 2 and line 4). The adult

C9ORF72-ALS sense and antisense established lines were used (Shaw *et al.*, 2018) for a cross with these transgenic lines.

2.3.2.4 Heat-shocking zebrafish embryos

At 24 hpf, zebrafish embryos were transferred from the 10 cm dishes and into 96-well, non-skirted PCR plates (Thermo Fisher Scientific, AB0600). Each embryo was pipetted into a single well of the 96-well plates and the plates were sealed using adhesive PCR plate seals (Thermo Fisher Scientific, AB0558). The plates were placed into a standard thermocycler (Bio-Rad Peltier Thermal Cycler PTC-100) and the embryos were heat-shocked for 3 hours using the incubation protocol in **Table 2.21**. Following heat shock, the embryos were transferred back into the 10 cm dishes and maintained in E3 medium. At 72 hpf, the embryos were screened and sorted under the fluorescence microscope according to DsRed fluorescence.

Table 2.21: Incubation programme for heat shocking zebrafish embryos

Incubating Temperature (°C)	Time (min:sec)	Cycles
23 37	30:00 30:00	3X
28	Infinite	

2.3.2.5 Zebrafish genotyping

Adult zebrafish at 3 months old or zebrafish embryos at 3 dpf were genotyped using the tail biopsy method (fin-clipping). Embryos at 5 dpf were genotyped as a whole. If fin-clipped, zebrafish embryos and adults were anaesthetised by treating with Tricane MS222 (4 g/L) diluted to 4.2ml per 100 ml aquarium water and moved to fresh water after biopsy. While anaesthetised, a small section from the end of the caudal fin was biopsied using a scalpel and transferred to a 96-well microliter plate containing 20 µl of QuickExtract™ solution (Epicentre Biotechnologies) for DNA extraction from the biopsy. 5 dpf embryos were directly placed into 20 µl QuickExtract™ for DNA extraction. Plates containing the biopsies or embryos were incubated at 65°C for 2 hours followed by incubation at 98°C for 2 min. Prior to genotyping by PCR using the DNA extracted from the fin clip (3-month-old adults only) and 5 dpf embryos, the DNA was diluted using nuclease-free water (1:3). A PCR reaction mix was used for

genotyping with the reaction containing a final concentration of 1X FIREPol® (Solis Biodyne, OÜ, Tartu, Estonia), 1 µM each of forward (Fv) and reverse (Rv) primers (**Table 2.22**), 3% DMSO and 1 µl of the diluted DNA template, made up to 10 µl with dH2O. The DNA templates were amplified using a 30X touchdown PCR programme (**Table 2.23**). The PCR products were separated on a 2% agarose gel electrophoresis by loading the entire 10 µl PCR mix. A 120V voltage was applied for 35 min to the gel and imaged on an SYNGENE G:box.

Primers were designed using the NCBI primer blast (**Table 2.22**). The plasmid sequences for pCDNA3.1-HC-ARFIP2 and the construct with C9ORF72 repeats (used to make the C9ORF72 transgenic zebrafish in (Shaw *et al.*, 2018)) were used as DNA templates. The HC-ARFIP2 (HC) primer pairs were designed so that they were able to amplify the region from before the 5' end of the HC-ARFIP2 sequence to a region near the 5' start of the HC-ARFIP2 sequence. The C9ORF72 (C9) primer pairs were designed to amplify a region from within the 3' end of the C9ORF72 repeat insert to a region flanking the 3' end of the insert. The primer pairs were checked against the genomes of zebrafish on the NCBI primer blast to ensure that there were no unintended products from PCR. Reference primers (E8, E2, and E4) were provided by Dr. Andrew Grierson and were used as positive controls.

Table 2.22: Primer sequences used for genotyping zebrafish

Gene		Sequence (5'->3')	Product Size
HC-ARFIP2 (HC)	Fv	TCAGCGTTTCCTTACTGTGTAGA	703
	Rv	TCCTAGCAGCGTTTCCCCATT	
C9ORF72 (C9)	Fv	CGGCCGCTGGGCAAG	280
	Rv	CTAGAAGGCACAGTCGAGGC	
GPAA1_e8 (E8)	Fv	AGTCCCTGACATTCCTTGTTGT	268
	Rv	TGTAAAACGACGGCCAGTTGTCCCTTTGAATCACGTCC	
PIGG_e2 (E2)	Fv	ATAAGAGTGTTTGTGTGTTTCTGC	254
	Rv	TGTAAAACGACGGCCAGTTGATTGGACTGCTTCTGACCT	
PRUNE_e4 (E4)	Fv	TGTAAAACGACGGCCAAGTTTGAGGACGCTGTTGTGGAA	353
	Rv	GACAAATGCACTTCTGGTAGTCAG	

Table 2.23: 30X touchdown PCR programme

Step	Temperature (°C)	Time (min:sec)	Cycles
Initial denaturation	94	3:00	
Touchdown **Denaturation** **Annealing** **Elongation**	 94 65-50 (-1°C/cycle) 72	 0:45 0:45 1:30	 15X
PCR **Denaturation** **Annealing** **Elongation**	 94 58 72	 0:30 0:45 1:30	 30X
Final elongation	68	10:00	
Hold	10	Infinite	

2.3.2.6 Fixing zebrafish embryos

Zebrafish embryos at 5 dpf were immobilised by immersion in Tricane/MS222 (4.2ml of 4 g/L per 100ml of E3). 40 to 60 embryos were collected in a 1.5ml Eppendorf. E3/Tricane solution was completely removed, and 1 ml of Fish Fix (0.1M PO Buffer, 1M $CaCl_2$, PFA powder, and 1g sucrose heated to 70°C and stored at -20°C) was added per tube and incubated overnight at 4°C. The following day, the Fish Fix was washed off three times for 5 min each in PBS (DEPC-PBS was used if the embryos were to be used for in situ hybridisation). For storing embryos, the embryos were dehydrated and rehydrated when needed for use (section 2.3.2.7).

2.3.2.7 Storing fixed embryos – Dehydration and rehydration

For storing fixed embryos, the embryos were dehydrated into 100% MeOH via a series of 25%, 50%, and 75% MeOH:PBS (DEPC-PBS was used if the embryos were to be used for in situ hybridisation) and were incubated for 15 mins each. The embryos in 100% MeOH were stored at -20°C for future use. On the day of the experiment, embryos were brought into water via a series of incubations in 75%, 50%, and 25% of MeOH:PBS (DEPC-PBS was used if the embryos were to be used for in situ hybridisation) for 15 min each.

2.3.2.8 Whole-mount zebrafish Indirect Immunofluorescence (IF)

All solutions used for zebrafish IF are listed in **Table 2.24** below. The embryos were fixed (section 2.3.2.6) and washed in PBT three times for 5 min each. To permeabilise or crack the embryos, embryos were incubated in 0.25% Trypsin/PBT

for 45 min at room temperature. This was followed by washing the embryos with PBT three times for 5 min each and then with PBDT for three times 10 min each. The samples were blocked in blocking buffer (PBDT+2% sheep serum) for 2 hours while on a rocker at room temperature. The primary antibody, Rabbit HA-tag monoclonal antibody (C29F4) (Cell Signaling Technologies, 3724), was diluted (1:500) in blocking buffer and added to embryos for incubation overnight at 4°C. The next day, the embryos were washed in PBDT at RT six times for 15 min each. A secondary antibody, goat anti-rabbit, Alexa Flour™488 (Life Technologies), was added at a concentration of 1:500 diluted in blocking buffer and incubated at 4°C overnight. The embryos were washed six times at RT in PBDT for 15 min each. The embryos were co-stained with Hoechst (1 µg/ml, Sigma, 861405) in PBDT for 1 hour at RT followed by 3 washes in PBT for 5 min each. The embryos were taken through 25% and 50% glycerol (diluted in dH_2O) series to 75% glycerol (15 min incubation in each). The embryos were then stored at 4°C until mounting or mounted immediately (section 2.3.2.12.1).

Table 2.24: Solution used for whole mount zebrafish IF

Buffer/Solution	Reagents
Phosphate-buffered saline (PBS)	Tablets (Sigma): 1 tablet dissolved in 200ml dH_2O
PBT	0.5% Triton X-100 (Alfa Aesar, A16046) in PBS
PBDT	PBS, 1% BSA, 1% DMSO and 0.5% Triton-X100

2.3.2.9 Fluorescence RNA in situ hybridisation (FISH) for zebrafish

All solutions used for zebrafish FISH are listed in **Table 2.25** below. The zebrafish embryos at 5 dpf were fixed (section 2.3.2.6) and washed 4 times for 5 min each in 1ml of PTW at room temperature. Embryos were incubated in 1ml of 20 µg/ml Proteinase K in PTW for 45 min at RT followed by re-fixation of embryos in 1ml of Fish Fix for 20 min at room temperature. Embryos were washed 5 times for 5 min each in 1ml PTW at room temperature. Embryos were then rinsed briefly for 5 min in a 300 µl cheap hybridisation solution (cheap hyb). The embryos were blocked in a 300 µl Full hybridisation solution (full hyb) for 2-3 hours at 68°C. A measured volume for an 80

nM final concentration of the LNA probes (mentioned in section 2.2.2.9) was incubated at 80°C for 75 seconds, then snap-cooled on ice for 5 mins to denature the secondary structure. Following this, the embryos were incubated with 80 nM of denatured RNA probes in a full hybridisation solution overnight at 68°C. The next day, the probe solution was removed, and the embryos were washed for 20 min at 65°C in 1ml 50:50 cheap hyb:2X SSC followed by a wash for 20 min at 65°C in 1ml 2X SSC. The embryos were washed two times for 1 hour each with 1 ml 0.2X SSC at 65°C. All wash solutions for day two were pre-heated at 65°C before adding to the embryos. Embryos were then washed for 10 min in 1ml 50:50 0.2X SSC:DEPC-PBS at room temperature. Embryos were incubated with Hoechst (1 µg/ml) in PTW for 40 min at RT and then washed in PTW three times for 5 min each. Embryos were transferred through a glycerol series into 75% glycerol by incubating in 25%, 50%, and then 75% glycerol (all diluted in DEPC-H2O) for 15 min each. The embryos were then stored at 4°C till mounting or mounted immediately (section 2.3.2.12.1).

Table 2.25: Solution used for zebrafish FISH

Buffer/Solution	Reagents
PTW	DEPC-PBS with 0.1% Tween 20
Cheap Hybridisation solution	50% formamide, 5X SSC, 0.5mg/ml tRNA, 0.05mg/ml Heparin, 0.1% Tween 20, 9.2 mM citric acid
Full Hybridisation solution	50% formamide, 5X SSC, 0.1% Tween 20, 9.2 mM citric acid

2.3.2.10 Swimming endurance test – swim tunnel

The swimming endurance of 6-month-old adult zebrafish was measured using a swim tunnel experiment. This assessment is the aquatic equivalent of a treadmill test. The critical swimming speed, Ucrit, which measures the maximum velocity a fish can maintain for a set period was measured. The experiment was conducted as previously described (Ramesh *et al.*, 2010b). Briefly, each adult zebrafish was introduced into the water flow tube of the swim tunnel. The fish was allowed to acclimatise for about 5 min. The water flow rate was gradually increased to 1.24 L/min and the fish was exposed to this for 5 min. Subsequently, the flow rate was increased

by 1.24 L/min increments every 5 mins until the fish could no longer endure this and continue swimming. The flow rate was reduced when the fish fell into a mesh at the end of the swim tunnel tube. At this point, the time was paused and was recorded as the first exhaustion time at fatigue velocity. Each fish was allowed two opportunities to continue swimming at the highest flow rate achieved by slowly increasing the flow rate back to the fatigue velocity. When the fish could not sustain swimming, the time was recorded as the second exhaustion time. The critical swimming speed was calculated based on the formula (Plaut, 2000; JR., 1964):

$$Ucrit = Ui + \left(\frac{(Uii \; x \; Ti)}{Tii} \right)$$

Where,

Ui=the highest velocity maintained for a whole interval (cm/sec),

Uii=the velocity increment (4.1 cm/sec),

Ti=the time elapsed at fatigue velocity (minutes),

Tii=the time interval (5 minutes)

When a strong swimming endurance deficit phenotype was observed, the fish were unable to make a quick recovery. To ensure animal welfare, the zebrafish was kept under observation after the swim tunnel experiment for ~45 min or until it was able to feed and swim freely to make full recovery. If it was not able to recover, the fish was culled following Schedule 1 methods (section 2.3.2.11).

2.3.2.11   Schedule 1 method for zebrafish

To ensure the welfare of the animals, zebrafish that were suffering were culled using appropriate methods of humane killing (Schedule 1). Zebrafish were culled by an overdose of anaesthetic (Tricaine/MS222) administered through immersion in water. The stock concentration of Tricaine used was 4 g/L (pH7.4) and was used at a working concentration of 1.33 g/L. Confirmation of death was carried out by confirming the onset of rigor mortis. The culls were logged and reported to the appropriate member of the aquarium staff.

2.3.2.12   Microscopy

2.3.2.12.1   Mounting zebrafish embryos for confocal imaging

Embryos in 75% glycerol were mounted onto slides using a dissecting microscope. The embryos were either mounted whole-mount in 75% glycerol by creating a stage using electrical tape with a window in the middle or the embryo heads and tails were separated, with the heads being used for genotyping (section 2.3.2.5) and the tails mounted in 75% glycerol for imaging. The embryos or tails were imaged on the SP5 confocal microscope system (Leica).

2.3.2.12.2   Confocal imaging for whole-mount zebrafish

Whole-mount zebrafish were imaged on the SP5 Confocal microscope system (Leica) with a 10x dry objective lens and analysed on ImageJ (NIH). The tail and trunk of the fish were imaged separately, and the images were taken as a z-stack with a step size of 3 μm, 1x zoom, frame average of 6, and line average of 1. Five individual fish were imaged per genotype.

2.3.2.12.3   Confocal imaging and analysis of Zebrafish for RNA foci

The zebrafish tails were used for imaging and heads were used for genotyping (section 2.3.2.5). Zebrafish tails were imaged on the SP5 Confocal microscope system (Leica) with a 63x 1.4 oil immersion objective lens. Five fields per zebrafish were taken along the tail as z-stacks with a step size of 1.51 μm, 2x zoom, frame average of 8, and line average of 1. Five individual fish were imaged per genotype. Images were analysed on ImageJ (NIH) while blinded. The nuclear threshold and total nuclear volume per field were quantified using an ImageJ macro made by Dr. David Burrows (Bankhead, 2014). The RNA foci in each field were counted manually while blinded through the z-stack per field. Bright, stained spots that were rounded and nuclear, and were completely absent in the control samples, were perceived as RNA foci. The total and the average number of RNA foci per field were normalised to the nuclear volume. An average of normalised foci from the five fields was taken and the total and the average number of foci per fish were graphed using GraphPad Prism 9.3.1.

2.4 Statistical analysis

All data in this thesis are presented as mean ± standard error of the mean (SEM) with 3 experimental repeats (unless otherwise stated). Data were analysed using GraphPad Prism 9.3.1 software (GraphPad Software, Inc., La Jolla, CA). Where needed, the distribution of the data was analysed with the Kolmogorov-Smirnov normality test for further statistical analysis. Appropriate statistical tests used are highlighted when used (stated in figure legends). The null hypothesis was rejected at a p-value less than 0.05 (*i.e.* $p < 0.05$ was considered to be statistically significant). The actual power and sample size (n) calculations were performed using the G*Power software for each experimental paradigm: comparing the means from the different sample groups to identify the sample size that corresponds to a power of 0.9 with $\alpha = 0.05$.

3. CONSTRUCTION AND VALIDATION OF PLASMIDS CONSTRUCTS AND VALIDATION OF THE *IN VITRO* MODEL

This chapter focuses on the construction and validation of plasmids used in the project for *in vitro* and *in vivo* experiments, as well as the validation of the *in vitro* model (iPSC-derived motor neurons). Plasmids made for *in vitro* experiments were used in the production of lentiviral vectors for efficient delivery of transgene into iPSC-derived motor neurons (MNs). Plasmids made for *in vivo* experiments were used to inject zebrafish at the one-cell stage to develop HC-ARFIP2 expressing stable transgenic zebrafish lines.

3.1 Introduction

The Arfaptin-2, ADP Ribosylation Factor Interacting Protein 2, (*ARFIP2*) protein is ubiquitously expressed, with a high expression seen in several tissues including the brain and skeletal muscles (Bastian *et al.*, 2021). The precise function of ARFIP2 is currently unclear, nevertheless, its structure and protein-protein interactions indicate possible functions (Mohammedeid, 2015). Domains in its structure indicate that it may be involved in membrane binding, sensing membrane curvature, and inducing membrane tubulation (Peter *et al.*, 2004; Ambroggio *et al.*, 2013). It may also play a vital role in the biogenesis of secretory storage granules (Cruz-Garcia *et al.*, 2013). Studies into ARFIP2 in neurodegenerative disorders have added further to its proposed functions. It has been shown to be involved in regulating protein aggregation and exerting neuroprotective effects, specifically through its C-terminus (Peters *et al.*, 2002; Rangone *et al.*, 2005). This highlighted the therapeutic potential of ARFIP2 in neurodegenerative disorders. To assess the effects of targeting ARFIP2 and expressing its C-terminus in models of ALS used later in the project, we developed and validated different plasmids for efficient delivery and expression in the models of ALS (*in vitro* and *in vivo*).

Suitable vectors for the delivery of transgenes to target cells or organisms need to be considered. Transfection is a powerful analytical tool to transfer various types of nucleic acid into cells. It is a widely used technique in cell culture to enable the study of gene functions in cells. Various transfection methods utilising different approaches have been developed over the years and the use of each method depends on the cell type, the type of nucleic acid, and experimental requirements (Boyer *et al.*, 2015; Borawski *et al.*, 2007; Mirska *et al.*, 2005; Sork *et al.*, 2016; Shi *et al.*, 2018; Fus-Kujawa *et al.*, 2021b). The transfection methods have been broadly classified into three groups – biological, chemical, and physical (Fus-Kujawa *et al.*, 2021a). A common biological method is virus-mediated transfection, also known as transduction. This is most appropriate for targeting post-mitotic cells and is widely researched as a potential tool for gene therapy. Chemical transfection methods involve the use of polycations, like polyethylenimine (PEI), to accelerate the transfer of DNA across the membrane. Microinjections and electroporation are examples of physical transfection methods (Kim and Eberwine, 2010). It is worth noting that not all of these techniques are applicable to all types of cells or all applications. Therefore, the ideal method is

determined based on the cell type and experimental requirements, and it should have a high transfection efficiency, low cell toxicity, minimum impacts on normal physiology, and be simple to use and repeat.

iPSC-derived motor neurons (MNs) were used as the *in vitro* model for ALS. Efficient transfection of post-mitotic cells, like differentiated neurons, remains a challenge in research owing to the neurons' exceptional sensitivity to physical stresses, temperature fluctuations, pH shifts, and osmolarity variations. Furthermore, chemical transfection methods (using PEI or lipofectamine) take advantage of dividing cells to introduce nucleic acids into cells and hence are not suitable for post-mitotic, non-dividing differentiated neurons. Previously, many transfection strategies for neuronal cells have been explored (Karra and Dahm, 2010), however, viral vectors have gained attention and have been shown to be effective agents for gene transfer for neuronal models.

In *in vitro*, viral vectors like lentiviruses (LV) are utilised to produce sustained genomic integration and inducible transgenic expression, and effectively infect post-mitotic mature neurons. One of the most significant benefits of utilising viral-based vectors is their low toxicity since they are inherently able to enter cells without generating physical stress. LV vectors, derived from the human immunodeficiency virus (HIV), are now commonly used as a gene therapy tool because they can infect both dividing and non-dividing cells. In August 2017, the first LV-based gene therapy was authorised in the United States by the Food and Drug Administration (Kymriah) (Sadelain, 2017) and several others are now in late-phase clinical trials. Though the theoretical possibility of insertional oncogenesis is associated with LV vectors, no cases of this potential complication have been recorded with the use of LV in gene therapy. With its extensive use in research, LV vectors, carrying the LV-based plasmids expressing the genes of interest, were made in this project for efficient transduction of iPSC-derived MNs to express the transgenes.

The microinjection technique for gene transfer is widely employed in scientific and therapeutic settings, such as the production of transgenic animals (Fus-Kujawa *et al.*, 2021a). The zebrafish model is a quick and inexpensive model and an attractive organism for simulating neurological disorders because of its conserved and simple

nervous system that is comparable to humans. It also has the capacity to generate transgenic animals (Lieschke and Currie, 2007). Previously, effective and successful zebrafish transgenic models for ALS carrying the SOD1 and C9ORF72 mutations have been generated that are able to model key aspects of the pathology (Shaw *et al.*, 2018; Ramesh *et al.*, 2010b). In zebrafish, microinjecting transgenic DNA constructs into the one-cell stage embryo is the conventional overexpression approach which results in ubiquitous overexpression, and the concentration of microinjections can be adjusted until a specific overexpression phenotype is observed with minimal toxicity. Therefore, in this project, plasmid constructs expressing the genes of interest were made for microinjecting embryos to generate transgenic zebrafish models to cross with the established C9ORF72-ALS model (Shaw *et al.*, 2018) and assess the effects of HC-ARFIP2 on ALS phenotypes.

A previous study looking into ARFIP2 in the neurodegenerative disorder Huntington's disease (HD) showed that ARFIP2 was widely expressed in healthy mice brains. However, ARFIP2 expression was elevated in the striatum, cortex, and cerebellum of the HD mouse model (Peters *et al.*, 2002). With its increase in expression levels in a neurodegenerative background, a therapeutic strategy would be to assess the effects of depleting the elevated endogenous expression levels of the protein on disease phenotypes. RNA interference (RNAi) is a potent tool for the manipulation of gene expression. RNAi is a phenomenon in which RNA fragments, like microRNAs (miRNAs) and short, interfering RNAs (siRNAs), can interrupt the translation of proteins by attaching to the messenger RNAs (mRNAs) that encode proteins, thereby depleting their levels (Kim and Rossi, 2008). RNAi can be triggered via a DNA-based approach in which siRNAs are produced by processing longer hairpin transcripts like short-hairpin RNAs (shRNAs). RNAi or depletion of protein can be achieved by promoter-based expression of specific shRNAs and their downstream processing (Scherer and Rossi, 2003; Hannon and Rossi, 2004). Therefore, to assess the effects of depleting the potentially elevated levels of ARFIP2 in ALS, shRNA-expressing plasmids were made in this project.

3.2 Aim

This chapter aims to describe the construction and validation of plasmids that were used for *in vitro* and *in vivo* experiments of the project. Firstly, it shows the

validation of the LV-based plasmids and the subsequent production of lentiviral vectors to efficiently transduce iPSC-derived MNs for delivery of HC-ARFIP2. The expression constructs were first validated in the simpler HEK293T cell model before validation in the more complex human-relevant iPSC model. Secondly, the successful construction of the plasmid expressing HC-ARFIP2 to generate stable transgenic zebrafish lines is shown. Finally, the chapter also describes the construction and preliminary validation of ARFIP2-shRNA plasmids to deplete the expression levels of endogenous ARFIP2 *in vitro*. Furthermore, this chapter describes the validation of the *in vitro* model that was used in chapter 4.

3.3 Results

3.3.1 Lentiviral-based plasmids successfully express either FL-ARFIP2, HC-ARFIP2, or GFP

3.3.1.1 Validation of pLV-FL and pLV-HC

The Sin-PGK-cPPT Lentiviral (LV) based plasmids expressing either FL-ARFIP2 (pLV-FL) or HC-ARFIP2 (pLV-HC) were made by Dr. Aida Mohammedeid during her Ph.D. project (Mohammedeid, 2015). Prior to using these plasmids for further experiments, they were verified by restriction enzyme digest and sequencing to ensure that the plasmids successfully carried the desired inserts.

The LV plasmids were transformed into DH5α competent cells (2.1.2.3) for maximised transformation efficiency. Individual colonies from the transformed bacterial cultures were picked and the plasmids were isolated using plasmid preparation kits (miniprep and midiprep). The LV plasmids were validated using BamHI-HF and XhoI restriction enzymes (2.1.2.6) and visualised on a 1% agarose gel (2.1.2.11) (**Figure 3.1 A**). A successful restriction digest of pLV-FL and pLV-HC showed that both plasmids carry the genes of interest (FL-*ARFIP2* at 1059 bp and HC-*ARFIP2* at 552 bp) with the LV backbone intact (LV backbone at ~10kbp). Following validation of the plasmids to confirm the presence of inserts, the plasmids were further investigated by DNA sequencing (2.1.2.15) to confirm that no spontaneous mutations occurred during plasmid transformation and isolation. The pLV-HC and pLV-FL sequences were used to generate the representative plasmid map (**Figures 3.1 B and C**). Sequences of pLV-FL and pLV-HC can be found in **Appendix 1**.

B

Sin-PGK-cPPT-HC-ARFIP2 (pLV-HC)
9360 bp

C

Sin-PGK-cPPT-HC-ARFIP2 (pLV-FL)
9867 bp
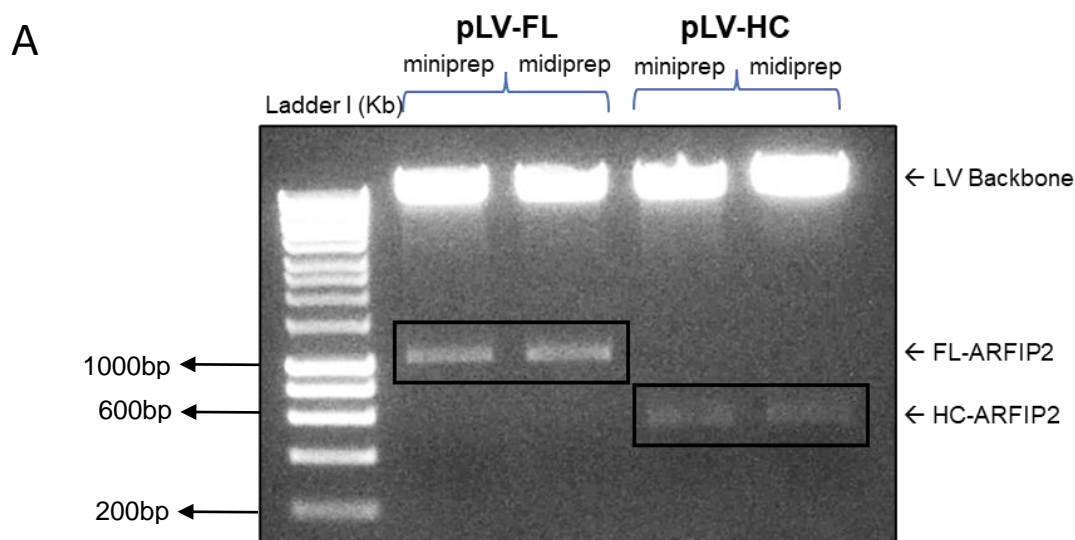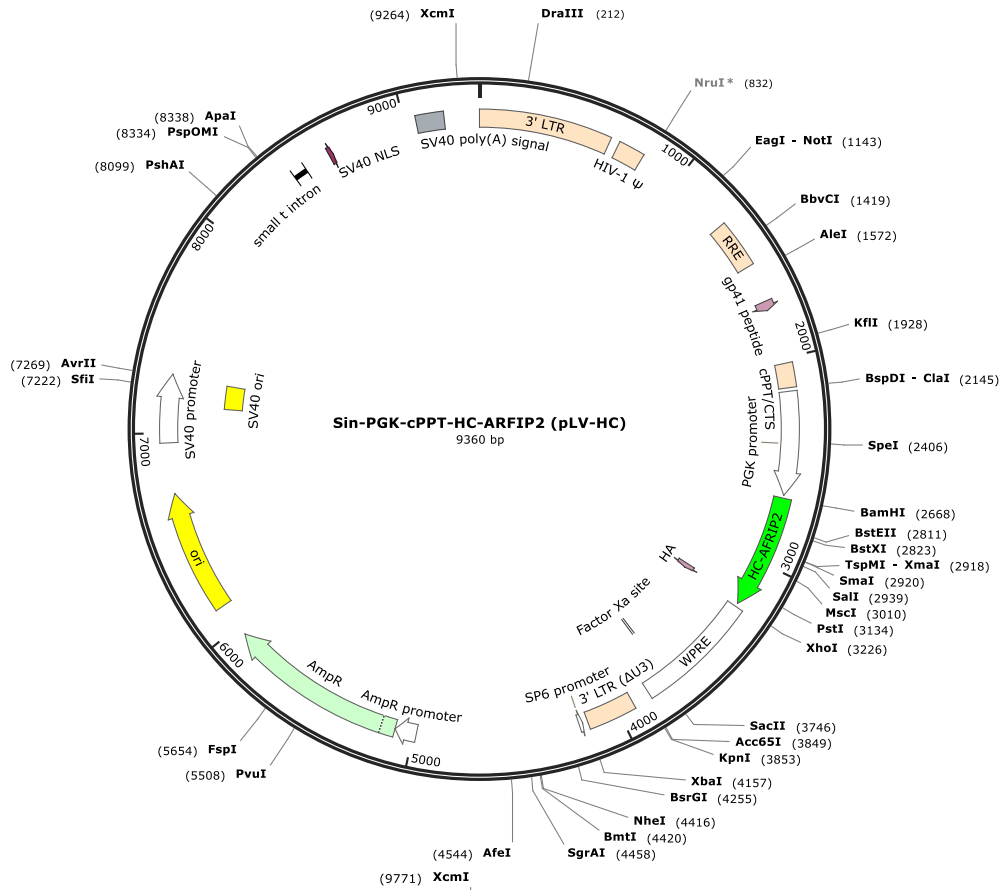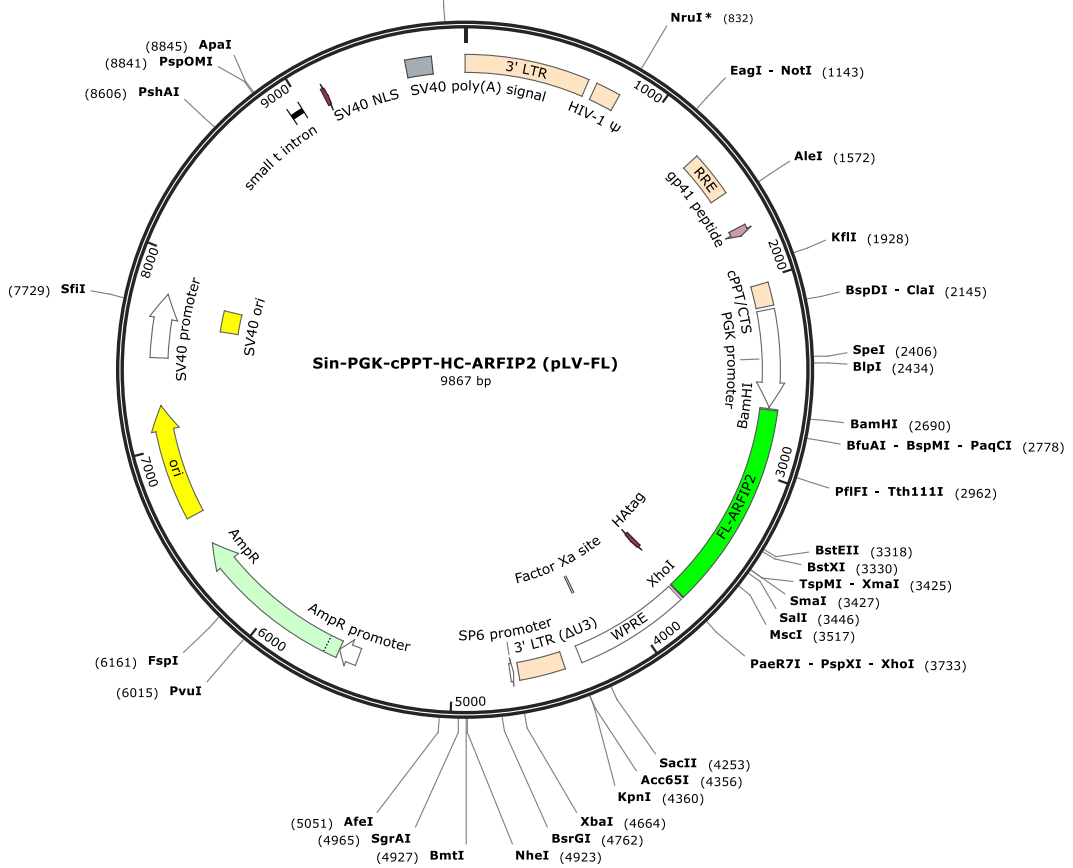
Figure 3.1: Digestion of pLV-FL and pLV-HC to validate the plasmids and pLV-HC plasmid map

(A) Successful digestion of the plasmids by BamHI-HF and XhoI. Plasmids were transformed and isolated from both miniprep and midiprep. The LV-backbone size is ~ 10 kbp. The FL-ARFIP2 insert is 1 kb and the HC-ARFIp2 is ~0.6 kbp. **(B-C)** Detailed plasmid map for pLV-HC and pLV-FL with unique restriction sites. The maps were generated using the SnapGene software.

Primary validation of LV-based plasmids was done using HEK293T cells, a human line, before moving onto human-relevant neuronal cell lines (iPSC-derived MNs). pLV-HC was validated in HEK293T cells to assess whether the plasmid is able to express the HC-ARFIP2 protein successfully and efficiently prior to lentiviral production. HEK293T cells were transfected with pLV-HC using the PEI transfection method (2.2.2.4) and the total protein was extracted 3 days post-transfection (2.1.2.17). Protein concentrations were determined (2.1.2.19), and the HC-ARFIP2 expression levels were detected on a western blot (2.1.2.20), using α-Tubulin as a loading control (**Figure 3.2**). The presence of the HC-ARFIP2 bands at 19kDa in transfected cell samples, compared to untransfected cell samples, shows the successful expression of the protein. The HC-ARFIP2 protein was detected using the anti-HA antibody. Results show that pLV-HC successfully expresses HC-ARFIP2 in HEK293T cells.
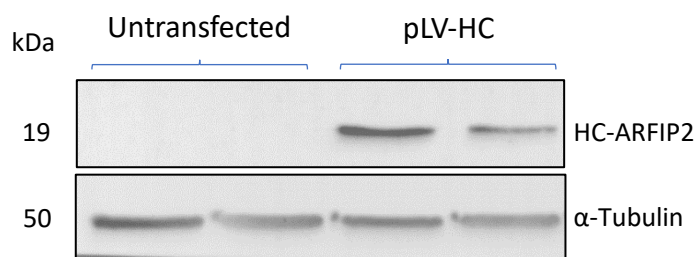
Figure 3.2: Validation of LV-based plasmids expressing HC-ARFIP2 in HEK293T cells

HEK293T cells were either left untransfected or transfected with pLV-HC. The expression of HC-ARFIP2 was detected on a western blot probed with anti-HA antibody (1:000). HC-ARFIP2 was detected at 19kDa. α-Tubulin at 50kDa was used as a loading control.

3.3.1.2 Validation of pLV-GFP

The Sin-PGK-cPPT LV-based plasmid expressing *GFP* (pLV-*GFP*) was provided by Professor Mimoun Azzouz's lab. The plasmid was validated by DNA

sequencing to ensure that the plasmid backbone was identical and comparable to the other LV plasmids and successfully carried the *GFP* reporter gene. The pLV-*GFP* sequence was used to generate the pLV-*GFP* representative plasmid map (**Figure 3.3 A**). The sequence of pLV-*GFP* can be found in **Appendix 1**.

The pLV-*GFP* plasmid was first validated in HEK293T cells to assess whether the plasmid is able to express the GFP reporter protein successfully and efficiently prior to lentiviral production. HEK293T cells were transfected with pLV-*GFP* using the PEI transfection method (2.2.2.4) and harvested 3 days post-transfection (2.1.2.17). GFP expression levels were detected on a western blot (2.1.2.20), using α-Tubulin as a loading control (**Figure 3.3 B**). The presence of the GFP bands at 27 kDa in transfected cell samples, compared to the untransfected cell sample, shows the successful expression of the GFP protein. The GFP protein was detected using the anti-GFP antibody. Results show that pLV-GFP successfully expresses GFP in HEK293T cells.

B

UT   pLV-GFP

kDa

27   GFP

50   α-Tubulin
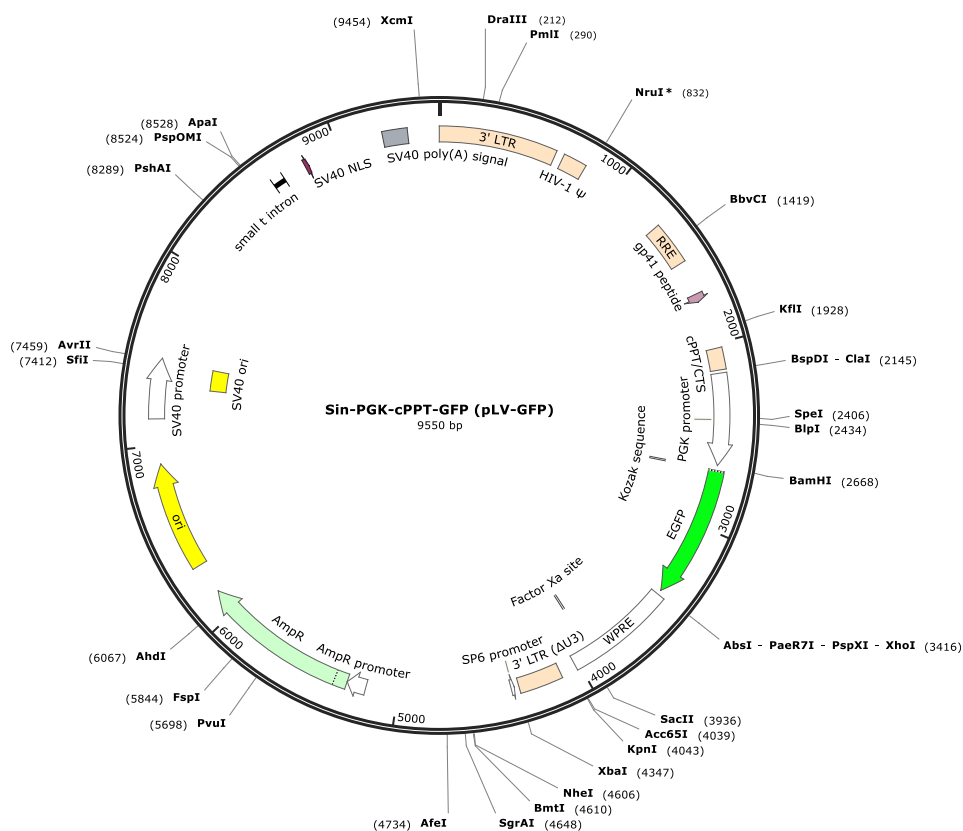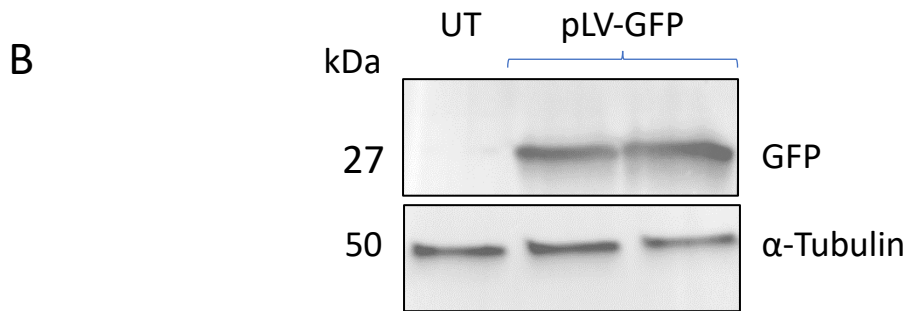
Figure 3.3: pLV-GFP plasmid map and validation in HEK293T cells
(A) Detailed plasmid map for pLV-GFP with the BamHI and XhoI restriction sites. The map was generated using the SnapGene software. **(B)** HEK293T cells were either left untransfected or transfected with pLV-GFP. The expression of GFP was detected on a western blot probed with anti-GFP antibody (1:500). GFP was detected at ~27kDa. α-Tubulin at 50kDa was used as a loading control. UT=Untransfected.

3.3.2 Human iPSCs successfully differentiate into functional mature motor neurons

iPSC-derived MNs were used as an *in vitro* model for this Ph.D. project. Human iPSCs were differentiated into MNs using an established differentiation protocol described in section 2.2.2.2. Following the protocol, human iPSCs went through neural induction (where an even neuroepithelial sheet forms) and caudalization till day 12 of the protocol, at which stage they formed neuron progenitor cells (NPCs). On day 12, the NPCs were fixed in 4% PFA (2.2.2.7) and validated at this stage by staining for NPC markers, Pax6 and Nestin (2.2.2.8) (**Figure 3.4 A**). The NPCs showed expression of the transient nuclear-paired box 6 (*Pax6*) transcription factor and cytoplasmic Nestin. At this stage, the cells no longer display the rounded iPSC morphology and instead exhibit the formation of neural rosettes. From day 12, with neuronal induction and maturation, the NPCs were differentiated into functional MNs (day 28). At the end of the protocol, these iPSC-derived MNs were fixed and validated by staining for MN markers, ChAT, MAP2 and Islet-1 (**Figure 3.4 B**). iPSC-derived MNs showed expression of choline acetyltransferase (*ChAT*) and microtubule-associated protein 2 (*MAP2*) in the cell bodies of the neurons and Islet-1 expression in the nucleus. At this stage, the mature MNs displayed long axonal projections from the cell bodies. At each stage, the nucleus was stained using Hoechst. Expression of specific markers at different stages of the differentiation protocol showed that the iPSCs were successfully differentiated into mature MNs and validated our *in vitro* model.
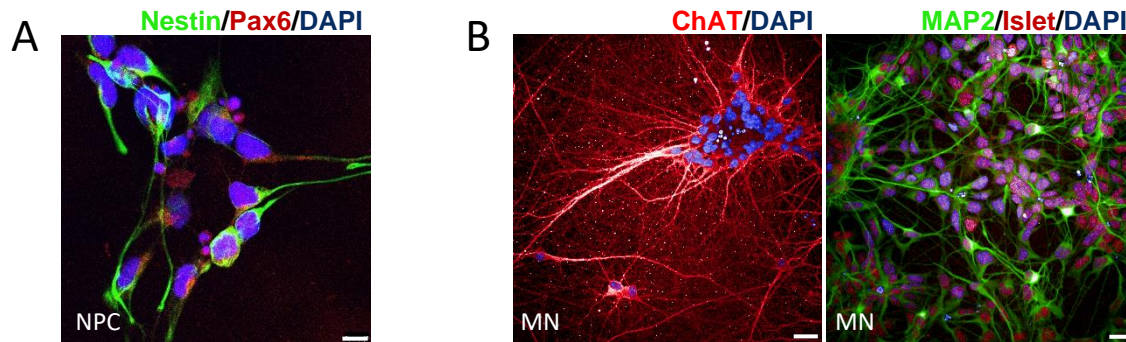
Figure 3.4: Validation of the in vitro model used in the project
(A) On day 12 of the differentiation protocol, the cells were fixed in 4% PFA and immunostained for the Pax6 and Nestin NPC markers using anti-Pax6 antibody (1:200) and anti-Nestin antibody (1:200). **(B)** On day 28 of the differentiation protocol, the cells were fixed in 4% PFA and immunostained for the ChAT, MAP2 and Islet MN markers using anti-ChAT antibody (1:100), anti-MAP2 antibody (1:1000) and anti-Islet-1 antibody (1:250). The nucleus was stained using Hoechst. Scale bars represent 10 μm. NPC= neuron progenitor cells; MN= iPSC-derived motor neurons.

3.3.3 Lentiviruses produced have high functional titres

Post-validation, pLV-FL, pLV-HC, and pLV-*GFP* were used in the production of corresponding lentiviral vectors (2.1.2.16). The lentiviral vector expressing FL-*ARFIP2* (LV-FL) was made to assess the effects of overexpressing the full-length protein on C9ORF72-ALS phenotypes. The lentiviral vector expressing HC-*ARFIP2* (LV-HC) was made for use in subsequent *in vitro* experiments to assess the effects of expressing HC-ARFIP2 on *C9ORF72*-ALS phenotypes in iPSC-derived MNs. The lentiviral vector expressing GFP (LV-*GFP*) was used as a control vector in subsequent *in vitro* experiments.

LV-FL, LV-HC, and LV-*GFP* were produced by transfecting HEK293T cells plated in 20 10 cm dishes with four plasmids using the calcium phosphate method (2.2.2.5). This method is cost-effective and widely used in large-scale LV production. Two packaging plasmids, one envelope plasmid, and the pLV carrying the gene of interest were transfected. The virus was harvested and purified 3 days post-transfection from the culture medium. The functional viral titres for each virus were determined, which ranged between $10^8 - 10^9$ transduction units per mL (TU/mL). **Table 3.1** below shows the titre achieved for each lentivirus produced.

Table 3.1: Titres of Lentiviral Vectors

Virus	Titre
LV-HC-ARFIP2	3.22×10^9 TU/mL
LV-FL-ARFIP2	1.42×10^8 TU/mL
LV-GFP	7.63×10^8 TU/mL

3.3.4 Lentiviral vectors generated express FL-ARFIP2 and HC-ARFIP2

3.3.4.1 Validation of lentiviral vectors expressing FL-ARFIP2 and HC-ARFIP2 in HEK293T cells

Prior to proceeding with the *in vitro* experiments using LV-HC, the efficiency of the virus was assessed to see if the virus is able to express the protein of interest. HEK293T cells were transduced at an increasing multiplicity of infection (MOI) of LV-FL or LV-HC (2.2.2.6.2) (**Figure 3.5**). Cells were harvested 4 days post-transduction (2.1.2.17) and the expression of the protein was determined using a western blot (2.1.2.20).

For both viruses, MOI 1, 3, and 10 were tested. These doses were chosen as MOI 5 has been reported to be the optimal LV dose for HEK293T cells (Davis, 2014). LV-FL was able to overexpress the full-length protein at all (low, medium, and high) MOIs, with the strongest expression seen at MOI 10 (**Figure 3.5 A**). The full-length protein was detected at 38.6 kDa in all transduced samples at each dose compared to the untransduced sample. LV-FL transduction at MOI 10 showed the highest expression with a 9.2-fold increase in expression when compared to MOI 1 and an 4.5-foldincrease at 3 (**Figure 3.5 B**). On the other hand, LV-HC was only able to show detectable expression of HC-ARFIP2 at the medium and high MOIs (MOI 3 and MOI 10), with the strongest expression seen at MOI 10 (**Figure 3.5 C**). LV-HC transduction at MOI 10 also showed the highest fold increase (83.1-fold) in HC-ARFIP2 expression when compared to MOI 1. At the lowest MOI 1, there was almost no difference in expression compared to untransduced cells. At MOI 3, there was a very low expression with a 14.7-fold increase in expression (**Figure 3.5 D**). HC-ARFIP2 was detected at 19 kDa in all transduced samples compared to the untransduced sample. Both the transductions with LV-FL and LV-HC showed a correlation between the increasing MOIs and the level of expression of transgene i.e., the level of expression

was dose-dependent. Results show that both viruses were able to efficiently express the desired proteins. Validation of the viruses in HEK293T cells was necessary to confirm their efficiency before proceeding with experiments involving the transduction of iPSC-derived MNs. At this stage, LV-GFP was not validated in HEK293T cells because of the limited amount of virus that was available.
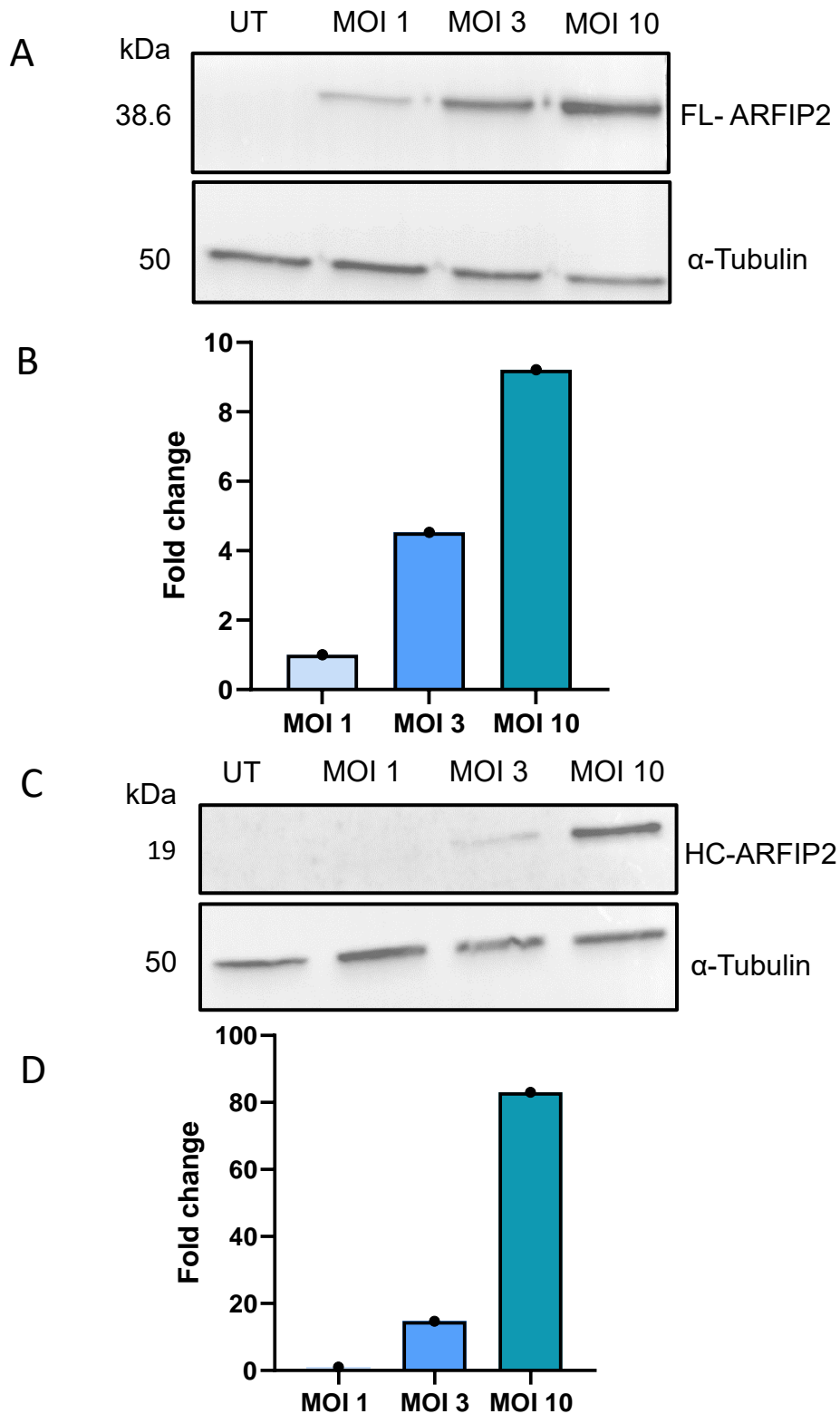
Figure 3.5: HEK293T cells transduced with LV-FL or LV-HC at increasing doses efficiently express FL-ARFIP2 or HC-ARFIP2
HEK293T cells were transduced with the lentivirus expressing **(A)** FL-ARFIP2 (LV-FL) or **(C)** HC-ARFIP2 (LV-HC). Protein was harvested 4 days post transduction. FL-ARFIP2 and HC-ARFIP2 expression was detected on a western blot using anti HA-tag antibody (1:1000). The band densities were measured using ImageJ to calculate the fold change in expression of **(B)** FL-ARFIP2 and **(D)** HC-ARFIP2 compared to the lowest dose. The blot and analysis show a dose response, i.e., with increasing MOI of virus, there is an increase in expression of transgene. FL-ARFIP2 was detected at 38.6kDa and HC-ARFIP2 was detected at 19kDa. α-Tubulin at 50kDa was used as a loading control. UT= untransduced.

3.3.4.2 LV-HC efficiently expresses HC-ARFIP2 in iPSC-derived motor neurons in a dose-dependent manner

LV-HC vector efficiency was further validated in control and *C9ORF72*-ALS iPSC-derived MNs. The MNs were transduced with LV-HC (2.2.2.6.1) at MOI 10, 20, and 40 on day 40 of the differentiation protocol. Following on from the results obtained from initial experiments in HEK293T cells, MOI 10 was chosen as the lowest dose to test due to the more complex nature of iPSC-derived MNs, followed by MOI 20 and MOI 40. Higher MOIs than MOI 40 were not tested due to the limited amount of virus. The protein was harvested (2.1.2.17), and the protein expression was detected on a western blot (2.1.2.20). Transgene expression was measured by densitometry. HC-ARFIP2 expression at 19 kDa was seen in all cell lines, the 3 controls and 3 *C9ORF72*-ALS lines (**Figure 3.6 A**). When normalised to the α-Tubulin loading control, it was observed that there was an increase in expression of HC-ARFIP2 in each cell line with increasing MOI of the virus. The HC-ARFIP2 expression at MOI 40 was the highest in all 6 cell lines compared to the lower MOIs (**Figure 3.6 B**) and there was no significant difference in the expression of HC-ARFIP2 between the control and ALS cell lines (**Figure 3.6 C**). LV-HC transduction at MOI 40 resulted in an average of a 5.1-fold significant increase in HC-ARFIP2 expression when compared to MOI 10. There was no significant fold increase in HC-ARFIP2 expression at the lower MOI of 20 (**Figure 3.6 D**). The fold increase in expression of HC-ARFIP2 with increasing MOI was in a dose-dependent manner (i.e. linear) (**Figure 3.6 E**). Overall, results show that LV-HC can successfully express HC-ARFIP2 at low and high MOIs in iPSC-derived MNs in a

dose-dependent manner, with the highest expression of HC-ARFIP2 at MOI 40 in all cell lines.

A
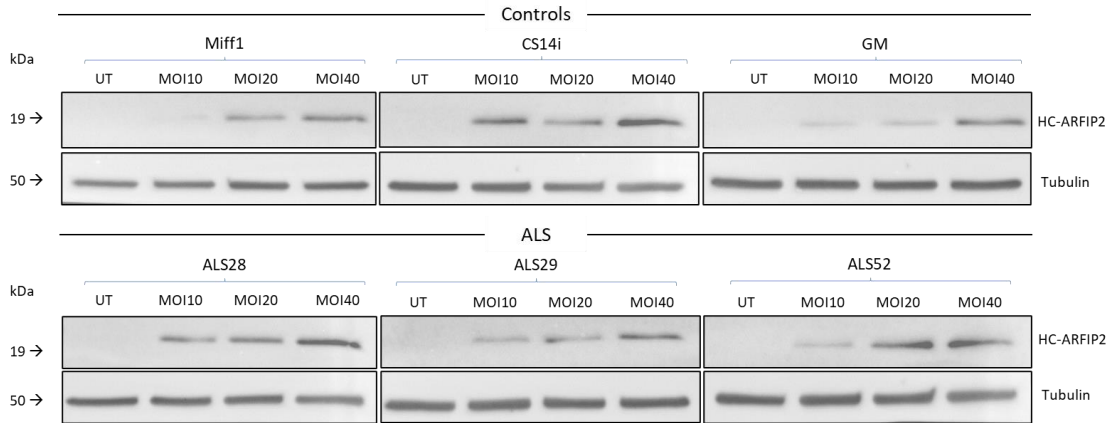
B
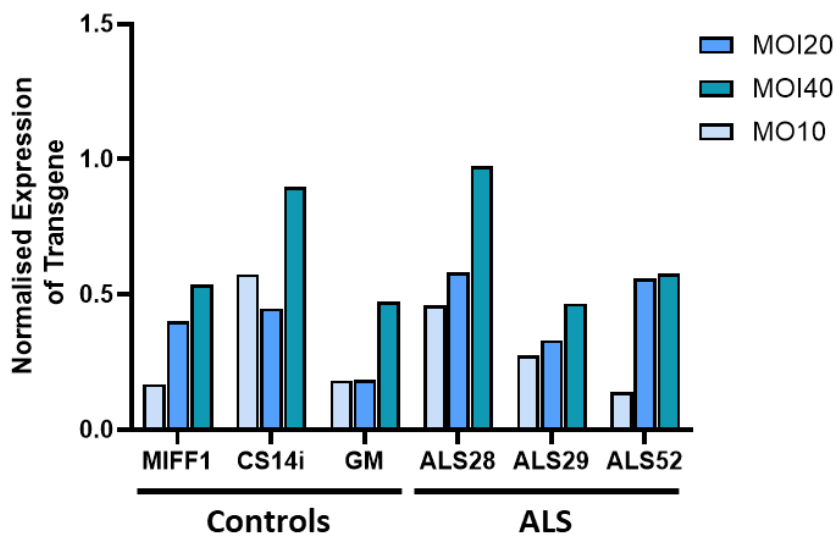
C
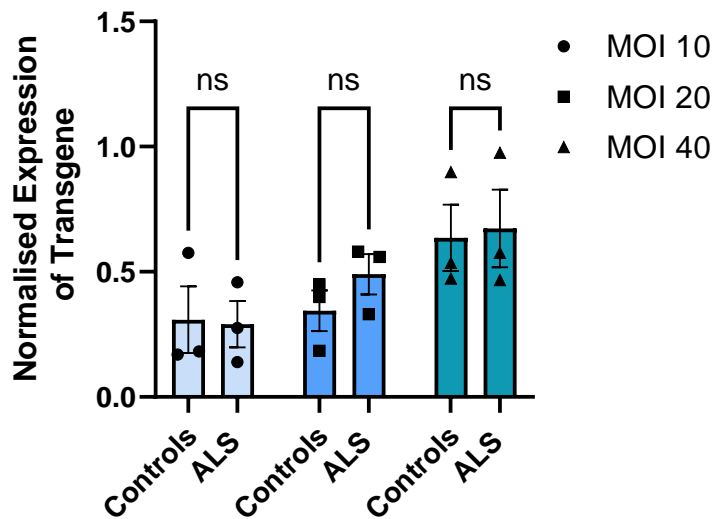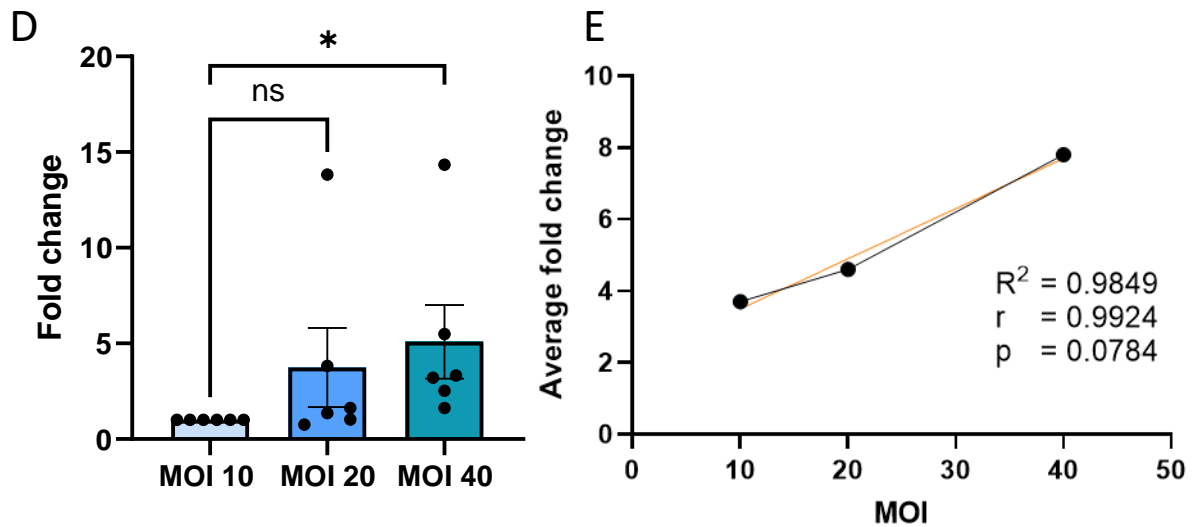
Figure 3.6: LV-HC transduces iPSC-derived motor neurons to efficiently express HC-ARFIP2

Healthy control and ALS iPSC-derived motor neurons (1x10^6 cells/well, n=3) were transduced with the lentivirus expressing HC-ARFIP2 (LV-HC). Protein was harvested 4 days post transduction. **(A)** HC-ARFIP2 expression was detected on a western blot using anti HA-tag antibody (1:1000). HC-ARFIP2 was detected at 19kDa. α-Tubulin at 50kDa was used as a loading control. **(B)** The expression level (mean ± SEM) was determined by densitometry of western blot using the GeneTools software. Transgene expression was normalised to α-Tubulin expression. With increasing MOI of virus, there was an increase in expression of HC-ARFIP2 in each cell line and **(C)** there is no significant difference in expression levels of HC-ARFIP2 between controls and ALS cell lines. Data were analysed using an ordinary two-way ANOVA with a Šídák's multiple comparisons test. **(D)** Compared to MOI10, there was a significant 5.1-fold increase in the expression of HC-ARFIP2 at MOI 40. There was no significant fold increase in HC-ARFIP2 expression at MOI 20 (MOI 20= 3.7-fold). The increase in expression of HC-ARFIP2 with increasing MOI was dose-dependent (i.e. linear). Data were analysed using the Wilcoxon Signed Rank Test. **(E)** The average fold change was plotted against the MOI of virus. There was a strong correlation observed. Simple linear regression and correlation analysis showed R^2=0.9849, r=0.9924 and p=0.0784. Error bars represent ± SEM. UT = untransduced.

3.3.5 LV-HC and LV-GFP efficiently transduce iPSC-derived MNs

3.3.5.1 Transduction efficiency of LV-HC

In order to select the optimal dose of LV-HC for subsequent *in vitro* experiments, the transduction efficiency (i.e., the percentage of infected cells) of LV-HC in iPSC-derived MNs was assessed. iPSC-derived MNs (n=6) were either left untransduced or transduced at MOI 10, 20, and 40 (2.2.2.6.1) on day 40 of the differentiation protocol (2.2.2.2). The cells were fixed in 4% PFA 4 days post-transduction (2.2.2.7). HC-ARFIP2 was stained using the anti-HAtag antibody and the nucleus was stained using Hoechst (2.2.2.8). Neuron-specific class III beta-tubulin was used as a neuron marker and was stained using the anti-beta-tubulin III (TUJ1) antibody. The cells were imaged using the Opera Phoenix High-Content imaging system (2.2.2.14.2) (**Figure 3.7 A**). Using the untransduced cells for each line to set a threshold for background, analysis of the HC-ARFIP2 stained cells showed that the average percentage of cells transduced (transduction efficiency) by LV-HC was 9.5% at MOI 10, 17.8% at MOI 20, and 30.6% at MOI 40 (**Figure 3.7 B**). The transduction efficiency of LV-HC was most significantly increased at MOI 40 when compared to untransduced cells. It was observed that at each dose, there was no significant difference in the transduction efficiency of LV-HC between the control and ALS cell lines (**Figure 3.7 C**). However, trends showed that at MOI 20 and 40, the ALS cell lines have a slightly higher transduction efficiency compared to controls. Further analysis showed that there was a strong and significant correlation between the average transduction efficiency and dose of LV-HC (r= 0.9977, p= 0.04), which highlighted that the transduction efficiency of the virus was dose-dependent (**Figure 3.7 D**). **Figure 3.7 E** shows the average number of cells/well for each cell line at each dose that was taken into account to measure the transduction efficiency of LV-HC at a single dose. Though there is no significant difference between the average number of cells/well at each condition, there is a trend to observe where, with an increase in the dose of LV-HC, there is an increase in the average number of cells/well. The intensity of the staining for HC-ARFIP2 was measured for each cell line to compare the intensity of staining at each dose (**Figure 3.7 F**). As expected, the average intensity of staining was higher at MOI 20 and 40 (303.5 and 303.1, respectively) and slightly lower at MOI 10 (264.2). However, there was no significant difference in the average intensity of HC-ARFIP2 staining between each dose of LV-HC. Overall, the results show that LV-HC can efficiently transduce iPSC-derived MNs in a dose-

dependent manner and that transduction at MOI 40 showed the highest transduction efficiency while the cells still looked healthy at 4 days post-transduction. Higher MOIs than MOI 40 were not tested because of the amount of virus that would be needed.

F
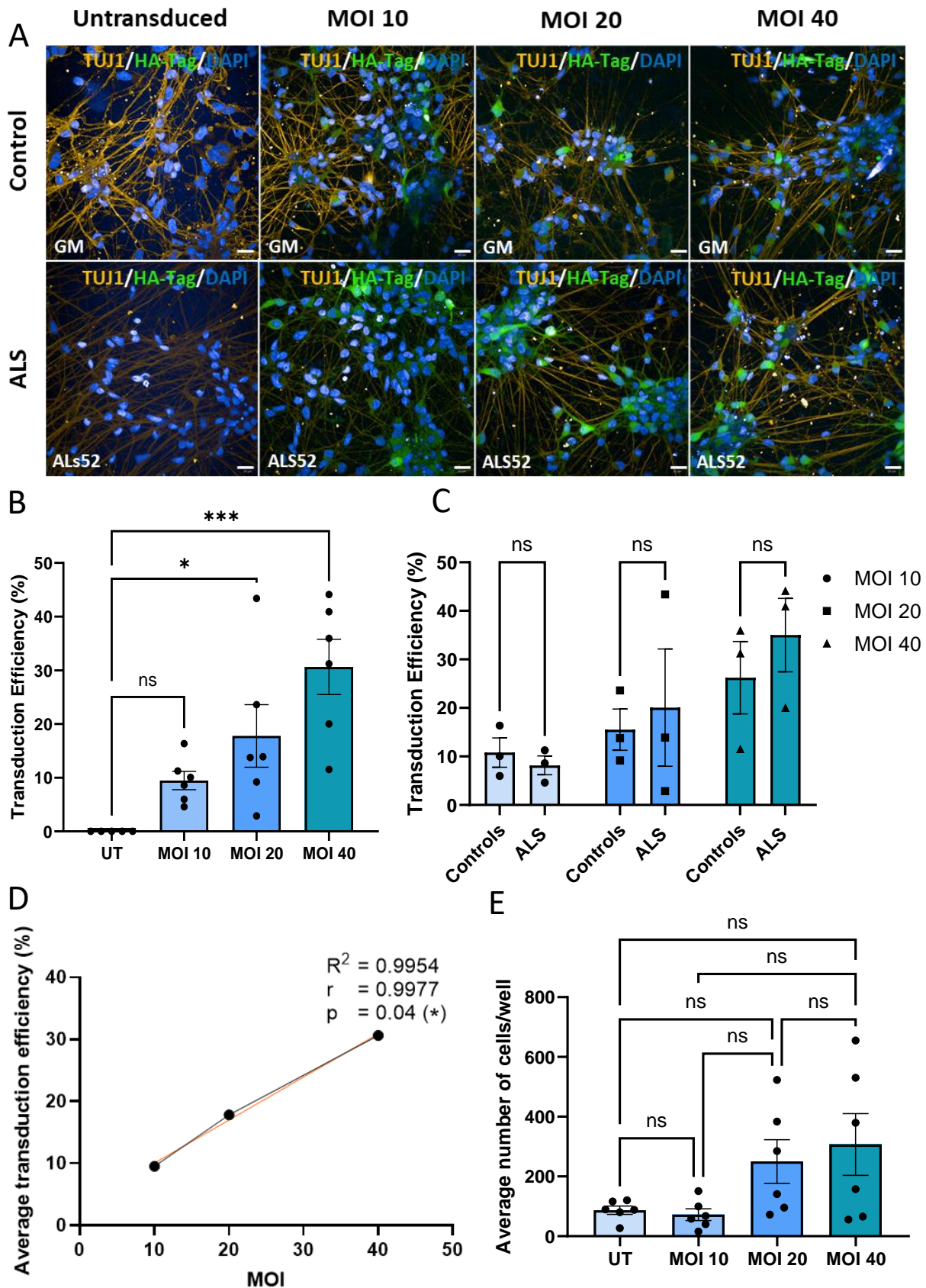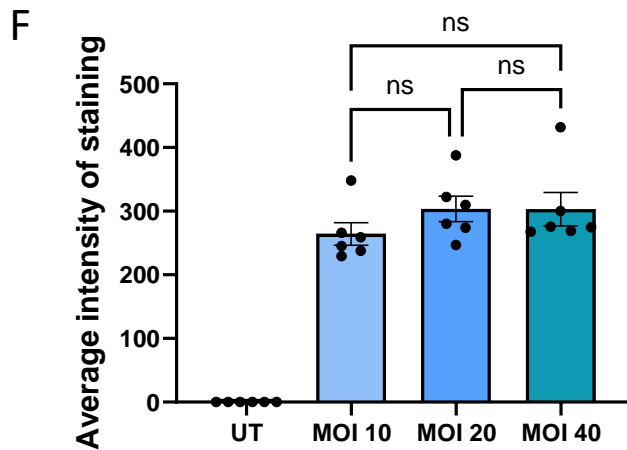
Figure 3.7: Transduction efficiency of LV-HC in iPSC-derived motor neurons with increasing MOI

Healthy controls (n=3) and ALS iPSC-derived motor neurons (n=3) at a density of 0.02 x10^6 cells/well were transduced with the lentivirus expressing HC-ARFIP2 (LV-HC) at increasing MOI of 10, 20 and 40 (1 experimental repeat). **(A)** The cells were fixed in 4% PFA 4 days post transduction and immunostained for HC-ARFIP2 using anti-HA-tag antibody (1:500) and for TUJ1 (1:200). Scale bars represent 20μm. **(B)** Transduction efficiency (n=6, mean ± SEM) was calculated at each MOI with comparison between transduced and untransduced. Transduction at MOI 40 showed the highest average transduction efficiency (30.62%) while the cells looked healthy 4 days post transduction. **(C)** When comparing the transduction efficiencies between healthy and disease cell lines, it was seen that there is no significant difference in the transduction efficiency of LV-HC between controls and ALS cell lines. Data were analysed using an ordinary two-way ANOVA with a Šídák's multiple comparisons test. **(D)** The average transduction efficiency was plotted against the MOI of virus. There was a strong significant correlation observed. Simple linear regression and correlation analysis showed R^2=0.9954, r=0.9977 and p=0.04. **(E)** The number of cells/well for each cell line under each condition was taken into account to obtain the average number of cells counted for each condition (UT n=87, MOI 10 n=72, MOI 20 n=250 and MOI 40 n=307.2). **(F)** The average intensity of staining for HC-ARFIP2 using the anti-HA-tag antibody was measured for each cell line at each dose. There were no significant differences in the intensity of staining for HC-ARFIP2 between each dose of LV-HC (MOI 10=264.2, MOI 20=303.5 and MOI 40=303.1). Data were analysed by an ordinary one-way ANOVA with a Šídák's multiple comparisons test. Error bars represent ± SEM. UT= untransduced.

3.3.5.2 Transduction efficiency of LV-GFP

As above, the transduction efficiency for LV-*GFP* at MOI 10, 20, and 40 was assessed to select the optimal dose for subsequent *in vitro* experiments (2.2.2.6.1).

The *C9ORF72*-ALS lines (n=3) were transduced. The nucleus was stained using Hoechst (2.2.2.8). The cells were imaged for GFP using the Opera Phoenix High-Content imaging system (2.2.2.14.2) (**Figure 3.8 A**). Using the untransduced cells to set the threshold for background, analysis of the cells expressing GFP showed that the average transduction efficiency for LV-*GFP* was 53.7% at MOI 10, 79.3% at MOI 20, and 27.1% at MOI 40 (**Figure 3.8 B**). The transduction efficiency of LV-*GFP* was significantly increased at all doses of LV-*GFP* (MOI 10, 20, and 40) when compared to untransduced cells, with the highest transduction efficiency at MOI 20. However, at the highest dose of MOI 40, the transduction efficiency drastically decreased and had the lowest significance when compared to untransduced cells. Correlation analysis showed that there was a weak, negative and non-significant correlation between the average transduction efficiency and dose of LV-*GFP* (r= -0.663 and p= 0.5386) (**Figure 3.8 C**). It was observed that at MOI 10 and 20, with an increasing dose of LV-GFP, the transduction efficiency increased. However, at the highest dose, the transduction efficiency was the lowest. The number of cells/well for each cell line at each dose that were used to analyse the transduction efficiency was measured (**Figure 3.8 D**). There was no significant difference between the average number of cells/well at each condition. However, there is a trend where with an increase in the dose of LV-*GFP*, there is a decrease in the average number of cells/well. The number of cells at MOI 40 (n= 90) were the lowest compared to all other conditions. The intensity of GFP expression was analysed at each dose (**Figure 3.8 E**). The average intensity of GFP increased with an increasing dose of LV-*GFP*. There was no significant difference in intensities between MOI 10 and 20 (1573.8 and 1974.5, respectively). The average intensity of GFP was highest at MOI 40 (2693.4), and this was significantly higher than the average intensities at both lower MOIs. Results show that LV-GFP can efficiently transduce iPSC-derived MNs and that transduction at MOI 20 showed the highest transduction efficiency while the cells still looked healthy at 4 days post-transduction. At the highest dose of MOI 40, the number of cells in the wells substantially decreased which indicated cell loss/death and toxicity, and therefore, higher MOIs than MOI 40 were not tested.
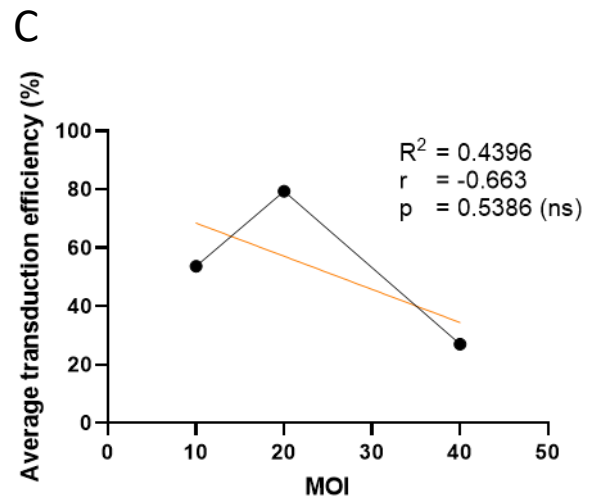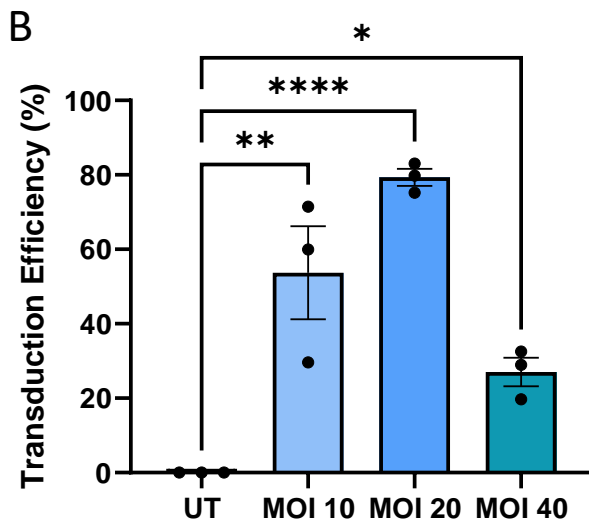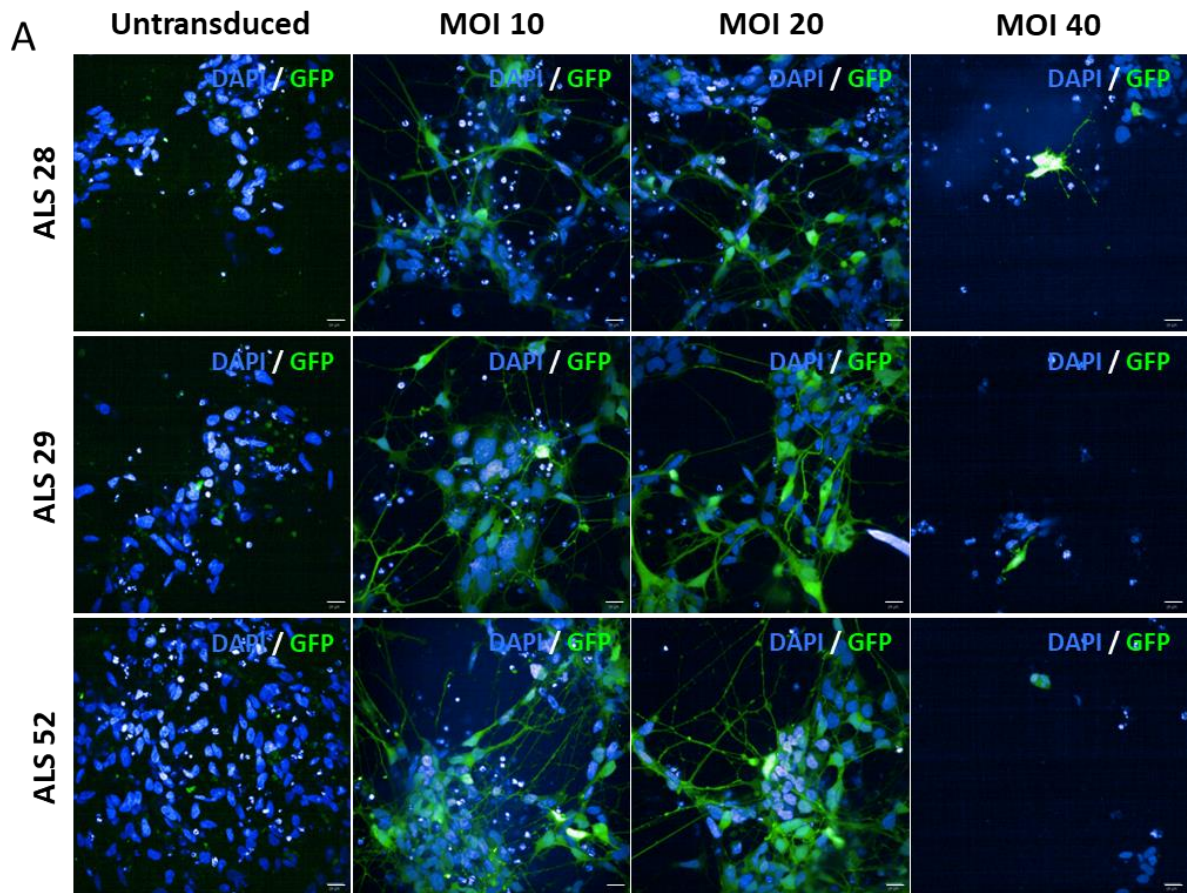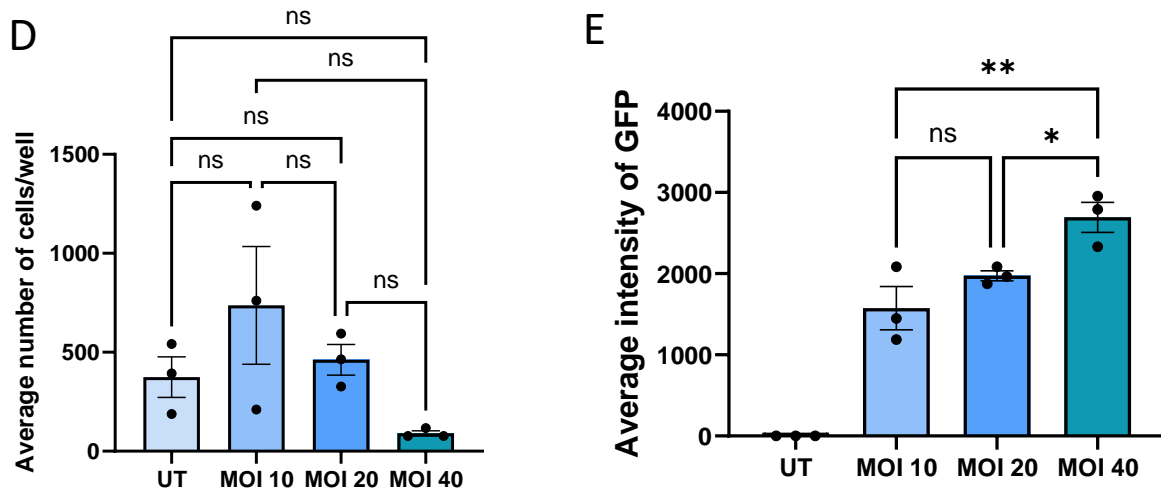
Figure 3.8: Transduction efficiency of LV-GFP in iPSC-derived motor neurons with increasing MOI

ALS iPSC-derived motor neurons (0.02 x10^6 cells/well) were transduced with the lentivirus expressing GFP (LV-GFP) at increasing MOI of 10, 20 and 40. **(A)** The cells were fixed in 4% PFA 4 days post transduction and imaged for GFP fluorescence. Scale bars represent 20µm. **(B)** Transduction efficiency (n=3, mean ± SEM) was calculated at each MOI with comparison between transduced and untransduced. Transduction at MOI 20 showed the highest average transduction efficiency (79.3%) while the cells looked healthy 4 days post transduction, however, at MOI 40 the transduction efficiency greatly reduced (27.1%). MOI 10 = 53.7%. **(C)** The average transduction efficiency was plotted against the MOI of LV-GFP tested. There was a weak, negative and non-significant correlation observed. Simple linear regression and correlation analysis showed R^2=0.0.4396, r=-0.663 and p=0.5386. **(D)** The number of cells/well for each cell line under each condition was taken into account to obtain the average number of cells counted for each condition (UT n=374.7, MOI 10 n=736.8, MOI 20 n=461.8 and MOI 40 n=90). **(E)** The average intensity of GFP expression was measured for each cell line at each dose. There were no significant differences in the intensity of GFP between MOI10 and 20 (MOI 10=1573.84 and MOI 20=1974.51). The average intensity of GFP was significantly higher at MOI 40 when compared to both lower doses (MOI 40=2693.4).Data were analysed by an ordinary one-way ANOVA with a Šídák's multiple comparisons test. Error bars represent ± SEM. UT = untransduced.

3.3.6 Construction of pcDNA3.1 plasmid with either FL-ARFIP2 or HC-ARFIP2 inserts to inject into zebrafish

A construct carrying the C-terminus of the protein was made to express HC-ARFIP2 in zebrafish and assess its effects *in vivo*. The FL-ARFIP2 construct was also made to analyse the effects of the full-length protein. To produce these plasmids, the desired genes were subcloned into the pcDNA3.1 mammalian expression vector backbone carrying all the necessary features.

3.3.6.1 FL-ARFIP2 and HC-ARFIP2 amplification by PCR and linearisation of the pcDNA3.1 vector

The FL-*ARFIP2* and HC-*ARFIP2* genes were subcloned from the sin-PGK-cPPT LV-plasmid into the pcDNA3.1 vector (2.1.2.1). First, the gene inserts were amplified by PCR. The FL-*ARFIP2* and HC-*ARFIP2* forward primers were designed to contain the HindIII restriction site at the 5' end. The reverse primer was common to both and was designed to contain the XhoI restriction site at the 3' end (**Table *2.1***). The PCR products were run and visualised on a 1% agarose gel (2.1.2.11), followed by gel extraction and purification of the desired bands (2.1.2.12). The expected sizes for the FL-ARFIP2 and HC-ARFIP2 PCR products were 1075 bp and 568 bp, respectively (**Figure 3.9**). **Figure 3.9 A** shows the desired band at the right height for HC-ARFIP2 only and not FL-ARFIP2. Subsequently, the PCR for FL-ARFIP2 was repeated with 3% DMSO to resolve any secondary structures (**Figure 3.9 B**).
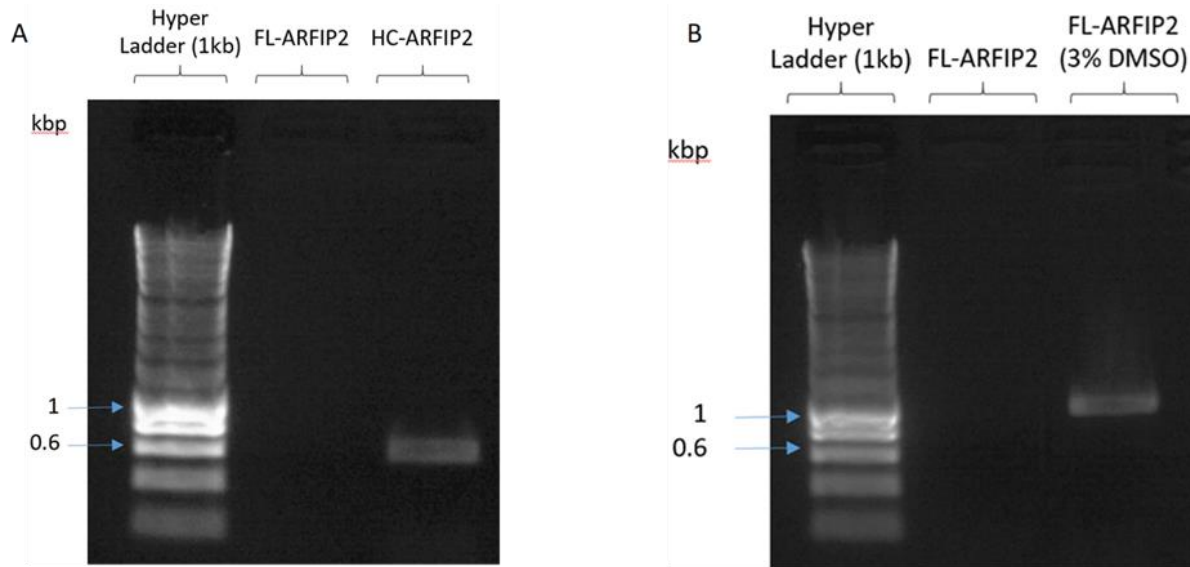
Figure 3.9: FL-ARFIP2 and HC-ARFIP2 PCR products for subcloning
PCR products for HC-ARFIP2 and FL-ARFIP2 were run on a 1% agarose gel. **(A)** The HC-ARFIP2 band is at 568 bp. There is no band for FL-ARFIP2 due to unresolved secondary structures. **(B)** PCR for FL-ARFIP2 was conducted in the presence of 3% DMSO to resolve the secondary structures. The FL-ARFIP2 band is at 1075 bp.

After PCR, the desired bands were excised, and the PCR products were purified to strip off the other ingredients of the PCR and then digested with HindIII and XhoI (2.1.2.6). The digested and purified PCR products were run on a 1% agarose gel and the corresponding bands were extracted and purified. Accordingly, the pcDNA3.1 plasmid was digested and linearised with the same restriction enzymes (HindIII and XhoI) to excise the sequence between these sites for later insertion of the desired inserts. The linearised pcDNA3.1 vector backbone was run on a 1% gel and the corresponding band was purified.

3.3.6.2 FL-ARFIP2 and HC-ARFIP2 inserts subcloned into the pcDNA3.1 vector backbone

The FL-*ARFIP2* or HC-*ARFIP2* insert was ligated into the pcDNA3.1 vector backbone between the HindIII and XhoI restriction sites (2.1.2.13). The ligation mix (with a 1:3 vector:insert molar ratio) was transformed into DH5α cells (2.1.2.3) and plated onto an ampicillin selection plate. 13 colonies from the pcDNA3.1-FL-*ARFIP2* transformed bacterial plate and 11 colonies from the pcDNA3.1-HC-*ARFIP2* transformed bacterial plate were picked and cultured. The corresponding plasmids

were isolated using a plasmid miniprep kit. Plasmid-preps expressing effective ligations were analysed by restriction enzyme digestion with HindIII and XhoI restriction enzymes (2.1.2.6) and visualised on agarose gels (2.1.2.11). **Figure 3.10** below shows the successful insertion of (**A**) FL-ARFIP2 and (**B**) HC-ARFIP2 into the vector backbone validated by digestion. The vector backbone can be seen at ~10kbp, FL-*ARFIP2* at 1059 bp, and HC-*ARFIP2* at 552 bp. Successful ligations are seen for most plasmid preps. The successfully ligated plasmid preps were further digested with NcoI-HF to confirm an intact vector backbone. Digestion with NcoI-HF (**Figure 3.10 C and D**) showed the expected bands at 7444bp, 2262bp, 1267bp, 735bp, and 420bp for FL-ARFIP2 (**C**) and at 6937bp, 2262bp, 1267bp, 735bp and 420bp for HC-ARFIP2 (**D**).

A

B
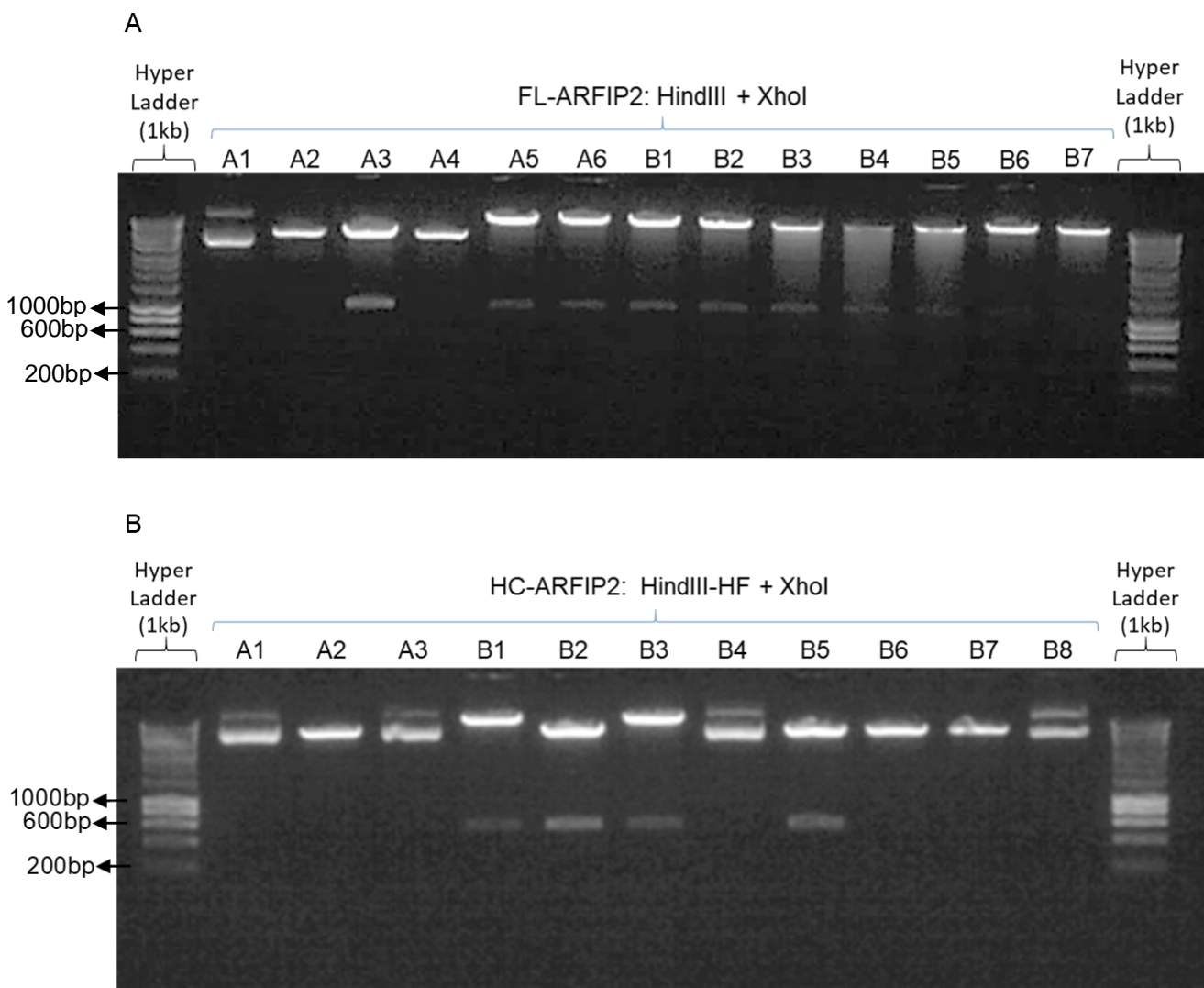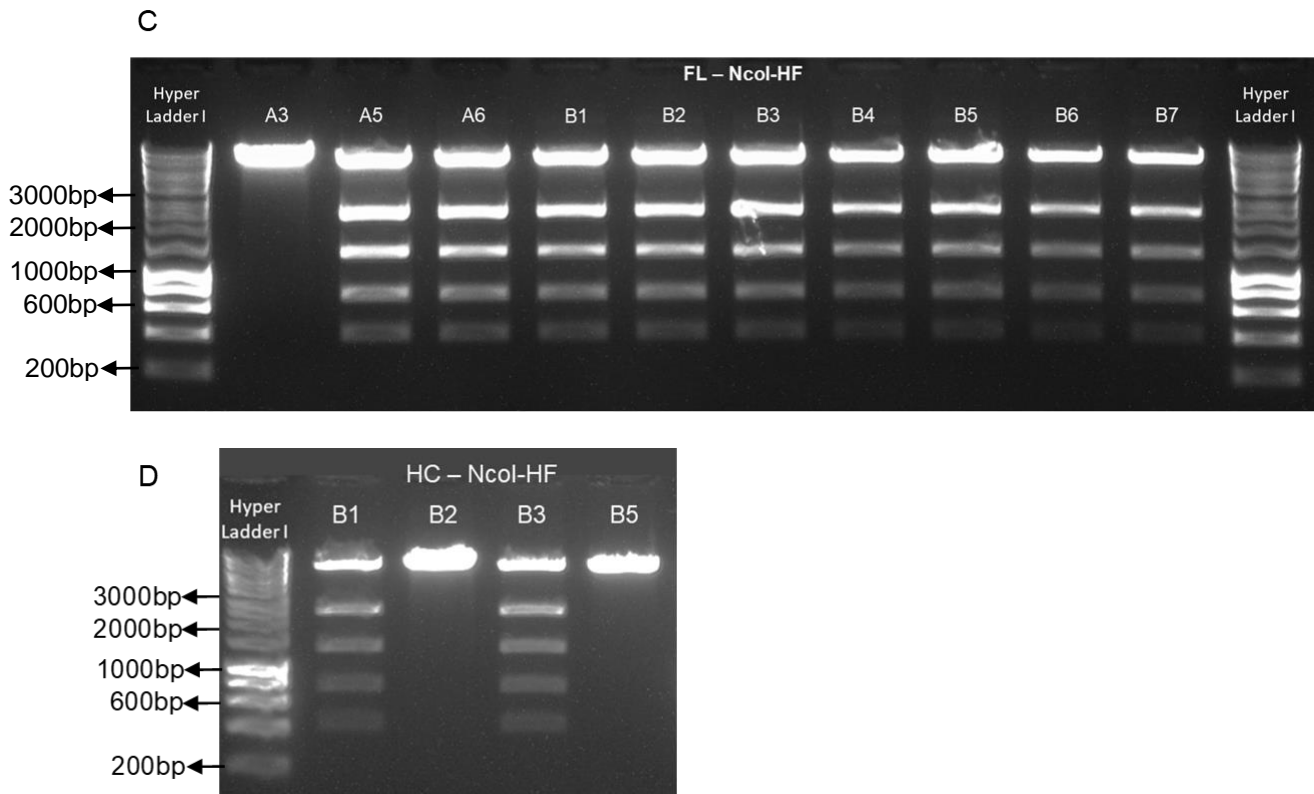
Figure 3.10: Insertion of FL-ARFIP2 and HC-ARFIP2 into pcDNA3.1 vector
The inserts were ligated into the vector and the plasmid-preps were digested with HindIII and XhoI (A-B). The gel images show successful ligations and digestion at the expected heights for some colonies. Vector backbone at ~10kbp, FL-ARFIP2 at 1kbp (A) and HC-ARFIP2 at ~0.6kbp (B). The plasmid-preps were digested with NcoI to check for an intact backbone (C-D).

One plasmid-prep for each of the two constructs was chosen for sequencing – A5 for FL-*ARFIP2* (FL-A5) and B3 for HC-*ARFIP2* (HC-B3). Before sequencing, the plasmids were re-transformed, colonies were picked, and plasmids were isolated using plasmid preparation kits. The plasmid-preps were digested with HindIII and XhoI to check for the presence of insert, as well as NcoI-HF to verify if the plasmid backbone was intact (**Figure 3.11 A**).

The gene of interest was downstream of the zebrafish ubiquitin (Ubi) promoter. The plasmid construct carried the DsRed reporter gene downstream of the zebrafish heat shock protein 70 (Hsp70) promoter that served as a readout for experiments *in vivo* and was utilised to track transgene expression. I-SceI restriction enzyme sites were used to linearise the DNA before injecting zebrafish embryos at the one-cell stage.

Figure 3.11 B shows the schematic representative plasmid maps of the vectors with the FL-*ARIP2* (pcDNA3.1-FL-*ARFIP2*) or HC-*ARFIP2* (pcDNA3.1-HC-*ARFIP2*).
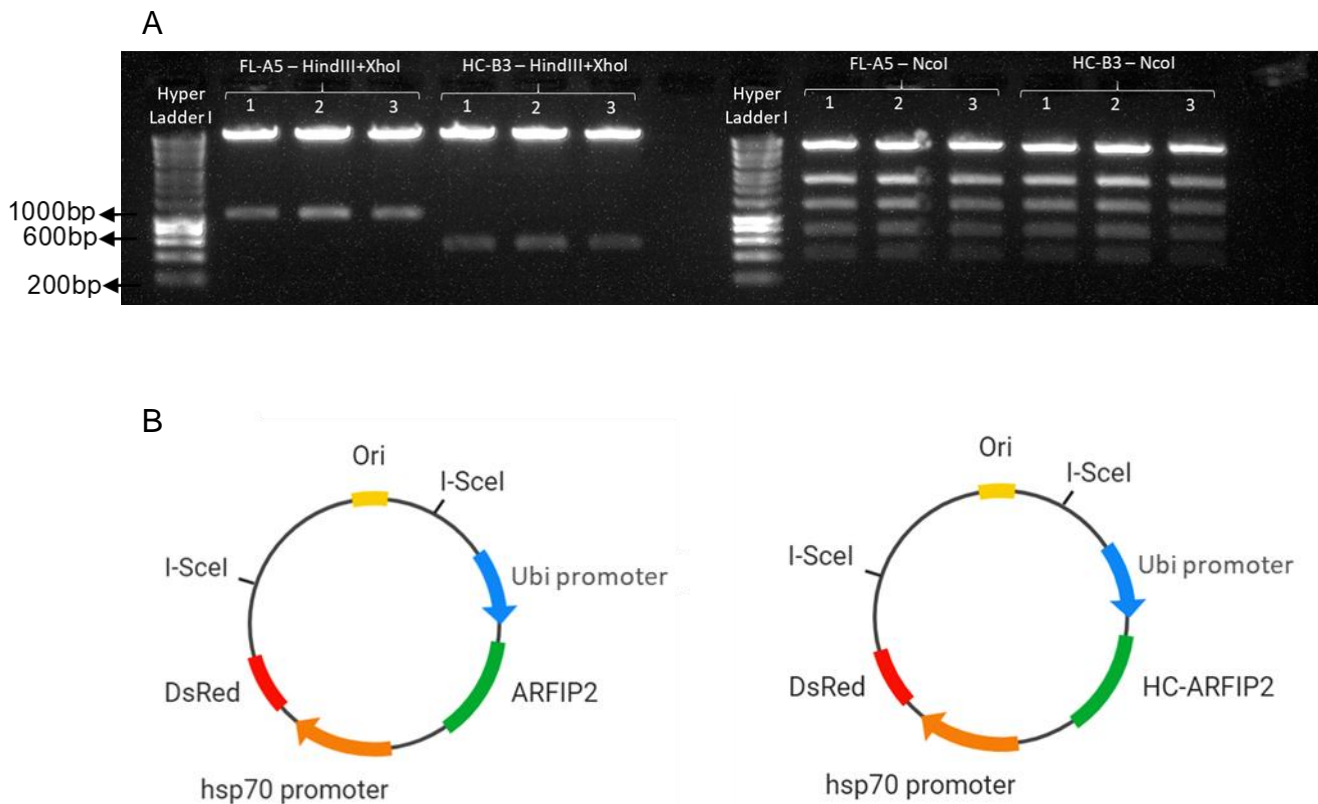
Figure 3.11: Plasmid map representation of the pcDNA3.1 vector with FL-ARFIP2 or HC-ARFIP2 insert and validation by digestion
One colony for each FL-ARFIP2 and HC-ARFIP2, A5 and B3 respectively, were chosen and expanded and isolated. **(A)** The plasmid-preps were digested with HindIII and XhoI as well as NcoI-HF. The gel images show successful cloning through the presence of the inserts at the expected heights (7444bp, 2262bp, 1267bp, 735bp, and 420bp for FL-ARFIP2 and at 6937bp, 2262bp, 1267bp, 735bp and 420bp for HC-ARFIP2). Digestion with NcoI-HF shows band at the expected heights and validated an intact vector backbone. **(B)** Schematic plasmid maps representing the pcDNA3.1-FLARFIP2 and pcDNA3.1-HC-ARFIP2.

DNA insertion and ligation were confirmed by DNA sequencing (2.1.2.15) of the whole plasmid using 12 primer pairs to confirm no spontaneous mutations had occurred in the insert and the vector backbone during the cloning process (primers used for sequencing are listed in **Table *2.4***). **Figure 3.12 A and B** show the sanger sequencing chromatograms for the FL-*ARFIP2* and HC-*ARFIP2* inserts, respectively. No

mutations (changes or missing base pairs) were observed in the insert sequences. The pcDNA3.1-FL-*ARFIP2* and pcDNA3.1-HC-*ARFIP2* whole plasmid sequences were used to generate the representative plasmid maps (**Figure 3.12 C and D**). Sequences of both constructs can be found in **Appendix 1**. Overall, results show the successful subcloning of FL-*ARFIP2* and HC-*ARFIP2* into the pcDNA3.1 mammalian expression vector backbone.

A
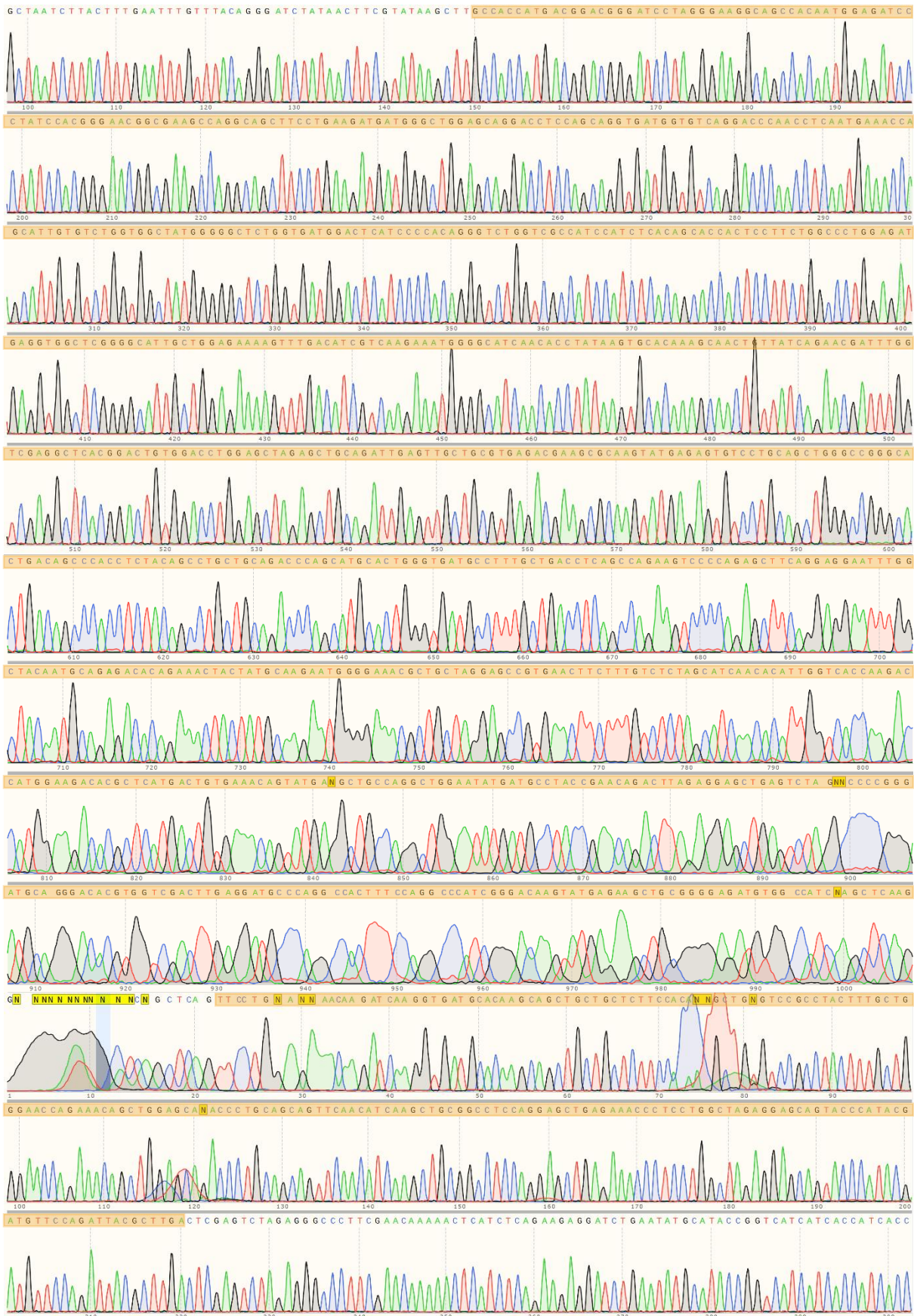
B
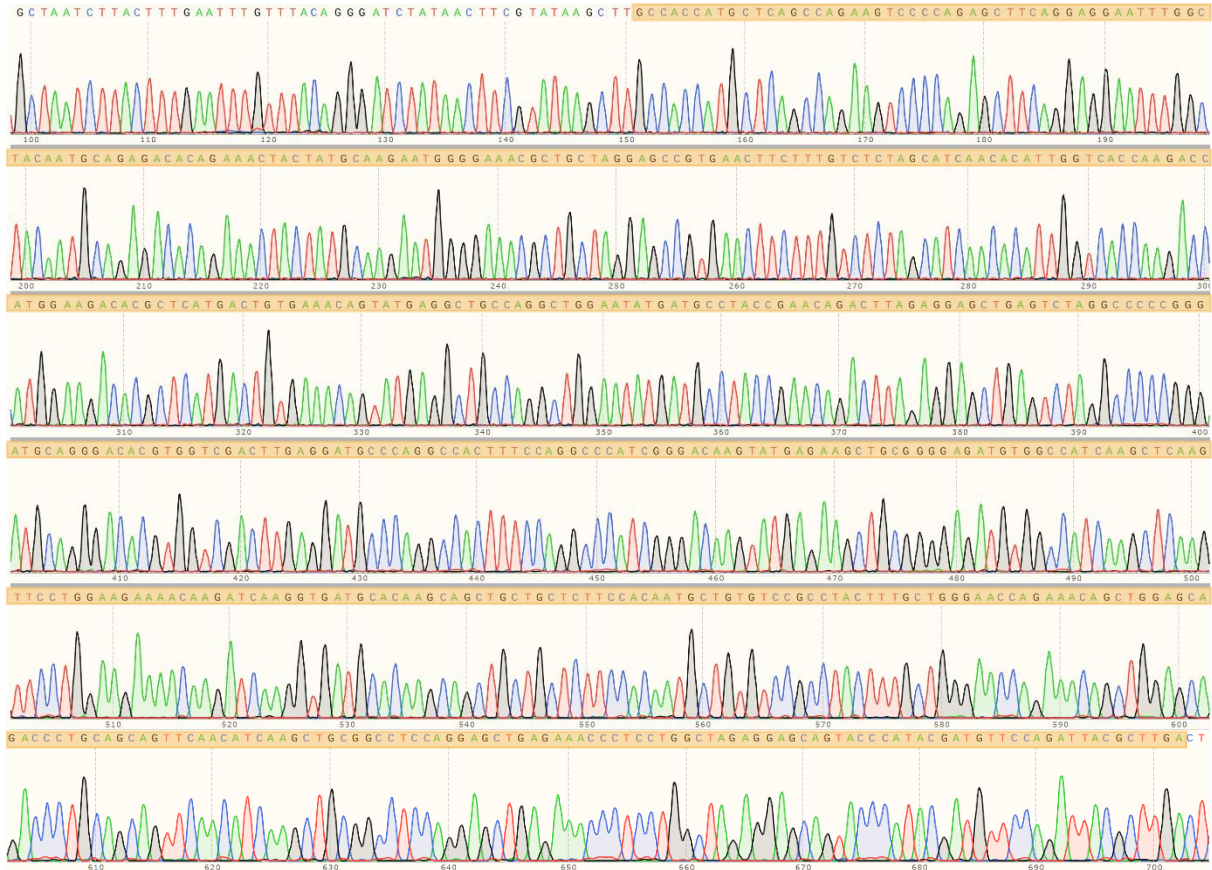
C

D
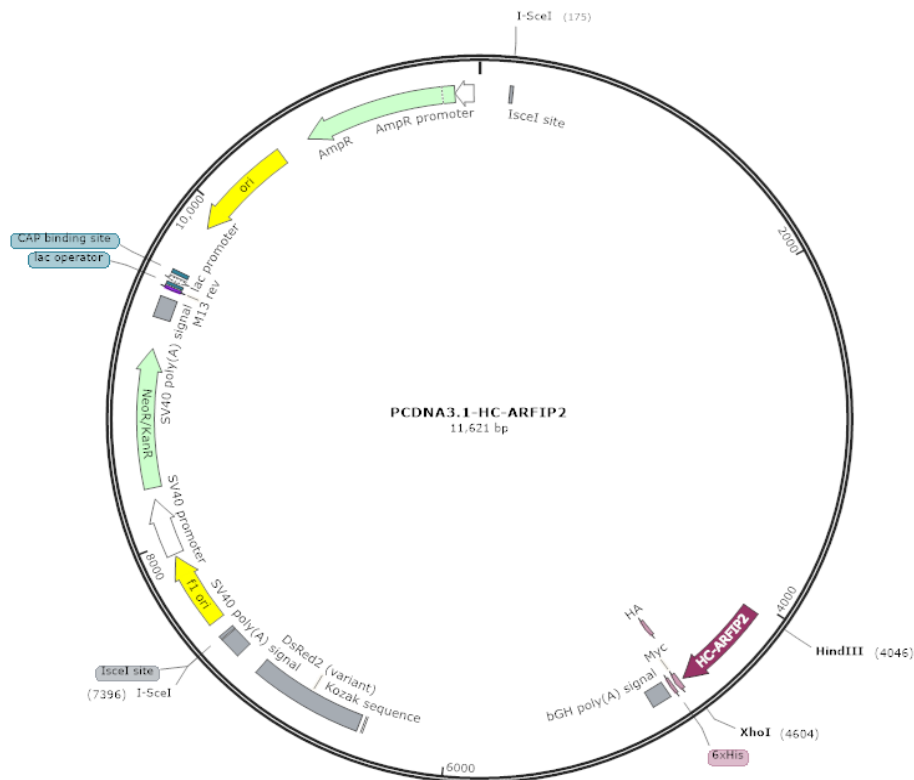
Figure 3.12: Sanger sequencing spectra and plasmid map representations of the pcDNA3.1 vector with FL-ARFIP2 or HC-ARFIP2 insert
Sanger sequencing chromatograms showing the sequence of **(A)** FL-ARFIP2 and **(B)** HC-ARFIP2 inserts in the pcDNA3.1 vector backbone. The sequences of the inserts are highlighted in orange. There were no mutations detected in the insert sequences. **(C)** Plasmid map for pCDNA3.1-FL-ARFIP2 and **(D)** pCDNA3.1-HC-ARFIP2. The maps were generated using the SnapGene software.

3.3.7 Construction and validation of ARFIP2-shRNA plasmids

3.3.7.1 Cloning of designed shRNA inserts into the pLVTHM vector backbone

shRNA constructs were designed (2.1.2.2) and made to assess whether depletion of ARFIP2 can affect disease pathogenesis. The pLVTHM plasmid was used for the vector backbone. The vector was digested with the ClaI and MluI-HF restriction enzymes to produce ends that were compatible with the shRNA oligo duplex (**Table 2.5** in chapter 2). The restriction enzyme digest of the vector was followed by dephosphorylation. **Figure 3.13** below shows the successful digestion and linearization of the pLVTHM vector. The band in lane 2 corresponds to the expected band of ~11 kbp for the digested vector backbone. This band was used as a reference to cut the gel (Lane 3) for gel extraction of the digested vector.
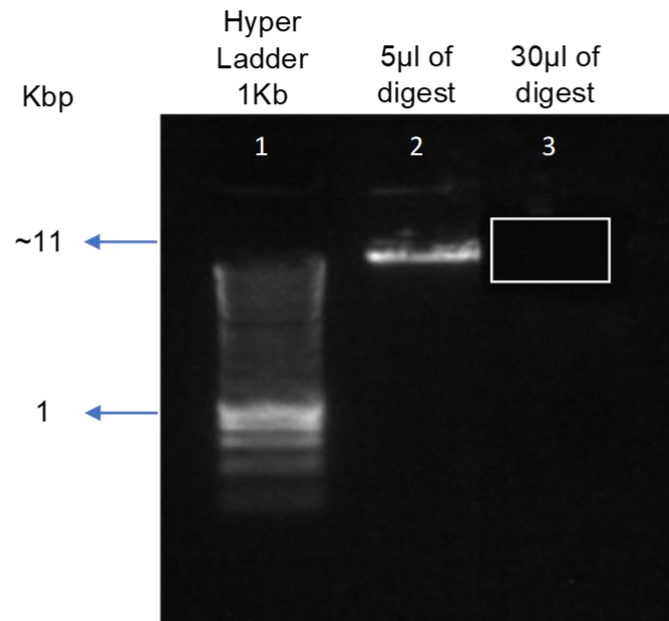
Figure 3.13: Gel image of pLVTHM digestion with MluI and ClaI
The pLVTHM plasmid was digested with MluI and ClaI. The vector backbone size is ~11Kb. It was extracted from the gel (Lane 3), purified and quantified with Nanodrop. Lane 1: 10µl Hyper Ladder I (1Kbp), Lane 2: 5µl of digest (used as reference to cut out gel at the right height). Lane 3: 30µl of digest (the white box indicates where the gel had been cut to excise the vector backbone band).

In order to produce the pLVTHM plasmid expressing the three (471, 480, and 749) desired shRNA sequences (**Table 2.5** in chapter 2), the shRNA oligos were annealed and phosphorylated. The oligo duplexes were then ligated into the digested and dephosphorylated pLVTHM vector backbone. Once ligated, the constructs were transformed into NEB-stable cells and individual colonies from the transformed bacterial culture plates were picked and the plasmids were isolated. The plasmids with the three individual shRNA sequences (471, 480, and 749) were sent for sequencing to check if the insert had ligated successfully in the desired locus and confirm that the inserts had no mutations. The Sanger sequencing chromatograms (**Figure 3.14**) show that the shRNA inserts (471, 480, and 749) were successfully cloned into the pLVTHM vector and that there were no mutations in the insert sequences. Digestion of the plasmid was not done to validate the successful insertion of the shRNA inserts into the backbone as the shRNA insert size was very small (67 bp) and difficult to visualise on a gel.

Figure 3.14: Sanger sequencing chromatograms for the shRNA inserts
Sanger sequencing chromatograms showing the sequence of **(A)** shRNA 471, **(B)** shRNA 480 and **(C)** shRNA 749 inserts in the pLVTHM vector backbone. The sequences of the shRNA inserts are highlighted in orange. There were no mutations detected in the insert sequences.

The three shRNA sequences used were referred to as 471, 480, or 749. **Figure 3.15** shows a representative schematic plasmid map of the pLVTHM vector with the shRNA insert as well as the plasmid maps for all three shRNA constructs. The plasmid sequences for all three plasmids can be found in **Appendix 1**.

A

B

C

pLVTHM+480
11,161 bp

D

pLVTHM+749
11,161 bp
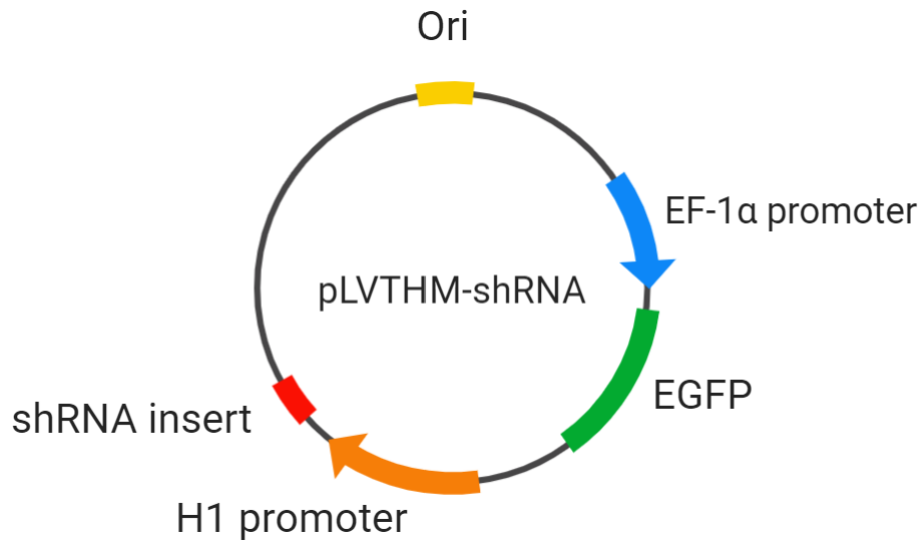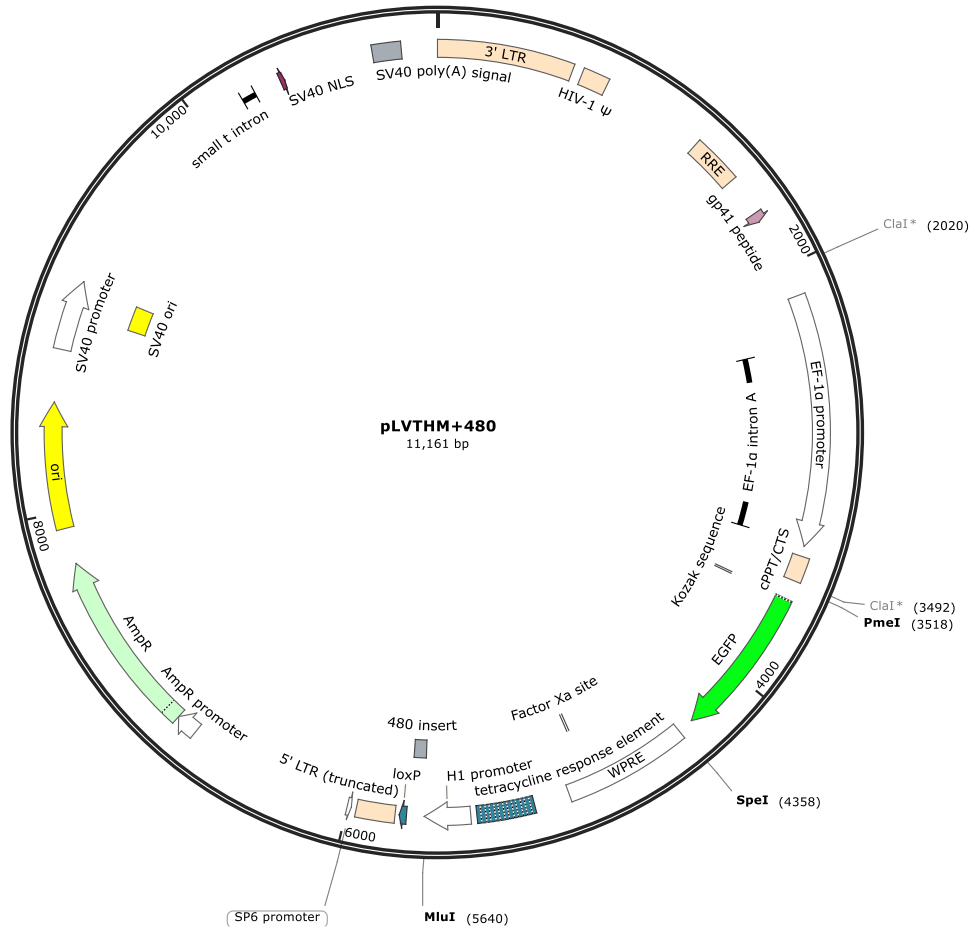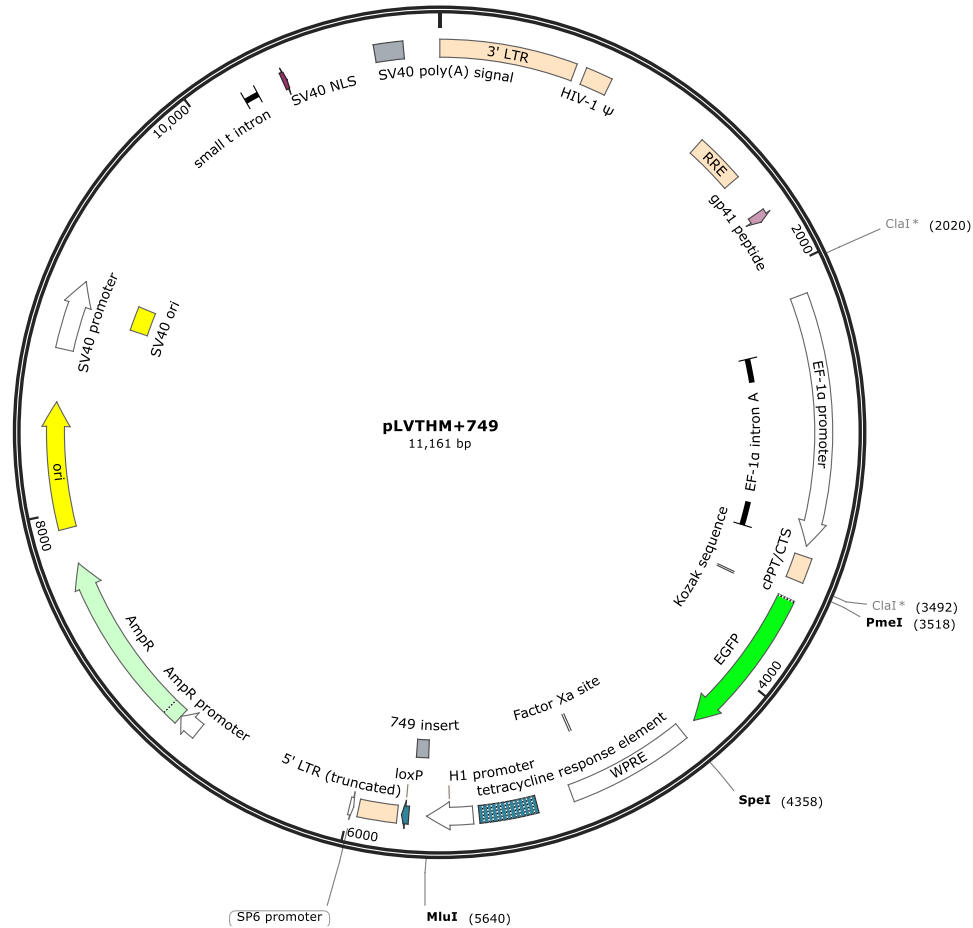
Figure 3.15: Representative plasmid maps of pLVTHM with shRNA inserts
(A) Schematic plasmid map to show a representative construct for pLVTHM-shRNA. **(B-D)** Plasmid maps representing pLVTHM plasmids with individual shRNA inserts (471, 480 and 749). The sequences of the inserts are given in Table 2.5 in chapter 2. The maps were generated using the SnapGene software.

3.3.7.2 Validation of ARFIP2-shRNA plasmids in HEK293T cells

To assess the efficiency of the *ARFIP2*-shRNA plasmids, primary validation was done using HEK293T cells. HEK293T cells were transfected with one or a combination of the three plasmids shown in **Figure 3.15**. The PEI transfection method was used (2.2.2.4), after which the proteins were extracted 3 days post-transfection (2.1.2.17) and quantified (2.1.2.19). Whether the plasmids were able to deplete the ARFIP2 expression level was determined by western blot (2.1.2.20) where α-tubulin was used as a loading control. Successful depletion of ARFIP2 was evaluated by identifying the reduced expression of a band corresponding to ARFIP2 at 38.6 kDa (**Figure 3.16**) in transfected samples compared to the untransfected control. ARFIP2 was detected using the rabbit anti-ARFIP2 antibody. **Figure 3.16** shows the results after transfection with **(A)** 2μg and 4μg of DNA or **(B)** with combinations of *ARFIP2*-shRNA plasmids (2 μg each). The western blot membrane shows that there is no depletion when transfected with 2μg or 4μg of DNA, however, there is depletion when transfected with combinations of the plasmids.
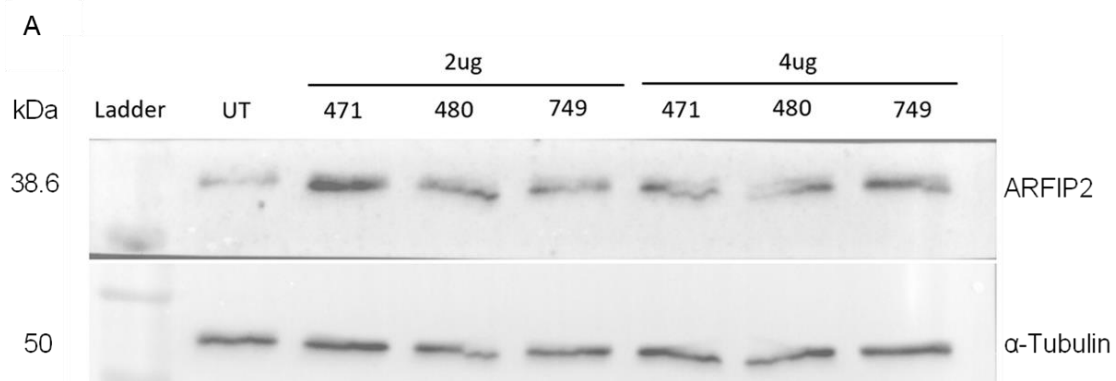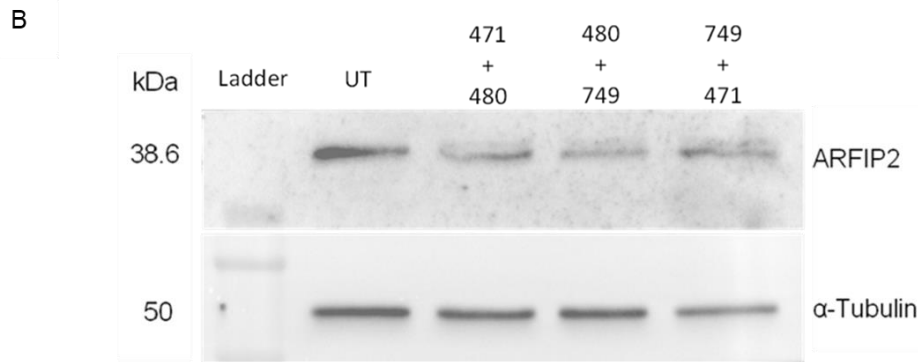
Figure 3.16: Validation of ARFIP2-shRNA plasmids for depletion of ARFIP2
HEK293T cells were transfected with the plasmids shown in Figure 3.15. For each condition, 20µg of total protein was loaded into each well. ARFIP2 expression was detected on a western blot by the rabbit anti-ARFIP2 antibody. ARFIP2 is detected at 38.6kDa. α-Tubulin (50kDa) was used as a loading control. **(A)** Cells were transfected with either 2µg or 4µg of each plasmid. **(B)** Cells were treated in combinations with 2µg of each plasmid. UT=untransfected.

Successful depletion of ARFIP2 was evaluated by comparison of band densities of transfected cells with untransfected cells (**Figure 3.17**). The ARFIP2 bands were normalised to the α-Tubulin loading control. The results show there was no depletion when treated with 2µg and 4µg of plasmid DNA, and rather there was an increase (**Figure 3.17 A**). However, when comparing the band densities for samples transfected in combinations of *ARFIP2*-shRNA with the untransfected sample, the results in **Figure 3.17 B** show successful depletion of ARFIP2 by approximately 50% in samples transfected with *ARFIP2*-shRNA plasmids with the 471 and 480 inserts and the 480 and 749 inserts.
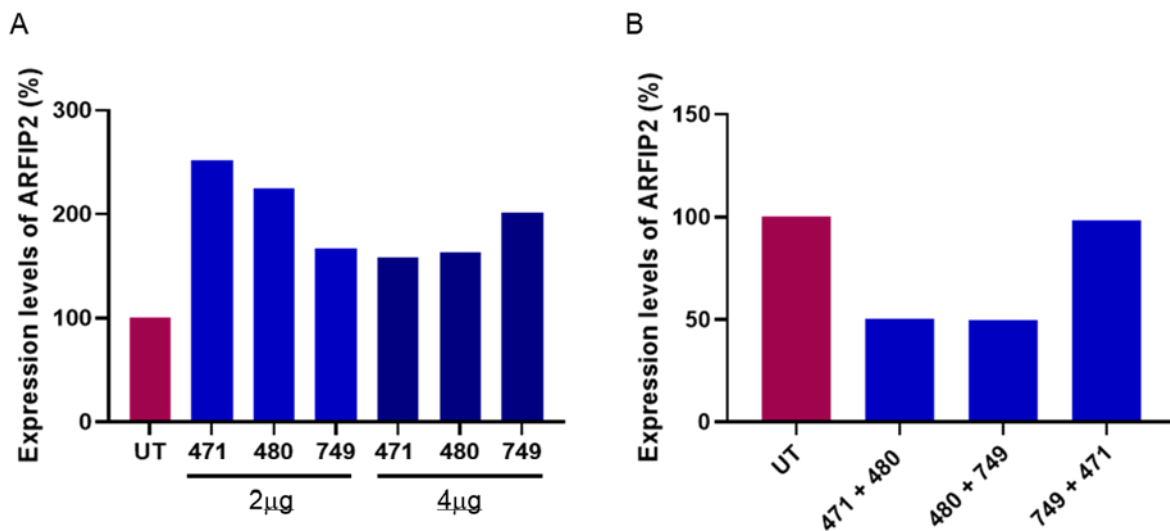
Figure 3.17: Preliminary comparison of band intensities between samples transfected with ARFIP2-shRNA plasmids and untransfected samples to determine depletion

The levels of ARFIP2 expression were compared between samples after normalisation with α-Tubulin. **(A)** Expression levels of ARFIP2 in samples transfected with 2μg or 4μg of each plasmid compared to the untransfected (UT) sample. No depletion in expression was seen, however, expression levels were increased (471 (2μg) = 252.41%, 480 (2μg) = 224.67%, 749 (2μg) = 166.51%; 471 (4μg) = 158.27%, 480 (4μg) = 163.34%, 749 (4μg) = 201.25%). **(B)** Expression levels of ARFIP2 in samples transfected in combinations with 2μg of each plasmid compared to the untransfected (UT) sample. There is a clear depletion when treated with 471+480 (50.18%) and 480+749 (49.97%). There was an no decrease in ARFIP2 levels when treated with 749+471 (98.21%). One experimental repeat.

HEK293T cells transfected with combinations of plasmids were used to test for transfection efficiency of the *ARFIP2*-shRNA plasmids. The transfected plate was imaged using the IN Cell Analyzer 2500 HS high content analysis (HCA) imaging system (GE Healthcare Life Sciences, 29240356) and analysed using the Columbus Image Data Storage and Analysis System (PerkinElmer). **Table 3.2** below summarises the data and shows a transfection efficiency of more than 80% for all combinations.

Table 3.2: Transfection efficiency for ARFIP2-shRNA plasmids in HEK293T cells

Plasmid	Total Number of Cells	Number of Cells Expressing GFP	Number of Analysed Fields	Transfection Efficiency (%)
471 + 480	3175	2634	4	83
480 + 749	3831	3472	4	91
749 + 471	2219	1988	4	90
LV-GFP	4728	4125	4	87
Untransfected	4273	0	4	N/A

3.4   Results Summary

This chapter describes the successful design, generation, and validation of the tools (lentiviruses, iPSC-derived MNs, and plasmids) that were used throughout the Ph.D. project. It was shown that:

- The pLV-FL and pLV-HC plasmids carry the FL-*ARFIP2* and HC-*ARFIP2* inserts.

- Using HEK293T cells for primary validation, the LV-based plasmids (pLV-HC and pLV-GFP) were able to successfully express the transgenes.

- The human iPSCs were successfully differentiated into mature and functional motor neurons which validated the *in vitro* model to be used in the project.

- The LV made with pLV-FL, pLV-HC, and pLV-GFP were produced with high functional viral titres.

- LV-HC and LV-GFP are able to express the transgenes in HEK293T cells and iPSC-derived MNs. Furthermore, LV-HC efficiently expressed HC-ARFIP2 in iPSC-derived MNs in a dose-dependent manner.

- LV-HC and LV-GFP efficiently transduce iPSC-derived MNs whereas only LV-HC transduces cells in a dose-dependent manner.

- LV-HC had the highest transduction efficiency at MOI 40, however, LV-GFP exhibited toxic effects at this dose and had the highest transduction efficiency at MOI 20.

- The HC-*ARFIP2* and FL-*ARFIP2* genes were successfully subcloned into the pCDNA3.1 backbone to generate plasmids for microinjecting zebrafish embryos.

- The three chosen shRNA oligos for endogenous ARFIP2 depletion were successfully cloned into the pLVTHM vector backbone and validated in HEK293T cells.

3.5 Discussion

Transfection is relatively difficult for cultured neurons and high transfection efficiencies are only possible with increasing toxicity (Li *et al.*, 2010). Therefore, viral vectors are effective agents for gene delivery to target cells, like neurons, which is crucial for evaluating therapeutic efficacy. Features of the viral vector expression cassettes can be manipulated to improve target selectivity and transgenic expression (Powell, Rivera-Soto and Gray, 2015). Despite the fact that virus-mediated transfections are very efficient and simple to employ, cytotoxicity and random integration are major drawbacks that can affect *in vitro* experiments and the downstream results. Viral vector introduction into the cell may trigger an inflammatory reaction. Furthermore, due to the random integration of the viral vector into the host genome, there is a risk of insertional mutagenesis which may cause tumour suppressor genes to be disrupted, oncogenes to be activated, and critical genes to be interrupted (Woods *et al.*, 2003). In addition to LV vectors, adeno-associated viral vectors (AAV) are also regarded as an increasingly useful gene therapy vector for the central nervous system. Various studies over the years show that the two vectors have been used to express transgenes in various brain regions and neuronal cell types (Schneider, Zufferey and Aebischer, 2008; Papale, Cerovic and Brambilla, 2009; Lim, Airavaara and Harvey, 2010; Manfredsson and Mandel, 2010). AAV is ideal and mostly used for *in vivo* gene delivery because of its low immunogenicity, however, evaluating AAV performance is not optimal with *in vitro* models. In most cases, AAV transduction observations *in vitro* are unable to match or predict the *in vivo* response (Bulcha *et al.*, 2021).

Therefore, in this project, we used the LV vector for efficient delivery of the transgenes into the iPSC-derived MN *in vitro* model. LV vectors are most suitable for gene transfer *in vitro* and are able to transduce non-dividing cells, including neuronal cells while enabling a stable long-term gene expression with integration into the host genome and with low immunogenicity (Liu and Berkhout, 2014; Lundstrom, 2019). Several studies have shown that many types of central and peripheral neurons have been successfully transduced using LV (Jakobsson *et al.*, 2003; Sawada *et al.*, 2010; Miletic *et al.*, 2004). The large packaging capacity of LV vectors (~9kb) enables large and multigene expression alongside the necessary components (Bulcha *et al.*, 2021). LV production and scale-up in a laboratory environment are faster and relatively trouble-free and

result in high titres. In this chapter, it was shown that the three LV vectors, LV-FL, LV-HC, and LV-GFP, have been produced at high titres ranging from 10^8-10^9 TU/mL.

Prior to LV production, the LV-based plasmids used to produce the LV were tested in HEK293T cells and this showed that the plasmids that carried the transgenes were able to successfully express the transgenes *in vitro*. Post LV production, LV-FL, and LV-HC were used to transduce HEK293T cells to evaluate the viruses produced before moving on to the more disease-relevant models – iPSC-derived MNs. HEK293T cells were transduced with LV-FL and LV-HC at an increasing dose of MOI 1, 3, and 10. These doses were chosen as MOI 5 has been reported to be the optimal LV dose for HEK293T cells (Davis, 2014). In addition to evaluating the virus, the aim of the experiment was also to get an estimate of the viral dose that would be ideal for each virus as a starting point for the range of doses that could be tested in the more sensitive iPSC-derived MNs. It was shown that LV-FL efficiently expressed the FL-ARFIP2 protein at all doses, with the highest efficiency and expression of the transgene seen at the highest dose of MOI 10. On the other hand, LV-HC showed the most efficient and highest expression of HC-ARFIP2 only at the MOI 10 in HEK293T cells. There was no discernible expression of the HC-ARFIP2 seen at the lowest dose of MOI 1 and a very faint expression at MOI 3. This difference in HC-ARFIP2 expression may be due to various reasons. Firstly, a reason for no or weak expression could be the low concentration of the antigen (HC-ARFIP2)at low doses of MOI 1 and 3. Since the detection of α-Tubulin showed equal loading, this demonstrated that the concentration of the sample itself was not the cause of the low/no signal. It highlighted that the HC-ARFIP2 levels might be too low and below the limit of detection at MOI 1 and weak at MOI 3. This supports that the virus itself is not inefficient, but that MOI 1 and 3 are insufficient doses for LV-HC. Secondly, reasons for no/weak signal might also be related to the western blot technique itself. There might be a low transfer efficiency for the membrane with LV-HC-treated samples. In the future, this could be more carefully confirmed with Ponceau S staining and optimising the transfer conditions for the target protein size. Moreover, prolonged washing can also decrease the signal of the low molecular weight proteins by reducing the antigen-antibody binding affinity, which can be reduced to a minimum in the future. Thirdly, the blocking solution (5% non-fat dry milk) might be masking the antigen. In the future, this could be avoided by using BSA or decreasing the amount of milk. Finally, a biological reason could be that HC-ARFIP2

is a truncated version of the ARFIP2 protein. HC-ARFIP2 is not normally expressed in cells and therefore at low doses, the truncated protein may have been degraded by the protein degradation pathways. LV-GFP was not validated at this stage due to the limited amount of virus. Regardless of the efficiency of transgene expression at different doses, both viruses were able to express the transgenes and showed the expression to be dose-dependent. Higher MOIs than MOI 10 were not tested at this stage because the aim of the experiment was achieved.

After the primary evaluation of LV-HC in HEK293T cells, LV-HC was further investigated in the iPSC-derived MNs *in vitro* model. Following on from the results obtained from initial experiments in HEK293T cells, MOI 10 was chosen as the lowest dose to test followed by MOI 20 and MOI 40. Higher MOIs than MOI 40 were not tested due to the limited amount of virus. Results showed that LV-HC successfully expressed HC-ARFIP2 at all MOIs in all iPSC-derived MN cell lines in a dose-dependent manner, with the highest and most significant expression of HC-ARFIP2 at MOI 40 in all cell lines. Furthermore, transgene expression was not altered under disease conditions when compared to healthy controls. Though the expression at MOI 10 was the highest in HEK293T cells, the expression at MOI 10 of HC-ARFIP2 was the lowest across all iPSC-derived MN lines. This may be attributed to the more complex nature of iPSC-derived MNs. This experiment showed that LV-HC was able to express the transgene in the *in vitro* model at the tested MOIs. However, since the western blot technique shows the expression across the entire sample, individual cell analysis was done to test the efficiency of the virus.

For future *in vitro* experiments, the transduction efficiencies of LV-HC and the control virus (LV-GFP) were tested to select the optimal doses with the most efficient gene transfer in conjunction with minimal toxicity in the *in vitro* model. Initially, the transduction efficiencies were assessed using confocal microscopy. Though the immunostaining was successful, imaging the various cell lines at different MOIs while capturing a good representative number of cells was time-consuming and demanding. With the protocol followed at the time, analysis of the images was done using the ImageJ software, and the positively transduced cells were manually counted This could lead to bias and give a false transduction efficiency. To overcome this, cells were immunostained and imaged using the Opera Phoenix High-Content Screening System

at high resolution. This was a high-throughput system and cut down the imaging time from weeks to hours, compared to the confocal method. Using the Opera system, the data were analysed using the Columbus Image Data Storage and Analysis System. The analysis was all automated with parameters that could be set to assess all conditions in an experiment in a consistent manner. Automation greatly reduces any bias.

Using the untransduced cells, a threshold for the background was set for the green channel that represented transduced cells. All cells were identified and counted and only the cells above the threshold were automatically identified and counted as positively stained to generate the transduction efficiencies at each MOI for both viruses. For LV-HC, the analysis showed that the virus transduces in a dose-dependent manner with transduction at MOI 40 showing the highest average transduction efficiency (30.62%). For this reason, MOI 40 was chosen as the optimal dose for LV-HC for subsequent *in vitro* experiments. On the other hand, the average transduction efficiency for LV-*GFP* was the lowest at MOI 40 (27.1%). Comparing the images at MOI 40 for both viruses, we can see that the cells transduced with LV-HC at MOI 40 show minimal toxicity as the cells still looked healthy 4 days post-transduction, whereas the cells did not look healthy and were sparse in the images of cells transduced with LV-*GFP* at MOI 40. This was supported by analysing the average number of cells/well that were counted at each dose for both viruses. While for LV-HC, the average number of cells/well was highest at MOI 40, for LV-*GFP* at MOI 40, it was the lowest. There was a drastic decrease in the average number of cells/well which indicated cell loss/death and therefore toxicity of LV-*GFP* at the highest dose of MOI 40. Therefore, MOI 40 for LV-*GFP*, which would have been the comparable dose for both viruses, could not be chosen for subsequent experiments. LV-GFP showed the highest transduction efficiency of 79.3% at MOI 20 and transduction efficiency of 53.7% at MOI 10. At both doses, the cells still looked healthy 4 days post-transduction. The decrease in transduction efficiency of LV-GFP from MOI 20 to MOI 10 was approximately 25% and so future work should include an assessment of MOI 5 since the transduction efficiency might have been more comparable to that of LV-HC. This was not assessed at the time of the project and therefore not used. Since the transduction efficiency of LV-*GFP* at MOI 10 was most comparable to the transduction efficiency of LV-HC at the chosen dose of MOI 40,

MOI 10 was chosen as the optimal dose of LV-*GFP* for subsequent experiments. This allowed us to account for any variation observed in experiments that may be caused by the presence of the lentivirus itself or its administration rather than transgene expression. This gave us a comparable percentage of transduced cells with LV-*GFP* and LV-HC-*ARFIP2*.

Looking at the data for LV-HC transduction, it can be suggested that some cells may have higher protein levels than others owing to possible multiple transductions in any one cell. The western blot data for LV-HC transduction shows higher expression at higher MOIs which could support this. However, looking at the transduction efficiencies, as the MOI increases, we do get more cells transduced, but the overall transduction efficiency is low in this cell type. Furthermore, the intensity of staining shows similar intensities of staining at MOI 20 and 40 and a non-significant difference in intensity at MOI 10. This suggests that, though there are more transduced cells, the protein levels at the different doses may be the same. This finding is consistent with multiple transductions in any one cell being a rare event. Moreover, an important observation in this chapter that warrants further study in the project was seen when comparing the expression levels of HC-ARFIP2 and transduction efficiency of LV-HC between control and patient lines. Results showed that the expression levels and transduction efficiency were not affected by the disease in ALS patients lines, compared to controls. This shows that the transgene and viral transduction were well tolerated and not altered by disease.

The *GFP* reporter gene has been extensively used in cells to label cells and measure gene expression. However, several reports have indicated that it may have detrimental effects on cell physiology (Ansari *et al.*, 2016; Baens *et al.*, 2006). GFP toxicity in cells was first reported in 1999 when they showed a link between GFP expression and apoptosis induction in a number of cell lines (Liu *et al.*, 1999). This link with apoptosis could explain the drastic decrease in average cells/well after LV-*GFP* transduction at MOI 40. This study later promoted several studies that looked into GFP cytotoxicity (Zhang *et al.*, 2003; Ansari *et al.*, 2016; Cabezas de la Fuente, 2021; Baens *et al.*, 2006). GFP-associated toxicity raises concerns over how useful LV-GFP is as a control virus for this study. The HC-*ARFIP2* transgene in LV-HC is tagged with an HA-tag for detection in the experiment. It would have been ideal if HC-ARFIP2 was tagged

with GFP which would allow the comparison of GFP toxicity across all treatment conditions. However, this was not possible because an HC-*ARFIP2-GFP* carrying construct was not available at the time and the HC-*ARFIP2-HA-tag* carrying construct was already in possession. To overcome GFP toxicity, a lentivirus expressing HA-tag only could be used but this was not possible as the protein sequence would be very short to be expressed on its own and would cause problems in detection. Therefore, LV-*GFP* was used as the control virus.

The FL-*ARFIP2* and HC-*ARFIP2* inserts from the LV-based plasmids were subcloned into the pCDNA3.1 backbone to generate the zebrafish constructs needed to make the corresponding transgenic zebrafish lines. *ARFIP2* is highly conserved, especially between humans and zebrafish, with 289 orthologs and a 69% identity between the zebrafish and human amino acid sequence. Most of the homology seen is in the c-terminus conserved region of the protein. With this information, the FL-*ARFP2* and HC-*ARFIP2* genes fused with the HA-tag epitope were subcloned downstream of the zebrafish ubiquitin (Ubi) promoter into the pCDNA3.1 backbone. The vector backbone (pcDNA3.1), provided by Dr. Tennore Ramesh, had been previously used to generate models of ALS (Ramesh *et al.*, 2010a; Shaw *et al.*, 2018). The *C9ORF72* ALS model was used as the *in vivo* model in this study. The same vector backbone (pcDNA3.1) that was used to generate the *in vivo* ALS model, was used to construct the FL-*ARFIP2* and HC-*ARFIP2* zebrafish constructs. Since the vector carried the DsRed reporter gene and all fluorescent proteins carry some associated toxicity, using the same vector backbone would allow easy comparison and consistency between the two models.

After the subcloning of the two inserts into pcDNA3.1, 13 individual pcDNA3.1-FL-*ARFIP2* colonies and 11 individual pcDNA3.1-HC-*ARFIP2* colonies were picked in order to assess whether the inserts were successfully ligated into the vector backbone and if the backbone was intact. For the pcDNA3.1-FL-*ARFIP2* construct, 10 out of 13 colonies showed the presence of the insert and 4 out of 11 colonies showed the presence of the insert for the pcDNA3.1-HC-*ARFIP2* construct. However, in all the positive colonies for both the constructs, it was interesting to see that there were variable vector backbone sizes, while some for pcDNA3.1-FL-*ARFIP2* showed smears. The smears could be indicative of DNAse contamination and therefore could also explain the different sizes of the band corresponding to the vector backbone in

the respective samples. Another reason for the variation might be that the inserts may be in different vectors. Following detecting the presence of the inserts in the constructs, all the positive colonies for both constructs were further digested with the NcoI-HF restriction to confirm an intact vector backbone. For the FL-*ARFIP2* construct, 9 out of 10 colonies showed bands at the expected heights and for the HC-*ARFIP2* construct, 2 out of 4 showed bands at the expected heights. One colony (i.e. plasmid-prep) for each of the two constructs was chosen for sequencing (A5 for FL-*ARFIP2* and B3 for HC-*ARFIP2*). They were chosen because they showed the presence of the inserts and displayed the bands for the vector backbone at the expected height initially.

Making the zebrafish constructs was a necessary step in the project to overexpress FL-ARFIP2 and also because HC-ARFIP2 is not endogenously expressed or a naturally occurring peptide that could be injected into adult zebrafish. The HC-ARFIP2 zebrafish construct made to express HC-ARFIP2 in the zebrafish transgenic line would allow us to naturally express HC-ARFIP2 with no need for an invasive procedure. Furthermore, using the zebrafish constructs, the generated transgenic model would allow us to assess the safety of overexpressing FL-ARFIP2 and expressing HC-ARFIP2 *in vivo* and also allow us to observe the long-term effects of the transgene expression.

The shRNA constructs were designed to investigate the effects of depleting ARFIP2 on disease pathogenesis in our ALS *in vitro* model. The pLVTHM plasmid (Wiznerowicz and Trono, 2003) was used for making the *ARFIP2*-shRNA constructs. It is an empty backbone lentivector that is widely used to express shRNA from the H1 promoter and also expressed GFP. This vector allows for the direct cloning of shRNA into lentiviral vectors. Results showed successful cloning of the shRNA inserts into MluI-ClaI in the pLVTHM vector backbone.

Evaluation of the three constructs in HEK293T cells showed that there was no depletion when the cells were transduced with 2μg or 4μg of each construct, compared to the untransfected control. However, when treated with 2μg of each construct in combination, there was a depletion of 50% with two combinations. This indicated that transducing with two shRNA constructs can deplete ARFIP2 to the desired levels. This may be because the multiple oligos binding to the mRNA might be more efficient for

translation suppression and thereby depleting endogenous ARFIP2. The two combinations with the presence of the 480-insert showed depletion. This may highlight that the 480 siRNA might be binding to an important site for translation. The 471 siRNA binds in the N-terminus region, 480 siRNA binds in the middle and the 749 siRNA binds in the C-terminus region. While the siRNAs tested in this chapter were designed according to the criteria mentioned in section 2.1.2.2.1 of chapter 2 and met the principles for success, whether they are able to deplete the target endogenous protein levels can only be determined empirically. The binding strength (considered by the GC content in the siRNA design) alone cannot accurately predict depletion. The location of the binding and other molecular surroundings also plays a key role.

Transfecting HEK293T cells with different amounts of the ARFIP2-shRNA construct and in combinations of two constructs was only preliminary (n=1), and therefore to significantly prove any depletion, this needed to be repeated. However, this was not done due to time constraints and project modifications as a result of the COVID-19 pandemic. In the future, transfecting with two constructs is not an efficient process. Therefore, when the most potent shRNA construct combination is validated and shows significant depletion in repeats, two shRNA sequences should be cloned into one construct, with one sequence directly downstream of the other for efficiency. Furthermore, both the sequences would be cloned under the same (H1) promotor to ensure similar expression levels of the two shRNA sequences. Once this is achieved, it would be followed by a viral prep for the construct and validation of the virus before assessing the effects of depletion of endogenous ARFIP2 in the ALS *in vitro* model.

4. ASSESSING THE EFFECTS OF EXPRESSING HC-ARFIP2 ON THE PATHOGENESIS OF C9ORF72-ALS USING IPSC-DERIVED MNS

This chapter focuses on assessing the expression and distribution of endogenous Arfaptin-2 (*ARFIP2*) in human brain tissue samples and iPSC-derived motor neurons. Additionally, it explores the effects of expressing the C-terminal of Arfaptin-2 (HC-*ARFIP2*) on disease characteristics in iPSC-derived motor neurons carrying the *C9ORF72* mutation. Three healthy control lines, three *C9ORF72*-ALS patient lines, and two isogenic control lines have been used. The effect of the HC-*ARFIP2* transgene expression on cell viability has been assessed. Furthermore, the effect of expressing HC-ARFIP2 on dipeptide repeat proteins (DPRs), RNA foci, TDP-43 mislocalisation, and proteasome activity have been evaluated.

4.1 Introduction

ALS is characterised by a high degree of aetiological heterogeneity, with various genetic factors, environmental influences and complex pathophysiologies in the multicellular environment of the ageing nervous system. Consequently, this results in substantial difficulties for *in vitro* modelling of ALS. With the development of human induced pluripotent stem cell (hiPSC) technology (Takahashi and Yamanaka, 2006; Takahashi *et al.*, 2007), our capacity to produce physiologically accurate *in vitro* models has undergone a significant transformation. Compared to other neuronal cell models, the iPSC model has a number of advantages. Firstly, since the iPSCs are reprogrammed from adult cells, during cell reprogramming, the genetic background of the donor is retained and therefore provides a great way to assess the effects of disease-causing mutations on the relevant cell type (neurons) without the need for overexpressing a transgene with the disease-causing mutation (Van Damme, Robberecht and Van Den Bosch, 2017). Secondly, with emerging gene editing technologies, it is possible to correct the ALS-causing mutations and generate isogenic control lines to directly compare the patient lines to their corrected counterpart and are effective controls to use in experiments (Kiskinis *et al.*, 2014). Finally, the iPSC model enables researching sporadic ALS cases *in vitro*, which has been difficult using other available model systems (Burkhardt *et al.*, 2013).

In this chapter, we used iPSC-derived motor neurons (MNs) as the *in vitro* model. Our model carries the *C9ORF72* hexanucleotide repeat expansion mutation, which is the most common genetic mutation in ALS. The precise function of C9ORF72 is not well known; however, evidence suggests its involvement in RNA processing, regulating endosomal trafficking, and autophagy in neurons (Farg *et al.*, 2014; Webster *et al.*, 2016; Webster *et al.*, 2018). The pathological mechanism by which the C9ORF72 mutation causes ALS remains unclear, however, haploinsufficiency of C9ORF72 (loss-of-function); translocation of HRE-containing RNA into the cytoplasm which generates RNA foci causing sequestration of RNA-binding proteins, and the production of dipeptide repeat proteins (DPRs) translated from the hexanucleotide expansion transcripts through non-ATG repeat-associated translation are all likely to contribute to mutation-dependent toxicity (van Blitterswijk, DeJesus-Hernandez and Rademakers, 2012; Robberecht and Philips, 2013).

While the exact cause of motor neuron loss is multifactorial, the formation of insoluble protein aggregates or intracellular cytoplasmic inclusions in neurons is a shared pathological feature in neurodegenerative diseases (Soto and Pritzkow, 2018). Although the exact role of these aggregates in the pathogenesis of neurodegenerative diseases is yet to be determined, researchers showed that these aggregates are indicators or causative factors in neurotoxicity. Aggregates form when the degradation capacity of a cell is exceeded or due to defects in the protein degradation pathways which include the ubiquitin-proteasome system (UPS) and autophagosome-lysosome pathway/autophagy (Johnston, Ward and Kopito, 1998; Takalo *et al.*, 2013). However, the exact mechanisms that lead to aggregation are still being studied. Therefore, maintaining cellular protein homeostasis and solubility might be key to treating these diseases. One way to achieve this balance is by maintaining proteasome activity.

Proteins found in aggregates tell us a lot about the pathogenesis of ALS and enable researchers to identify therapeutic targets for the disease. Over the years, several common constituents of protein aggregates have been identified. The TAR-DNA binding protein of 43 kDa (TDP-43) is a key component of aggregates in both sporadic and familial ALS cases (Arai et al., 2006; Neumann et al., 2006). Regardless of the disease onset and background, around 97% of all ALS cases exhibit TDP-43 positive aggregates, making it a distinguishing feature of the disease. Several studies over the years have indicated that ALS patients also almost universally display abnormal mislocalisation of the TDP-43 protein from the nucleus and into the cytoplasm, where it may build up into aggregates (Barmada et al., 2010; Suk and Rousseaux, 2020). It has been suggested that TDP-43 mislocalisation induces toxicity of its own through both loss and gain of function mechanisms. TDP-43 plays a key role in mRNA maturation processes in the nucleus and loss of these processes due to mislocalisation can have toxic cellular effects (Tollervey *et al.*, 2011; Klim *et al.*, 2019). TDP-43 also exhibits roles in the cytoplasm some of which include translation regulation, regulation of autophagy, and mRNA stability (Birsa, Bentham and Fratta, 2020). It is clear that TDP-43 mislocalization is toxic and can be a factor in several cellular characteristics seen in ALS. The processes determining TDP-43 mislocalization are, however, still largely unknown. To completely prevent TDP-43 pathology, it will be essential to uncover the processes that cause mislocalization, and strategies for the prevention of TDP-43 proteinopathy.

Previous studies have shown that a protein called ADP-ribosylation factor interacting protein 2 or Arfaptin-2 (*ARFIP2*) also colocalizes with proteins in aggregates and may be of therapeutic potential (Rangone *et al.*, 2005; Peters *et al.*, 2002). Peters et al. (2002) looked at ARFIP2 in Huntington's disease (HD) and reported that while the full-length ARFIP2 promoted aggregation, expressing only the c-terminal of the protein (HC-ARFIP2) inhibited aggregation, but not in the presence of a proteasome inhibitor. This suggested that ARFIP2 may be working by affecting the proteasome activity. Here it was shown that an increase in the concentration of ARFIP2, decreased the proteasome activity and that expressing HC-ARFIP2 dominantly rescues proteasomal activity, thereby exerting its neuroprotective effects. Rangone et al. (2005) showed that the neuroprotective effects of full-length ARFIP2 depend on its phosphorylation state, and its phosphorylation promotes survival and decreases aggregation. Although the two studies described different approaches by which ARFIP2 is neuroprotective, it was noteworthy that the phosphorylation site is on the C-terminal of the protein, therefore the pathway by which ARFIP2 is neuroprotective may be the same. Since protein aggregation is a common feature in neurodegenerative disorders, therapeutic strategies may be transferrable. Therefore, previously, our lab looked into ARFIP2 in ALS. Preliminary data in our lab showed that HC-ARFIP2 significantly enhanced survival in primary motor neurons (Mohammedeid, 2015).

Accordingly, based on the findings in previous studies that HC-ARFIP2 can have neuroprotective effects and can regulate proteasome activity, as well as the preliminary data from our lab in the context of ALS, we hypothesized that ARFIP2 may be a novel therapeutic target for ALS, targeting of which may ameliorate the disease phenotypes in *C9ORF72*-ALS iPSC-derived MNs.

4.2 Aim

The work described in this chapter aims to assess the effects of HC-ARFIP2 on the disease pathogenesis of *C9ORF72*-ALS in iPSC-derived MNs and identify whether HC-ARFIP2 has therapeutic potential *in vitro*. Firstly, the expression and distribution of endogenous ARFIP2 in ALS biosamples, which include human brain tissue sections and iPSC-derived MNs, have been measured. Secondly, since our *in vitro* model carries the *C9ORF72* genetic mutation, the effect of expressing HC-ARFIP2 on

pathological hallmarks of DPRs and RNA foci has been assessed. Finally, the chapter shows the effects of expressing HC-ARFIP2 on proteasome activity in iPSC-derived MNs as well as the effect of HC-ARFIP2 on the TDP-43 mislocalisation pathology widely seen in ALS. Furthermore, the effect of expressing HC-ARFIP2 on cell viability has been evaluated.

4.3 Results

4.3.1 Endogenous ARFIP2 levels are elevated in the ALS disease background

Previous literature exploring ARFIP2 in HD had shown that ARFIP2 expression levels were increased in disease when compared to controls. With this in perspective, the expression levels of ARFIP2 were measured in the cerebellum and frontal cortex sections from ALS and healthy control human brain tissue and iPSC-derived MNs.

4.3.1.1 Endogenous ARFIP2 levels in human brain tissue samples

The expression and distribution of endogenous ARFIP2 in the cerebellum and frontal cortex of control, sALS and *C9ORF72*-ALS cases was assessed by immunohistochemistry (2.1.2.22), blinded to relevant clinical information.

The omission of the primary antibody produced minimal nonspecific immunoreactivity in both regions of interest, thereby confirming the specificity of the immunoreactive profile observed with the ARFIP2 primary antibody (**Figure 4.1 A and E**).

Immunoreactivity for ARFIP2 was predominantly associated with cytoplasmic staining of Purkinje cells in the cerebellum (**Figure 4.1 A-B**) and pyramidal neurons in the frontal cortex (**Figure 4.1 E-H**). The immunoreactive profile was qualitatively assessed across the cohort. In the cerebellum of *C9ORF72*-ALS patients (**Figure 4.1 C**), ARFIP2 immunoreactivity appears more intense, especially in the Purkinje cells (as indicated by the black arrows in the panel) than compared to healthy controls (**Figure 4.1 B**). A similar pattern of staining in the cytoplasm and neuronal cell body of Purkinje cells is observed in the cerebellum of the sALS patients (**Figure 4.1 D**) compared to controls; however, the immunoreactivity appears to be less intense compared to the cerebellum of *C9ORF72*-ALS patients. Furthermore, the immunoreactivity of the granular layer of the cerebellum in *C9ORF72*-ALS cases also appears to be more intense compared to healthy controls and sALS cases, whereas the immunoreactivity of the granular layer in sALS cases is similar to healthy controls.

In the frontal cortex, ARFIP2 immunoreactivity was associated with cells morphologically resembling pyramidal neurons (indicated by the red arrows in the panel) and glial cells (indicated by blue arrows) as shown in **Figure 4.1 E-H**.

Qualitative assessment of the immunoreactive profile suggests that ARFIP2 expression is more extensive in the frontal cortex of C9ORF72 and sALS cases (**Figure 4.1 G and H**) compared to healthy controls (**Figure 4.1 F**). The immunoreactivity of ARFIP2 appears to be most intense in the nucleus and cytoplasm of pyramidal neurons as well as surrounding glia in sALS cases (**Figure 4.1 H**). The immunoreactive profile of ARFIP2 in *C9ORF72*-ALS cases (**Figure 4.1 G**) closely resembles that seen in sALS (**Figure 4.1 H**) but appears less intense.

Overall the results suggest that endogenous ARFIP2 is expressed at higher levels in the cerebellum of *C9ORF72*-ALS cases, particularly in the Purkinje cells and the granular layer, and in the pyramidal neurons and surrounding glia of the frontal cortex of both *C9ORF72*-ALS and sALS cases. These results offer support to our hypothesis for targeting this protein in ALS.

	Cerebellum	Frontal Cortex
Secondary Ab only	A	E
Healthy controls	B	F
C9ORF72-ALS	C	G
sALS	D	H
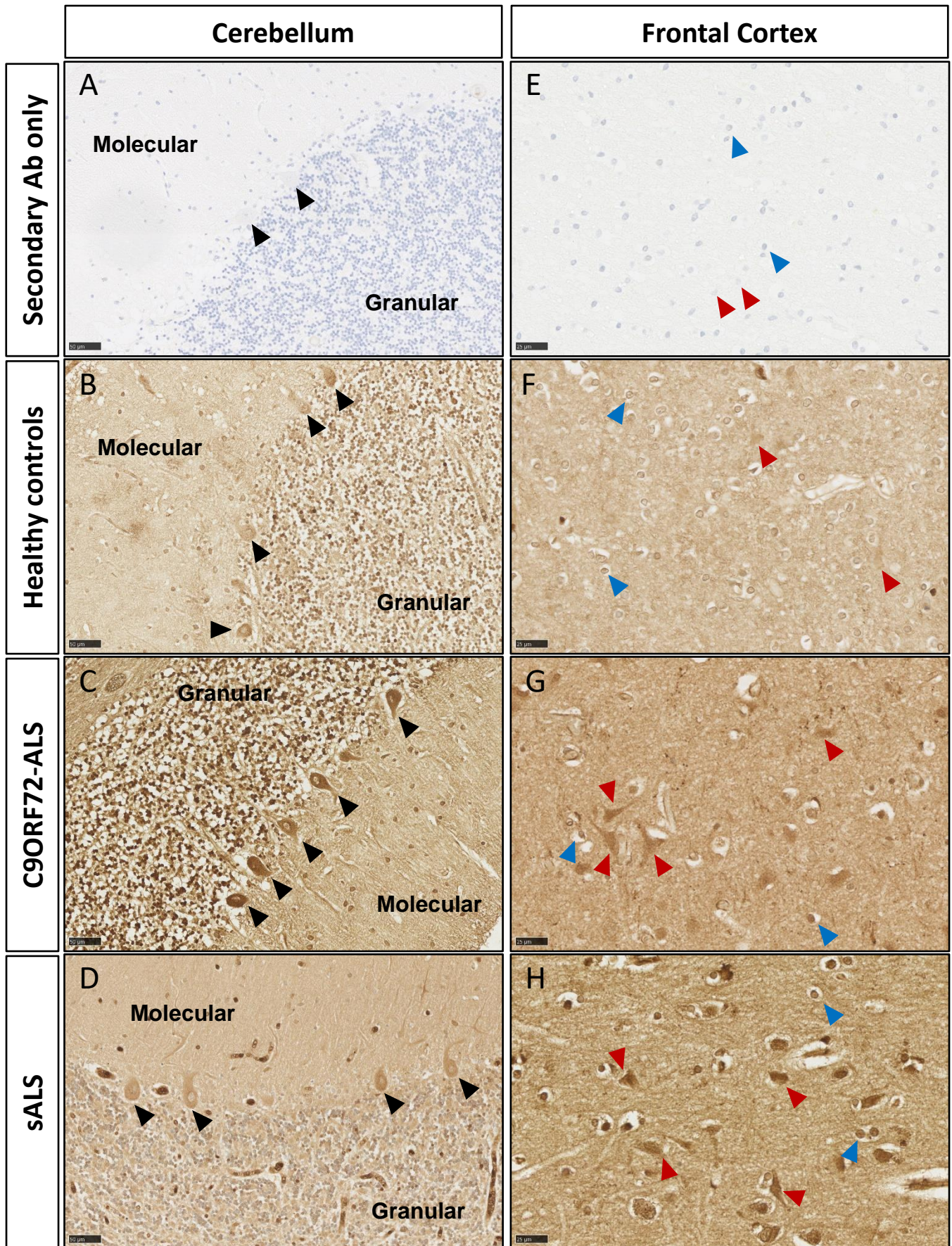
Figure 4.1: Representative images showing that endogenous ARFIP2 is more extensively expressed in the cerebellum of *C9ORF72*-ALS and the frontal cortex of sporadic-ALS cases

Human cerebellum (**A-D**) and neocortex – frontal lobe (**E-H**) tissue sections from healthy cases, sporadic ALS patients (n=6) and ALS patients with the C9ORF72 mutation (C9ORF72-ALS) (n=6) were assessed for endogenous ARFIP2 expression. The panel shows representative ARFIP2 staining of the cerebellum (40x magnification) and frontal cortex sections (80x magnification) across the cohort. The C9ORF72-ALS cerebellum (**C**) contained a higher level of ARFIP2 expression, especially in the cytoplasmic compartment of Purkinje cells (indicated by black arrows). ARFIP2 expression was associated with the pyramidal neurons (indicated by red arrows) and glial cells (indicated by blue arrows) in the frontal cortex and appeared to be more extensively expressed in the sALS cases (**H**). Scale bar=50 µm for cerebellum and 25 µm for frontal cortex.

4.3.1.2 Endogenous ARFIP2 levels in iPSC-derived MNs

With altered levels of endogenous ARFIP2 observed in ALS using human brain tissue, endogenous levels of ARFIP2 in iPSC-derived MNs were assessed to see if the result was replicated in the *in vitro* model of human motor neurons. Human iPSCs were differentiated into functional mature MNs (2.2.2.2), and the total protein was extracted on day 40 of the protocol (2.1.2.17). Using a western blot (2.1.2.20), the endogenous levels of ARFIP2 were detected in healthy control cell lines (n=3), *C9ORF72*-ALS cell lines (n=3), and one isogenic cell line which was the only available isogenic control line at the time (n=1) (**Figure 4.2**). The representative blot shows endogenous full-length ARFIP2 bands at 38.6 kDa and an α-Tubulin band at 50 kDa which was used as a loading control (**Figure 4.2 A**). The blot shows that the ARFIP2 bands for the ALS cell lines were increased in intensity compared to those for the healthy control lines and the isogenic control line, while the total amount of protein was equally loaded. Densitometric analysis of the western blots from 4 experimental repeats was undertaken and the endogenous ARFIP2 levels were normalised to the loading control (**Figure 4.2 B**). In the analysis, a measurement for one of the ALS lines (ALS 29) in one of the repeats was identified as a clear outlier and, hence, this data point was omitted from the analysis. Identification of outliers was conducted on GraphPad using the ROUT method with the lowest Q value of 0.1% (Q value measures the aggressiveness of ROUT and the lower the Q value, the stricter the identification and removal of definitive outliers). Analysis showed that the endogenous levels of

ARFIP2 were significantly increased in *C9ORF72*-ALS iPSC-derived MNs when compared to the healthy control cell lines and an isogenic control without the ALS disease background. The expression levels of endogenous ARFIP2 were normalised to the control cell lines to assess the percentage change in levels across the different conditions (**Figure 4.2 C**). Analysis showed significantly elevated endogenous ARFIP2 levels in *C9ORF72*-ALS iPSC-derived MNs compared to healthy controls and an isogenic control. The ARFIP2 levels were observed to be significantly elevated in ALS by approximately 86% compared to controls and by 112% compared to the isogenic control.
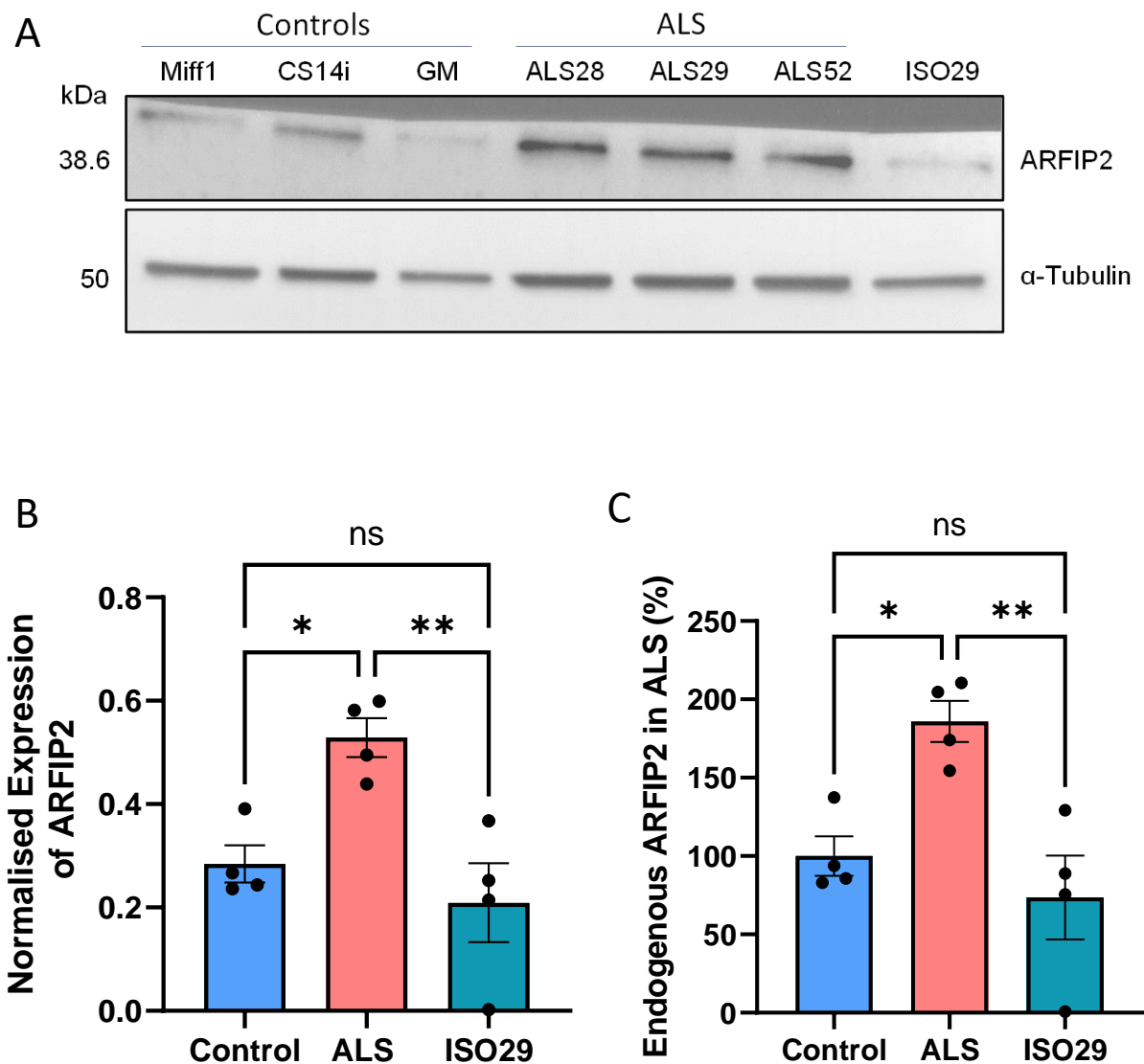
Figure 4.2: Endogenous ARFIP2 levels are elevated in iPSC-derived MNs from *C9ORF72*-ALS patients

iPSC derived MNs were lysed and endogenous ARFIP2 was detected on **(A)** a western blot probed with anti-ARFIP2 specific antibody (1:500). For each cell line, 20µg of total protein was loaded into each well. α-Tubulin was used as a loading control (50 kDa). **(B)** The ARFIP2 expression levels were compared between samples after normalisation with α-Tubulin. Densiometric analysis showed that the endogenous ARFIP2 levels in C9ORF72-ALS iPSC-derived MNs (mean=0.53, n=3) were significantly increased compared to healthy controls (mean=0.28, n=3). When compared to the isogenic control without the ALS background (0.21, n=1), a significant decrease in ARFIP2 levels was observed when compared to the C9ORF72-ALS iPSC-derived MNs. **(C)** The expression levels were normalised to the controls. Analysis showed significantly elevated endogenous ARFIP2 in C9ORF72-ALS iPSC-derived MNs (mean=186%; n=3) compared to healthy controls (mean=100%; n=3) and an isogenic control for an ALS line (74%, n=1). The ARFIP2 levels were significantly elevated in ALS by approximately 86% compared to healthy controls. There was a significant increase in endogenous ARFIP2 levels by 112% in ALS cell lines compared to the isogenic control. N=3 with 4 experimental repeats. Data were analysed using an ordinary one-way ANOVA with Tukey's multiple comparisons test; Data presented as mean ± SEM; $p<0.05$. ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$. ISO29 = Isogenic control for ALS29 line. Total sample size needed=12 and actual power=0.9510229.

4.3.2 Effect of HC-ARFIP2 expression on cell viability in C9ORF72-ALS iPSC-derived MNs

We investigated cell viability in our *in vitro* model upon transduction with LV-HC and LV-*GFP* using a resazurin reduction assay (2.2.2.12) to assess the effects of expressing HC-ARFIP2 on cell viability in *C9ORF72*-ALS iPSC-derived MNs. Cell viability was assessed over two weeks post-transduction on day 40 of the differentiation protocol (2.2.2.2) (**Figure 4.3**). For each cell line from each condition, the relative fluorescence unit (RFU) was measured. A low RFU is indicative of low metabolic activity and therefore low cell viability as living cells are metabolically active. The RFU for each cell line under each condition was plotted against days post-transduction (**Figure 4.3 A-E**). The graphs for each condition show the cell viabilities of the cell lines over the period of time of the experiment. Results showed that the untransduced ALS cells and the ALS cells transduced with LV-*GFP* had the lowest cell viability across the measurements taken at all time points when compared to the control cell lines. However, when the cells were treated with LV-HC, they were

healthier over time than the untransduced ALS cells or those treated with the control virus. All the measurements were normalised to the first time point (4 days post-transduction) to see the percentage change of the cell viabilities over time (**Figure 4.3 F**). Results showed that the untransduced ALS cell lines showed the steepest decline in cell viability compared to other conditions. Whereas the LV-HC transduced ALS cell lines showed stable cell viability across all measurements with a slight increase over time. This trend, as well as measurements observed for LV-HC transduced ALS cell lines, were consistent with untransduced controls and LV-HC transduced control cell lines. **Figure 4.3 G** shows the statistical analysis for the percentage change of cell viabilities over time. There is a significant difference in cell viability between untransduced controls and untransduced ALS cell lines. With LV-HC transduction the cell viability of ALS cell lines was improved, and results showed a significant difference in cell viability between ALS untransduced cells and those transduced with LV-HC. There was no significant difference between untransduced controls and LV-HC transduced ALS cell lines. Overall the results showed that HC-ARFIP2 expression improved cell viability in ALS cell lines.
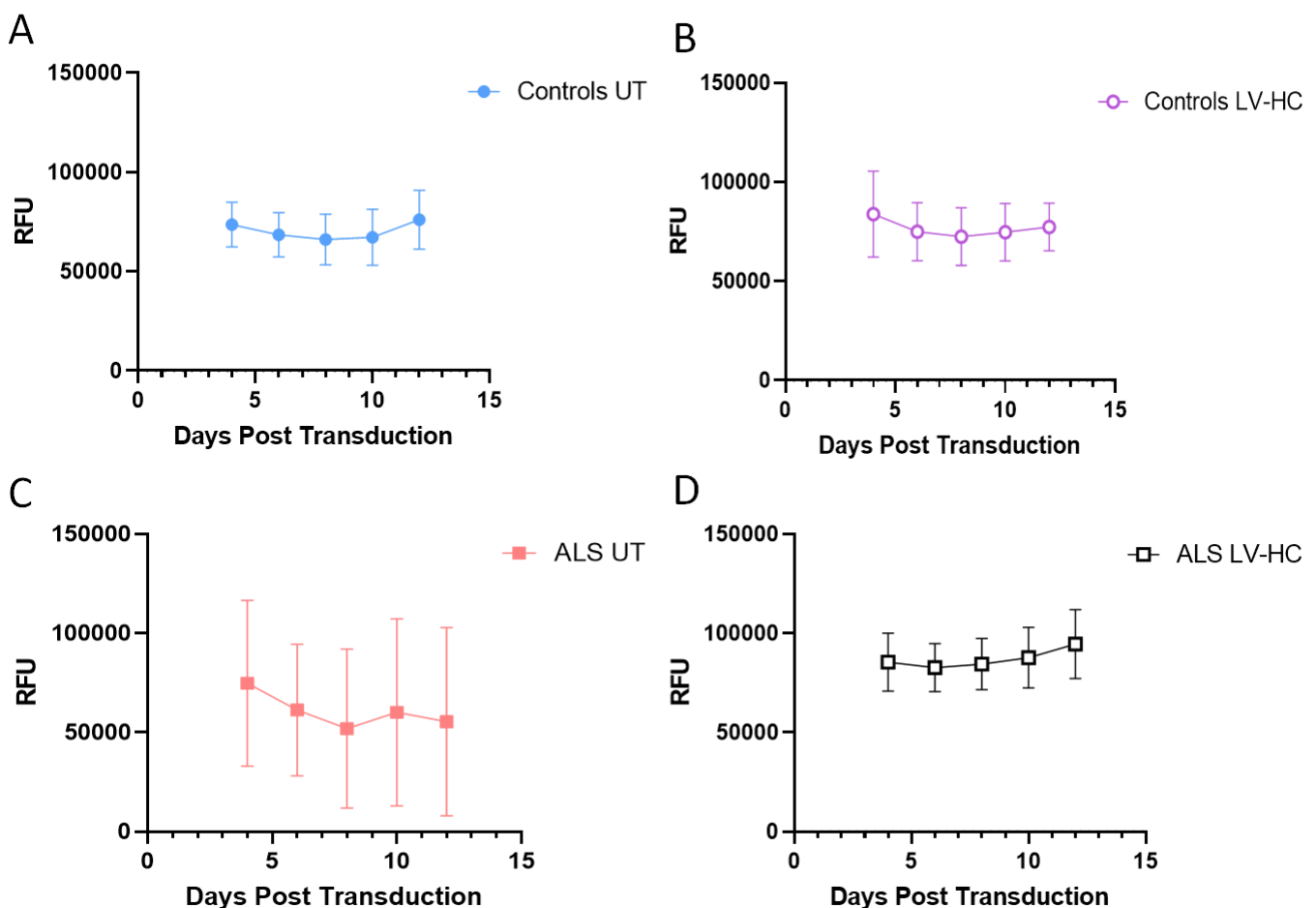
E
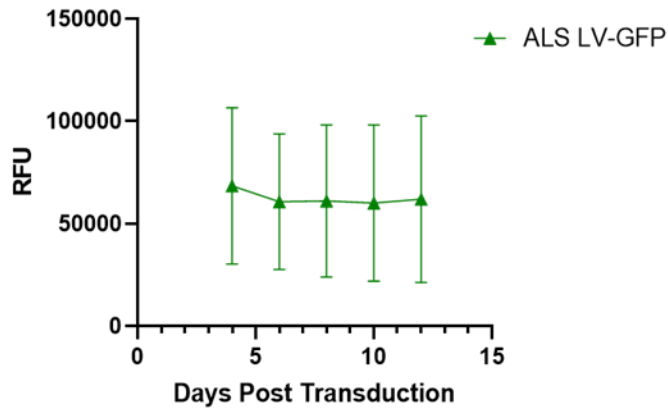
F
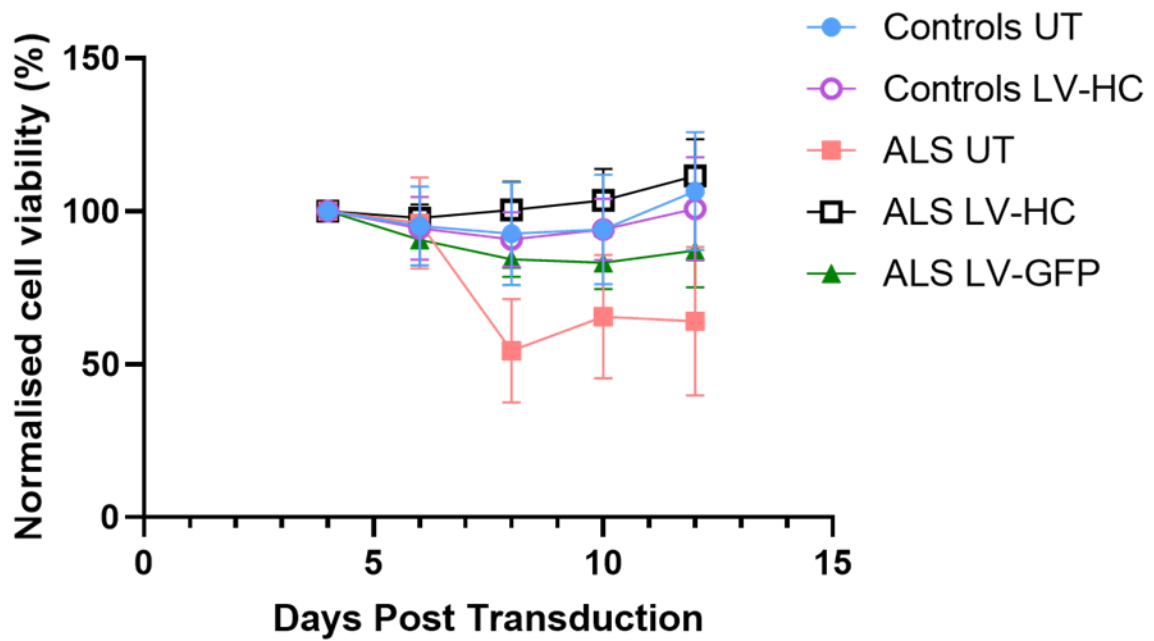
G

Controls UT vs. Controls LV-HC	ns
Controls UT vs. ALS UT	*
ALS UT vs. ALS LV-HC	*
Controls UT vs. ALS LV-HC	ns
ALS UT vs. ALS LV-GFP	ns

Figure 4.3: Effect of expressing HC-ARFIP2 on cell viability in iPSC-derived MNs

iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC or LV-GFP. Post transduction, the cells were assessed for cell viability using a resazurin reduction assay. Measurements were taken across a span of two weeks after transduction. **(A-E)** The graphs show the measured relative fluorescence units (RFU) for all cell lines in each condition over time. A phenotype of ALS cell lines having a lower cell viability compared to control untransduced cell lines was seen. It was shown that the ALS cell lines transduced with LV-HC are healthier over time than the untransduced ALS cells or those treated with the control LV-GFP virus. **(F)** The measurements taken over the two weeks were normalised to the first measurement taken at 4 days post transduction. This graph showed how the cell viability was affected over time. It was observed over two weeks that there was no significant difference in cell viability between untransduced controls and transduced controls, and untransduced ALS cell lines and ALS cell lines transduced with LV-GFP. ALS untransduced cells and those treated with LV-GFP had the lowest cell viability. However, when treated with LV-HC, there was an increase in cell viability and a significant difference between untransduced ALS cells and transduced ALS cells was observed. **(G)** The table shows the significance for the graph where the measurements over time are normalised to the first time point. Data were analysed using a two-way ANOVA with a Tukey's multiple comparisons test. N=3 with 1 experimental repeat. Data presented as mean ± SEM; $p < 0.05$. ns = $p > 0.05$, * = $p ≤ 0.05$, ** = $p ≤ 0.01$, *** = $p ≤ 0.001$.

After assessing metabolic activity as a proxy for cell viability after transduction with LV-HC, a live/dead assay (2.2.2.13) was conducted to assess the toxicity of transduction and HC-ARFIP2 expression (**Figure 4.4**). Post conducting the resazurin reduction assay on iPSC-derived MNs, after the last measurement on day 12 post-transduction, the cells were stained with SYTO9 to stain for live cells and propidium iodide (PI) to stain for dead cells (2.2.2.13) and the cells were imaged (2.2.2.14.1). The panel with all cell lines under each condition shows that there is minimal cell death, even two weeks after transduction with LV-HC or LV-*GFP*.
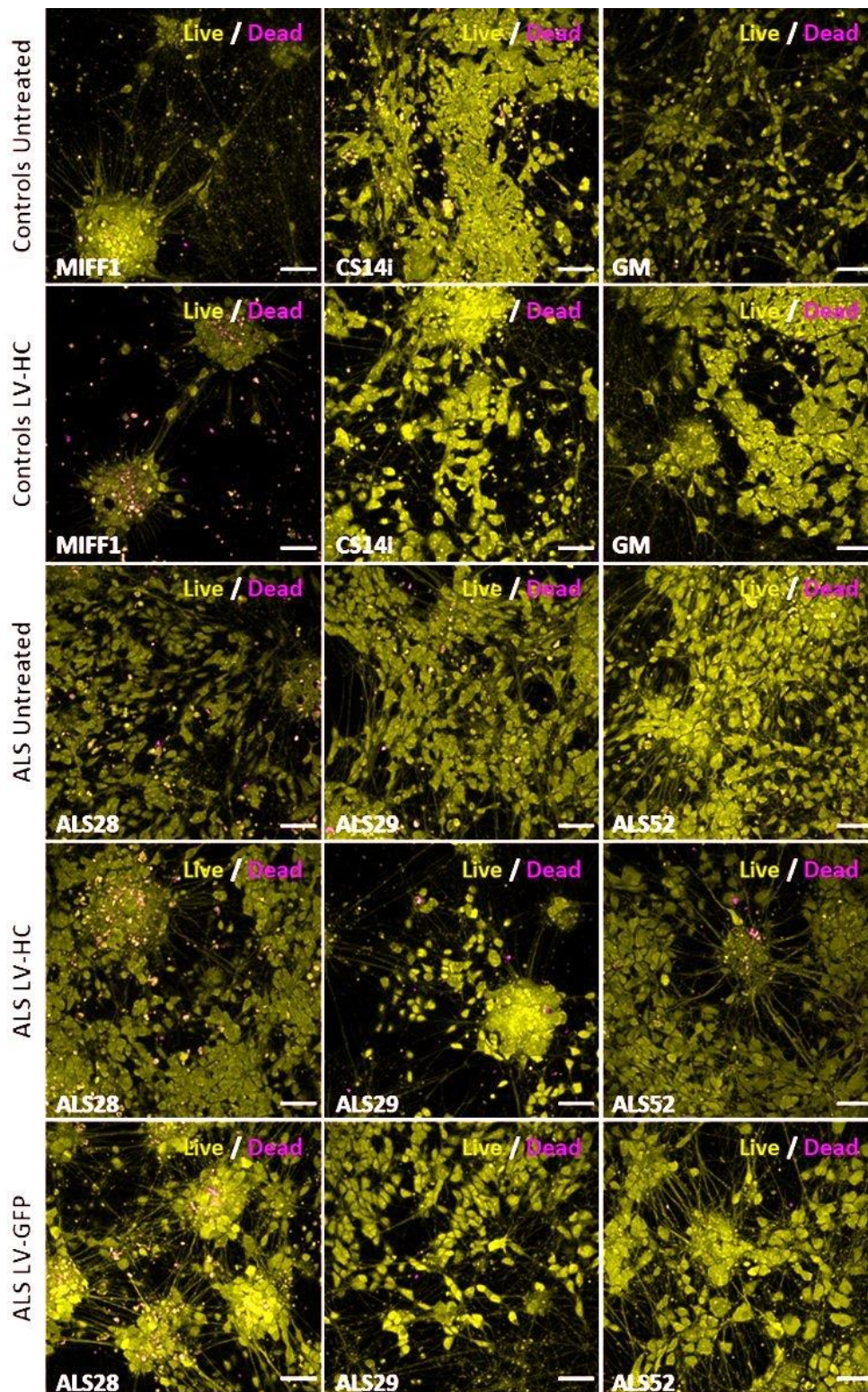
Figure 4.4: Effect of expressing HC-ARFIP2 on cell viability in iPSC-derived MNs two weeks after transduction

Post conducting the resazurin reduction assay on iPSC derived MNs, after the last measurement on day 12 post transduction, the cells were stained with 0.02% SYTO9 to stain for live cells and propidium iodide (PI) to stain for dead cells. The panel shows a representative image for each cell line in each condition. The yellow channel shows live cells, while the magenta channel shows dead cells. Scale bar = 50 µm.

4.3.3 HC-ARFIP2 expression reduces poly-GP DPR levels

The iPSC-derived MNs used in the project carry the *C9ORF72* mutation, and a characteristic of this mutation is the generation of five different pathological dipeptide repeats (DPRs) from both the sense and antisense RNA transcripts. The poly-GP DPR was chosen for assessment as this DPR is generated from both the sense and antisense transcripts. Poly-GP DPRs, therefore, serve as a proxy for both the sense and antisense DPRs. To assess the effects of HC-ARFIP2 on poly-GP levels, iPSC-derived MNs were cultured until day 40 (2.2.2.2) at which point they were either left untransduced or were transduced with the LV-HC or LV-*GFP* as a control virus (2.2.2.6.1). The cells were harvested 4 days post-transduction (2.1.2.17) and a poly-GP electrochemiluminescence (MSD) ELISA was conducted (2.1.2.21) to determine the levels of the poly-GP DPRs in *C9ORF72*-ALS iPSC-derived MNs (**Figure 4.5**). Standard curves were generated using known concentrations of standards and their raw response values (**Figure 4.5 A**). The standard curves were used to interpolate the amounts of poly-GP DPRs for each condition (**Figure 4.5 B**). It was observed that the poly-GP DPRs were significantly detected in ALS untransduced cells (0.04 ng/ml) compared to controls (0.01 ng/ml). Since the controls do not have the expansion mutation and therefore any DPRs, any signal or amount interpolated for control lines was inferred as assay background. Therefore, all results were normalised to the untransduced controls (**Figure 4.5 C**). It was observed that with HC-ARFIP2 expression, the poly-GP DPR levels in LV-HC transduced ALS lines were significantly decreased and returned to the levels in control cell lines. The reduction in poly-GP DPRs seen with transduction of LV-HC in the ALS cell lines, after normalising to the assay background, represented an approximately 92% decrease compared to ALS untransduced lines (**Figure 4.5 D**). Results from the 4 experimental repeats showed that expressing HC-ARFIP2 reduces the levels of the poly-GP DPR in *C9ORF72*-ALS iPSC-derived MNs.

A

B

C

D
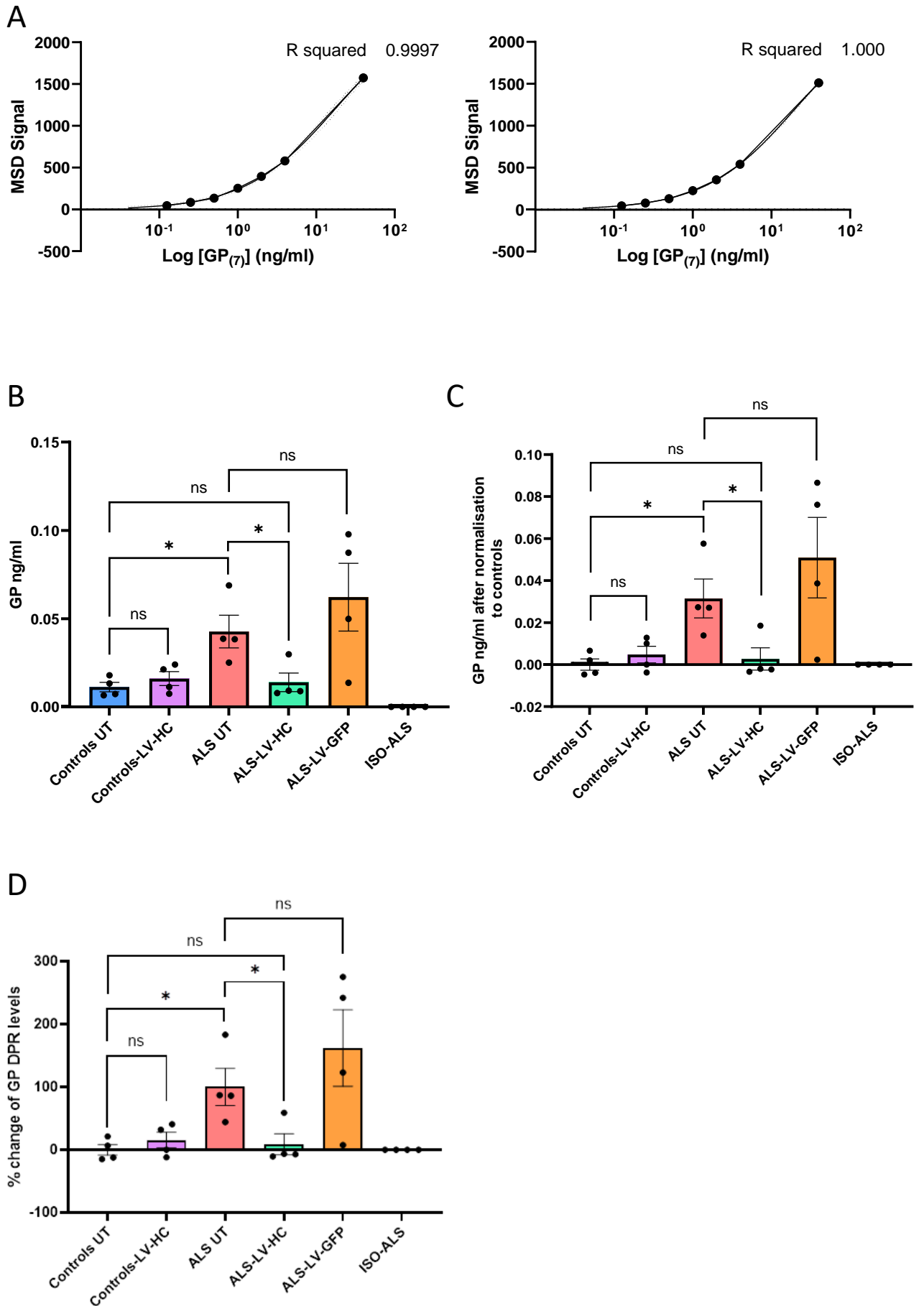
Figure 4.5: Expression of HC-ARFIP2 reduces poly-GP DPR levels in *C9ORF72*-ALS iPSC-derived MNs.
iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC-ARFIP2 or LV-GFP. MNs were lysed, and total cell lysate was used to conduct a poly-GP MSD ELISA. **(A)** Raw response values were plotted against known concentrations of standards to generate standard curves. **(B)** The amounts of poly-GP DPRs for each condition were interpolated from the standard curves. A significant amount of poly-GP DPRs was observed in ALS untreated lines (mean=0.04 ng/ml) compared to untreated controls (mean=0.01 ng/ml). With treatment of the ALS lines with LV-HC-ARFIP2, a significant decrease in poly-(GP) DPR levels was observed (mean=0.01 ng/ml), highlighting a therapeutic potential of HC-ARFIP2. There was an increase in poly-GP DPRs with treatment with LV-GFP (mean=0.06 ng/ml), though this increase was not significant. **(C)** Since poly-GP DPR levels were detected in untransduced controls, all results were normalized to the untransduced controls. It was observed that with LV-HC expression, the poly-GP DPR levels were returned to control levels. **(D)** The percentage decrease between the ALS untreated line and ALS lines treated with LV-HC is around 92%. N=3; data shown from 4 experimental repeats. Data were analysed using an unpaired t-test. Data presented as mean ± SEM; $p < 0.05$. ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.UT = untransduced. ISO ALS = Isogenic controls for ALS lines.

4.3.4 Effect of expressing HC-ARFIP2 on RNA foci distribution

The ALS iPSC lines used have been reported to show the expected nuclear RNA foci pathology in *C9ORF72*-associated ALS. Using in situ hybridisation (2.2.2.9), the presence of RNA foci in our *in vitro* model was assessed. iPSC-derived MNs were cultured until day 40 (2.2.2.2) when they were either left untransduced or transduced with either LV-HC or LV-*GFP* (2.2.2.6.1) (**Figure 4.6**). The cells were fixed 4 days post-transduction (2.2.2.7) and probed with LNA probes specific for either sense (**Figure 4.6 A**) or antisense (**Figure 4.6 B**) RNA foci. An anti-HA-tag antibody was used to select cells transduced with LV-HC and an anti-MAP2 antibody was used to visualise the MAP2 neuronal marker. The plates were imaged on the Opera Phoenix High Content Imaging system (2.2.2.14.2) and the images are presented in highlight mode. The panels (**Figure 4.6 A and B**) show the presence of nuclear RNA foci in all ALS lines in the untransduced, LV-HC transduced and LV-*GFP* transduced conditions (as shown in the insets where the white arrows point to the nuclear RNA foci and the red arrows point to examples of background staining).

The detection of RNA foci was analysed using set parameters (detailed in **Appendix 2**) and the percentage of cells with sense or antisense nuclear RNA foci were compared (**Figure 4.7A and C**). Additionally, the breakdown and burden by RNA foci per cell was measured (**Figure 4.7 B and D**). RNA foci analysis was conducted with the help of Allan Shaw. RNA foci analysis showed that in the untransduced controls, isogenic control lines and the RNAse-treated ALS cell lines, at least 1% of the cells were identified to have RNA foci present. Therefore 1% of the cells with RNA foci were considered as background. Stricter parameters were not used as those that were clearly foci were not being identified and, therefore, exclusion of background was not possible.

Analysis of the sense RNA foci showed that there were no significant differences between the different conditions (**Figure 4.7 A**). The ALS lines did not show the presence of sense RNA foci. The percentage of cells detected with sense RNA foci were all either equal to or below the 1% threshold for background. The graph showing the sense RNA foci breakdown and burden (**Figure 4.7 B**) showed similar results, where the detection of sense RNA foci was below the threshold for background.

Analysis of the percentage of cells with antisense RNA foci showed that the antisense RNA foci were more abundant and detectable in the ALS cell lines than the sense RNA foci (**Figure 4.7 C**). There was a significant increase observed in the percentage of cells with antisense foci in ALS untransduced cells (4%), LV-HC ALS transduced cells (4.3%) and LV-*GFP* ALS transduced cells (4.8%), compared to the untransduced controls (1%). There were no significant differences between the three ALS conditions. The isogenic lines showed that 1.2% of the cells had antisense RNA foci and the RNAse-treated ALS cell lines showed that 0.9% of the cells had antisense RNA foci, which is similar to the control background threshold (1%). This offered support to results, as well as to the foci detection methods and parameters used. The analysis demonstrated that the cells displayed the antisense RNA foci phenotype in the ALS cell lines. The graph showing the antisense RNA foci breakdown and burden showed that all ALS cell lines were identified to have antisense RNA foci and that most cells typically had 1 antisense RNA foci (**Figure 4.7 D**). When comparing the foci burden of 1, there was a significant difference in the percentage of cells with 1 antisense RNA foci between the untransduced controls (1.2%) and all the conditions with ALS cell

lines (ALS UT = 2.8%, ALS LV-HC = 2.9% and ALS LV-*GFP* = 2.9%), except for when the ALS cell lines were treated with RNAse (0.83%). There were no significant differences between the three ALS conditions – ALS UT, ALS LV-HC, and ALS LV-*GFP*. For antisense RNA foci, a foci burden of 2-4 or 5 or more antisense foci was only seen for the three ALS conditions and there were no significant differences between them. 2-4 antisense RNA foci were detected in 0.9%, 1.1%, and 1.4% of the cells, and 5 or more antisense foci were detected in 0.3%, 0.4%, and 0.4% of the cells in untransduced ALS cells, LV-HC transduced ALS cell and LV-*GFP* transduced ALS cells, respectively. The average number of cells for each condition that were analysed are depicted in **Figure 4.7 E** and **F** for sense and antisense RNA foci, respectively. There were no significant differences observed in the average cells/well that were analysed between the different conditions for sense and antisense RNA foci.

These results show that sense RNA foci were not detected and only the antisense RNA foci were detectable and quantifiable in the *C9ORF72*-ALS *in vitro* model used.

A
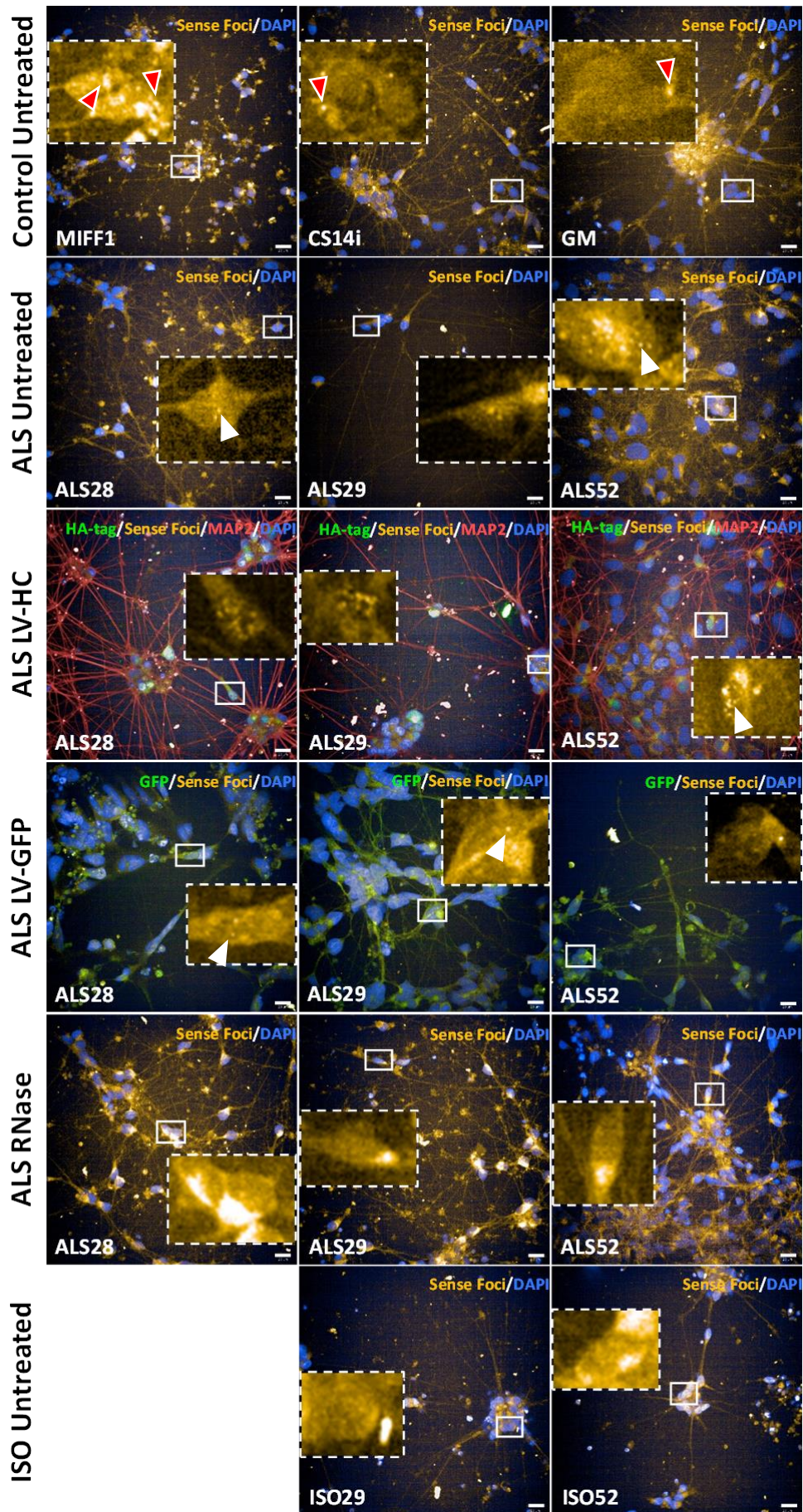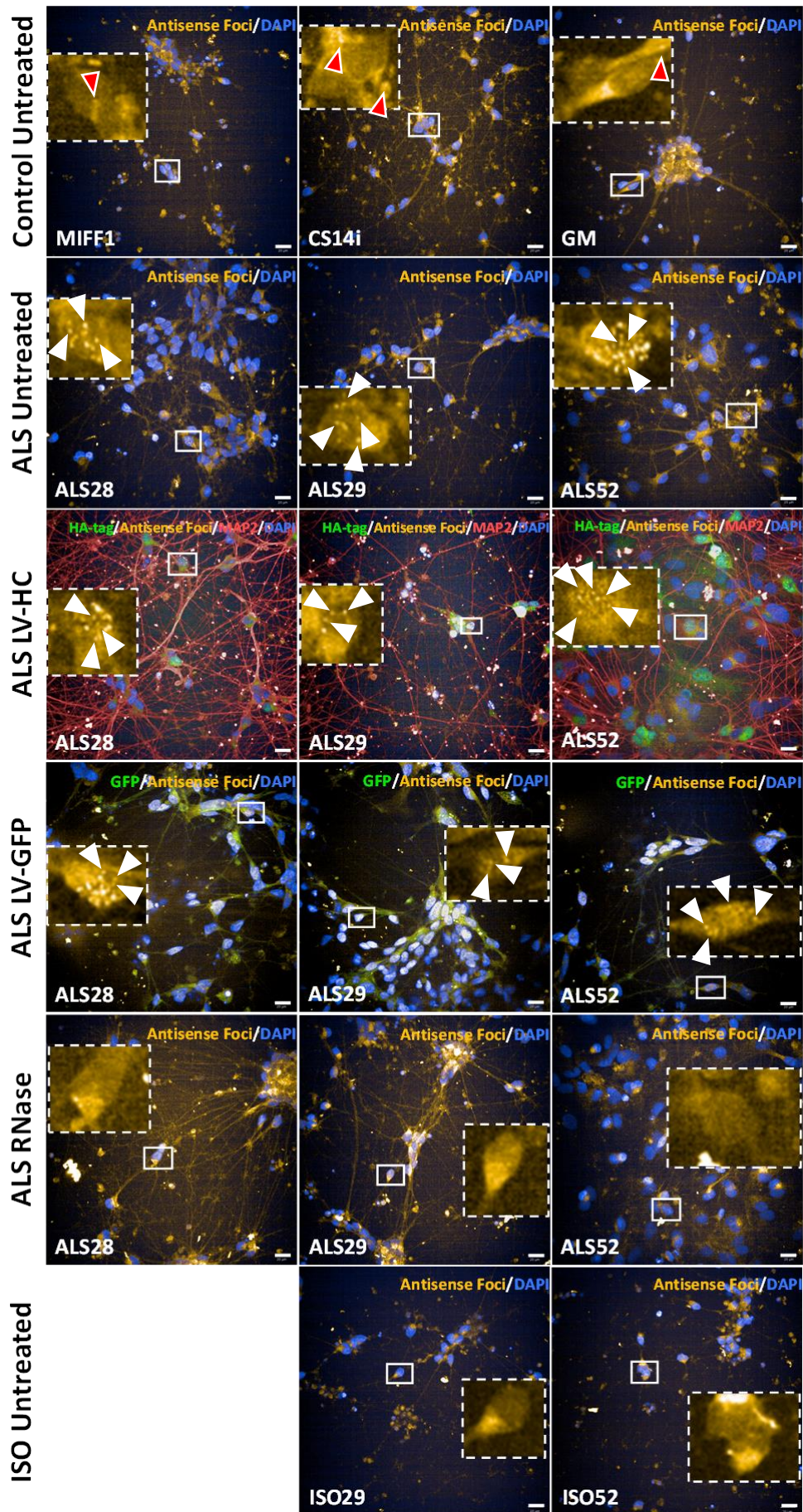
Figure 4.6: Detection of sense and antisense RNA foci in *C9ORF72*-ALS iPSC-derived MNs.

iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC or LV-GFP. MNs 4 days post transduction were fixed in 4% PFA and probed with locked nucleic acid (LNA) probes specific for **(A)** sense and **(B)** antisense RNA foci (1:400) using in situ hybridisation. An anti-HA-tag antibody (1:500) was used to select cells transduced with LV-HC and an anti-MAP2 antibody (1:1000) was used to visualise the MAP2 motor neuron marker. **(A-B)** The panels show representative images for each cell line under each condition. The nucleus was stained using Hoechst. Insets show zoomed images of one cell where white arrows show RNA foci and red arrows show background. Scale bar = 20µm.

E

F
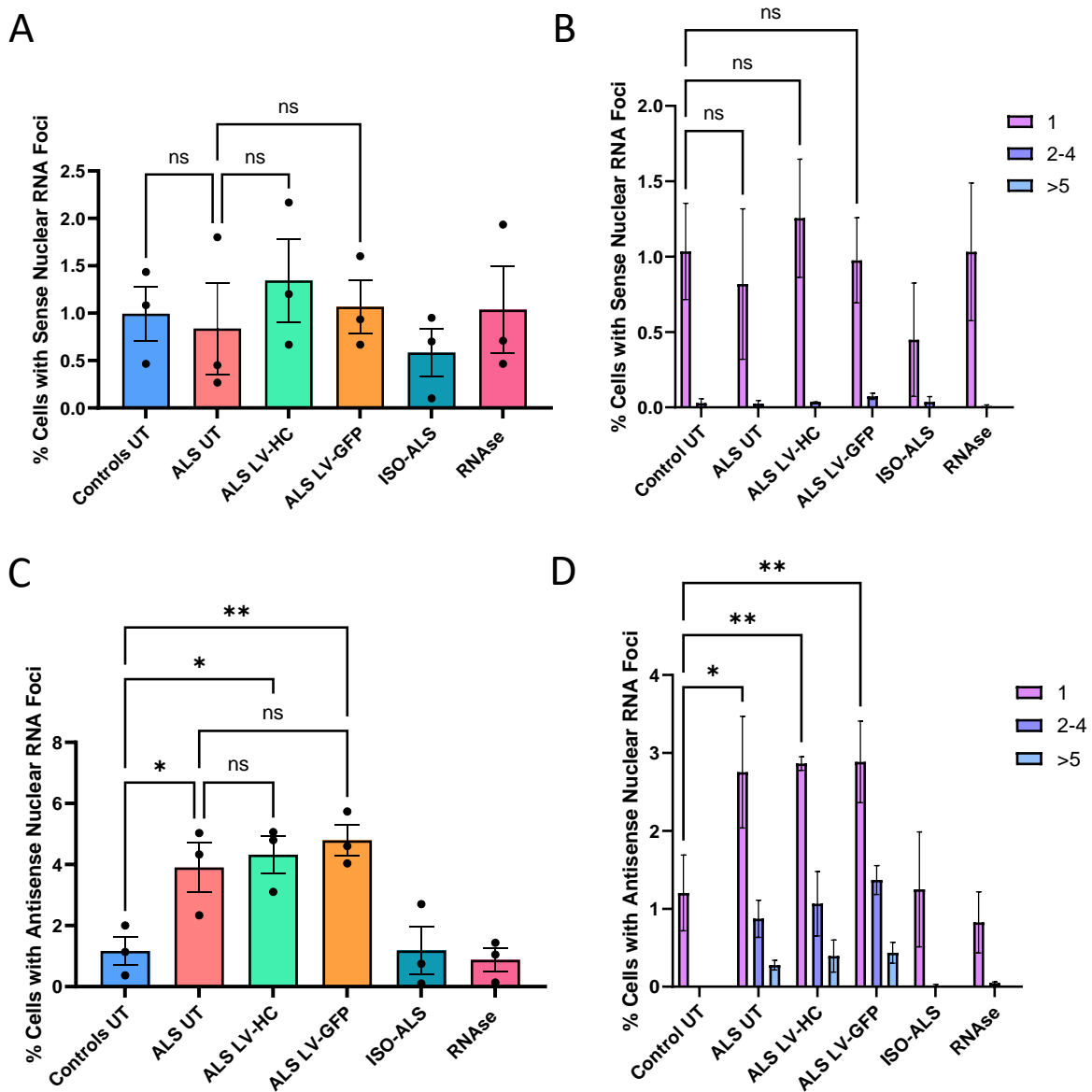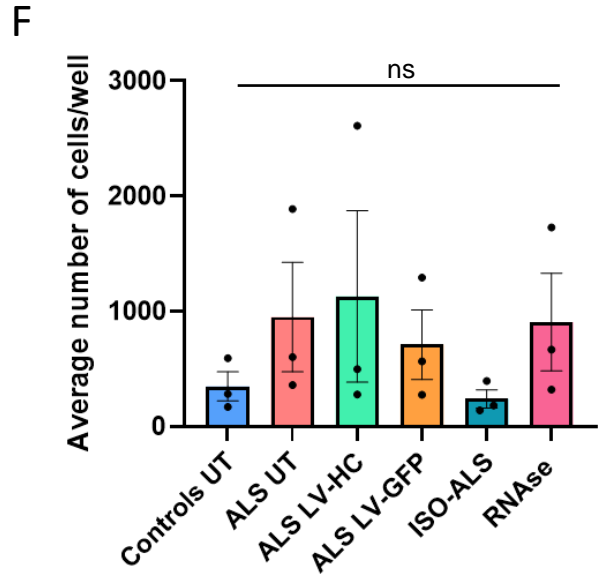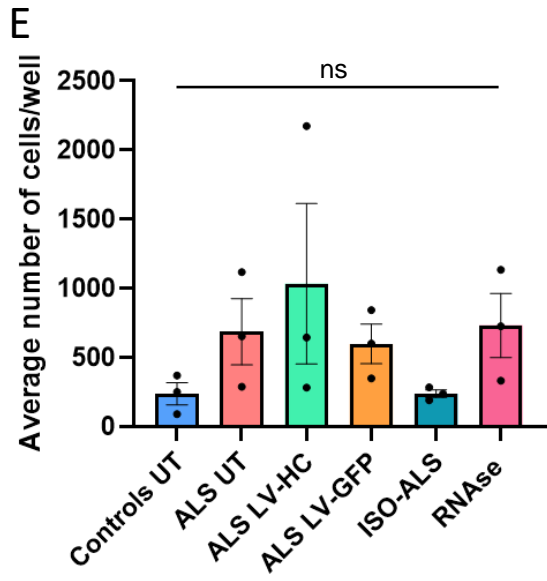
Figure 4.7: The effect of expressing HC-ARFIP2 on sense and antisense RNA foci in *C9ORF72*-ALS iPSC-derived MNs.

iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC or LV-GFP. **(A)** The graph shows the quantification of the percentage of cells with sense foci for every condition. There is no significant difference in the percentage of cell with sense foci between each condition (Controls UT = 0.96%, ALS UT = 0.84%, ALS LV-HC = 1.34%, ALS LV-GFP = 1.07%, ISO-ALS = 0.58%, RNAse = 1.04%). **(B)** Histogram showing the breakdown and burden by sense foci per cell. There is no significant difference in sense foci burden (cells with 1, 2-4 or 5+ foci) between each condition. **(C)** The graph shows the quantification of the percentage of cells with antisense foci for every condition. There is no significant difference in the percentage of cell with antisense foci between each condition (Controls UT = 0.89%, ALS UT = 3.9%, ALS LV-HC = 4.32%, ALS LV-GFP = 4.79%, ISO-ALS = 1.18%, RNAse = 0.87%). **(D)** Histogram showing the breakdown and burden by antisense foci per cell. There is a significant difference in the antisense foci burden (only between cells with 1 foci) between untransduced controls (0.93%) and the three ALS conditions (ALS UT = 2.75%, ALS LV-HC = 2.86% and ALS LV-GFP = 2.88%). There is no significant difference between the ALS conditions. **(E)** The graph shows the average number of cells analysed for detection of sense RNA foci for untransduced controls (mean=237.3), untransduced ALS lines (mean=686.8), LV-HC transduced ALS lines (mean=1034), LV-GFP transduced ALS lines (mean=597.3), isogenic ALS lines (mean=237.8) and RNAse treated ALS lines (mean=731.4). **(F)** The graph shows the average number of cells analysed for detection of antisense RNA foci for untransduced controls (mean=349.9), untransduced ALS lines (mean=951.4), LV-HC transduced ALS lines (mean=1130), LV-GFP transduced ALS lines (mean=711.8), isogenic ALS lines (mean=240.6) and RNAse treated ALS lines (mean=907.1). There were no significant differences observed in the average cells/well that were analysed between the different conditions for sense and antisense RNA foci. N=3 with 3 experimental repeats. Data were analysed using an ordinary one-way ANOVA with a Šídák's (or Tukey's for average cells/well) multiple comparisons test for the percentage of cells with foci. Data were analysed using a two-way ANOVA with Dunnett's multiple comparisons test for the histograms showing foci burden. Data presented as mean ± SEM; $p < 0.05$. ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$. UT = untransduced. ISO ALS = Isogenic controls for ALS lines.

In the panels showing sense and antisense RNA foci (**Figure 4.6 A and B**), it was observed in the LV-HC transduced cells that only a fraction of the cells were transduced (as seen by the HA-tag staining). This can be explained due to the low transduction efficiency of LV-HC in the *in vitro* model used (shown in section 3.3.5.1 of chapter 3). Since only a proportion of the cells were transduced and the effect of HC-ARFIP2 might have been masked by the untransduced cells, the LV-HC transduced ALS cells were analysed by dividing the cells in the LV-HC transduced wells into untransduced and transduced populations (**Figure 4.8**). The percentage of cells that were shown to have the presence of RNA foci in the ALS LV-HC condition with RNAse treatment ranged from 0.6% to 1.8%. Therefore 2% of the cells with RNA foci were considered as background.

Analysis of the sense RNA foci in LV-HC transduced cells showed that there were no significant differences in the percentage of cells with sense RNA foci between the untransduced (1.8%) and transduced (2%) populations in the ALS-HC condition (**Figure 4.8 A**). The percentage of cells detected with sense RNA foci were all below the 2% threshold for background and therefore the ALS LV-HC transduced cells did not show a significant presence of sense RNA foci. Analysis of the sense RNA foci burden (**Figure 4.8 B**) showed similar results. These results reinforced the concept that sense RNA foci were not detectable in our *in vitro* model.

Analysis of the antisense RNA foci in LV-HC transduced cells showed that there was no significant difference in the percentage of cells with antisense RNA foci in the successfully LV-HC transduced population (8.6%) when compared to the untransduced population (7.4%) in the ALS-HC condition (**Figure 4.8 C**). Analysis of the antisense RNA foci burden (**Figure 4.8 D**) showed no significant differences between the cells with 1, 2-4 and 5 or more antisense foci between the untransduced (5.4%, 1.5% and 0.6%, respectively) and transduced (5.6%, 2% and 1%, respectively) populations.

Overall the results show that sense RNA foci were not detected and only the antisense RNA pathology was observed in our *in vitro* model. On assessing the effects of HC-ARFIP2 on RNA foci, our results showed that there were no measurable effects in *C9ORF72*-ALS iPSC-derived MNs.
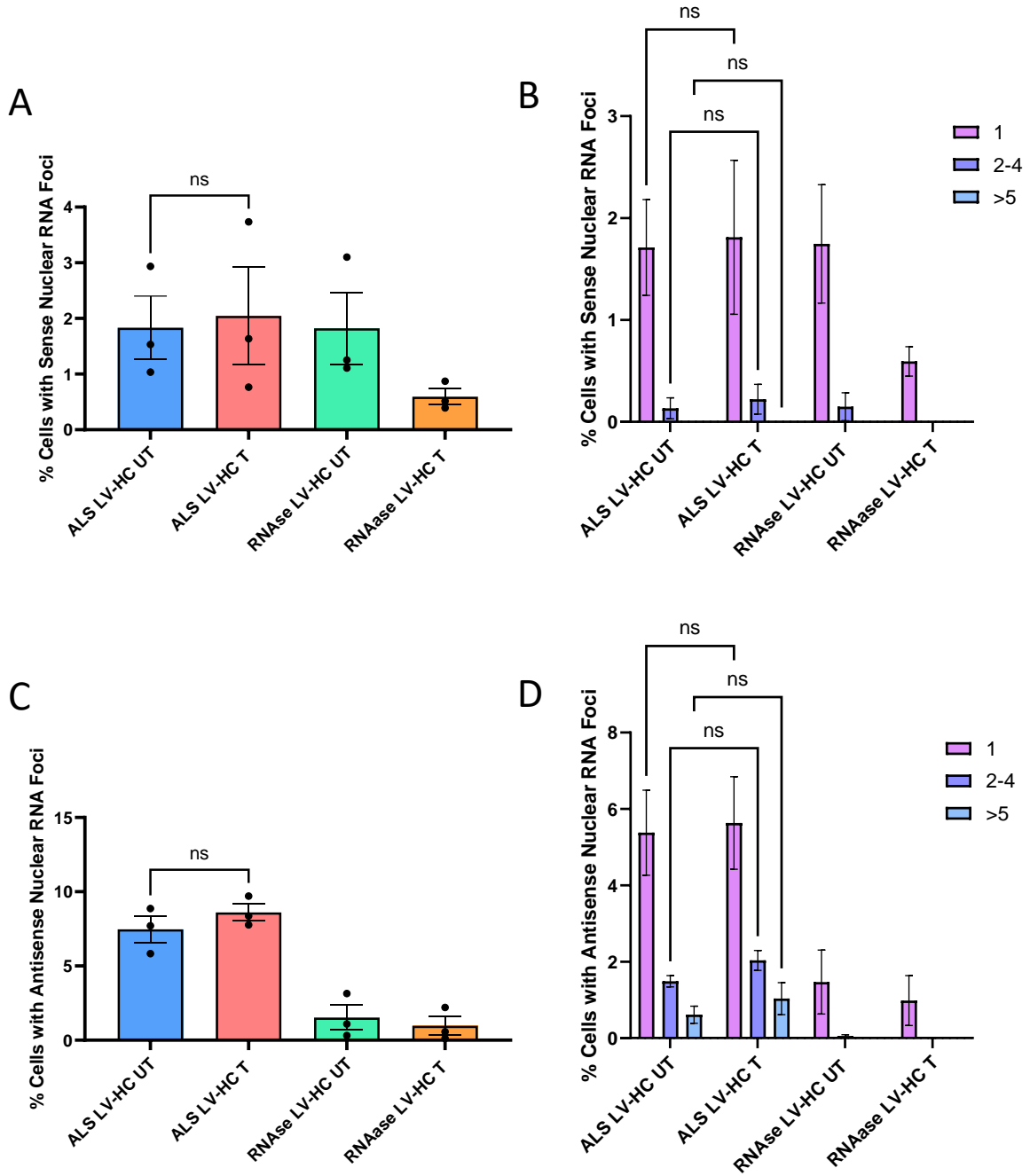
Figure 4.8: Expression of HC-ARFIP2 has no effect on RNA foci in iPSC-derived MNs

iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC or LV-GFP. MNs 4 days post transduction were fixed in 4% PFA and probed with locked nucleic acid (LNA) probes specific for sense or antisense RNA foci (1:400) using in situ hybridisation. **(A)** The graph shows the quantification of the percentage of ALS cells in the LV-HC transduced condition with sense RNA foci that were untransduced, transduced and both conditions treated with RNAse. There is no significant difference in the percentage of cells with sense foci between ALS LV-HC UT = 1.8% and ALS LV-HC T = 2% (RNAse LV-HC UT = 1.8%, RNAse LV-HC T = 0.6%). **(B)** Histogram showing the breakdown and burden by sense foci per cell. There is no significant difference in sense foci burden (cells with 1, 2-4 or 5+ foci) between the untransduced (1 = 1.7%, 2-4 = 0.13%, 5+ = 0%) and transduced cells (1 = 1.8%, 2-4 = 0.2%, 5+ = 0%). **(C)** The graph shows the quantification of the percentage of cells with antisense foci for cells that were untransduced, transduced and both conditions treated with RNAse. There is no significant difference in the percentage of cell with antisense foci between ALS LV-HC UT = 7.5% and ALS LV-HC T = 8.6% (RNAse LV-HC untransduced = 1.5%, RNAse LV-HC transduced = 1%). **(D)** Histogram showing the breakdown and burden by antisense foci per cell. There is no significant difference in antisense foci burden (cells with 1, 2-4 or 5+ foci) between the untransduced (1 = 5.4%, 2-4 = 1.5%, 5+ = 0.6%) and transduced cells (1 = 5.6%, 2-4 = 2%, 5+ = 1%). N=3 with 3 experimental repeats. Data were analysed using an unpaired t-test for the percentage of cells with foci. Data were analysed using a two-way ANOVA with Dunnett's multiple comparisons test for the histograms showing foci burden. Data presented as mean ± SEM; p<0.05. ns = p > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001. UT = untransduced. T= transduced. ISO ALS = Isogenic controls for ALS lines.

4.3.5 Effect of HC-ARFIP2 expression on nuclear TDP43 mislocalisation

A characteristic observed in ALS is the abnormal mislocalisation of the TDP-43 protein from the nucleus and into the cytoplasm. To assess whether this mislocalisation phenotype was displayed in our *in vitro* model and the effects of HC-ARFIP2 on TDP-43 mislocalisation, iPSC-derived MNs were either left untransduced or transduced with LV-HC on day 40 of the differentiation protocol (2.2.2.2, 2.2.2.6.1). The cells were fixed (2.2.2.7) and stained with the anti-TDP-43 antibody (Proteintech 18280-1-AP) to assess TDP-43 mislocalisation and anti-HA-tag antibody to select cells transduced with LV-HC (2.2.2.8). The iPSC-derived MNs were imaged on the Opera Phenix High Content Imaging system (2.2.2.14.2) (**Figure 4.9 A**) to save time as well

as to increase the number of cells analysed. In addition to TDP-43 and HA-tag, the MNs were stained with the anti-MAP2 antibody to visualize the neuronal marker. Using the high throughput imager, a TDP-43 mislocalisation phenotype in the ALS untransduced cells was observed again. However, this seemed to be partial as a considerable proportion of the cells in the ALS untransduced cell lines showed nuclear TDP-43 staining. Analysis of the staining from 3 experimental repeats (**Figure 4.9 B**) showed that in the controls, approximately 90% of the cells in the untransduced control group retained nuclear TDP-43. There was no significant mislocalisation of TDP-43 in ALS cell lines, however, there was reduced nuclear expression as it was shown that only 80% of the cells retained nuclear TDP-43. With LV-HC transduction of ALS cell lines, there was no significant difference seen in the TDP-43 mislocalisation. The intensity of the nuclear TDP-43 staining was also measured (**Figure 4.9 C**). A similar trend relating to the percentage of cells with nuclear TDP-43 was observed and there were no significant differences in the intensity of the nuclear TDP-43 staining between the different conditions. **Figure 4.9 D** shows the average total number of cells/well for each cell line under each condition that were analysed to measure the percentage of cells that were nuclear TDP-43 positive. There was no significant difference between the average number of cells/well under each condition and the number of cells counted was approximately equal (n=~1000 cells) across all conditions. Overall, the results show a trend towards TDP-43 mislocalisation in the C9ORF72-ALS iPSC-derived MNs, however, this observation did not reach statistical significance and more repeats would be needed to confirm whether observed trends

are significant. Furthermore, the expression of HC-ARFIP2 had no effect on the ALS-associated TDP-43 mislocalisation phenotype in *C9ORF72*-ALS iPSC-derived MNs.
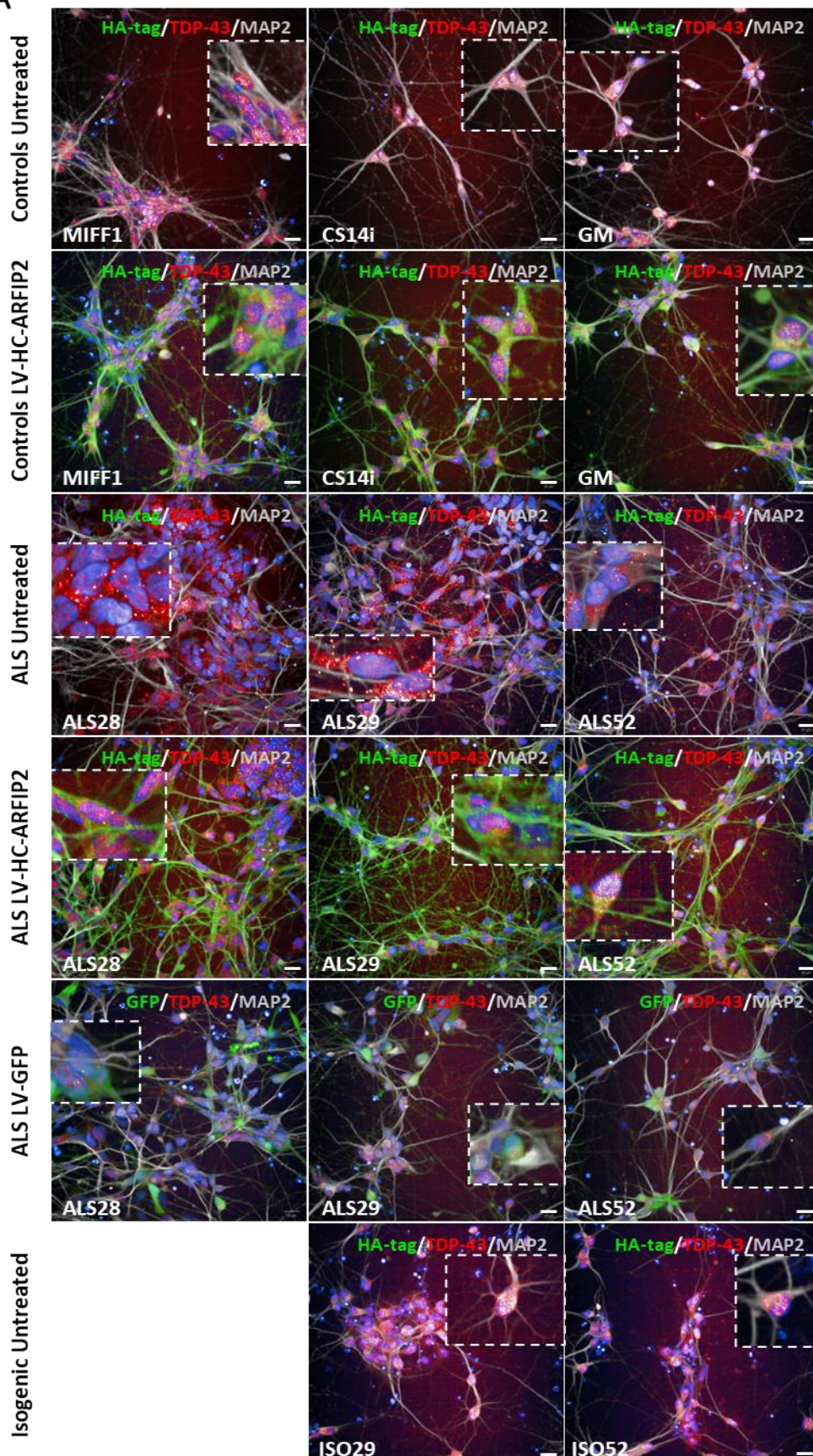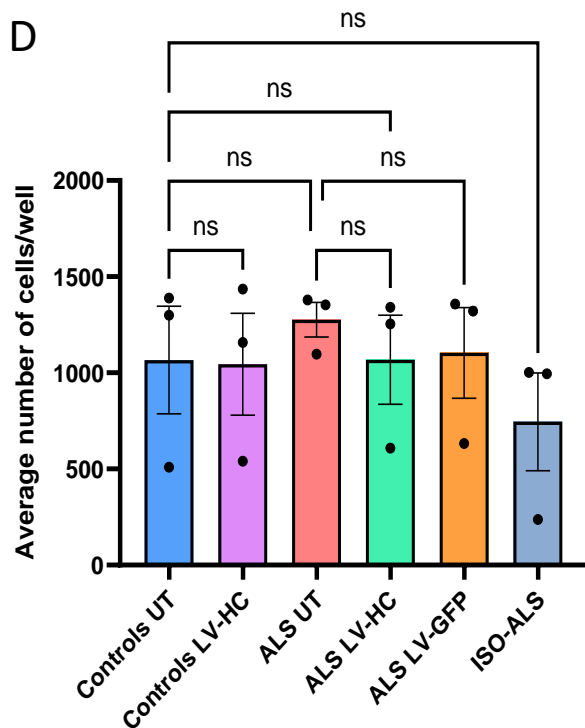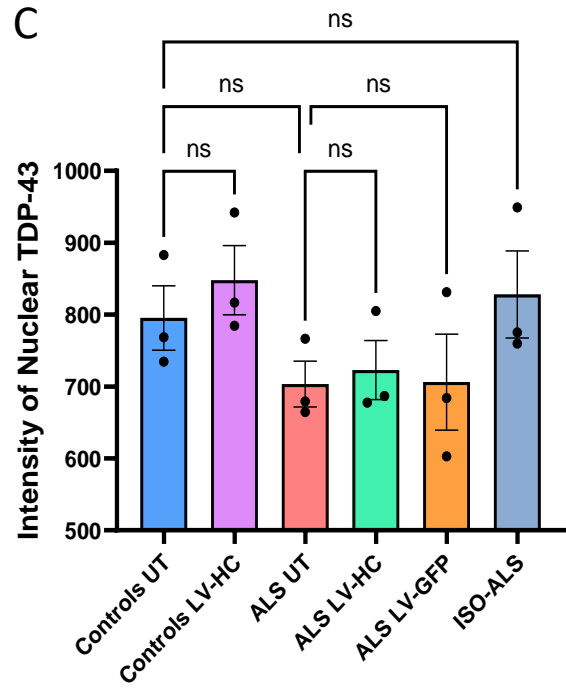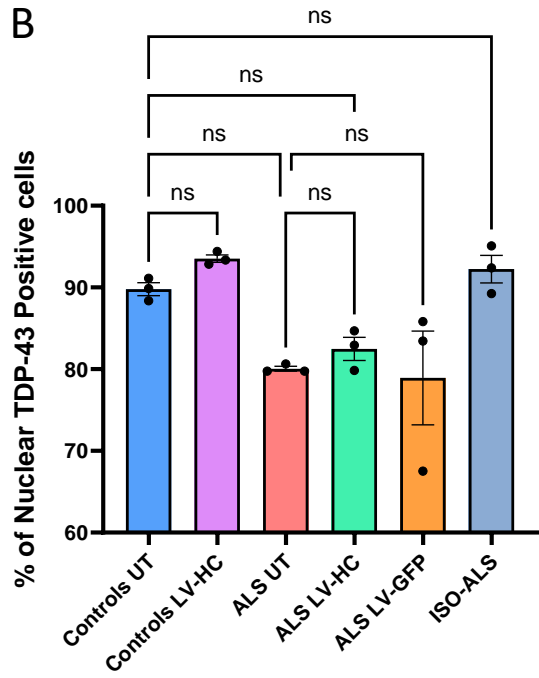
Figure 4.9: Expression of HC-ARFIP2 may potentially rescue the TDP-43 mislocalisation phenotype in *C9ORF72*-ALS iPSC-derived MNs.
iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC or LV-GFP. MNs 4 days post transduction were fixed and stained with anti-TDP-43 antibody to assess TDP-43 mislocalisation, anti-HA-tag antibody to select cells transduced with LV-HC-ARFIP2 and anti-MAP2 antibody to visualize the neuronal marker. **(A)** The panel shows representative images for each cell line under each condition. Insets show zoomed images of one cell. The cells were imaged on the Opera Phoenix High Content Imaging System. Scale bar = 20μm. **(B)** MNs stained for TDP-43 mislocalisation were imaged and averages for each cell line over 3 experimental repeats were analysed. Though not significant, a partial phenotype of TDP-43 mislocalisation was observed in ALS untreated lines compared to untreated controls. There was no significant decrease in nuclear TDP-43 positive cells in the ALS untransduced condition (mean=80%) compared to controls untransduced (mean=90%). With LV-HC transduction of ALS cells, there was no effect of HC-ARFIP2 expression on the TDP-43 mislocalisation (mean=82%). **(C)** The intensity of the nuclear TDP-43 staining was measured. There were no significant differences seen between each of the conditions (controls UT =795.4, controls LV-HC = 847.9, ALS UT = 703.5, ALS LV-HC = 723.0, ALS LV-GFP = 706.1 and ISO-ALS = 828.1). **(D)** The number of cells/well for each cell line under each condition was taken into account to obtain the average number of cells counted for each condition (controls UT n=1066, controls LV-HC n=1045, ALS UT n=1276, ALS LV-HC n=1068, ALS LV-GFP n =1104 and ISO-ALS n=745.1). N=3 with 3 experimental repeats. Data were analysed using an ordinary one-way ANOVA with a Šídák's multiple comparisons test. Data presented as mean ± SEM; $p<0.05$. ns = $p > 0.05$, * = $p ≤ 0.05$, ** = $p ≤ 0.01$, *** = $p ≤ 0.001$.UT = untransduced.

4.3.6 Effect of HC-ARFIP2 expression on proteasome activity

To determine possible factors that could be responsible for the reduction in poly-GP DPR levels as well as the improvement in cell viability with HC-ARFIP2 expression, the proteasome activity of the cells was assessed. iPSC-derived MNs were cultured until day 40 (2.2.2.2) of the protocol at which stage they were either left untransduced or were transduced with LV-HC or LV-GFP (2.2.2.6.1). The cells were lysed for the proteasome activity assay and the proteasome activity for each cell line was determined using a proteasome activity assay kit (2.2.2.11) and graphed (**Figure 4.10**). Analysis of proteasome activity across 3 experimental repeats showed that there was no significant difference in the average proteasome activity of the cells across all conditions. However, observed trends showed that there was decreased average proteasome activity in the ALS untransduced cells (0.003 mU x $1x10^6$ cells)

compared to the untransduced controls (0.04 mU x 1 x10^6 cells). When transduced with LV-HC, the average proteasome activity seems to be increased (0.047 mU x 1x10^6 cells) compared to the ALS untransduced cells. However, there was a lot of variability between the repeats, and this observed increase did not reach statistical significance. On the other hand, with the LV-GFP transduction, results showed that there was a negative proteasome activity (-0.04 mU x 1x10^6 cells) when compared to untransduced ALS cells and untransduced controls cells, which would be interpreted as no proteasome activity in the sample. The results indicate that ALS cell lines appear to have decreased proteasome activity, though not reaching statistical significance.

While expressing HC-ARFIP2 may increase the proteasome activity in some cell lines, overall there was no significant effect of expressing HC-ARFIP2 on proteasome activity and there was a lot of variability seen in the results. Therefore, data variability prevents drawing any conclusions about the effect of HC-ARFIP2 on proteasome activity.
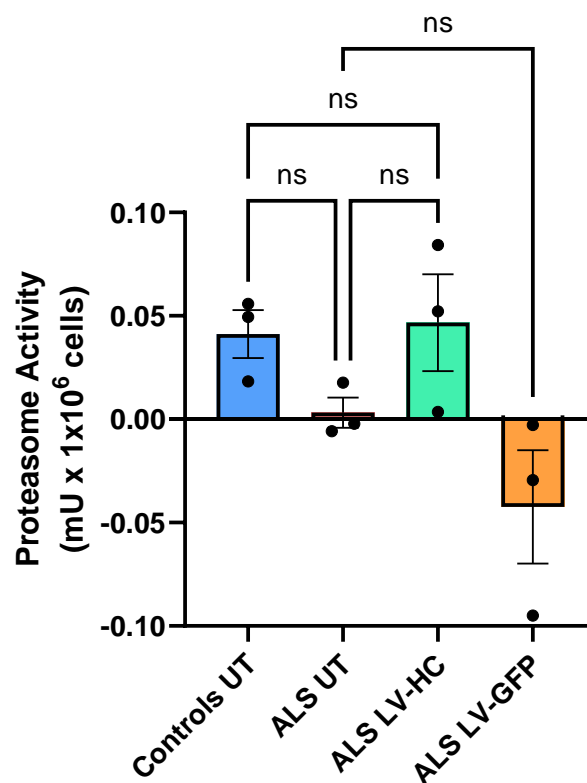
Figure 4.10: Effect of expressing HC-ARFIP2 on proteasome activity in *C9ORF72*-ALS iPSC-derived MNs

iPSC-derived motor neurons were cultured until day 40 and were either left untransduced or were transduced with LV-HC or LV-GFP. 4 days post transduction, the cells were lysed using NP-40 without proteasome inhibitors, and total cell lysate was used to conduct a proteasome activity assay. There were no significant differences observed between each condition for proteasome activity (controls UT = 0.038 mUx1x10^6, ALS UT = 0.003 mUx1x10^6, ALS LV-HC = 0.047 mUx1x10^6 and ALS LV-GFP = -0.042 mUx1x10^6). However, there seems to a positive trend when ALS lines were transduced with LV-HC. N=3 with 3 experimental repeats. Data were analysed using an ordinary one-way ANOVA with a Šídák's multiple comparisons test. Data presented as mean ± SEM; $p<0.05$. ns = $p > 0.05$, * = $p \leq$ 0.05, ** = $p \leq 0.01$, *** = $p \leq 0.001$.UT = Untransduced.

4.4   Results Summary

This chapter shows the effects of expressing HC-ARFIP2 on the disease pathophysiology of ALS in *C9ORF72*-ALS iPSC-derived MNs. It was shown that:

- The endogenous ARFIP2 levels appear to be more highly expressed in the cerebellum of *C9ORF72*-ALS cases, particularly in the Purkinje cells and the granular layer, and in the pyramidal neurons and surrounding glia of the frontal cortex of both *C9ORF72*-ALS and sALS cases.

- There was a significant increase in endogenous levels of ARFIP2 in the presence of the *C9ORF72*-ALS disease background in the iPSC-derived MNs compared to healthy controls and isogenic controls.

- HC-ARFIP2 expression improved cell viability in ALS cell lines and LV-HC was well tolerated over time at the selected transduction dose.

- Expressing HC-ARFIP2 significantly reduced the levels of the poly-GP DPRs in *C9ORF72*-ALS iPSC-derived MNs by approximately 92%.

- Sense RNA foci were not detectable and only the antisense RNA foci were detectable and quantifiable in the *C9ORF72*-ALS *in vitro* model used.

- On assessing the effects of HC-ARFIP2 on RNA foci, results showed that there was no measurable effect on RNA foci in *C9ORF72*-ALS iPSC-derived MNs

- There was a trend towards TDP-43 mislocalisation in the *C9ORF72*-ALS iPSC-derived MNs, however it was not statistically significant and the expression of HC-ARFIP2 had no effect on TDP-43 mislocalisation.

- The results are indicative of a decreased proteasome activity in the ALS cell lines but this was not significant and while expressing HC-ARFIP2 may increase the proteasome activity in some cell lines, there was a lot of variability which prevents us currently from drawing any conclusions about the effect of HC-ARFIP2 on proteasome activity.

4.5   Discussion

The ARFIP2 protein is ubiquitously expressed, with high levels of expression seen in several tissues including the brain and skeletal muscles (Stelzer *et al.*, 2016; Safran *et al.*, 2022; *Human Protein Atlas* ; Uhlén *et al.*, 2015). A previous study looked at the distribution of ARFIP2 in different regions in the brains of mice in a neurodegenerative background, specifically in Huntington's disease (HD) (Peters *et al.*, 2002). They showed that ARFIP2 was ubiquitously expressed in normal mice and transgenic mice brains, however, the levels of ARFIP2 were increased in the striatum, cortex, and cerebellum of transgenic mice which were models of HD. Since there are no other previous studies looking at ARFIP2 in the context of neurodegenerative disorders and neurodegenerative disorders share pathologies, we investigated ARFIP2 levels in ALS. In this chapter, we assessed the expression and distribution of ARFIP2 in healthy and disease biosamples to examine whether the endogenous levels were altered by disease in ALS, as seen in HD.

With an increase of ARFIP2 levels seen in the cerebellum and cortex, the human brain cerebellum and frontal cortex sections were qualitatively assessed to determine the distribution of ARFIP2 in ALS. Results from assessing the expression and distribution of ARFIP2 in human brain tissue were consistent with the previously published finding (Peters *et al.*, 2002) which offers support to our results. It was observed that the endogenous ARFIP2 levels appear to be increased in the cerebellum of *C9ORF72*-ALS cases, particularly in the Purkinje cells and granular layer, and in the pyramidal neurons and surrounding glia of the frontal cortex of *C9ORF72*-ALS and sALS cases compared to healthy control cases. However, a quantitative assessment is required to confirm this observation. Purkinje cells, found in the cerebellum, regulate and coordinate motor movements (Paul and Limaiem, 2022) which are primarily affected in ALS. Previous research has implicated the selective loss and degeneration of Purkinje cells in ALS (Afshar *et al.*, 2017; Tan *et al.*, 2016). Pyramidal neurons are the most prevalent type of neuron in the cortex, suggesting their role in processing external signals and motor control. Immunostaining studies have previously shown the loss and degeneration of cortical pyramidal cells in ALS (Nihei, McKee and Kowall, 1993; Maekawa *et al.*, 2004; Mochizuki *et al.*, 2011). An observed increased immunoreactivity of ARFIP2 in Purkinje cells and pyramidal neurons might be indicative of their involvement in ALS pathology and supports our hypothesis for

targeting ARFIP2 in ALS. In ALS, it is expected to see neuropathology in the motor cortex and spinal cord (Saberi *et al.*, 2015). Since the only other previous work on ARFIP2 in neurodegenerative disorders was undertaken in rodent models of HD and highlighted the cerebellum and cortex, the cerebellum and frontal cortex sections were chosen for pilot investigation. It was planned to extend the investigation to include the more relevant regions where we expect to see ALS pathology, however, due to the delays in obtaining the human CNS tissue because of the COVID-19 pandemic and the subsequent time constraints of the Ph.D., this could not be done and would be ideal to investigate in the future. The staining of the sections investigated was done using the DAB staining method and though this method is effective, it is semi-quantitative. If time permitted, we could be quantitative in assessment to confirm the results observed seen by using the fluorescence detection approach. This experiment was conducted with one repeat and, while there appeared to be more ARFIP2 immunoreactivity in ALS, for a true representation of the expression and distribution of ARFIP2 in ALS, more repeats of the staining would be required. The control for IHC in this experiment included a no primary antibody control. Though this confirmed the specificity of the secondary antibody, there was no confirmation of the specificity of the primary antibody. In the future it would be worthwhile to include an isotype control which will enable this and provide a measure of non-specific binding. Nevertheless, the observations from human brain tissue offered support for further investigation of ARFIP2 in ALS.

Using our *in vitro* model of *C9ORF72*-ALS iPSC-derived MNs, the expression of ARFIP2 was assessed in control and ALS cell lines as well as an isogenic control line. Analysis showed that there was a significant increase in ARFIP2 in ALS cell lines compared to controls and the isogenic control line used, where the ALS-causing mutation was corrected. Only one isogenic line (ISO29) was used in this experiment as this was the only isogenic control available at the time. ISO29 was the disease-corrected control line for the patient-derived ALS29 line. When compared specifically to its true control, there was an approximately 250% significant increase in the levels of ARFIP2 in the ALS29 patient line, across 4 repeats. With establishing more isogenic control lines for the patient-derived cell lines used, it would be interesting to see whether this significant increase in endogenous ARFIP2 levels can be observed in other ALS-isogenic control pairs. This would also further increase the power of the

experiment. The results show some variation in the data which can be attributed to natural variation in patients as the *in vitro* models used are real patient-derived lines. Nonetheless, the significant increase of endogenous levels of ARFIP2 in the presence of the *C9ORF72*-ALS disease background warrants further investigation into the effects of targeting ARFIP2 on the pathogenesis of ALS in *C9ORF72*-ALS iPSC-derived MNs.

Post transduction of iPSC-derived MNs, cell viability was evaluated using the resazurin reduction assay to assess the effects of HC-ARFIP2 in ALS and any possible transduction-related toxicity. The resazurin reduction assay was chosen as it is a very mild assay and allowed us to conduct a longitudinal study of LV-HC transduction on cell viability with exposing the cell to minimal stress. The resazurin assay is an easy, rapid and sensitive measurement for cytotoxicity and viability that has been widely used (O'Brien *et al.*, 2000; Ahmed, Gogal and Walsh, 1994; Page, Page and Noel, 1993). In metabolically active cells, the non-fluorescent dye is reduced to a fluorescent compound, the level of expression of which is proportional to the number of viable cells, allowing for quantitative measurement of cell viability (O'Brien *et al.*, 2000). Results indicated that ALS untransduced cell lines showed a steep decline in metabolic activity and therefore cell viability over time. However, with HC-ARFIP2 expression, the cell viability in ALS cell lines was stable over time. Since living cells are metabolically active and can reduce resazurin to resorufin, the low fluorescence output of the untransduced ALS cell lines measured over time using the resazurin reduction assay may be indicative that the metabolic activity is impaired in ALS. A growing body of evidence has suggested that metabolic abnormalities in ALS are correlated with disease development (Vandoorne, De Bock and Van Den Bosch, 2018; Tefera *et al.*, 2021; Hor *et al.*, 2021). With HC-ARFIP2 expression, the fluorescence output of the ALS cell lines was improved and comparable to the output for control cell lines. Therefore, it can be inferred that HC-ARFIP2 affects cell metabolism and therefore cell viability in ALS which highlights its therapeutic potential in ALS. Post transduction, the live/dead staining showed that there were minimal dead cells observed in all transduced cell lines, similar to that seen in the untransduced cell lines. In the untransduced ALS cell lines, where there was a steep decline in cell viability as determined by the resazurin reduction assay, the live/dead staining assay showed very little cell death. This may indicate that the assay was more indicative of the density

of live cells than measuring the amount of live or dead cells. This could have been because the dead cells had been washed away during the assay. Since the density of live cells for each cell line was similar between transduced cell lines and untransduced cell lines, the results suggest that the LV-HC lentivirus was well tolerated at the selected dose and did not show any transduction-associated toxicity over time. Furthermore, HC-ARFIP2 transgene expression was not toxic in iPSC-derived MNs.

A poly-GP MSD ELISA was conducted to assess the effects of HC-ARFIP2 on DPRs in ALS. After interpolating the amount of poly-GP DPRs and normalising for the assay background, it was interesting to see a non-specific signal in the LV-HC transduced control cell lines. Since DPRs are not generated in healthy controls, this suggests that any amount of poly-GP detected must be due to LV-HC transduction. Lentiviruses are positive-sense single-stranded RNA viruses, and it has been shown that the viral DNA has a mean GC content of 49.8% (Auewarakul, 2005). This GC-rich component may explain the signal in LV-HC transduced controls. Looking at the LV-HC transduction of ALS lines, results showed that expressing HC-ARFIP2 significantly decreased the levels of the poly-GP DPR in *C9ORF72*-ALS iPSC-derived MNs and demonstrated a therapeutic potential of HC-ARFIP2 in ALS. In the LV-HC transduced ALS lines, very low amounts of poly-GP DPRs were detected compared to the untransduced ALS lines which can be attributed to the GC content of the viral genome. Therefore, it can be inferred that HC-ARFIP2 expression substantially decreases poly-GP DPRs, returning the expression to the non-specific levels seen in the healthy control lines. Though the results are promising, and the poly-GP DPRs were detectable in ALS, there was a lot of variability in the samples which can be attributed to the low response on the standard curve. The poly-GP DPRs were detected but the raw values were towards the lower part of the standard curve thus approaching the limits of detection. To overcome this and add more power to the results, in the future a standard curve with lower concentrations could be generated in the experiment. Furthermore, this experiment could be repeated using cell lysates from cells plated at higher densities. While poly-GP may be a good representative of the DPRs generated from both sense and antisense strands, it is classed as the least toxic along with poly-PA (Lee *et al.*, 2016). On the other hand, the arginine-rich DPRs, poly-GR, and poly-PR, are shown to have associated toxicity in several *in vitro* models (Lee *et al.*, 2016; Wen *et al.*, 2014; Tao *et al.*, 2015). The poly-GP MSD ELISA was established in-house using a

custom-made antibody and adapted from previously described methods (Simone *et al.*, 2018). At the time of the project, the poly-GP ELISA was the only assay established, however, all the other DPR MSD ELISAs have now been established. In the future, assessing the effects of HC-ARFIP2 on all other DPRs, especially the more toxic poly-GR and poly-PR, will provide further insight into the therapeutic potential of HC-ARFIP2 in ALS. Furthermore, it would be interesting to see whether DPRs specifically from the sense and antisense strands are detectible in the sample as well as to investigate whether the effects of HC-ARFIP2 expression on DPR reduction are sense or antisense specific; or affect DPRs more generally.

The pathogenic role of DPR-associated toxicity in the pathophysiology of ALS is largely unclear. Additionally, it is unknown how much of a pathogenic contribution any specific DPR makes. Research has shown that poly-GP is unable to aggregate by itself and therefore has fewer intracellular protein interactions compared to other DPRs (Lee *et al.*, 2016). This suggests that poly-GP is most likely the least toxic of the DPR species. In contrast, one study reported that overexpression of poly-GP interferes with the UPS and demonstrated its associated cytotoxicity via UPS dysfunction (Yamakawa *et al.*, 2015). With this in mind, two possibilities are probable. Either the reduction in poly-GP DPR levels due to expression of HC-ARFIP2 would affect or rescue proteasome activity or HC-ARFIP2 may affect proteasome activity and therefore result in the reduction of poly-GP DPR levels. Previously, increasing the concentration of ARFIP2 has been shown to downregulate proteasome activity and expression of HC-ARFIP2 has been shown to rescue proteasomal activity, thereby exerting its neuroprotective effects (Peters *et al.*, 2002). Furthermore, phosphorylation of ARFIP2 at a site on the C-terminal has been shown to prevent the inhibition of the proteasome (Rangone *et al.*, 2005). Therefore, the effect of HC-ARFIP2 on proteasome activity was evaluated. Results showed a lot of variability in the outcome of the assay which prevented drawing any conclusion regarding the effect of HC-ARFIP2 on proteasome activity. This could be attributed to various reasons. A proteasome activity kit was used which recommended the use of a particular microplate for the experiment. While this may be optimised, it has been reported that there are many important factors that need to be taken into consideration when selecting a microplate. Protein orientation in plate wells is unpredictable, and the intricate process of binding between enzymes and substrates is influenced by a wide

range of factors (Cui, Gilda and Gomes, 2014). The microplate type used might have affected the measurement of proteasome activity and therefore further optimisation of the proteasome activity kit used with different microplates could be useful in attaining conclusive results. Furthermore, the eukaryotic proteasome consists of three pairs of proteolytic active sites specific for distinct substrates – chymotrypsin-like, trypsin-like and caspase-like (Heinemeyer *et al.*, 1997). Though the kit used takes advantage of the chymotrypsin-like activity (most important in protein degradation) and had a straightforward protocol to follow, the subsequent calculations to determine the proteasome activity were quite complicated. Furthermore, it was very time consuming to optimise the assay to determine the dilutions to be used for the samples as well as the time of the first reading (T1) and time of the second reading (T2) in the appropriate linear range to measure the difference in fluorescence intensity between the two time points. If time permitted, using alternative ways to measure proteasome activity (especially for measuring each of the three activities of the proteasome) could be explored (Maher, 2014).

Yamakawa et al. (2015) assessed the protein degradation system of DPRs in *C9ORF72*-ALS. They reported that poly-GP, along with poly-GR and poly-PR, are degraded by both UPS and autophagy, while the UPS pathway does not contribute to degrading poly-GA. Recently, ARFIP2 has been shown to be involved in autophagy by regulating the amino acid starvation-dependent distribution of ATG9A vesicles which then deliver the PI4-kinase, PI4KIIIβ, to the autophagosome initiation site (Judith *et al.*, 2019). Therefore, in the future, it would be interesting to assess the effects of expressing HC-ARFIP2 on autophagy markers which could help explain the pathway by which HC-ARFIP2 expression reduces the levels of the poly-GP DPR in *C9ORF72*-ALS iPSC-derived MNs.

Another characteristic of *C9ORF72*-ALS is the presence of RNA foci. A study from the suppliers of the iPSCs that were used in the project reported that the cellular model of *C9ORF72*-ALS iPSC-derived MNs has an increased transcription of the hexanucleotide repeat in the *C9ORF72*-ALS motor neurons compared to controls (Sareen *et al.*, 2013). This was consistent with previous reports which showed a relationship between an increased GC content and an increased transcription level (Kudla *et al.*, 2006; Bauer *et al.*, 2010), perhaps as a result of reduced transcriptional

pausing (Zamft *et al.*, 2012). Additionally, it was reported that the increase in transcription resulted in the accumulation of repeat-containing RNA foci in the *C9ORF72*-ALS iPSC-derived MN cultures (Sareen *et al.*, 2013). The foci detected were mainly in the nucleus, with cells having 1-3 foci and cytoplasmic foci were also sometimes seen. Results on RNA foci in this chapter showed that nuclear RNA foci were detected in our motor neuron cultures. In our cells, we detected only the antisense nuclear RNA foci. These results were in support of the previous work focusing on RNA foci in the cell model used (Sareen *et al.*, 2013) as they only reported the presence of antisense RNA foci. Since we were unable to detect sense RNA foci, assessment of the effects of HC-ARFIP2 on sense RNA foci were inconclusive. With HC-ARFIP2 expression, the results showed no measurable effect on antisense RNA foci in *C9ORF72*-ALS iPSC-derived MNs. This can be attributed to the sensitivity of the assay and detection method used to assess RNA foci as well as the high background levels seen in the experiment. The fluorescence in situ hybridisation (FISH) assay is not as sensitive compared to the other assays used to measure the effect on other disease characteristics, for example, the MSD ELISA assay. Further optimisation of FISH, focusing on minimising background and improving the sensitivity of the assay would be desirable. With optimisation, the sense RNA foci may be detectable and there may be an effect of HC-ARFIP2 on RNA foci that can be measurable.

The previous study showing the presence of RNA foci in the cell model used showed that they detected foci in ~20% of the cells while we were able to detect foci in only ~5% of the cells. While the probes used and the fluorescence in situ hybridisation (FISH) and FISH/immunocytochemistry staining methods were the same, differences were in the iPSC differentiation protocol and the detection methods used. The differentiation protocol used in this project (2.2.2.2) follows a well-established protocol (Du *et al.*, 2015) which has a protocol length of 28 days. On the other hand, Sareen et al. (2013) use the protocol described in the study which spans 75 days. While the motor neuron differentiation was validated and were cultured until day 40 for use in experiments shown in this chapter, the length of culture and maturity of the generated motor neurons may not be enough to develop disease characteristics and might explain the low percentage of cells with RNA foci in our MN cultures. We detected the RNA foci in MN cultures using the Opera Phenix High-Content Imaging System

whereas Sareen et al. (2013) used confocal microscopy for the detection method. Using the Opera Phoenix High Content Imaging system provides high resolution, unsupervised image acquisition, high throughput, allows for analysis of more cells and avoids spectral crosstalk. However, the exclusion of background in the analysis was not possible and any minimisation of noise/background is critical when assessing for small characteristics such as RNA foci. Therefore, repeating the experiment with a more stringent washing regime might be able to overcome this. Furthermore, using the confocal microscope for the detection of RNA foci could be a more sensitive method. In the future, it would be interesting to see whether using confocal microscopy can detect antisense foci in a larger proportion of cells as well as the detection of sense RNA foci in the MN cultures.

A previous study reported that the antisense foci were associated with TDP-43 mislocalisation, a defining feature of ALS neurodegeneration, in motor neurons, which are the primary disease target in ALS (Cooper-Knock *et al.*, 2015). Previously, the effects of TDP-43 mislocalisation unrelated to mutations or aggregation have been looked at. Upon manipulation of the nuclear localization sequence (NLS), it was reported that there was a depletion in the endogenous TDP-43 protein from the nucleus and increased cytoplasmic insoluble aggregates (Winton *et al.*, 2008). In a study involving mice, it was reported that with TDP-43 mislocalisation, the mice also showed elevated levels of TDP-43 aggregates (Walker *et al.*, 2015). The study described that the endogenous nuclear TDP-43 protein depletion due to TDP-43 mislocalisation was followed by the formation of aggregates. This implies that TDP-43 mislocalisation is possibly upstream of aggregation. It has been suggested that prolonged TDP-43 mislocalisation promotes aggregation which can be overwhelming for the degradation pathways, leading to eventual cell death (Suk and Rousseaux, 2020). Targeting the aggregation may be too late and therefore early therapeutic strategies that rescue TDP-43 mislocalisation are important. Since the presence of antisense foci was observed in our motor neuron cultures and the importance of developing therapeutic strategies targeting TDP-43 mislocalisation, we looked into the TDP-43 mislocalisation pathology in our model and assessed the effects of HC-ARFIP2 on this ALS phenotype. Though the panel in the results (4.3.5) was indicative by the observation that there may be a TDP-43 mislocalisation phenotype that may be rescued, these observations did not reach statistical significance. The analysis did not

show a statistically significant phenotype but showed that there was a trend towards the TDP-43 mislocalisation phenotype in the *C9ORF72*-ALS iPSC-derived MNs. Furthermore, with HC-ARFIP2 expression, there was no effect on the ALS-associated TDP-43 mislocalisation pathology, although again a positive trend was observed.

The TDP-43 mislocalisation phenotype was not significantly detected in our *C9ORF72*-ALS iPSC-derived MN model. Even though TDP-43 pathology (mislocalisation and aggregation) is a key component of *C9ORF72*-ALS at post-mortem, very few models to date have robustly shown this in hiPSC-derived MNs. It is unknown if this means that the TDP-43 pathology in *C9ORF72*-ALS can be disassociated from other characteristics of disease pathology, or if further cell maturation is needed for the TDP-43 pathology to develop (Giacomelli *et al.*, 2022). The MN differentiation protocol used is able to differentiate iPSCs into mature MNs in 28 days and the MNs used in the experiment were cultured until day 40 for further cell maturation. Since it has been suggested that TDP-43 mislocalisation may be upstream of aggregation (Suk and Rousseaux, 2020; Walker *et al.*, 2015), it is likely that the cells may exhibit the TDP-43 aggregation pathology. If time permitted, it would be interesting to assess TDP-43 aggregation and whether expressing HC-ARFIP2 has any effect on TDP-43 aggregation. Additionally, it would be interesting to assess the TDP-43 mislocalisation phenotype at an earlier stage of the differentiation. While further cell maturation may be required to see the TDP-43 pathology, the use of another age-associated model may also be key. TDP-43 pathology, in the absence of a direct genetic background, might be primarily driven by biological aging. It can be argued that iPSCs are embryonic due to reprogramming-associated cellular rejuvenation which might erase many of the age-related features (Studer, Vera and Cornacchia, 2015). Therefore, this means they would not display age-related phenotypes. It has been shown that direct conversion or transdifferentiation strategies can retain the age-dependent features of donor cells, which are lost in iPSC reprogramming (Gatto *et al.*, 2021; Son *et al.*, 2011; Meyer *et al.*, 2014; Mertens *et al.*, 2015). These strategies focus on bypassing the iPSC stage and directly reprogramming donor fibroblast cells into relevant cell types. If time permitted, transdifferentiated tripotent-induced neural progenitor cells (iNPCs) or iNPC-derived MNs (iMNs) would be a useful model to use to model TDP-43 pathology and assess the effects of HC-ARFIP2.

5. ASSESSING THE EFFECTS OF EXPRESSING HC-ARFIP2 ON THE PATHOPHYSIOLOGY OF C9ORF72-ALS USING ZEBRAFISH

This chapter focuses on assessing the effects of expressing the C-terminal of Arfaptin-2 (HC-ARFIP2) on the characteristics of ALS using the *C9ORF72*-ALS zebrafish model. HC-ARFIP2 expressing heterozygous transgenic zebrafish lines have been established and crossed with the *C9ORF72*-ALS sense and antisense heterozygous lines. The crosses generated four distinct genotype experimental populations which have been used in the work described in this chapter. The effect of the HC-ARFIP2 transgene expression on motor performance, dipeptide repeats (DPRs), and RNA foci in zebrafish has been assessed.

5.1   Introduction

Using *in vitro* models for research has many advantages including higher throughput, control of the physical and chemical environments, and analysis using a less complex system which allows for understanding specific cellular mechanisms. However a major drawback of *in vitro* research experiments is that they fail to capture the conditions of the cells in their physiological environment. Cellular models of ALS cannot display the complex interactions occurring between the motor neurons (MNs) and the surrounding environment (Van Damme, Robberecht and Van Den Bosch, 2017). This in turn limits the value of the data from *in vitro* experiments to predict *in vivo* behaviour, and therefore *in vitro* studies are less translatable to humans. To overcome these shortcomings and demonstrate the innate complexity of the organ systems, *in vivo* models are necessary for the cross-validation of research and assessing the potential of therapeutic strategies.

Since the first ALS-linked genetic mutation in the *SOD1* gene was identified in the 1990s (Rosen *et al.*, 1993), over the decades, several animal models have been generated to study the pathophysiology of ALS including rodents, *Drosophila melanogaster* (fruit flies), *Danio rerio* (zebrafish), *Caenorhabditis elegans* (worms), *Saccharomyces cerevisiae* (yeast) extensively reviewed in (Bonifacino *et al.*, 2021). The first transgenic animal model for ALS was a mouse model based on the G93A SOD1 mutation in ALS (Gurney *et al.*, 1994) and since then many mouse models of ALS have been developed, making mice the most common *in vivo* system to model ALS (Philips and Rothstein, 2015; Lutz, 2018; Picher-Martel *et al.*, 2016). With the chromosome 9 open reading frame 72 (*C9ORF72*) genetic mutation being the most common mutation in ALS (Renton *et al.*, 2011; DeJesus-Hernandez *et al.*, 2011), mouse models have been generated carrying this mutation (Xu *et al.*, 2013). Mouse models of this mutation have been shown to exhibit neuromuscular junction damage, cognitive impairment, and gait deficit (Herranz-Martin *et al.*, 2017). Furthermore, the mouse models showed further ALS-associated characteristics including TDP-43 aggregate pathology, DPRs, loss of Purkinje and cortical neurons, and motor deficits (Schludi *et al.*, 2017; Liu *et al.*, 2016; Chew *et al.*, 2015b). Mice models are preferred and considered the gold standard for validating mechanisms involved in disease pathogenesis and providing preclinical data on therapeutic molecules under consideration. However, to date, therapeutic strategies developed using mouse

models, and other preclinical *in vivo* models, have failed to translate into successful interventions in humans with ALS. Additionally, mice have high costs of maintenance, slow growth of offspring, and age-related phenotypes that develop slowly. Furthermore, it is relatively time and manpower-consuming to develop transgenic mouse models.

To overcome the limitations of mouse models, small-animal models are now widely used to model ALS. Small animal models are inexpensive to maintain, generated quickly, and suitable for genetic or compound screening with high throughput (Van Damme, Robberecht and Van Den Bosch, 2017). Zebrafish are a promising and useful vertebrate small animal model. Most human genes have a zebrafish homologue and ~70% homology in the protein-coding sequence (Phillips and Westerfield, 2014; Howe *et al.*, 2013). Additionally, advantages of using zebrafish include the high number of offspring, quick maturation to adulthood compared to mice, and its transparent embryonic development which enables *in vivo* imaging (Rafferty and Quinn, 2018).

Zebrafish are increasingly being used for studying neurological disorders (Babin, Goizet and Raldúa, 2014; Braems, Tziortzouda and Van Den Bosch, 2021; Van Damme, Robberecht and Van Den Bosch, 2017; Fortier, Butti and Patten, 2020) because of the conserved vertebrate nervous system, relative ease of genetic modification to make transgenics and knockout models and the ability to generate genetic mosaics (Meng *et al.*, 2008; Carmany-Rampey and Moens, 2006; Lieschke and Currie, 2007; Doyon *et al.*, 2008). Zebrafish offer accessible MNs for manipulation *in vivo* for motor neuron disease studies (Kabashi *et al.*, 2010; Boon *et al.*, 2009; Lemmens *et al.*, 2007). Furthermore, all the key domains of the mammalian brain are found in the zebrafish central nervous system and both species produce the same neurotransmitters including glutamate, GABA, serotonin, dopamine, and acetylcholine (Babin, Goizet and Raldúa, 2014; Panula *et al.*, 2006). Though there are clear differences in the size of zebrafish brains, cognitive processing, and sensory pathways share homology with humans (Tropepe and Sive, 2003). Various genes associated with human neurodegenerative disorders including ALS have been identified in zebrafish. Moreover, the motor system in both zebrafish and humans share similarities (Stil and Drapeau, 2016; Madgwick *et al.*, 2015), making zebrafish a good model for the study of motor neuron diseases (McWhorter *et al.*, 2003; Wood *et al.*, 2006).

Previously, stable transgenic *SOD1* mutant zebrafish lines have been generated which display the pathological characteristics of ALS including an age-related decrease in swimming endurance, muscle atrophy, and sometimes early death (Ramesh *et al.*, 2010b; McGown *et al.*, 2013; Sakowski *et al.*, 2012). More recently, a stable *C9ORF72* transgenic model was generated and characterized (Shaw *et al.*, 2018). The transgenic lines exhibited an accumulation of RNA foci and DPRs. Furthermore, the zebrafish recapitulated motor deficits, muscle atrophy, MN loss, and death in early adulthood. This *C9ORF72*-ALS transgenic model was used in the work described in this chapter to assess the effects of expressing HC-ARFIP2 in *C9ORF72*-ALS.

ARFIP2 is highly conserved between humans and Zebrafish, with a 69% identity between the zebrafish and human amino acid sequence. Most of the homology seen is in the c-terminus conserved region. With this knowledge, in this chapter, we have generated stable HC-*ARFIP2* transgenic lines to naturally express human HC-ARFIP2 without the need for an invasive procedure to inject an artificially synthesised peptide.

Following on from the therapeutic potential of expressing HC-ARFIP2 on ALS disease characteristics using *C9ORF72*-ALS iPSC-derived MNs highlighted in chapter 4, we hypothesized that HC-ARFIP2 may exert therapeutic effects *in vivo* using the *C9ORF72*-ALS transgenic zebrafish model.

5.2   Aim

The work described in this chapter aimed to assess the effects of HC-ARFIP2 on the disease pathophysiology of *C9ORF72*-ALS in zebrafish models and identify whether HC-ARFIP2 has therapeutic potential *in vivo*. Firstly, stable transgenic lines expressing HC-ARFIP2 have been generated and established. Secondly, these lines were then crossed with the *C9ORF72*-ALS sense and antisense lines to establish experimental populations of non-transgenics, fish expressing the HC-ARFIP2 transgene only, *C9ORF72*-ALS mutants, and double mutants which express HC-ARFIP2 in the *C9ORF72*-ALS disease background. Finally, these genotyped zebrafish populations were used to assess the effects of HC-ARFIP2 transgene expression on motor performance, DPRs and RNA foci in zebrafish.

5.3 Results

5.3.1 Successful generation of HC-ARFIP2 transgenic lines

To assess the effects of HC-ARFIP2 on *C9ORF72*-ALS phenotypes in zebrafish models of ALS, stable zebrafish transgenic lines expressing HC-ARFIP2 were generated.

5.3.1.1 Microinjections of zebrafish embryos at the one-cell stage and transient expression assay

The pCDNA3.1-*HC-ARFIP2* plasmid (**Figure 5.1 A**) was linearised (2.3.1.1.1) (**Figure 5.1 B**) and injected into the one-cell stage zebrafish embryos (2.3.2.3.1). The embryos were heat-shocked (2.3.2.4). Transgene expression was tracked using *Discosoma sp.* red (DsRed) fluorescent protein expression driven by the zebrafish heat shock protein 70 (hsp70) promoter. A widespread DsRed distribution was seen as strong transgenics, distribution in fewer areas was seen as moderate transgenics, and expression in patchy areas was seen as weak transgenics. About 600 embryos were injected and at least 90 embryos showed weak expression and 150 embryos showed moderate expression at 72 hours post fertilisation (hpf). No strong expression of DsRed was observed and thus the DsRed expressing embryos were sorted according to the mosaic distribution of DsRed fluorescence into weak or moderate expressing transgenics. Those with a visible mosaic DsRed expression were raised to 5 days post fertilisation (dpf) and the protein was extracted from some of the sorted embryos (2.1.2.18) to detect transgene expression (**Figure 5.1 C**). HC-ARFIP2 expression was detected in both weak and moderate DsRed expressors, compared to the non-transgenic controls (NTG) and the two G4C2 repeat containing positive control (2-mer). The observed intensities of the bands were consistent with the visualised DsRed expression, while all the samples were equally loaded. Some of the injected embryos were immunostained at 5 dpf (2.3.2.8) with an anti-HA-tag antibody to visualise the distribution of the transgene in the zebrafish embryos (**Figure 5.1 D**). The HC-ARFIP2 injected embryos showed detectable expression of HC-ARFIP2 and a stable expression of HC-ARFIP2 as seen in fish 5 dpf, compared to the NTGs. The embryos were chimeras and hence the expression was not seen everywhere. HC-ARFIP2 expression was localised in the muscle fibres and the gut, as seen in the tail and trunk of the embryos, respectively. Results showed that the embryos were

successfully injected with the linearised DNA containing the HC-*ARFIP2* transgene and the injected embryos were able to stably express the transgene.
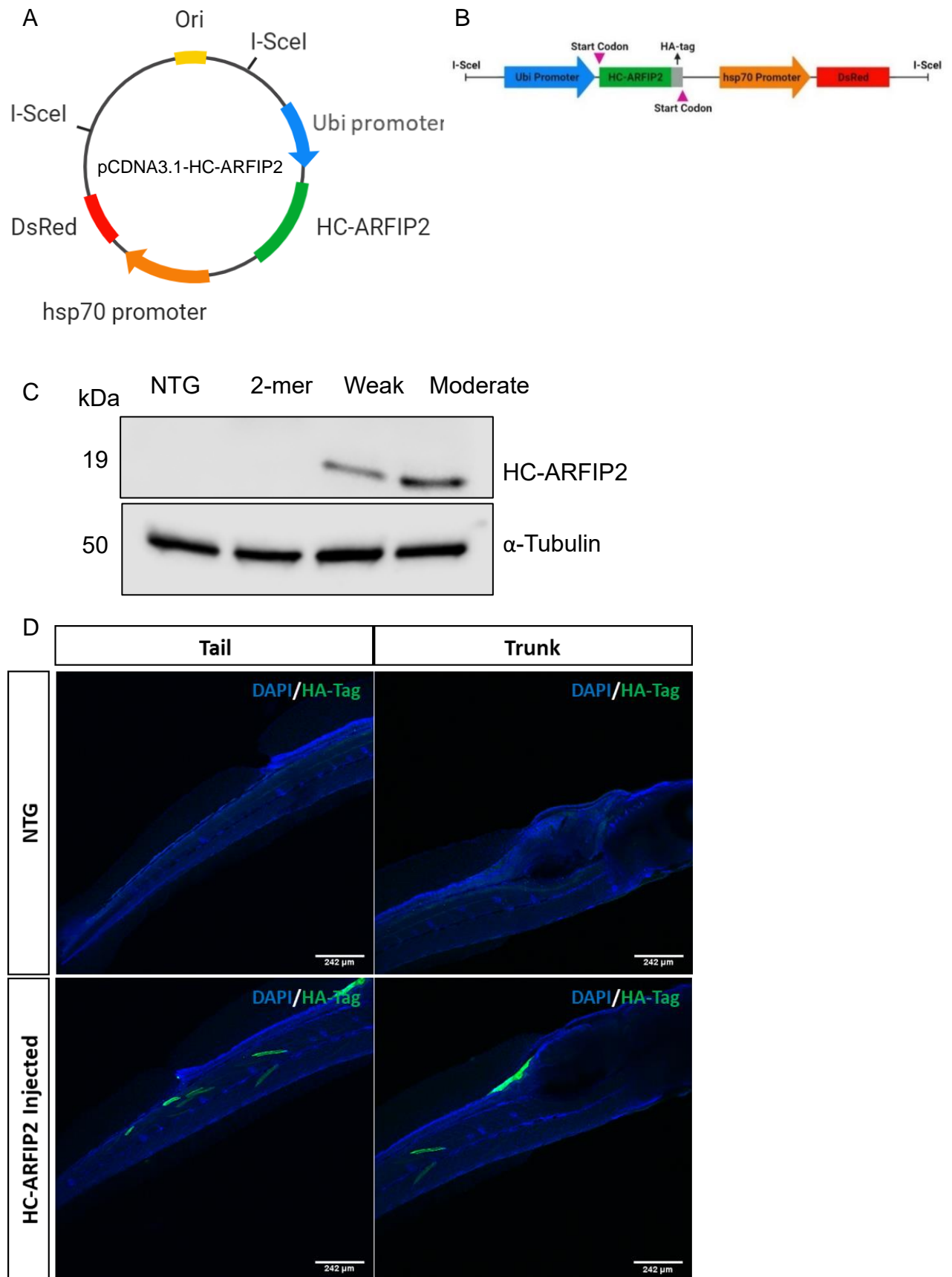
Figure 5.1: Generation of HC-ARFIP2 expressing transgenic zebrafish
(A) Schematic plasmid map representing the pcDNA3.1-*HC-ARFIP2* plasmid used to generate HC-ARFIP2 expressing transgenic zebrafish. **(B)** The plasmid was linearised using I-SceI and the DNA was made up to a final concentration of 40 ng/nl. **(C)** The linearised DNA (1 nl) was used for microinjecting one-cell stage zebrafish embryos. Injected zebrafish embryos were raised to 5 dpf and sorted into weak or moderate expressors according to the mosaic distribution of DsRed fluorescence. At 5dpf, the protein was extracted from 70 weak expressing and 20 moderate expressing embryos. Transgene expression was detected using a western blot. HC-ARFIP2 expression (19kDa) was detected using anti-HAtag antibody (1:1000) and α-Tubulin (50kDa) was used as a loading control. **(D)** Zebrafish embryos injected with the pCDNA3.1-*HC-ARFIP2* linearised plasmid at the one-cell stage which showed moderate expression were raised to 5 days post fertilisation (dpf) and fixed in fish fix. They were stained to see the distribution of HC-ARFIP2 using the anti-HAtag antibody (1:500). Hoechst (1 µg/ml) was used to stain nuclei. Zebrafish show a stable expression of HC-ARFIP2. The distribution of expression is cytosolic and is expressed in the muscle fibers (tail) and gut (trunk) in injected fish, compared to non-transgenics (NTG). Scale bar = 242µm. NTG = non-transgenic. 2-mer = positive control embryos that contain 2 G4C2 repeats and were generated using the same vector backbone.

5.3.1.2 Identification of positive founder fish

The moderate DsRed expressing embryos were raised to adulthood to screen for positive founder fish (F0) which would have the HC-*ARFIP2* transgene integrated into the genome of germ cell precursors. With the transgene integration into the germline being random, the injected embryos raised to adulthood were screened by outcrossing them (2.3.2.2) with AB wild-type fish to identify those that had integration of the construct into their germline. Approximately 31 injected adults were screened, and 4 positive founders (2, 4, 7, and 24 – highlighted in green) were identified (**Table 5.1**). Though the outcrosses involving F0 7 and 24 were able to produce transgenic embryos, the embryos did not develop properly and were not viable. Furthermore, the percentages of transgenic embryos produced for F0 7 and 24 were very low (average of 2.4% and 1.6%, respectively). Out of those identified, F0 line 2 and line 4 were able to produce viable embryos at high percentages (average of 13.5% and 7.1%, respectively). F0 2 and 4 were chosen to generate and establish transgenic lines expressing HC-ARFIP2.

Table 5.1: Screening of injected zebrafish that were grown to adulthood to identify positive founders (F0) that have HC-*ARFIP2* integrated into the genome of germ cell precursors.

Line	Gender	F0 (Y/N)	Number of Transgenic F1	Total number of embryos	Percentage of Transgenic embryos (%)
1	F	N	0	128	0.00
2	M	Y	6	117	5.13
			20	104	19.23
			10	76	13.16
			20	122	16.39
3	F	N	0	150	0.00
4	M	Y	1	170	0.59
			9	110	8.18
			4	32	12.50
5	F	N	0	105	0.00
6	F	N	0	123	0.00
7	F	Y	1	41	2.44
			0	54	0.00
			0	52	0.00
8	F	N	0	129	0.00
9	F	N	0	131	0.00
10	F	N	0	135	0.00
11	F	N	0	74	0.00
12	M	N	0	143	0.00
13	M	N	0	138	0.00
14	M	N	0	86	0.00
15	F	N	0	58	0.00
16	M	N	0	92	0.00
17	M	N	0	98	0.00
18	M	N	0	74	0.00
19	M	N	0	116	0.00
20	M	N	0	174	0.00
21	F	-	-	-	-
22	F	N	0	122	0.00
23	F	N	0	128	0.00
24	F	Y	1	94	1.06
			2	92	2.17

			0	60	0.00
25	M	N	0	172	0.00
26	M	N	0	96	0.00
27	M	N	0	130	0.00
28	M	N	0	94	0.00
29	M	N	0	150	0.00
30	M	N	0	222	0.00
31	M	N	0	168	0.00

The cells highlighted in **orange** indicate adult fish or embryos that were non-viable. The cells highlighted in **green** indicate the fish that were identified as positive founder fish (F0).

5.3.1.3 Establishment of HC-ARFIP2 transgenic lines

F0 2 and 4 were crossed (2.3.2.2) with AB wild-type fish to produce heterozygous HC-ARFIP2 expressing embryos (F1). The F1 embryos were heat-shocked (2.3.2.4) and visualised and sorted according to DsRed expression. At 5 dpf, the F1 embryos were analysed for transgene expression and distribution (**Figure 5.2**). The protein was extracted from the F1 embryos of F0 2 (2.1.2.18), and transgene expression was detected. A strong band for the transgene was seen in F1 embryos from F0 2, while no bands were detected for the NTG embryos and the positive control embryos (**Figure 5.2 A**). The F1 embryos from both F0 2 and 4 were immunostained to assess the distribution of HC-ARFIP2 (2.3.2.8). The embryos showed a stable and ubiquitous expression of the transgene as the staining was visible all over, compared to the NTGs (**Figure 5.2 B**). The intensity of the transgene expression in F1 embryos from F0 4 appeared to be higher than that observed in embryos from F0 2. The F1 embryos carrying the transgene (DsRed positive embryos) were raised to adulthood and stable heterozygous HC-*ARFIP2* transgenic lines were established (now referred to as Line 2 and Line 4).

Overall the results showed the successful generation of HC-ARFIP2 expressing transgenic zebrafish lines. Furthermore, the transgene expression was stable and ubiquitous, as seen in embryos at 5 dpf.
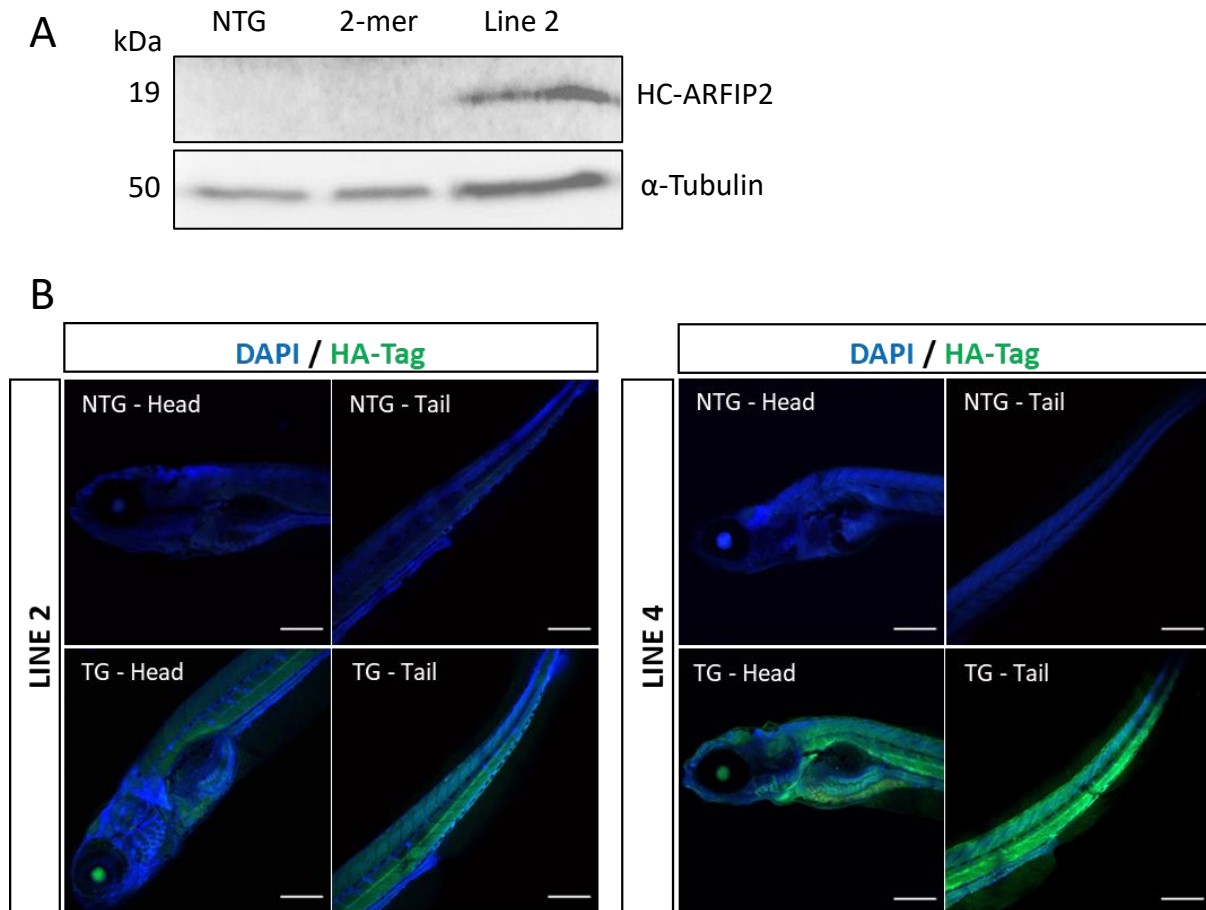
Figure 5.2: Expression and distribution of HC-ARFIP2 in embryos (F1) from crossing identified positive founders (F0) and wild-types.
Line 2 and Line 4 positive founders (F0) were crossed with AB wild-type zebrafish to produce heterozygous *HC-ARFIP2* transgene expressing embryos (F1). Embryos that showed ubiquitous DsRed fluorescence were sorted and used to assess HC-ARFIP2 expression and distribution. **(A)** At 5 days post fertilisation (dpf), the embryos from line 2 F0 were deyolked and the protein was extracted and assessed on a western blot. HC-ARFIP2 expression was detected using anti HA-tag antibody (1:1000). HC-ARFIP2 was detected at 19kDa. α-Tubulin at 50kDa was used as a loading control. **(B)** Zebrafish embryos from crossing F0 (line 2 and line 4) with AB wild-type zebrafish were fixed in fish fix at 5dpf. The embryos were stained to see the distribution of HC-ARFIP2 using the anti-HAtag antibody (1:500). Hoechst (1 µg/ml) was used to stain nuclei. Zebrafish showed a stable expression of HC-ARFIP2 and there was a ubiquitous expression of the transgene. Scale bar = 250µm. NTG = non-transgenic. 2-mer = positive control embryos that contain 2 G4C2 repeats and were generated using the same vector backbone.

5.3.2 Crossing HC-ARFIP2 transgenic fish to C9ORF72-ALS sense and antisense lines resulted in four genetically distinct experimental populations

The heterozygous HC-*ARFIP2* transgenic lines were crossed to the heterozygous *C9ORF72*-ALS sense (G4C2) and antisense (C4G2) lines (2.3.2.2) to generate zebrafish populations that would express HC-ARFIP2 in the *C9ORF72*-ALS background and would allow assessment of the effects of HC-ARFIP2 on *C9ORF72*-ALS associated characteristics. The offspring from the crosses were viable and generated all four possible genotypes.

5.3.2.1 Breeding HC-ARFIP2 transgenics with C9ORF72-ALS zebrafish lines produce three populations with distinct DsRed expression phenotypes

After crossing the transgenic lines with the *C9ORF72*-ALS lines, the embryos were heat-shocked (2.3.2.4). Both the *C9ORF72* heterozygous and transgenic heterozygous lines carry the DsRed reporter gene and when crossed, though the cross was expected to give four different genotypes (**Figure 5.3 A**), the zebrafish fry displayed three DsRed phenotypes with different fluorescence intensities under the fluorescence microscope – bright, faint and no DsRed expression (**Figure 5.3 B**). The DsRed expression was earlier observed to be faint in the *HC-ARFIP2* transgenic embryos compared to the bright DsRed expression in *C9ORF72*-ALS embryos. With this knowledge, it was predicted that the zebrafish embryos with the *C9ORF72*-ALS background would display a very bright DsRed expression (**+/+[1]** and **+/-[1]**), regardless of the presence of the HC-*ARFIP2* transgene. The faint DsRed expressing embryos would be the HC-*ARFIP2* transgenic embryos (**+/-[2]**) and the no DsRed expressing embryos would be the non-transgenics (**-/-[3]**). The yolk sac of the embryos (indicated using yellow arrows) for all three DsRed phenotypes were observed to have a bright fluorescence due to autofluorescence. The zebrafish embryos from each cross (with sense and antisense lines) were sorted into the three distinct populations based on the DsRed phenotypes and raised to adulthood for later genotyping.

A

C9ORF72

X	+	−
+	**+/+¹**	**+/-²**
−	**+/-¹**	**-/-³**

HC-ARFIP2

B
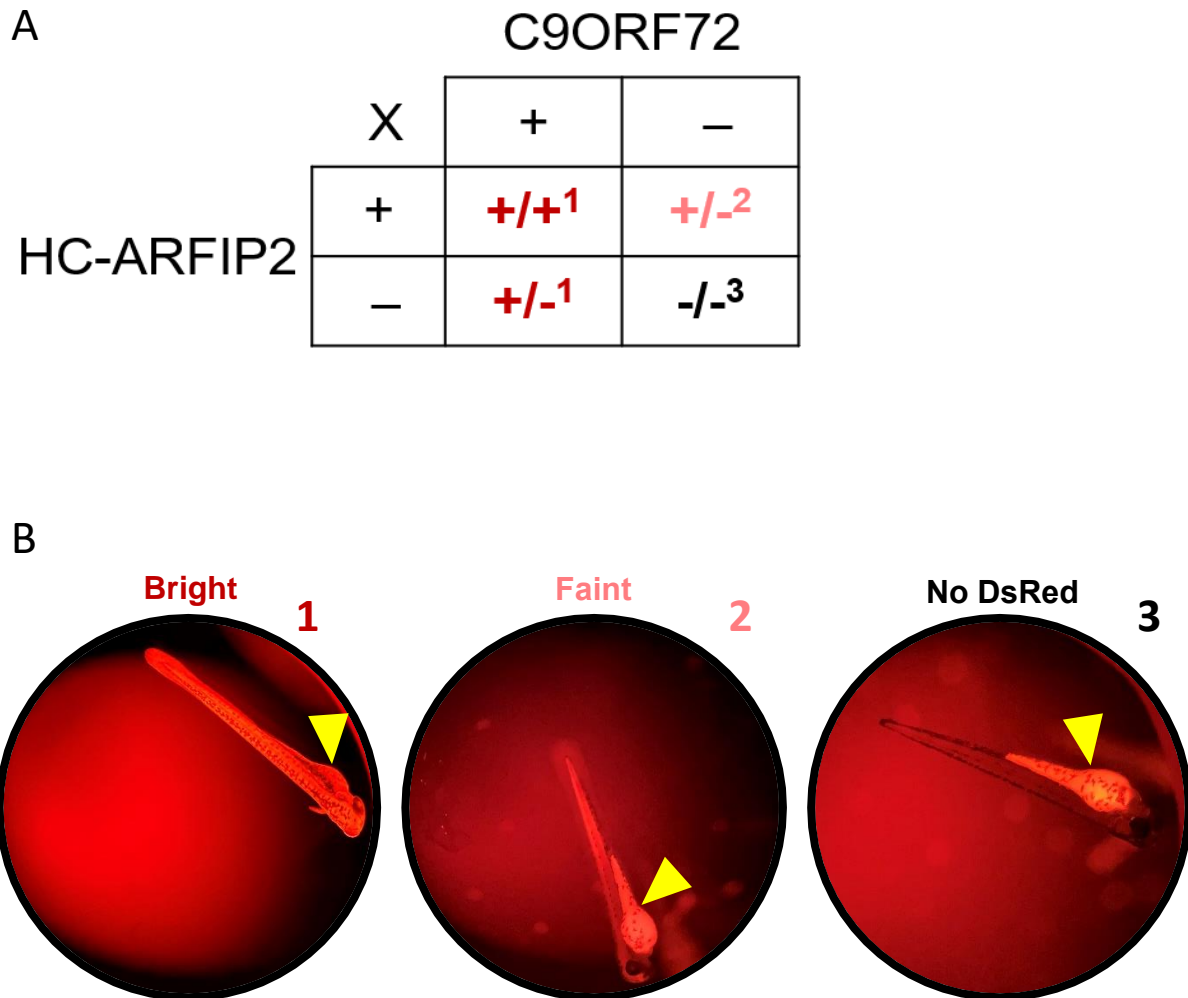
Figure 5.3: Expected genotypes and observed phenotypes from crossing HC-ARFIP2 F1 with C9ORF72 sense and antisense lines

(A) Punnett square showing the expected genotypes when *HC-ARFIP2* line 2 and line 4 F1 heterozygous (+/-) zebrafish were crossed to *C9ORF72*-ALS sense (G4C2) and antisense (C4G2) heterozygous (+/-) lines. The two lines were crossed with an expectation to see 25% double mutant (HC/C9) zebrafish which would have both HC-ARFIP2 and C9ORF72 genes (+/+), 25% C9ORF72 (C9) only expressing zebrafish (+/+), 25% HC-ARFIP2 (HC) only expressing zebrafish (+/-) and 25% non-transgenics (NTG) (-/-). **(B)** Representative pictures of the zebrafish under the optical microscope used for sorting the embryos. The zebrafish fry was sorted into three populations based on DsRed fluorescence phenotype – Bright (50%), Faint (25%) and No DsRed (25%). The bright DsRed phenotype corresponds to a mixed population of double mutants with both the *HC-ARFIP2* insert and C9 repeats (HC/C9); and fish with C9 repeats but no *HC-ARFIP2* insert (C9). The faint DsRed phenotype corresponds *HC-ARFIP2* transgenics with no C9 repeats (HC). The no DsRed phenotype corresponds non-transgenics with no *HC-ARFIP2* or C9 repeats (NTG).

5.3.2.2 Validation of primers for genotyping zebrafish

Before genotyping adult zebrafish with unknown genotypes, the primers designed for genotyping (**Table *2.22***) were validated using 5 dpf embryos of known genotype (**Figure 5.4**). Non-transgenic embryos (NTGs), embryos from each HC-*ARFIP2* transgenic line (lines 2 and 4), and embryos from each of the *C9ORF72*-ALS sense (C9-G4C2) and antisense (C9-C4G2) lines were used. Three embryos for each genotype were used and genotyped (2.3.2.5). The non-transgenic embryos did not show a band for the *HC-ARFIP2* transgene (Cter), however, there was a band for *C9ORF72* (C9) but at the incorrect height. The band corresponding to the *C9ORF72* gene was expected to be at 280 bp. The two HC-*ARFIP2* transgenic lines (line 2 and line 4) showed the presence of the HC-*ARFIP2* transgene with the presence of bands at 703 bp (Cter). As seen in NTGs, there was a band for C9 but at the incorrect height. Both the sense and antisense *C9ORF72*-ALS lines showed no band for Cter and hence indicated the absence of the HC-*ARFIP2* transgene. Both *C9ORF72*-ALS lines showed bands for C9 at the right height of 280 bp and there was no band at the incorrect height that was seen in other genotypes. All genotypes showed bands for the E8, E2, and E4 reference primers at the correct heights (268, 254, and 353 bp, respectively). These reference primers were used as positive controls during the genotyping process. A no-DNA template sample for each primer pair was used as a negative control. No bands were observed for the no DNA template samples. Overall, the results showed that all the primers designed and validated were suitable for genotyping zebrafish.
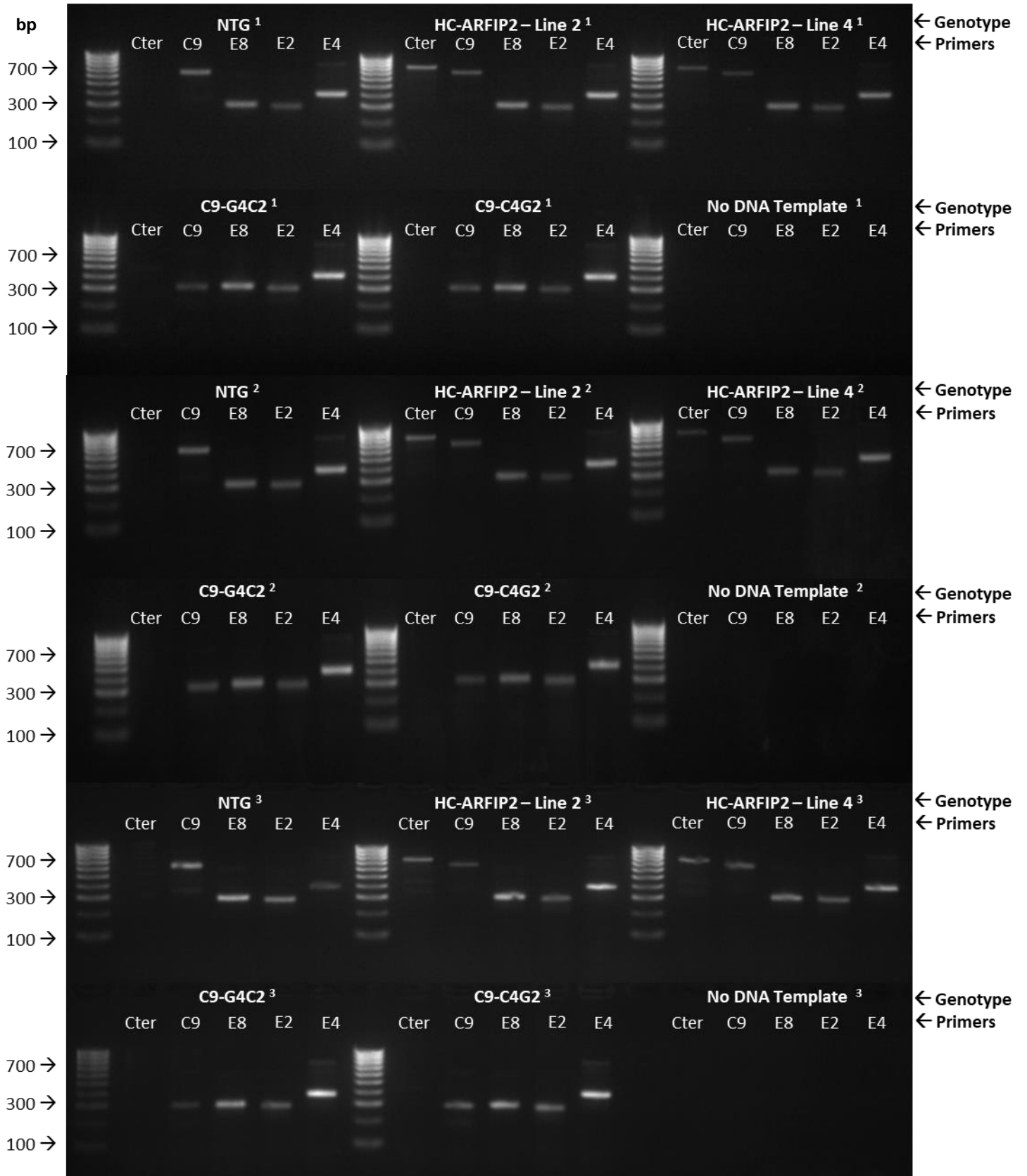
Figure 5.4: Validating primers for genotyping using 5 dpf zebrafish embryos with known genotypes

Selected primers were designed and validated to genotype zebrafish with known genotypes. Primers specific for *HC-ARFIP2* (Cter), *C9ORF2* sense (G4C2) and antisense (C4G2) repeats (C9) and zebrafish reference genes (E8, E2 and E4) as positive controls were used. Zebrafish genotypes used were non-transgenics (NTG), line 2 and line 4 F1, and *C9ORF2* sense and antisense embryos (n=3). A no DNA template was used as negative control. Hyper ladder IV (100bp) was used. Expected bands (bp) – Cter (703), C9 (280), E8 (268), E2 (254) and E4 (353).

After validating the primers using known genotypes, the primers were used to genotype embryos at 5 dpf of unknown genotypes from crossing each HC-*ARFIP2* transgenic line (lines 2 and 4) with the sense and antisense *C9ORF72*-ALS lines (**Figure 5.5**) before genotyping the adults raised from the crosses. The embryos were heat-shocked (2.3.2.4) and sorted according to the DsRed fluorescence phenotype (bright, faint, or no DsRed expression). From each cross, 3 no DsRed expressing, 3 faint DsRed expressing, and 6 bright DsRed expressing embryos were chosen at random for genotyping (2.3.2.5). As seen previously, in the absence of the C9 insert, there was a band for C9 at the incorrect height. Crossing HC-*ARFIP2* line 2 with the sense (**Figure 5.5 A**) and antisense (**Figure 5.5 B**) *C9ORF72*-ALS lines showed that all the embryos with no DsRed expression did not have the HC-*ARFIP2* insert or the *C9ORF72* repeats and were NTGs. The faint DsRed expressing embryos were shown to be HC-*ARFIP2* transgenics as they showed the presence of the HC-*ARFIP2* insert only. A mixed population of double mutants, with both the HC-*ARFIP2* insert and sense or antisense *C9ORF72* repeats and embryos with the sense or antisense *C9ORF72* repeats but no HC-*ARFIP2* insert, was seen in the bright DsRed expressing embryos. In each cross, 50% were double mutants and 50% were embryos with only the *C9ORF72* repeat in the bright DsRed expressing embryos. When crossing HC-*ARFIP2* line 4 with the sense (**Figure 5.5 C**) and antisense (**Figure 5.5 D**) *C9ORF72*-ALS lines, comparable results were seen. However, in the embryos with no DsRed expression from crossing line 4 with the antisense line, 2 out of 3 embryos were HC-*ARFIP2* transgenics which may be attributed to an error in visually sorting the embryos according to DsRed phenotypes. Furthermore, in line 4 cross with the antisense C9 line, 4 embryos were double mutants, and 2 embryos were those with only the *C9ORF72* repeats. The E4 reference primer was chosen as the positive control. Genotyping for embryos of known genotypes (**Figure 5.5 E**) was undertaken along

with genotyping of unknown genotypes for the correct interpretation of the results. Overall, the results showed that the primers used were able to successfully genotype zebrafish embryos, resulting in four genetically distinct experimental populations.
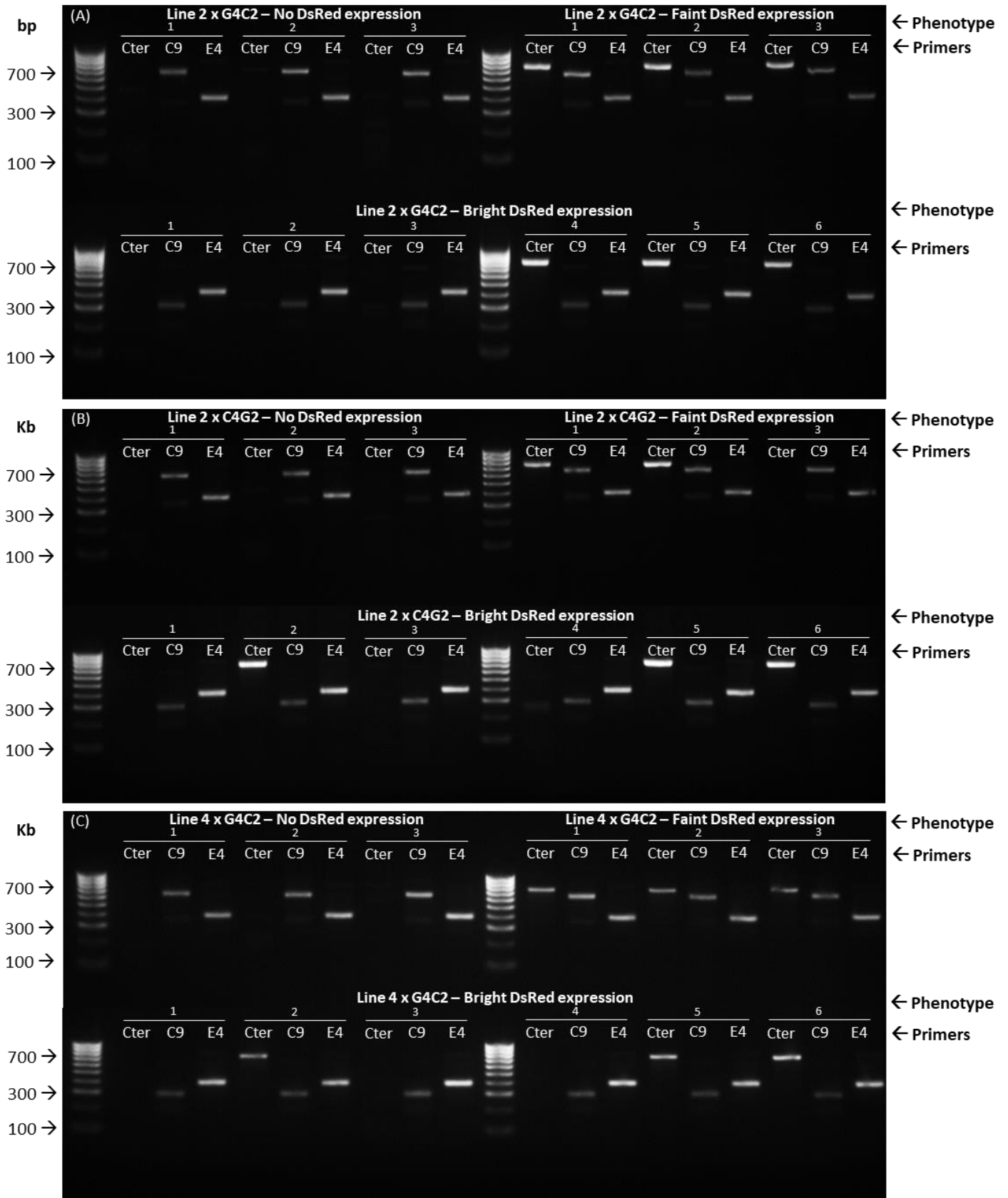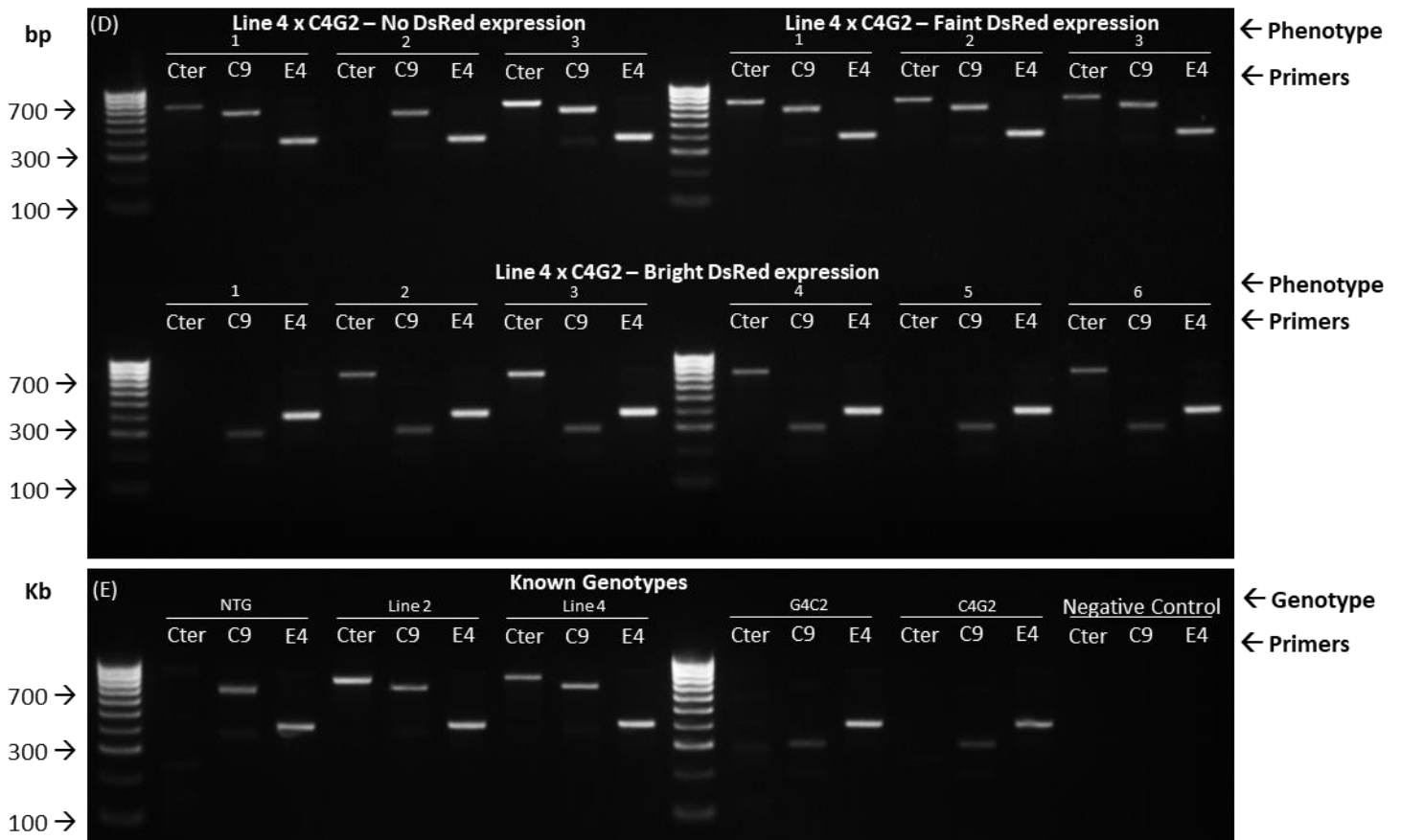
Figure 5.5: Genotyping zebrafish fry with unknown genotypes obtained from crossing HC-*ARFIP2* transgenics with sense and antisense *C9ORF72*-ALS lines

HC-ARFIP2 heterozygous transgenics were crossed with sense (G4C2) and antisense (C4G2) *C9ORF72*-ALS heterozygous lines. The zebrafish fry was heat-shocked and separated into three populations according to DsRed phenotypes – no DsRed expression, faint DsRed expression and bright DsRed expression. The fry was then genotyped using primers specific for *HC-ARFIP2* (cter), *C9ORF2* sense and antisense repeats (C9) and a reference gene *(E4)* as positive control. **(A)** *HC-ARFIP2* line 2 crossed with sense (G4C2) *C9ORF72* heterozygous fish. **(B)** *HC-ARFIP2* line 2 crossed with antisense (C4G2) *C9ORF72* heterozygous fish. **(C)** *HC-ARFIP2* line 4 crossed with sense (G4C2) *C9ORF72* heterozygous fish. **(D)** *HC-ARFIP2* line 4 crossed with antisense (C4G2) *C9ORF72* heterozygous fish. **(E)** known genotype fish and negative controls. No DsRed expressors (n=3) are non-transgenics with no *HC-ARFIP2* or C9 repeats (NTG). Faint DsRed expressors (n=3) are *HC-ARFIP2* transgenics with no C9 repeats (HC). The bright DsRed expressors (n=6) are a mixed population of double mutants with both the *HC-ARFIP2* insert and C9 repeats (HC/C9); and fish with C9 repeats but no *HC-ARFIP2* insert (C9). Expected bands (bp) – Cter (703), C9 (280), and E4 (353).

5.3.2.3 Genotyping of adult zebrafish via fin clipping

The zebrafish fry obtained from crossing the two HC-*ARFIP2* transgenic lines with the sense and antisense *C9ORF72*-ALS lines were sorted according to DsRed phenotypes at 3 dpf, raised to adulthood, and genotyped at 3 months via fin clipping (2.3.2.5) to further sort the zebrafish according to genotypes. The adults were genotyped using primers specific for HC-*ARFIP2* (Cter), *C9ORF72* sense and antisense repeats (C9), and a reference gene (*E4*) as the positive control, as previously shown. **Figures A1 – A11** in **Appendix 3** show the results of genotyping of the adult fish. Results showed that there were 18 double mutants (now referred to as HC/C9), 16 adults that only carry the *C9ORF72* repeats (now referred to as C9), 33 adults that only carry the HC-*ARFIP2* transgene (now referred to as HC) and 22 NTGs from crossing line 2 with the sense *C9ORF72*-ALS line; 15 HC/C9, 33 C9, 21 HC, and 37 NTGs from crossing line 2 with the antisense *C9ORF72*-ALS line; 29 HC/C9, 18 C9, 41 HC, and 42 NTGs from crossing line 4 with the sense *C9ORF72*-ALS line and 54 HC/C9, 17 C9, 7 HC and 24 NTGs from crossing line 4 with the antisense *C9ORF72*-ALS line. Overall, the results showed that the primers used were able to genotype adult zebrafish at 3 months old via fin clipping and four distinct genotype populations were generated as expected. Post-genotyping, the zebrafish adults were sorted into the four genetically distinct populations – HC/C9, C9, HC, and NTGs, and allowed to develop further.

5.3.2.4 Comparison of observed phenotypes and genotypes with expected phenotypes and genotypes

After genotyping the adults, the observed phenotype (at 3 dpf) and genotype ratios (at 3 months old) of the zebrafish were compared with the expected phenotype ratios (50% bright DsRed, 25% faint DsRed, and 25% no DsRed) and genotype ratios (25% for all expected genotypes) from crossing heterozygous HC-*ARFIP2* transgenic lines with heterozygous *C9ORF72*-ALS lines (**Figure 5.6 and 5.7**). This comparison assessed whether a particular genotype was more favourable than the others. For the HC-*ARFIP2* line 2 crossed with the sense and antisense *C9ORF72*-ALS lines, the observed phenotype ratios from both crosses were not significantly different (ns) from the expected phenotype ratios (**Figure 5.6 A**). The observed genotype ratios from line 2 crossed with the sense (G4C2) line were not significantly different (ns) from the expected genotype ratios. However, the observed genotype ratios from line 2 crossed

with the antisense (C4G2) line were significantly different from the expected genotype ratios (**Figure 5.6 B**). In this cross, the NTG genotype population was approximately 10% higher than the expected percentage. The observed phenotype ratios from crossing line 4 with sense (G4C2) and antisense (C4G2) were both significantly different from the expected phenotype ratios (**Figure 5.7 A**). In both crosses, the no DsRed expression population was a lot higher than expected. In line 4 crossed with the sense line, the faint DsRed expression population was also higher than expected. The observed genotype ratios from line 4 crossed with the sense (G4C2) line were significantly different from the expected genotype ratios where NTG and HC were approximately 7% higher than the expected percentage and the C9 population was 12% lower than expected (**Figure 5.7 B**). Analysis for line 4 crossed with the antisense (C4G2) line could not be performed as the faint phenotype tank (HC) was not available for genotyping due to contamination of water quality. Results highlighted that where significant differences between observed and expected phenotypes and genotypes were seen, the no DsRed expressing or NTG population was higher than expected. This shows that the NTG genotype was more favourable than the genetically modified genotypes (transgenic or disease mutation). A decrease in the total number of zebrafish analysed from phenotype to genotype can be seen in both **Figures 5.6 and 5.7**. This is because the phenotypes were recorded at 5 dpf and genotypes at 3 months old. As the zebrafish are raised to adulthood, a reduction in the total number of embryos raised is always observed.

A

Line 2 x G4C2 Phenotype

- 24.53% No DsRed
- 33.96% Faint DsRed
- 41.51% Bright DsRed

NS

Total=106

Line 2 x C4G2 Phenotype

- 29.71% No DsRed
- 24.64% Faint DsRed
- 45.65% Bright DsRed

NS

Total=138

B

Line 2 x G4C2 Genotype

- 24.72% NTG
- 37.08% HC
- 20.22% HC/C9
- 17.98% C9

NS

Total=89

Line 2 x C4G2 Genotype

- 34.91% NTG
- 19.81% HC
- 14.15% HC/C9
- 31.13% C9

**

Total=106
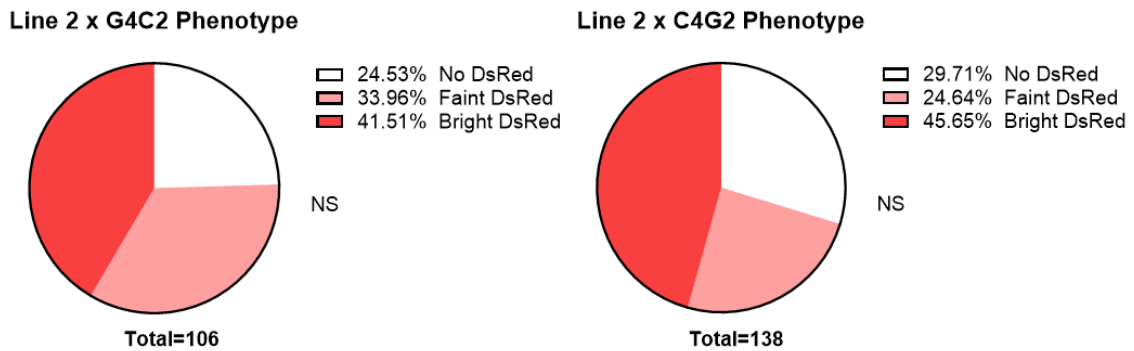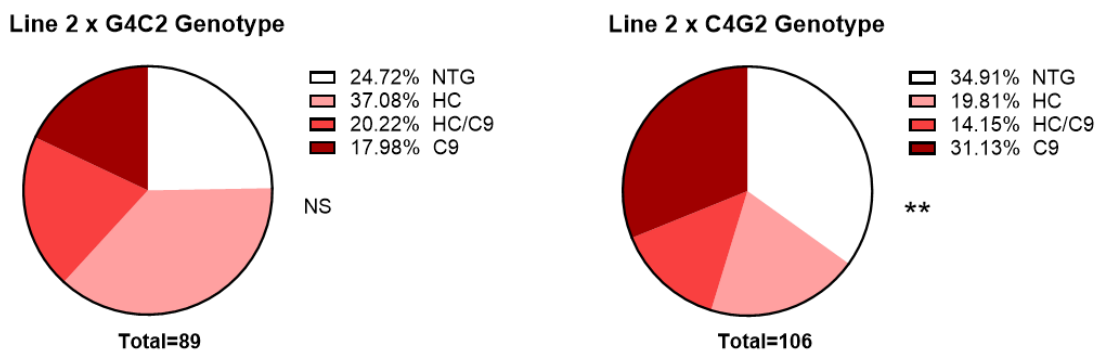
Figure 5.6: Comparison of phenotypes and genotypes recorded in zebrafish populations obtained from crossing HC-*ARFIP2* line 2 with *C9ORF72* sense and antisense lines

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with sense (G4C2) and antisense (C4G2) *C9ORF72* heterozygous lines. The phenotypes and genotypes of each population were recorded to compare the numbers against expected mendelian ratios (Figure 5.3 A). **(A)** After crossing, the zebrafish fry was separated into three phenotypes according to DsRed expression – Bright, Faint and No DsRed. The observed phenotype ratios from both crosses were not significantly different (ns) to expected phenotype ratios (50% bright DsRed, 25% faint DsRed and 25% no DsRed). **(B)** The assorted zebrafish were raised to adulthood and genotyped at 3 months to identify non-transgenics with no *HC-ARFIP2* or *C9ORF72* repeats (NTG), *HC-ARFIP2* only transgenics with no *C9ORF72* repeats (HC), double mutants with both the *HC-ARFIP2* insert and *C9ORF72* repeats (HC/C9) and those with *C9ORF72* only repeats but no *HC-ARFIP2* insert (C9). The observed genotype ratios from line 2 crossed with the sense (G4C2) line were not significantly different (ns) to expected genotype ratios (25% for all 4 expected genotypes). However, the observed genotype ratios from line 2 crossed with the antisense (C4G2) line were significantly different (**, p=0.0078) to expected genotype ratios (25% for all 4 expected genotypes). Statistical analysis to compare observed distribution with expected was done using a chi-squared test for goodness of fit. p<0.05. ns = p > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001.

A

Line 4 x G4C2 Phenotype

30.51% No DsRed
29.94% Faint DsRed
39.55% Bright DsRed

*

Total=177

Line 4 x C4G2 Phenotype

32.50% No DsRed
26.88% Faint DsRed
40.63% Bright DsRed

*

Total=160

B

Line 4 x G4C2 Genotype

32.31% NTG
31.54% HC
22.31% HC/C9
13.85% C9

**

Total=130

Line 4 x C4G2 Genotype

23.53% NTG
6.86% HC
52.94% HC/C9
16.67% C9

#

Total=102
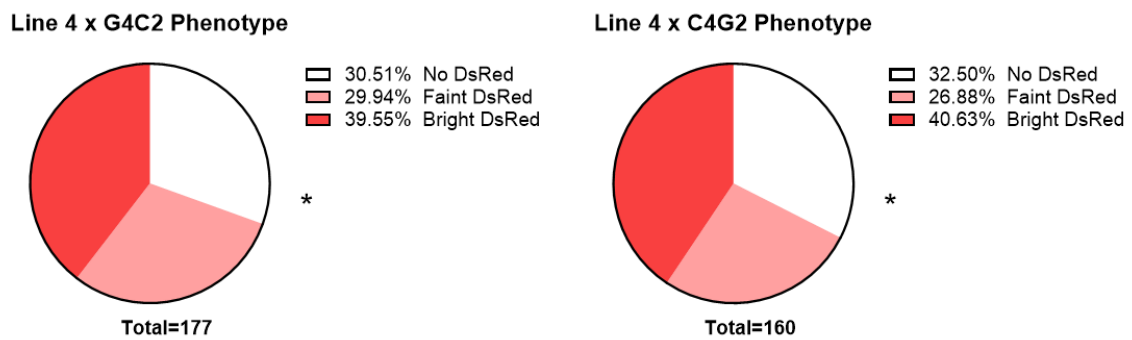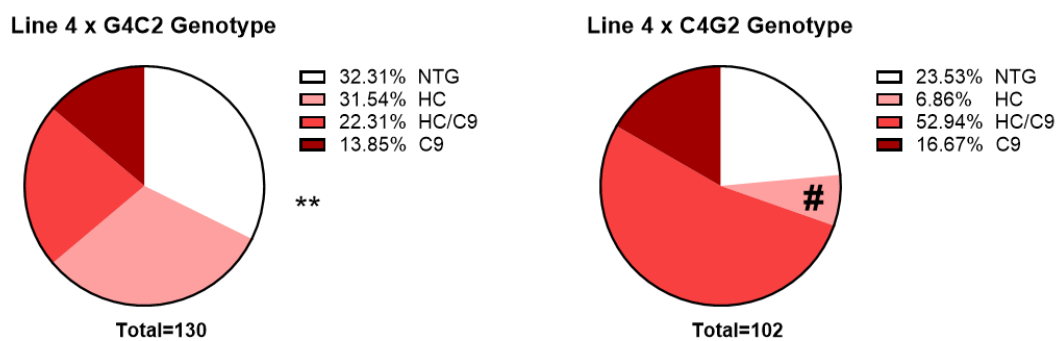
Figure 5.7: Comparison of phenotypes and genotypes recorded in zebrafish populations obtained from crossing HC-*ARFIP2* line 4 with *C9ORF72* sense and antisense lines

HC-ARFIP2 Line 4 heterozygous transgenics were crossed with sense (G4C2) and antisense (C4G2) *C9ORF72* heterozygous lines. The phenotypes and genotypes of each population were recorded to compare the numbers against expected mendelian ratios (Figure 5.3 A). **(A)** After crossing, the zebrafish fry was separated into three phenotypes according to DsRed expression – Bright, Faint and No DsRed. The observed phenotype ratios from crossing line 4 with sense (G4C2) and antisense (C4G2) were both significantly different (*, p=0.028 and p=0.0362 respectively) to the expected phenotype ratios (50% bright DsRed, 25% faint DsRed and 25% no DsRed). **(B)** The assorted zebrafish were raised to adulthood and genotyped at 3 months to identify non-transgenics with no *HC-ARFIP2* or *C9ORF72* repeats (NTG), *HC-ARFIP2* only transgenics with no *C9ORF72* repeats (HC), double mutants with both the *HC-ARFIP2* insert and *C9ORF72* repeats (HC/C9) and those with *C9ORF72* only repeats but no *HC-ARFIP2* insert (C9). The observed genotype ratios from line 4 crossed with the sense (G4C2) line were significantly different (**, p=0.0079) to expected genotype ratios (25% for all 4 expected genotypes). Analysis for line 4 crossed with the antisense (C4G2) line could not be performed as the faint phenotype tank (HC) was not available. Statistical analysis to compare observed distribution with expected was done using a chi-squared test for goodness of fit. p<0.05. ns = p > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001.

5.3.3 Effect of HC-*ARFIP2* expression on motor performance in *C9ORF72*-ALS zebrafish lines

5.3.3.1 HC-*ARFIP2* rescues the swimming endurance phenotype in the *C9ORF72*-ALS antisense background

The embryos from crossing HC-*ARFIP2* transgenic line 2 with the *C9ORF72*-ALS sense (G4C2) and antisense (C4G2) were raised to adulthood and genotyped by fin clipping at 3 months (2.3.2.5) and sorted accordingly into the four genotypes – NTG, C9, HC, and HC/C9. The zebrafish were raised further. At 6 months, the neuro-muscular integrity of the fish and the effect of HC-*ARFIP2* on motor performance in zebrafish was assessed by conducting a swim tunnel experiment (2.3.2.10) to measure the swimming endurance of each fish (n=10 per genotype). **Figure 5.8** shows the percentage of fish still swimming as the flow rate of water (L/min) increased over time and therefore shows the swimming endurance for each genotype population. In the cross with the sense *C9ORF72*-ALS line (**Figure 5.8 A**), the C9-S fish failed to continue swimming at earlier time points and at a lower flow rate than their NTG, HC, and HC/C9-S clutch mates. The C9-S fish had the lowest median of 22.3 mins where 50% of the fish were still swimming. There was a significant difference in the swimming endurance between the C9-S fish and the NTG fish. There was a significant difference between C9-S and HC/C9-S, where the swimming endurance was improved in the HC/C9-S population. This highlighted that with the expression of HC-ARFIP2, the *C9ORF72*-ALS-associated motor performance phenotype in zebrafish was improved in the presence of the disease background. The HC/C9-S fish swimming endurance was well aligned with their NTG clutch mates and showed no significant differences. Comparable results were seen when HC-*ARFIP2* line 2 was crossed with the antisense *C9ORF72*-ALS line (**Figure 5.8 B**).

A L2 x G4C2

Comparison	Gehan-Breslow-Wilcoxon test
NTG vs C9-S	**
NTG vs HC	*
NTG vs HC/C9-S	ns
C9-S vs HC	*
C9-S vs HC/C9-S	*
HC vs HC/C9-S	ns

B L2 x C4G2

Comparison	Gehan-Breslow-Wilcoxon test
NTG vs C9-A	***
NTG vs HC	ns
NTG vs HC/C9-A	ns
C9-A vs HC	****
C9-A vs HC/C9-A	**
HC vs HC/C9-A	ns
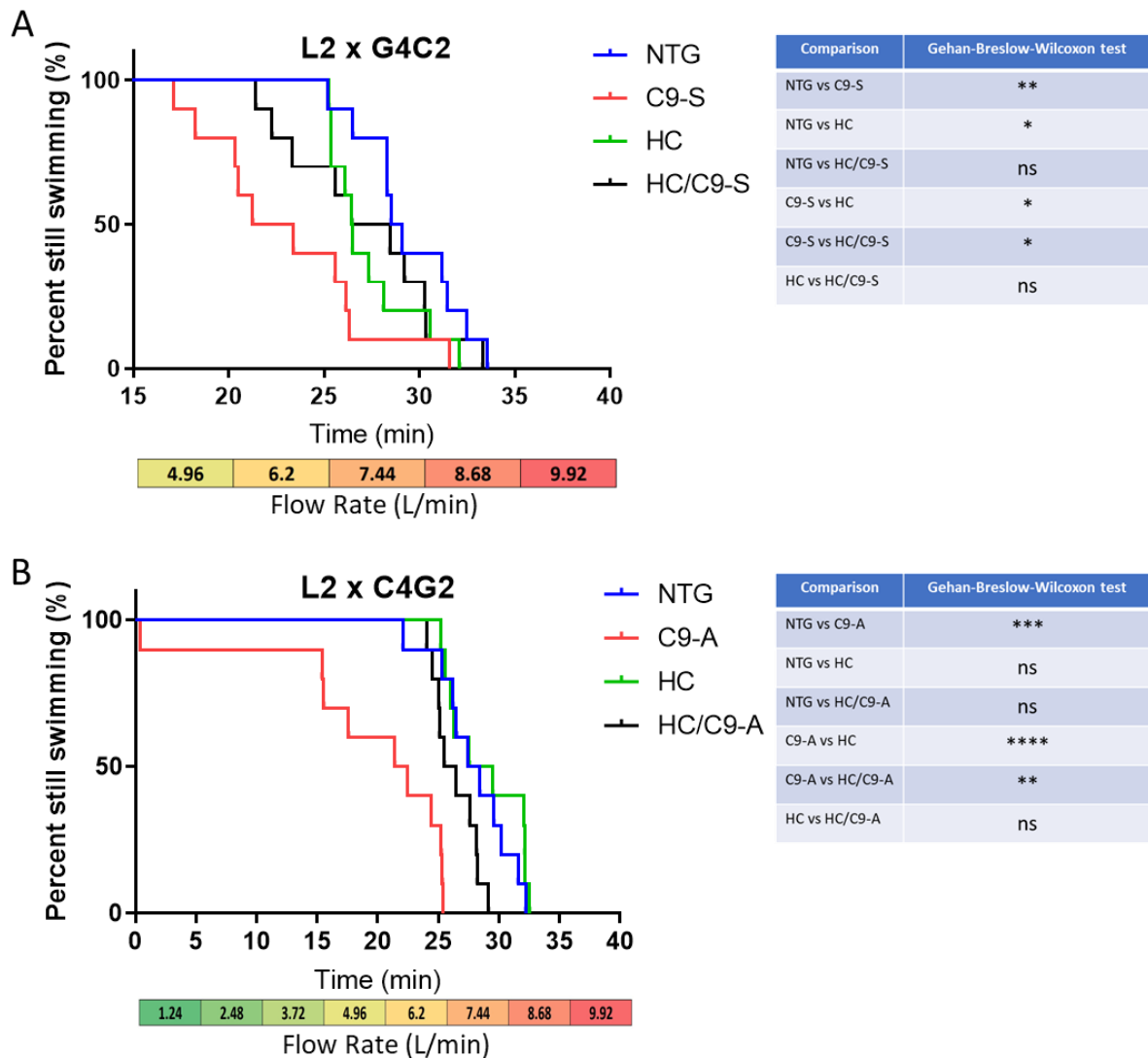
Figure 5.8: Effect of expressing HC-ARFIP2 on swimming endurance in sense and antisense *C9ORF72* zebrafish crossed with HC-*ARFIP2* line 2

Kaplan-Meier curves of swimming endurance for genotyped adults at 6 months of age from *HC-ARFIP2* line 2 crossed with the sense and antisense *C9ORF72* lines. **(A)** *HC-ARFIP2* line 2 crossed with the sense (G4C2) C9ORF72 line. At 6 months old, the C9-S fish (median=22.33 mins) failed to continue swimming at earlier time points than their NTG (median=28.79 mins), HC (median=26.465 mins) and HC/C9-S (median=27.45 mins) clutch mates. There was a significant difference between C9-S and NTG, C9-S and HC/C9-S; and C9-S and HC. The HC/C9-S fish swimming endurance was well aligned with their NTG clutch mates and showed no significant differences. N=10 per genotype. **(B)** *HC-ARFIP2* line 2 crossed with the antisense (C4G2) *C9ORF72* line. At 6 months old, the C9 fish (median=21.935 mins) failed to continue swimming at earlier time points compared to their NTG (median=27.93 mins), HC (median=28.525 mins) and HC/C9 (median=25.985 mins) clutch mates. There was a significant difference between C9 and NTG, C9 and HC/C9 and C9 and HC. The HC/C9 fish swimming endurance was well aligned with their NTG clutch mates and showed no significant differences. N=10 per genotype. A stronger positive impact of HC-ARFIP2 co-expression was seen in the C9 antisense line than in the C9 sense line. Data were analysed using a Gehan-Breslow-Wilcoxon test. ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

The swimming endurance tested was quantified by calculating the critical swimming velocity (Ucrit) (**Figure 5.9**). In the HC-*ARFIP2* line 2 crossed with both the sense (**Figure 5.9 A**) and antisense lines (**Figure 5.9 B**), a significant motor endurance deficit phenotype in the C9-S and C9-A genotyped was observed, where the C9-S and C9-A genotypes had significantly lower swimming velocities (mean=5.3 cm/s and mean=4.6 cm/s, respectively) compared to NTGs (mean=6.7 cm/s and mean=6.5 cm/s, respectively) and HC clutch mates (mean=6.6 cm/s and mean=6.5 cm/s, respectively). The antisense C9-A zebrafish displayed a stronger phenotype of swimming velocity deficits than the C9-S zebrafish. In the antisense background (**Figure 5.9 B**), the HC/C9-A population showed a significant rescue of this phenotype where their swimming velocity was similar to that of the NTGs (mean=6.0 cm/s and mean=6.5 cm/s, respectively), even though a much stronger phenotype of motor velocity deficit was seen in the antisense background. Though a significant rescue of the motor deficit phenotype was not observed in the sense background (**Figure 5.9 A**), the swimming endurance of the HC/C9 population (mean=6.2 cm/s) was similar to the swimming endurance of the NTGs (mean=6.8 cm/s), where the difference between the two populations was not significant. Furthermore, as seen in the HC genotypes in both the sense and the antisense background, the expression of the HC-*ARFIP2* transgene did not have any adverse effects on the motor performance of the zebrafish.

In addition to a more severe motor performance deficit observed in the C9-A population, it was interesting to see that immediately following the swim tunnel testing, most of the C9-A fish were unable to spontaneously swim around (data not shown). When placed back into the tanks, it was observed that the C9-A fish would sit at the bottom of the tank and were unable to move for a period of time. On the other hand, though the HC/C9-A fish would also initially sit at the bottom immediately after the swim tunnel test, they were able to make some movements and eventually make full recovery. Video evidence of this observation is available.

Overall, these results show that the 6-month-old adult zebrafish with the *C9ORF72*-ALS mutation, both in the sense and antisense background, show a strong and significant phenotype of swimming endurance and velocity deficits. The C9-A genotype zebrafish show a much stronger and more severe motor performance deficit phenotype compared to the C9-S genotype zebrafish. The preliminary data here

showed that expressing HC-ARFIP2 in the ALS disease background (particularly in the antisense background) significantly rescues this phenotype and therefore the therapeutic potential of HC-ARFIP2 in *C9ORF72*-ALS zebrafish is encouraging but requires further validation. Furthermore, HC-*ARFIP2* transgene expression does not show any adverse effects on motor performance.
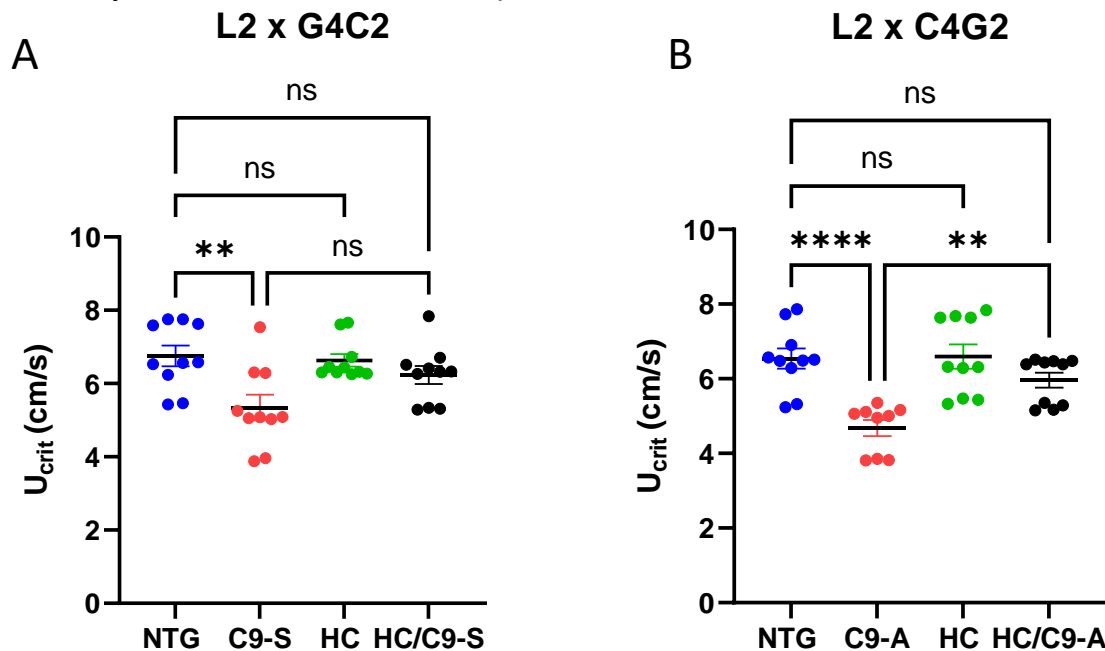
Figure 5.9: Adult C9 zebrafish show significant reduction in their swimming velocity at 6 months of age compared to their clutch mates and HC-ARFIP2 expression can potentially rescue this phenotype

Swimming velocity was determined by calculating the Critical Swimming Velocity (U_{crit}) of matched adults by conducting a swim tunnel procedure. Measurements from the procedure allowed for calculations of the U_{crit} for each fish. **(A)** U_{crit} of 6-month-old adult genotyped populations from the *HC-ARFIP2* line 2 crossed with the sense (G4C2) *C9ORF72* line. There is a significant difference between C9-S (mean=5.347 cm/s) and NTG (mean=6.752 cm/s) which displays a clear phenotype in the C9-S population. HC mean=6.633 cm/s and HC/C9-S mean=6.231 cm/s. N=10 per genotype. Total sample size=40 and actual power=0.9292795. **(B)** U_{crit} of 6-month-old adult genotyped populations from the HC-ARFIP2 line 2 cross with the antisense (C4G2) C9ORF72 line. There is a significant difference between C9-A (mean=4.679 cm/s) and NTG (mean=6.539 cm/s) which displays a clear phenotype in the C9-A population. This phenotype is significantly rescued in the HC/C9-A (mean= 5.964 cm/s) population. HC mean=6.596 cm/s. N=10 per genotype. Total sample size=20 and actual power=0.9025828. The antisense C9-A zebrafish display a stronger phenotype of swimming velocity deficits compared to C9-S zebrafish. Data were analysed using the ordinary one-way ANOVA with a Šídák's multiple comparisons test. Error bars represent ± SEM; p<0.05. ns = p > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001.

5.3.3.2 Effect of gender on motor performance in zebrafish

The effect of gender on the motor performance of zebrafish in each genotype population was assessed to see the impact of gender. Gender-specific analysis was carried out to observe how the genotypes compare with each other in specific genders and whether the effect of HC-ARFIP2 is gender-specific (**Figures 5.10 and 5.11**). In the genotype populations obtained from crossing HC-*ARFIP2* line 2 with the sense *C9ORF72*-ALS line (**Figure 5.10**), the females failed to continue swimming at earlier time points than their male clutch mates in all genotypes (**Figure 5.10 A-D**). There was a significant difference in swimming endurance between males and females in each genotype (NTG, C9-S, HC, and HC/C9-S). **Figure 5.10 E and F** show the gender-specific analysis. Results showed that at 6 months old, both C9-S males (**Figure 5.10 E**) and females (**Figure 5.10 F**) had the lowest swimming endurance compared to the other genotypes. The C9-S males (median=26.13 mins) significantly failed to continue swimming at earlier time points than their male NTG (median=31.45 mins) clutch mates. There were no significant differences in the swimming endurances between the males in other cohorts. There was no significant rescue of the swimming endurance deficit in the HC/C9-S male population. The C9-S females (median=20.33 mins) significantly failed to continue swimming at earlier time points than their female NTG (median=28.3 mins), HC (median=25.36 mins), and HC/C9-S (median=23.32 mins) clutch mates. The HC/C9-S female swimming endurance was well aligned with their female NTG clutch mates and showed no significant differences. There was a significant rescue of the swimming endurance deficit in the HC/C9-S female population.

Figure 5.10: Effect of gender on swimming endurance and gender-specific analysis of the effect of HC-ARFIP2 on swimming endurance in genotypes obtained from crossing HC-*ARFIP2* line 2 with the sense *C9ORF72* zebrafish line

Kaplan-Meier curves of swimming endurance for genotyped male and female adults at 6 months of age from *HC-ARFIP2* line 2 crossed with the sense (G4C2) *C9ORF72* lines. The females failed to continue swimming at earlier time points than their male clutch mates in the **(A)** NTG genotype cohort (median=28.3 mins and 31.45 mins, respectively), **(B)** C9-S genotype cohort (median=20.33 mins and 26.13 mins, respectively), **(C)** HC genotype cohort (median=25.36 mins and 28.11 mins, respectively) and in the **(D)** HC/C9-S genotype cohort (median=23.32 mins and 30.25 mins, respectively). There was a significant difference in swimming endurance between males and females in each genotype (NTG, C9, HC and HC/C9) obtained from crossing *HC-ARFIP2* line 2 with the sense (G4C2) line. Gender-specific analysis was conducted to observe how the genotypes compare with each other for specific genders and whether the effect of HC-ARFIP2 is gender-specific. **(E)** At 6 months old, the C9-S males (median=26.13 mins) failed to continue swimming at earlier time points than their male NTG (median=31.45 mins) clutch mates. There were no significant differences in the swimming endurance between the males in other cohorts (HC median=28.11 mins, HC/C9-S median=30.25 mins). **(F)** At 6 months old, the C9-S females (median=20.33 mins) failed to continue swimming at earlier time points than their female NTG (median=28.3 mins), HC (median=25.36 mins) and HC/C9-S (median=23.32 mins) clutch mates. There was a significant difference between C9-S females and NTG females, C9-S females and HC/C9-S females; and C9-S females and HC females. The HC/C9-S female swimming endurance was well aligned with their female NTG clutch mates and showed no significant differences. The HC-ARFIP2 effect on swimming endurance in the sense *C9ORF72* background seems to be female-specific. N = 5 per gender for each genotype. Data were analysed using a Gehan-Breslow-Wilcoxon test. p<0.05. ns = p > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001.

Comparable results to those seen in the sense background were observed in the antisense background (**Figure 5.11**). The females failed to continue swimming at earlier time points than their male clutch mates in all genotypes (**Figure 5.11 A-D**). There were significant differences in swimming endurance between males and females in each genotype (NTG, C9-A, HC, and HC/C9-A). The gender-specific analysis (**Figure 5.11 E and F**) showed that at 6 months old, both C9-A males and females had the lowest swimming endurance compared to the other genotypes. The C9-A males (median=25.26 mins) and females (median=15.55 mins) significantly failed to continue swimming at earlier time points than their NTG male (median=30.19

mins) and female (median=26.22 mins), HC male (median=32.12 mins) and female (median=26.03 mins) and HC/C9-A male (median=27.02 mins) and female (median=24.27 mins) clutch mates. In both males and females, the swimming endurance in the HC genotype was well aligned with their NTG clutch mates and showed no significant differences. There was a significant rescue of the swimming endurance deficit in the HC/C9-A male and female populations.

Overall the results show that in all genotypes from the sense and antisense crosses, the females have a lower swimming endurance than their male clutch mates. Furthermore, in the sense background, the preliminary results hinted that the effect of HC-ARFIP2 expression on swimming endurance was female-specific as there was a significant rescue of the swimming endurance deficit only in the HC/C9-S female population, although a non-significant rescue trend was observed in the male population. On the other hand, there was a significant rescue of the motor performance phenotype in both male and female HC/C9-A populations in the antisense background and therefore the effect here did not appear to be gender-specific.

Figure 5.11: Effect of gender on swimming endurance gender-specific analysis of the effect of HC-ARFIP2 on swimming endurance in genotypes obtained from crossing HC-*ARFIP2* line 2 with the antisense *C9ORF72* zebrafish line

Kaplan-Meier curves of swimming endurance for genotyped male and female adults at 6 months of age from *HC-ARFIP2* line 2 crossed with the antisense (C4G2) *C9ORF72* lines. The females failed to continue swimming at earlier time points than their male (median=) clutch mates in the **(A)** NTG genotype cohort (median=26.22 mins and 30.19 mins, respectively), **(B)** C9-A genotype cohort (median=15.55 mins and 25.26 mins, respectively), **(C)** HC genotype cohort (median=26.03 mins and 32.13 mins, respectively) and **(D)** in the HC/C9-A genotype cohort (median=24.27 mins and 27.02 mins, respectively). There was a significant difference in swimming endurance between males and females in each genotype (NTG, C9, HC and HC/C9) obtained from crossing *HC-ARFIP2* line 2 with the antisense (C4G2) line. Gender-specific analysis was conducted to observe how the genotypes compared with each other for specific genders and whether the effect of HC-ARFIP2 was gender-specific. **(E)** At 6 months old, the C9-A males (median=25.26 mins) failed to continue swimming at earlier time points than their male NTG (median=30.19 mins), HC (median=32.12 mins) and HC/C9-A (median=27.02 mins) clutch mates. There was a significant difference between C9-A males and NTG males, C9-A males and HC/C9-A males; and C9-A males and HC males. The HC male swimming endurance was well aligned with their male NTG clutch mates and showed no significant differences. **(F)** At 6 months old, the C9-A females (median=15.55 mins) failed to continue swimming at earlier time points than their female NTG (median=26.22 mins), HC (median=26.03 mins) and HC/C9-A (median=24.27 mins) clutch mates. There was a significant difference between C9-A females and NTG females, C9-A females and HC/C9-A females; and C9-A females and HC females. The HC female swimming endurance was well aligned with their female NTG clutch mates and showed no significant differences. The HC-ARFIP2 effect on swimming endurance in the antisense C9ORF72 background does not seem to be gender-specific. N = 5 per gender for each genotype, except for HC/C9-A where N = 2 for females and N = 8 for males. Data were analysed using a Gehan-Breslow-Wilcoxon test. p<0.05. ns = p > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001.

5.3.3.3 Effect of weight on motor performance in zebrafish

The correlation between weight (g) and time swum (mins) was assessed to see the effect of weight on swimming endurance in zebrafish (**Figure 5.12**). In the sense background (**Figure 5.12 A**), a significant negative correlation was observed between the weight and time swum which showed that as the weight increases, the time swum and therefore the swimming endurance of the fish decreases. However, this correlation was a very weak relationship and therefore showed a weak effect of weight.

In the antisense background (**Figure 5.12 B**), an extremely weak non-significant correlation was observed between the weight and time swum. Therefore it was inferred that there was no effect of the weight on swimming endurance in the antisense background.
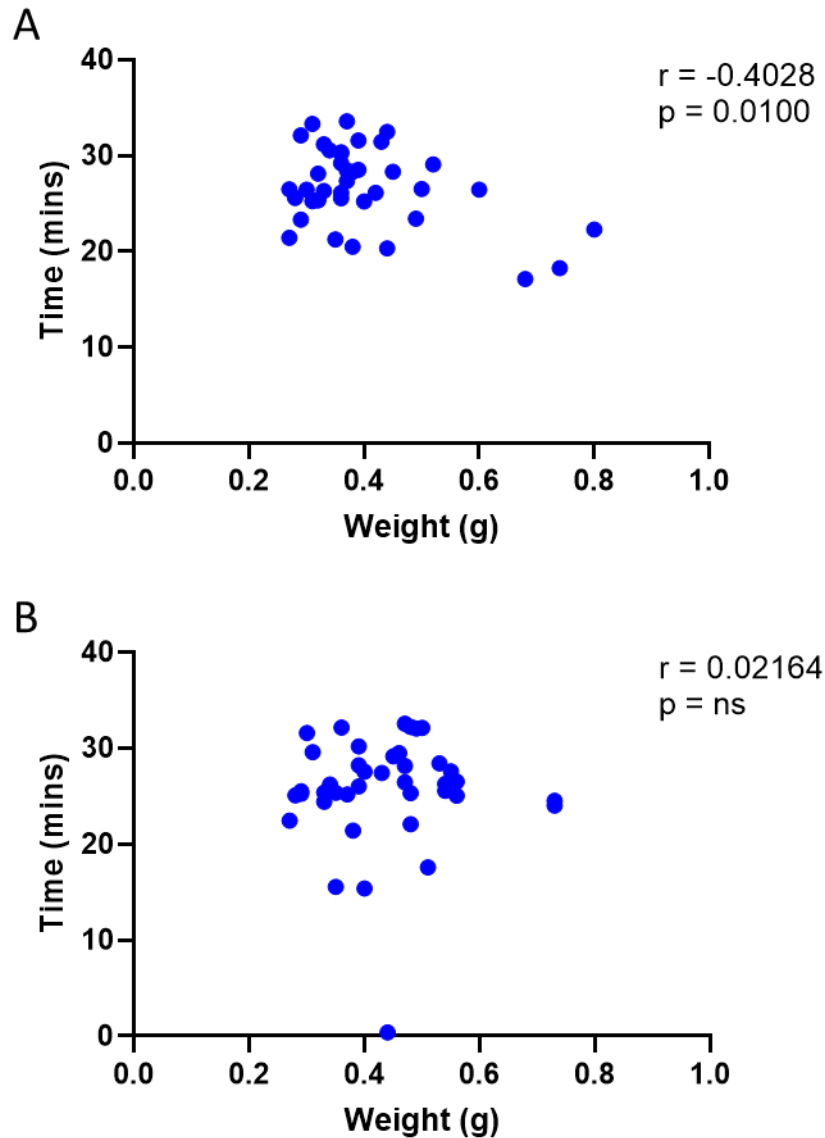
Figure 5.12: Effect of weight on swimming endurance in genotypes obtained from crossing HC-*ARFIP2* line 2 with the sense and antisense *C9ORF72*-ALS zebrafish lines

Correlation between weight (g) and time (mins) to assess effect of weight on swimming endurance in zebrafish. **(A)** Scatter diagram shows the relationship between weight and time swum for adults from the *HC-ARFIP2* line 2 crossed with the sense (G4C2) *C9ORF72*-ALS line. It shows a weak but significant negative correlation (r=-0.4028, p=0.0100) between weight and time swum, i.e. swimming endurance. N=40. **(B)** Scatter diagram shows the relationship between weight and time swum for adults from the *HC-ARFIP2* line 2 crossed with the antisense (C4G2) *C9ORF72*-ALS line. It shows a weak non-significant correlation (r=0.02164, p=ns) between weight and time swum, i.e. swimming endurance. N=40.

5.3.4 Effect of expressing HC-ARFIP2 on poly-GP DPRs in *C9ORF72*-ALS zebrafish

The *C9ORF72*-ALS zebrafish model was used in the project and a characteristic of the *C9ORF72* mutation is the presence of pathological DPRs. To assess the effects of HC-ARFIP2 on DPRs *in vivo*, the HC-*ARFIP2* line 2 transgenic line was crossed with the sense (G4C2) and antisense (C4G2) *C9ORF72*-ALS lines. Zebrafish embryos at 3 dpf were genotyped (2.3.2.5), homogenised at 5 dpf and the whole lysate was used to conduct a poly-GP electrochemiluminescence (MSD) ELISA (**Error! Reference source not found.**) (**Figure 5.13**). The poly-GP DPR was chosen for assessment as this DPR is generated from both the sense and antisense transcripts. Poly-GP DPRs, therefore, serve as a proxy for both the sense and antisense DPRs. Standard curves were generated using known concentrations of standards and their raw response values (**Figure 5.13 A**). The raw response values for the zebrafish embryos in each genotype from the sense and antisense crosses were graphed (**Figure 5.13 B**). The standard curves were used to interpolate the amounts of poly-GP DPRs for each genotype population (**Figure 5.13 C**). Results showed the presence of the poly-GP DPRs at significantly elevated levels in the C9-S genotypes (412.9 ng/ml) and C9-A (8.7 ng/ml) in the sense and antisense backgrounds, respectively. Higher levels of the poly-GP DPR were detected in the sense background (C9-S) compared to the antisense background (C9-A). In the NTG and HC genotypes, no poly-GP DPRs were observed as expected. It was observed that with HC-ARFIP2 expression in the disease background (HC/C9-S and HC/C9-A), there were no significant differences in the levels of poly-GP DPRs compared to C9-S and C9-A genotypes (**Figure 5.13 B and C**). This showed that HC-ARFIP2 expression has no effect on poly-GP DPRs in *C9ORF72*-ALS zebrafish.

The HC-*ARFIP2* line 4 transgenic line was crossed with the sense (G4C2) and antisense (C4G2) *C9ORF72*-ALS lines to assess the effects of HC-ARFIP2 on DPRs in the embryos from each genotype obtained from the cross. The raw response values for the embryos in each genotype from the sense and antisense crosses were graphed (**Figure 5.13 D**). The standard curves (**Figure 5.13 A**) were used to interpolate the amounts of poly-GP DPRs for each genotype from line 4 (**Figure 5.13 E**). As above, significantly elevated levels of poly-GP DPRs were detected in the C9-S (52.3 ng/ml) and the C9-A (24.3 ng/ml) genotypes in the sense and antisense backgrounds,

respectively, and higher levels were again detected in the sense background (C9-S). In the NTG and HC genotypes, no poly-GP DPRs were observed. While it was observed that there was no difference in poly-GP DPR levels between C9-A (24.3 ng/ml) and HC/C9-A (26.7 ng/ml) in the antisense background, it was interesting to see a significant reduction of about 33% in the sense background between C9-S (52.3 ng/ml) and HC/C9-S genotypes (34.9 ng/ml) obtained from the cross from HC-*ARFIP2* line 4.

Overall, the results show that the *C9ORF72*-ALS *in vivo* model recapitulates the disease characteristic of DPRs in both the sense and antisense background. It was observed that the levels of the poly-GP DPRs are higher in the sense background than in the antisense background. The expression of HC-ARFIP2 does not have any effect on the poly-GP DPRs in the cross with HC-*ARFIP2* line 2, however, in the cross with HC-*ARFIP2* line 4, there was a reduction in the amounts of poly-GP in only the sense background. However, further validation of these preliminary results is required.
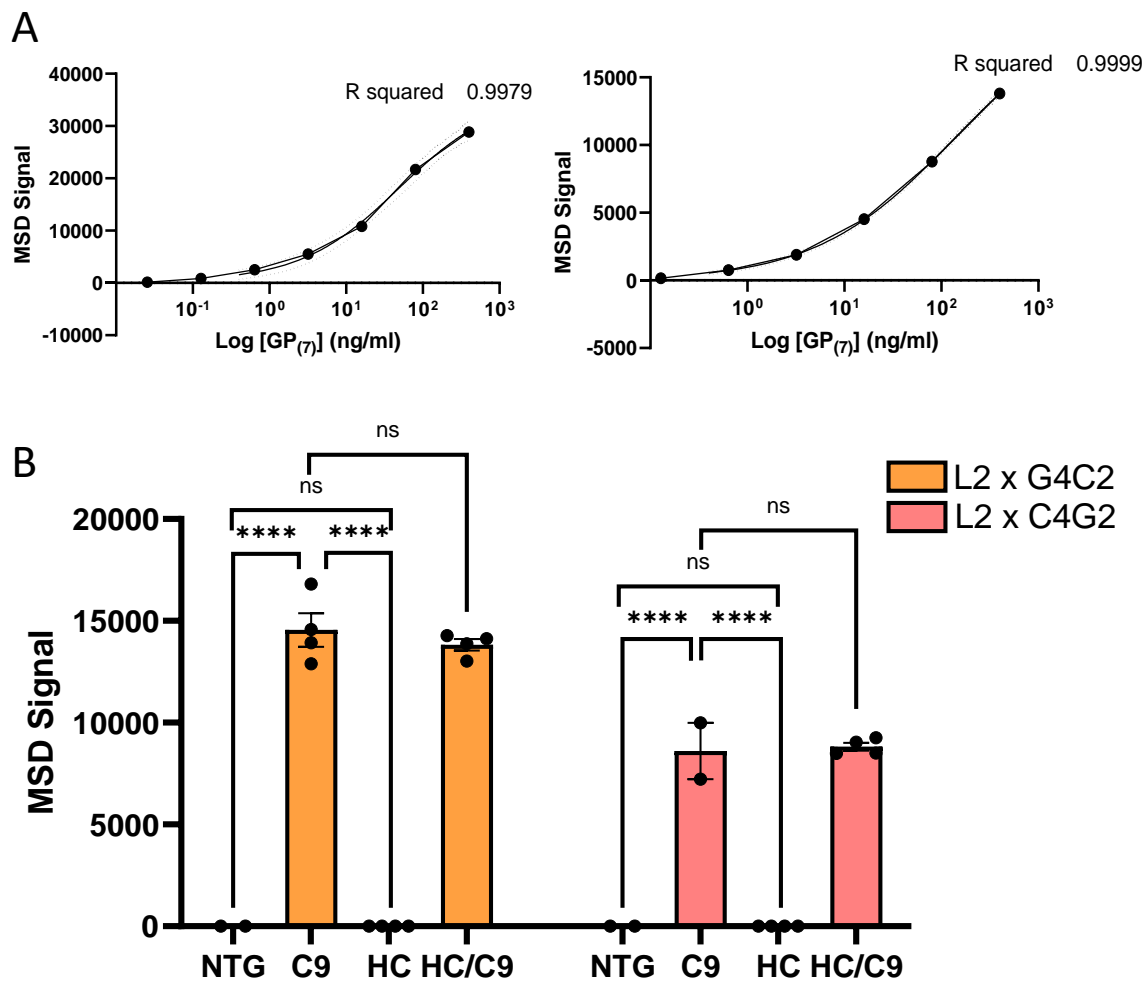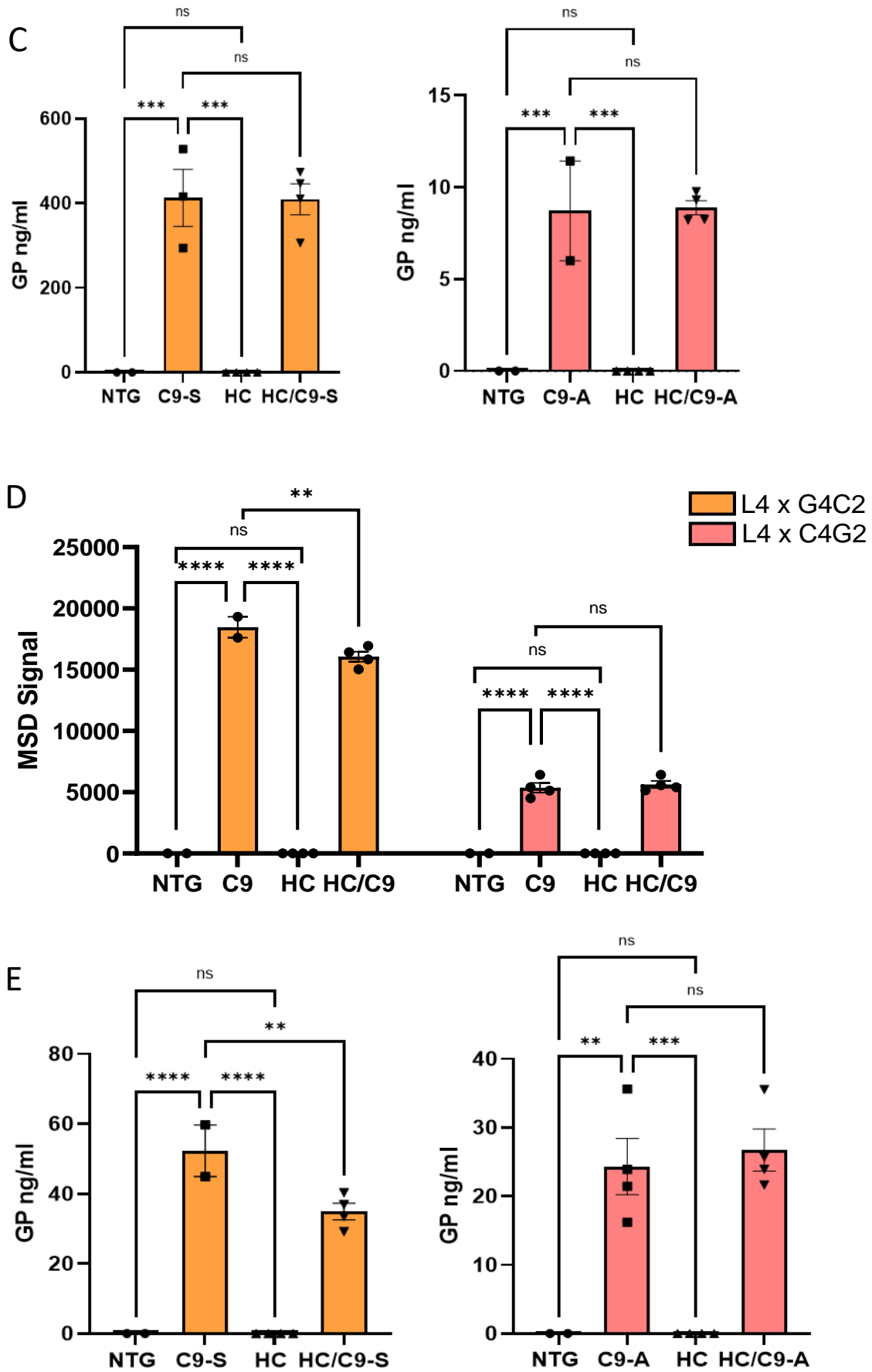
Figure 5.13: Effect of expressing HC-ARFIP2 on poly-GP DPRs in sense and antisense *C9ORF72* zebrafish crossed with HC-*ARFIP2* line 2 and line 4
HC-ARFIP2 line 2 and line 4 were crossed with sense (G4C2) and antisense (C4G2) *C9ORF72*-ALS lines. The zebrafish embryos were genotyped at 3 days post fertilisation (dpf) and 2 embryos per genotype were pooled into one sample to be assayed on 5dpf. **(A)** Standard curves were generated and used to interpolate and determine the amount of GP DPRs in sample. **(B)** Raw response values from the assay were plotted against genotypes of embryos obtained from crossing line 2 with sense (G4C2) and antisense (C4G2) lines. NTG n=2, C9-S n=4, HC n=4 and HC/C9-S n=4 for G4C2 and NTG n=2, C9-A n=2, HC n=4 and HC/C9-A n=4 for C4G2. There was no response for NTG and HC groups, as expected. A significant amount of poly-GP DPRs was detected in the C9-S/A group to support the DPR phenotype in C9ORF72-ALS, as expected. There was no significant difference between the C9-S/A and HC/C9-S/A groups in the sense (mean=14553 and mean=13836.5, respectively) and antisense cohorts (mean=8602 and mean=8822.25, respectively). **(C)** Raw response values were used to determine the amount of poly-GP DPRs using the standard curves. No significant differences in the amount of DPRs were seen between the C9-S/A and HC/C9-S/A groups in the sense (mean=412.9 ng/ml and mean=409.5 ng/ml, respectively) and antisense cohorts (mean=8.7ng/ml and mean=8.9 ng/ml, respectively). Total sample size=8 and actual power= 0.9999656. (**D**) Raw response values from the assay were plotted against genotypes of embryos obtained from crossing line 4 with sense (G4C2) and antisense (C4G2) lines. NTG n=2, C9-S n=2, HC n=4 and HC/C9-S n=4 for G4C2 and NTG n=2, C9-A n=4, HC n=4 and HC/C9-A n=4 for C4G2. There was no response for NTG and HC groups, as expected. A significant amount of poly-GP DPRs was detected in the C9-S/A group to support the DPR phenotype in *C9ORF72*-ALS, as expected. There was a significant difference between the C9-S and HC/C9-S groups in the sense cohort (mean=18467.5 and mean=16063, respectively), however, there was no significant difference in the antisense cohort (mean=5368.75 and mean=5641.25, respectively). (**E**) Raw response values were used to determine the amount of poly-GP DPRs using the standard curves. A significant difference in the amount of DPRs were seen between the C9-S and HC/C9-S groups in the sense cohort (mean=52.3 ng/ml and mean=34.9 ng/ml, respectively), however, there was no significant difference in the antisense cohort (mean=24.3 ng/ml and mean=26.7 ng/ml, respectively). Total sample size=8 and actual power= 1. One experimental repeat. Data were analysed by an ordinary one-way ANOVA with a Šídák's multiple comparisons test. Error bars represent ± SEM. p<0.05. ns = P > 0.05, * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001.

5.3.5 Expression of HC-ARFIP2 has no effect on RNA foci in *C9ORF72*-ALS zebrafish

The C9ORF72-ALS zebrafish model used has been shown to have the presence of nuclear RNA foci in the nuclei of muscle cells (Shaw *et al.*, 2018). To assess the effects of HC-ARFIP2 on RNA foci *in vivo*, in situ hybridisation (2.3.2.9) was used to detect sense and antisense RNA foci in 5 dpf genotyped zebrafish embryos from the HC-*ARFIP2* transgenic line 2 crossed with the *C9ORF72*-ALS sense (G4C2) and antisense (C4G2) lines. Embryos were imaged on the confocal microscope (2.3.2.12.3) and the foci were imaged and counted while blinded to the genotype of the fish. Additional blinded RNA foci quantification was also performed by Dr. Dave Burrows as a comparison for accuracy to show the conservation of trends and to account for any bias (**Figure A12** in **Appendix 3**). The outcome was consistent with the quantification seen in the sense and antisense background (**Figure 5.14**) and similar trends were observed. This validated the observations and the results showing the effects of HC-ARFIP2 on RNA foci in zebrafish. RNA foci were observed in the sense background as indicated by the white arrows in both the C9-S and HC/C9-S genotypes (**Figure 5.14 A**), but none in the HC and NTG genotypes as expected. Images were analysed using ImageJ (2.3.2.12.3) and the average and total sense RNA foci were normalised to the nuclear volume (microns^3) (**Figure 5.14 B**). There was no significant difference in sense RNA foci between C9-S and HC/C9-S populations. In the antisense background (**Figure 5.14 C**), antisense RNA foci were observed, similar to those in the sense background. There was no significant difference in antisense foci between C9-A and HC/C9-A (**Figure 5.14 D**). Interestingly, it was observed that the sense RNA foci were detected at a higher number than the antisense RNA foci in the zebrafish models.

Overall, the results show that the *C9ORF72*-ALS associated RNA foci characteristic was detected in the C9-S/A genotypes. It was observed that the sense RNA foci were detected at higher levels than the antisense RNA foci. Furthermore, HC-ARFIP2 expression does not have an effect on sense and antisense RNA foci in the *C9ORF72*-ALS zebrafish models.

D
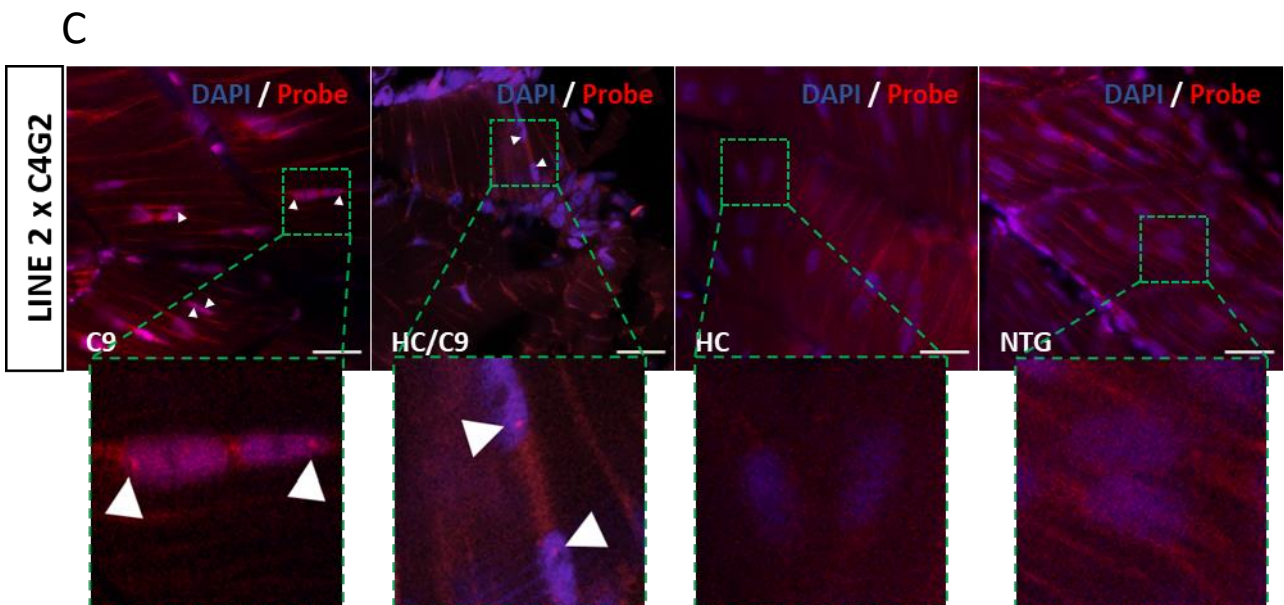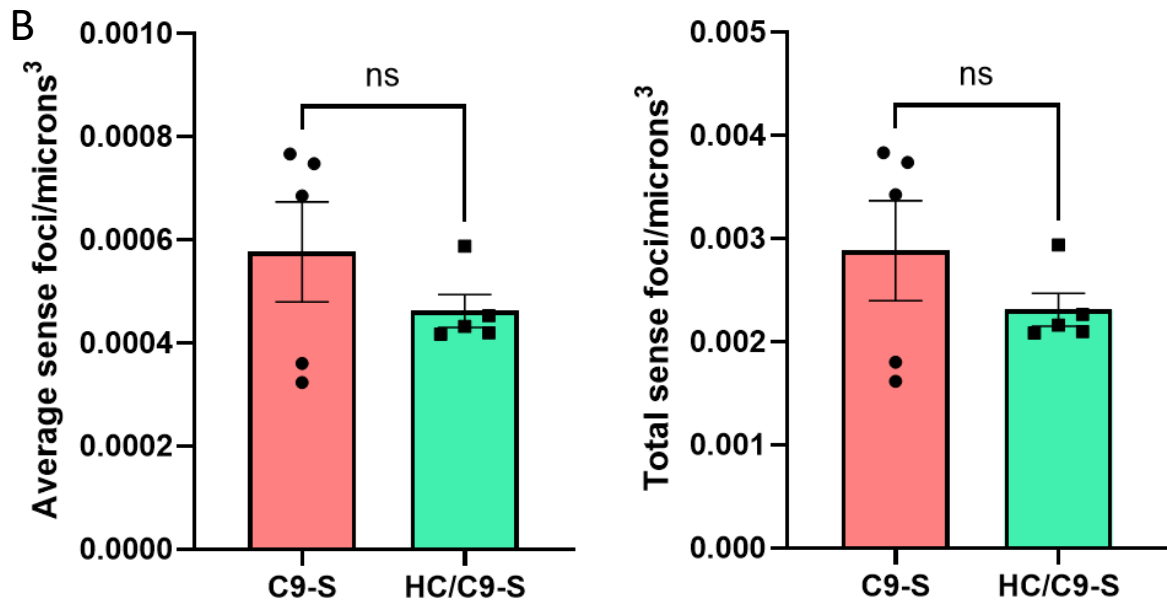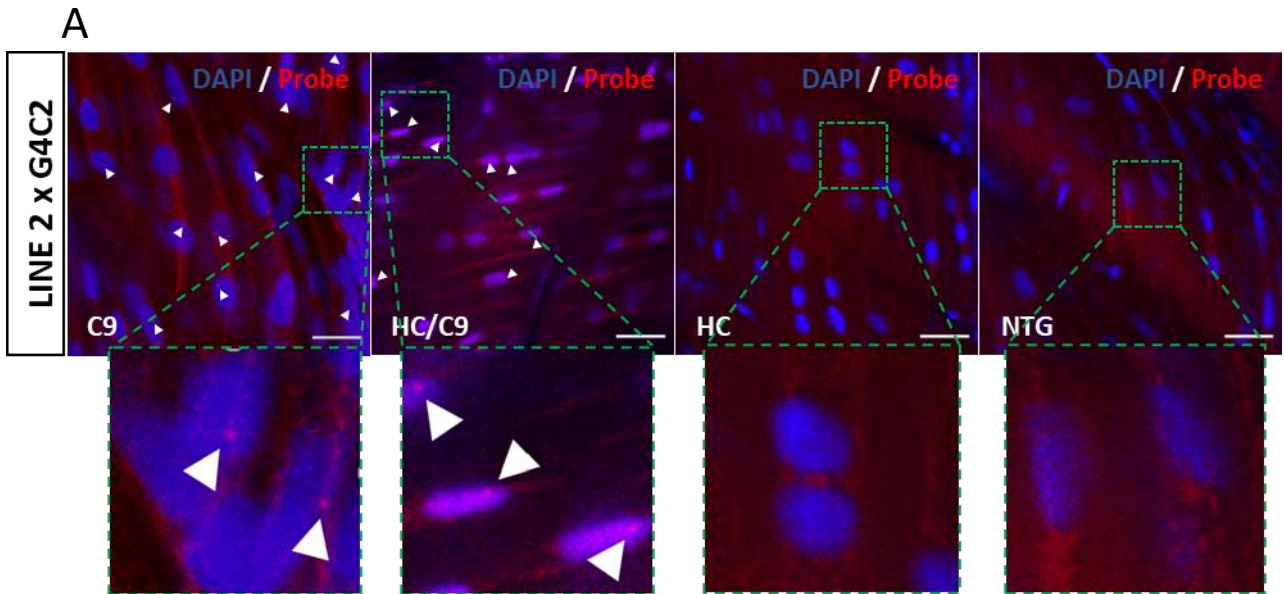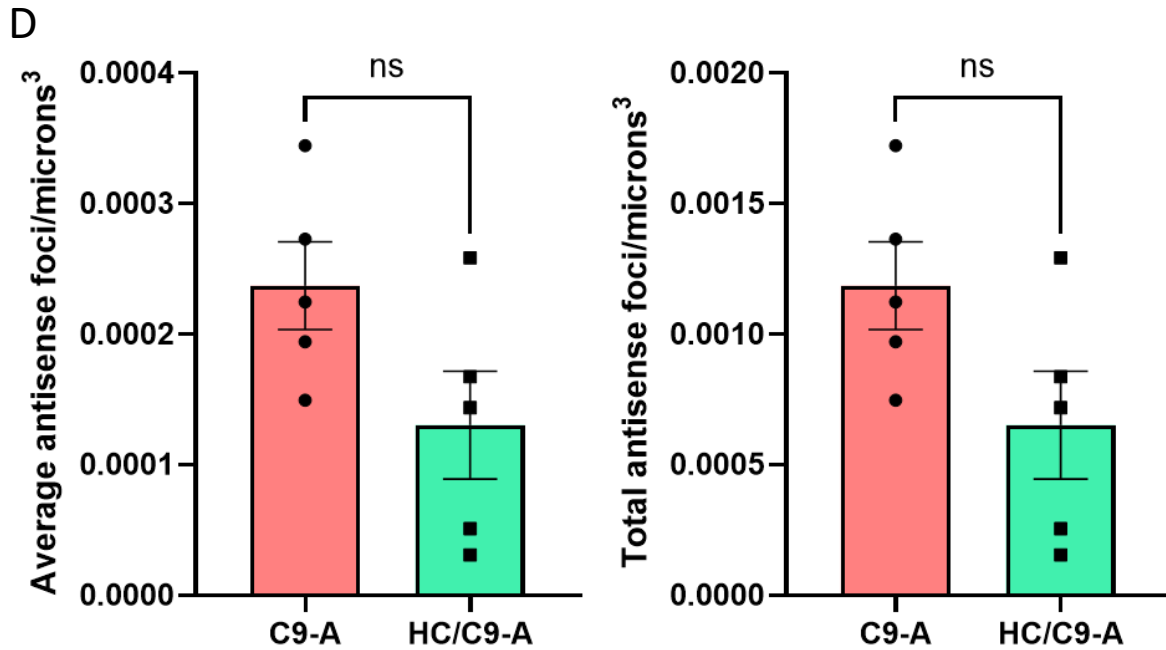
Figure 5.14: Effect of expressing HC-ARFIP2 on sense and antisense RNA foci in zebrafish

Zebrafish from crossing *HC-ARFIP2* Line 2 with sense and antisense *C9ORF72* lines were stained with locked nucleic acid (LNA) probes specific for sense and antisense RNA foci at 5 days post fertilisation (dpf). White arrows and insets show RNA foci. **(A)** *HC-ARFIP2* Line 2 was crossed with the sense *C9ORF72* line. Sense RNA foci were more abundant in the C9-S population than in the HC/C9-S double mutant population. No sense RNA foci were observed in the HC and NTG populations. **(B)** Average number of sense RNA foci observed in the C9-S population was 0.00058 (n=5) and 0.00046 (n=5) in the HC/C9-S double mutant population. Total number of sense RNA foci observed in the C9-S population was 0.0029 (n=5) and 0.0023 (n=5) in the HC/C9-S double mutant population. There was no significant difference in sense RNA foci between C9-S and HC/C9-S populations. **(C)** *HC-ARFIP2* Line 2 was crossed with the antisense *C9ORF72* line. Antisense RNA foci were more abundant in the C9-A population than in the HC/C9-A double mutant population. No antisense RNA foci were observed in the HC and NTG populations. **(D)** Average number of antisense RNA foci observed in the C9-A population was 0.00024 (n=5) and 0.00013 (n=5) in the HC/C9-A double mutant population. Total number of antisense RNA foci observed in the C9-A population was 0.0012 (n=5) and 0.00065 (n=5) in the HC/C9-A double mutant population. There was no significant difference in antisense RNA foci between C9-A and HC/C9-A populations. Data were analysed using an unpaired t-test. Error bars represent ± SEM; $p<0.05$. ns = $p > 0.05$, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

5.4   Results Summary

This chapter shows the effects of expressing HC-ARFIP2 on the disease characteristics of *C9ORF72*-ALS in zebrafish. It was shown that:

- The zebrafish embryos were successfully injected at the one-cell stage with the linearised DNA containing the HC-*ARFIP2* transgene and the injected embryos were able to stably express the transgene at 5 dpf.

- Four positive founders were identified and out of those identified, F0 line 2 and line 4 were able to produce viable embryos at high percentages and were chosen to generate and establish transgenic lines expressing HC-ARFIP2.

- HC-ARFIP2 expressing transgenic lines were established and the transgene expression was stable and ubiquitous, as seen in embryos 5 dpf.

- When the HC-*ARFIP2* transgenic lines were crossed with the *C9ORF72*-ALS lines, three distinct populations based on the DsRed phenotype were produced. The populations were genotyped and resulted in four genetically distinct experimental populations.

- All the primers designed and validated were suitable for genotyping zebrafish.

- When there were significant differences between observed and expected phenotypes and genotypes, the no DsRed expressing or NTG population was higher than expected which shows that the NTG genotype was more favourable than the genetically modified genotypes (transgenic or disease mutation).

- 6-month-old adult zebrafish with the *C9ORF72*-ALS mutation, both in the sense and antisense background, show a strong phenotype of swimming endurance deficits, and the phenotype was more severe in the antisense background.

- Preliminary data showed that expressing HC-ARFIP2 in the disease background (particularly in the antisense background) significantly rescues the swimming endurance deficit phenotype and therefore highlights the therapeutic potential of HC-ARFIP2 in *C9ORF72*-ALS zebrafish. Though encouraging, this finding needs further validation.

- HC-ARFIP2 expression does not show any adverse effects on motor performance.

- The females have a lower swimming endurance than their male clutch mates in all genotypes.

- In the sense background, the preliminary results suggested that the effect of HC-ARFIP2 expression on swimming endurance was observed to be female-specific whereas there was no gender-specific effect in the antisense background.

- There was no effect of weight on swimming endurance in the antisense background and a very weak effect in the sense background.

- The *C9ORF72*-ALS *in vivo* model recapitulates the disease characteristic of DPRs in ALS, and the poly-GP levels are higher in the sense background than in the antisense background.

- Expression of HC-ARFIP2 does not have any effect on the poly-GP DPRs in the cross with HC-*ARFIP2* line 2, however, in the cross with HC-*ARFIP2* line 4, there was a reduction in the amounts of poly-GP in only the sense background.

- The *C9ORF72*-ALS associated RNA foci characteristic was detected in the *in vivo* model and higher levels of RNA foci were detected in the sense background.

- HC-ARFIP2 expression does not have an effect on sense and antisense RNA foci in the *C9ORF72*-ALS zebrafish model.

5.5 Discussion

The pCDNA3.1 mammalian expression plasmid allows for a high level of stable expression which was seen in the transgenic embryos and adult zebrafish described in this chapter. Into the plasmid construct, I-SceI meganuclease sites aided to linearise the construct and increased the rate of transgenesis (Grabher, Joly and Wittbrodt, 2004; Liu, Jenkins and Copeland, 2003). In any model organism, a ubiquitous promoter to drive transgene expression is an essential component of a comprehensive transgenesis toolbox. One study identified and showed that the Ubi promoter in zebrafish is an endogenous ubiquitous transgene driver (Mosimann *et al.*, 2011). Similar to previous observations in the study, it was observed that the Ubi promoter was able to drive visibly strong and ubiquitous expression from within 72 hours post-injection to adulthood, as well as a stable transgene expression through multiple generations which suggested that transgene silencing is unlikely. The zebrafish heat shock protein 70 (hsp70) promoter driving the expression of the DsRed fluorescent protein in the construct has been previously used to track the transgene expression and acts as a drug screening system previously utilised in the zebrafish models of ALS (Shaw *et al.*, 2018; Ramesh *et al.*, 2010b). The heat-shock-reporter gene system allowed for the differentiation of transgenic from non-transgenic fish.

To establish HC-*ARFIP2* transgenic lines, a method established by two independent groups was followed (Higashijima *et al.*, 1997; Long *et al.*, 1997). Post microinjection, the DNA-injected embryos were analysed using the transient expression assay. It was observed that the transgene expression occurs in a mosaic manner even though the transgene was downstream of the Ubi promoter. This is because only a subset of the cells would have the foreign DNA integrated into the genome in the injected embryos. The DNA-injected embryos raised to adulthood were analysed for positive founders. It has been reported that the frequency of obtaining positive founders with the method used for generating transgenic zebrafish has been 5-20% (Higashijima, 2008). Our results showed a frequency of 13% (4 positive founders out of 31 DNA-injected fish) which was consistent with the previously reported frequency and supported the DNA injection method used.

Although four positive founders were identified, only two (line 2 and line 4) were able to establish transgenic lines. Crossing line 7 and line 24 produced very low

percentages of transgenics and the transgenic embryos were not able to survive. This may be because either the transgene might have integrated and disrupted a part of the genome essential in the development and viability of the embryos or the HC-*ARFIP2* transgene expression was toxic for development. However, the latter was unlikely because the crosses from line 2 and line 4 produced transgenic embryos at high percentages which were able to be raised to adulthood. The fish did not show any signs of disrupted growth and development which showed that the HC-*ARFIP2* transgene expression was safe and not toxic for development. This suggests that the non-viability of the embryos from lines 7 and 24 was because of the site of insertion of the transgene causing the lethal effect, thereby either disrupting an essential gene for development or overexpressing a detrimental gene due to proximity to the Ubi promoter. Furthermore, the one transgenic embryo from line 7 showed an exceptionally high DsRed fluorescence. It has been shown that in response to cellular stress, the Hsp70 promoter is enhanced to drive the expression of the downstream protein (Mayer and Bukau, 2005). Since DsRed is downstream of the Hsp70 promoter in the transgenic lines, the extremely bright expression might be indicative of cellular stress and could explain why the embryo was not able to survive. However, the high DsRed expression was not seen in the embryos from line 24.

Embryos at 5 dpf of the two established HC-*ARFIP2* transgenic lines showed a stable and ubiquitous expression. For line 2 F1 embryos, the transgene expression was detectable on a western blot as well as observed using immunostaining. However, for line 4 F1 embryos, transgene expression was only observed using immunostaining. Several western blots were attempted to detect the presence of transgene expression in line 4 embryos, but it was difficult to do so. Repeated western blots for line 2 embryos also showed difficulty in detecting transgene expression. This can be attributed to the high proportion of yolk proteins in the early embryo which can interfere with the protein detection of the target protein from the cells of the embryo proper. It has been shown that without deyolking of the embryos, less abundant proteins fall below the limit of detection, whereas after deyolking there were strong specific signals for the less abundant proteins (Link, Shevchenko and Heisenberg, 2006). Though this was reported, the HC-*ARFIP2* embryos were not deyolked as the yolk-to-cell mass ratio is already decreased at 5 dpf. Since embryos from both lines showed transgene expression using immunostaining and validated the generation of transgenesis,

deyolking embryos before western blotting was not attempted. In the future, deyolking methods could be applied for proteomic analysis of early embryos.

Both transgenic lines were crossed with the sense and antisense *C9ORF72*-ALS zebrafish lines and the zebrafish fry were sorted according to phenotypes and genotypes. Though genotypes from both transgenic lines were validated, the genotypes from line 2 (in the sense and antisense background) were chosen for further experiments to assess the effects of HC-ARFIP2 in ALS. This was due to three reasons. Firstly, there were no significant differences between the observed and expected phenotypes from the sense and antisense lines as well as in the genotypes from the sense line. This showed that no particular phenotypes or genotypes were favoured, and where significant differences were seen (in the genotypes from the antisense lines), the NTG population was favoured. Whereas with line 4 crosses, there were significant differences in all the observed and expected phenotypes and genotypes from sense and antisense lines. Secondly, experiments with all genotypes from the line 4 crosses would not be possible. The tank housing the faint DsRed phenotype embryos (HC genotype) from the antisense line was lost when the fish were one month old due to contamination of water quality. At this stage, the fish are quite fragile and susceptible to changes in water quality. Finally, though the line 4 F1 embryos showed a more intense transgene expression in the immunostaining analysis, transgene expression from line 2 F1 embryos was detected through both western blotting and immunostaining.

Genotyping zebrafish was successful with the designed primers and enabled further sorting of the zebrafish clutch from the crosses into experimental populations. While the results showed that all the primers were suitable for detecting target genes, the primers for the *C9ORF72* gene showed a band at an incorrect height in the absence of the *C9ORF72* gene. This can be attributed to the annealing of the primers and amplifying of a region of the zebrafish genome with possible complementarity to the primer sequence. To overcome this, touchdown PCR was used for genotyping. A touchdown PCR increases sensitivity and specificity by decreasing non-specific and off-target bindings of the primer to DNA, reducing the primer-dimer formation, and is also effective in higher GC content DNA templates (Green and Sambrook, 2018). Despite this, bands at an incorrect height were seen for the C9 primers. Since the

band at the incorrect height was only observed in the absence of the *C9ORF72* gene and was not seen in the presence of the *C9ORF72* gene, the C9 primers were used for genotyping. Bands at the correct height corresponding to the *C9ORF72* gene were seen in the presence of the *C9ORF72* gene. Furthermore, the sizes of the bands disparate in size allowed differentiation between the bands. Reference primers were used alongside the primers for HC-*ARFIP2* and *C9ORF72* as positive controls. After validation of the primers, the E4 reference primers were used for genotyping as the expected band was more distinguishable from the band for the *C9ORF72* gene at the correct height.

The HC-*ARFIP2* transgenic lines and the *C9ORF72*-ALS model, both carry the same DsRed reported genes to track transgene expression. While the DsRed expression allowed for transgene expression tracking in independent lines, when crossed together, genotyping of the zebrafish clutch was essential to differentiate between the populations because the same reporter gene was expressed. Though genotyping was successful, the process was needed for all adults and embryos before any experiments. This was time-consuming, laborious, and involved various techniques and reagents. The use of a different reporter gene like *GFP* for generating HC-*ARFIP2* transgenic lines would help overcome this issue. However, it has been suggested that the cytotoxicity and immunogenicity of GFP can confound the results of *in vivo* experiments (Ansari *et al.*, 2016). Since all fluorescent proteins potentially carry some associated toxicity, support for using DsRed to track transgene expression in HC-*ARFIP2* transgenic lines lies in the fact that DsRed has been used in established zebrafish ALS models and has not shown any related toxicity (Shaw *et al.*, 2018; Ramesh *et al.*, 2010b). It allows for consistency and a direct comparison of any toxicity observed between the HC-*ARFIP2* transgenic model and the *C9ORF72*-ALS model used in the project. Furthermore, all characteristics of DsRed have been identified and addressed in the established ALS models.

The *C9ORF72*-ALS model and HC-*ARFIP2* transgenic lines both express DsRed downstream of the hsp70 promoter. It has been reported that in the presence of abnormal cellular proteins, heat shock proteins are upregulated (Bukau, Weissman and Horwich, 2006; Dedmon *et al.*, 2005). The embryos from the *C9ORF72*-ALS lines were observed to have a brighter DsRed expression phenotype. This is attributed to

the previous reports which showed that in the *C9ORF72*-ALS zebrafish model, the *C9ORF72* repeat expansion triggers the heat shock response (HSR) (Shaw *et al.*, 2018). It was shown that the DsRed expression positively correlated with disease severity, progression, and the presence of aberrant proteins like DPRs resulting in the zebrafish displaying a high intensity of DsRed. This phenotype of brighter DsRed expression in *C9ORF72* zebrafish aided in initially sorting the zebrafish clutch obtained from crossing HC-*ARFIP2* transgenics with the *C9ORF72*-ALS model into populations with distinct phenotypes and predicted genotypes. Though genotyping was time-consuming, brighter DsRed expression helped to save time during the genotyping process. An increase in the drive of the Hsp70 promoter is induced and upregulated in response to cellular stress from various insults (Mayer and Bukau, 2005) and the upregulation of Hsp70 has previously been reported in neurodegenerative disorders (Perez *et al.*, 1991; Mansilla *et al.*, 2014), including ALS (Shaw *et al.*, 2018). Since DsRed expression is downstream of the Hsp70 promoter, the bright DsRed expression can be used as a readout for cellular stress. If time permitted, it would have been interesting to assess whether expressing HC-ARFIP2 affects DsRed expression levels and therefore cellular stress in ALS.

The *C9ORF72*-ALS zebrafish models used in the study have been previously reported to show clear adult-onset ALS-like motor performance deficits (Shaw *et al.*, 2018). Our results were consistent with this. It was shown that at 6 months old, the C9 zebrafish in the sense and antisense background (C9-S and C9-A) show a significant phenotype of swimming endurance deficit. A significant rescue of this phenotype with HC-ARFIP2 expression was seen which highlights its therapeutic potential in ALS, however, this was only seen in the antisense background. This may suggest that the effect of HC-ARFIP2 on swimming endurance might be due to specific effects on the antisense pathology, however, this was not the case. Comparing the swimming endurance for all genotypes in the sense and antisense background, it was observed that the swimming endurance for HC/C9-S and HC/C9-A were similar and in line with the swimming endurances of their NTG clutch mates. Therefore it can be inferred that any effect observed of HC-ARFIP2 in the antisense background can be explained by the more extreme phenotype seen in the antisense background (C9-A) compared to the sense background (C9-S). A good representation of this severe phenotype can be observed in one of the female fish in the antisense (C9-A) background that could swim for only

0.39 seconds at the lowest flow rate of water in the swim tunnel. Prior to the experiment, the fish was healthy and swam around normally and also showed healthy swimming movements in the swim tunnel during the acclimatisation period of 5 mins before starting the water flow. Furthermore, the motor performance analysis described in this chapter was conducted at 6 months of age, and it has been reported that the swimming ability of zebrafish at maximum speed continued to decrease as the fish aged to 12 months of age (Shaw *et al.*, 2018). With the development of a possible stronger phenotype in the C9 genotype as the fish age, if time permitted, a significant rescue of the swimming endurance deficit could be achieved in the sense (C9-S) background using the same swim-tunnelled cohorts of zebrafish at 12 months of age. In the addition to highlighting the therapeutic effect of HC-ARFIP2, results showed that there were no effects or changes in the swimming endurance of the zebrafish in the HC-*ARFIP2* transgenic genotypes (HC-S and HC-A) which were comparable to the swimming endurance of their NTG clutch mates. This showed that transgene expression was safe and does not have any adverse effects on swimming endurance.

Though the results showing the effects of HC-ARFIP2 on motor performance in zebrafish are promising, further validation of the experiment is required. The swim tunnel experiment could not be conducted while being blinded to the genotype. This could have increased the risk of bias in the results. Therefore further validation by repeating the swim tunnel experiment in a blinded manner is necessary to increase the credibility and reliability of the effect of HC-ARFIP2 on motor performance. Moreover, conducting more repeats using zebrafish clutches from further repeated crosses would highlight the reproducibility as well as the validity of the results. More experimental repeats would also confirm whether the effects of the transgene are due to the expression of HC-ARFIP2 or any other unmeasurable events like homologous recombination.

The effects of gender and weight on motor performance were assessed to confirm that the rescue of the motor performance deficit phenotype in the C9 genotypes was due to the expression of HC-ARFIP2 and not influenced by other external factors like gender and weight. Results showed that in both the sense and antisense background, the females had lower swimming endurance than their male clutch mates. Therefore the samples for each genotype included equal numbers for each gender to ensure that

the results were not biased, except for in HC/C9-A where this was not possible due to only a few numbers of fish available for the genotype and the sample population for only HC/C9-A included more males than females. It can be argued that the therapeutic effect of HC-ARFIP2 seen in the antisense background (HC/C9-A) might be affected by this as the males generally showed higher swimming endurances. However, the gender-specific analysis offered support for the therapeutic potential of HC-ARFIP2. It was observed that in the antisense background, the effect of HC-ARFIP2 was not gender-specific and a significant rescue of the motor performance phenotype was seen in both males and females in the HC/C9-A genotype cohort. In the sense background, the effects of HC-ARFIP2 were seen to be female-specific at 6 months of age. Since it was observed that the C9-S population did not have a strong motor performance phenotype at 6 months, analysis at 12 months of age might show significant differences in motor performance in both males and females and the effect of HC-ARFIP2 may no longer be female-specific in the sense background. Furthermore, the sample population only consisted of 5 fish per gender which is a small cohort. Increasing the cohort size as well as conducting more experimental repeats might show that there are no gender-specific effects in the sense background as well.

It was reported that there was a significant decrease in the total weight of *C9ORF72*-ALS zebrafish compared to NTGs and the *C9ORF72*-ALS fish appeared to be smaller during the early stages of development (Shaw *et al.*, 2018). With this knowledge, the zebrafish in each sample population were size-matched to reduce the effects of external factors. Before the swim tunnel experiment, the zebrafish were weighed to assess the effects of weight on motor performance. Results showed that there was no effect of weight on swimming endurance in the antisense background and a very weak effect in the sense background. This highlights that the rescue of the swimming endurance phenotype in *C9ORF72*-ALS fish was solely due to the expression of HC-ARFIP2 and not influenced by the external factors measured. This offers support to the therapeutic potential of HC-ARFIP2 in ALS.

In *C9ORF72*-ALS zebrafish models, the mutation-associated characteristic of DPRs has been previously reported (Ohki *et al.*, 2017; Shaw *et al.*, 2018). Our results showed that the *C9ORF72*-ALS model used showed significant levels of the poly-GP DPR in

both the sense and antisense background. It was reported that in the *C9ORF72*-ALS patient samples where a considerable DPR load was observed, the hsp70 protein levels were increased, showing that DPRs can cause heat shock response activation (Shaw *et al.*, 2018; Mordes *et al.*, 2018). The significant levels of the poly-GP DPR can explain why the C9 genotype embryos had a stronger DsRed expression. Both the established *HC-ARFIP2* transgenic lines were crossed with the sense and antisense *C9ORF72*-ALS lines. Though a strong effect of HC-ARFIP2 expression on poly-GP DPRs was seen *in vitro* as described in chapter 4 (4.3.3), in the embryos of the line 2 genotypes, there was no effect of HC-ARFIP2 on poly-GP DPRs. The failure to show any efficacy can be attributed to the inherent complexity of organ systems *in vivo* and the various processes occurring. Therefore, it is unlikely that *in vitro* results will translate in *in vivo*. On the other hand, in the line 4 genotypes, an effect of HC-ARFIP2 expression on DPRs was observed, however, only in the sense background. This may offer some support to the therapeutic potential of HC-ARFIP2. It was observed that the staining intensity of HC-ARFIP2 expression in line 4 F1 transgenic embryos was higher than in line 2 F1 transgenic embryos. Therefore the effect of HC-ARFIP2 on poly-GP DPRs in line 4 may be due to a higher expression of HC-ARFIP2 in the transgenic lines. If time permitted, it would be essential to assess the expression level of HC-ARFIP2 in the transgenic lines. The assessment of the effect of HC-ARFIP2 using line 4 was only conducted for DPRs. Since the transgene expression was observed to be more intense in line 4, if time permitted, analysis of the effects of HC-ARFIP2 on all other read-outs using line 4 would be interesting and provide further insights into the therapeutic potential of HC-ARFIP2 as well as the level of expression needed to see neuroprotective effects *in vivo*.

Though the preliminary results showed a significant decrease in poly-GP DPR levels due to HC-ARFIP2 expression, further validation is needed. Further repeats are required to determine the reliability and validity of the results. More experimental repeats for the experiment could not be conducted due to the time constraints of the Ph.D. following COVID-19 and the time-consuming process of genotyping embryos needed before the experiment. Furthermore, while genotyping randomly selected embryos from a particular phenotype population, it was not always possible to get enough embryos for a particular genotype. If time permitted, in addition to conducting more repeats of the experiment, genotyping could be done for more embryos to

increase the sample size in each genotype population and therefore increase the validity and power of the experiment. ALS is an adult-onset disorder where disease symptoms develop in individuals at 40-60 years of age (Masrori and Van Damme, 2020). The data generated for DPRs in zebrafish used embryos at 5 dpf. While the embryos were able to recapitulate the DPR disease characteristic, this was not a good representation of the adult-onset disorder. A therapeutic benefit of HC-ARFIP2 on the motor performance phenotype was seen at 6 months of age. Therefore, assessing the effect of HC-ARFIP2 on DPRs in various regions of the brain and spinal cord at 6 months or using zebrafish at an older age might provide further insight into the therapeutic potential of HC-ARFIP2 expression on DPRs. Moreover, the preliminary results showing the decrease in poly-GP levels due to HC-ARFIP2 expression warrant further investigation on the effects of HC-ARFIP2 on other DPRs expressed in *C9ORF72*-ALS, especially the more toxic arginine-rich DPRs.

Several different zebrafish *C9ORF72*-ALS models have displayed the RNA foci phenotype (Lee *et al.*, 2013; Ohki *et al.*, 2017; Swinnen *et al.*, 2018; Shaw *et al.*, 2018). Consistent with this, our results showed that both sense and antisense RNA foci were observed in the C9-S and C9-A genotypes, respectively. Furthermore, HC-ARFIP2 did not affect RNA foci levels in zebrafish. However further validation is required to draw definite conclusions because the data shown was generated from one experimental repeat using embryos at 5 dpf. Though embryos at 5 dpf displayed the RNA foci phenotype, assessing the effect of HC-ARFIP2 on RNA foci in adult zebrafish would be a better representative for the adult-onset disorder, as well as comparing the foci load in various regions of the brain and spinal cord in adult zebrafish. Furthermore, the gender of the embryos analysed for RNA foci (as well as for DPRs) could not be identified at the embryo stage. The gender identification was only possible after 3 months of age. With differences in swimming endurance seen between males and females, using adult zebrafish would enable us to see gender-specific effects on RNA foci and DPRs as well. Additionally, the sample population for each genotype was low. Only five embryos per genotype were analysed due to the time constraints of the Ph.D. Analysing more embryos per genotype would increase the power of the experiment, however, genotyping and imaging each embryo on the confocal microscope was time-consuming. Moreover, though the conservation of trends across different individual quantifiers was observed and validated the results, the quantification of RNA foci was

done manually because the size of RNA foci was small which made it difficult to filter out the background and identify RNA foci in an automated manner. Manual quantification required human interpretation of what foci are, causing high variability, inaccuracy and high chances of error between quantifiers. In the future, different methods of automated quantification could help overcome these issues. Automated quantification would reduce the source of human error, allow for reproducible and repeatable results, and would be a faster method.

Results from the analysis of DPRs and RNA foci in zebrafish showed that the DPRs and RNA foci were more abundant in the sense background compared to the antisense background. This observation in our zebrafish model is consistent with previously reported findings that showed that sense RNA foci are more abundant in patient cells than antisense RNA foci and RAN translation is more effective from the sense transcripts than from the antisense transcripts which suggested that most of the DPR expression relates to sense transcripts (Mackenzie *et al.*, 2015; Mann *et al.*, 2013; Mizielinska *et al.*, 2013; Zu *et al.*, 2013). This consistency offers support to the results as well as to the *C9ORF72*-ALS zebrafish model used.

Though the DPRs and RNA foci were increased in the sense background, it was interesting to see that the C9 genotype in the antisense background (C9-A) displayed a stronger motor performance deficit phenotype. The poor swim tunnel performance of C9-A can be attributed to muscle atrophy. It was reported that swim tunnel-tested zebrafish which showed significantly poorer swim tunnel performance than their NTG clutch mates had clear signs of severe muscular atrophy (Shaw *et al.*, 2018). In the future, it would be interesting to assess muscular atrophy in both the sense and antisense backgrounds to see whether it is more severe in the antisense background compared to the sense background. Furthermore, assessing the effects of HC-ARFIP2 on muscular atrophy in zebrafish will allow us to further evaluate the therapeutic potential of HC-ARFIP2 in ALS.

6. GENERAL DISCUSSION

Amyotrophic Lateral Sclerosis (ALS) is a rare progressive adult-onset neurodegenerative disorder characterised by motor neuron degeneration and loss in the brain, brainstem, and spinal cord. It is marked by fasciculations, muscle stiffness and progressive weakness and leads to challenges in speaking, swallowing and breathing (Feldman *et al.*, 2022). With the heterogeneity of the disease at the clinical, neuropathological and genetic levels, ALS is now regarded as a multisystem disorder (van Es *et al.*, 2017; Masrori and Van Damme, 2020). The pathogenesis of motor neuron degeneration in ALS is highly varied and multifactorial. Pathological processes arise from genetic mutations in around 40 known ALS genes due to toxic gain of function or loss of function mechanisms (Nguyen, Van Broeckhoven and van der Zee, 2018). Protein aggregation, a pathological hallmark of most neurodegenerative disorders including ALS, also confers toxicity. Pathophysiological processes can be classed into five main groups: impaired RNA metabolism, altered proteostasis, mitochondrial dysfunction, oxidative stress and cytoskeletal or trafficking defects (Feldman *et al.*, 2022). Despite extensive research, the complete molecular pathways leading to ALS are yet to be fully elucidated.

Currently, there is no known cure for ALS, however, various disease-modifying interventions have been approved. Until recently, only two pharmacological interventions, Riluzole and Edaravone, were approved and available in some countries for prolonging survival and slowing down the disease progression, respectively (Andrews *et al.*, 2020; Witzel *et al.*, 2022). The anti-glutamate agent, Riluzole, prolongs patient survival only by a few months, however, it is debated whether this prolongation happens at all stages of ALS or only at advanced disease stages. Recently, a retrospective analysis of data from a dose-ranging study investigated the effects of Riluzole across different stages of the disease (Fang *et al.*, 2018). Analysis of the data from the original dose-ranging trial (Lacomblez *et al.*, 1996) showed that Riluzole prolonged survival in the last clinical stage (stage 4) of ALS compared with the placebo, whereas, there was no difference in progression from stages 2 or 3 to later stages or death between Riluzole treated groups and placebo. This suggests that the disease-modifying effects of Riluzole occur at advanced disease stages. However, the earliest stage (stage 1) of ALS was not analysed in the study and therefore findings from this study need further validation. Furthermore, several open-label non-randomised trials have indicated that Riluzole is most effective at earlier stages of ALS

and suggested early disease modulation (Traynor *et al.*, 2003; Zoing *et al.*, 2006; Riviere *et al.*, 1998; Geevasinga *et al.*, 2016). Keeping these findings in mind and given the lower likelihood that any treatment could confer considerable therapeutic effect within a depleted population of motor neurons in the advanced stage of ALS, earlier efficacy seems more probable than later effects. Moreover, the mechanism of action of Riluzole has remained quite elusive. Initially, Riluzole was recognised as a glutamatergic antagonist and works through its anti-glutamatergic activity, even though its direct interaction with glutamate receptors has been debated (Doble, 1996). Various pathways have since been suggested, in addition to the anti-glutamatergic regulation of excitotoxic pathways, including effects on mitochondrial function, alteration to fat metabolism, effects on persistent sodium channel function and potentiation of calcium-dependent potassium currents (Cheah *et al.*, 2010; Jaiswal, 2016; Thakor *et al.*, 2021; Chowdhury *et al.*, 2008; Thompson *et al.*, 2018; Deflorio *et al.*, 2014). Previously, it has been shown that protein kinase CK1 isoform δ (CK1δ) plays a key role in the phosphorylation of TDP43, promoting mislocalisation and accumulation of TDP43 (Nonaka *et al.*, 2016). It has also been shown that Riluzole has inhibitory effects on the catalytic activity of CK1δ, which could prevent TDP43 pathology (Bissaro *et al.*, 2018). This further adds to the mechanism of action of Riluzole. The finding that Riluzole may benefit at later stages of ALS (Fang *et al.*, 2018) also highlights that Riluzole might affect different pathways dependent on disease stage.

Edaravone, also known as Radicava, is an antioxidant drug marketed to treat ALS. Its therapeutic benefit may be due to its antioxidant properties which can affect oxidative stress, a pathological mechanism associated with ALS (Cruz, 2018; Cunha-Oliveira *et al.*, 2020; Ito *et al.*, 2008). Previously, the drug was available only intravenously and recently the FDA has approved an oral formulation of Edaravone called Radicava ORS (Dorst and Genge, 2022; Aschenbrenner, 2022). In Japan, Edaravone was approved in 2015 and in 2017 in the U.S. This approval was largely based on a single trial conducted in Japan involving early-stage ALS patients (Cruz, 2018; Abe *et al.*, 2014; The Writing Group, 2017). While the trials of the antioxidant Edaravone show efficacy (The Writing Group, 2017; Edaravone (MCI-186) ALS 16 Study Group, 2017), there was a lack of generalisability to the wider population of ALS patients and there are

concerns regarding the associated safety and benefits. Therefore, Edaravone has not yet been approved worldwide and was declined by the European Medicines Agency.

Recently, it was announced that an investigational drug called Tofersen (also known as BIIBO67) is under the peer review period by the FDA. Tofersen is an antisense oligonucleotide (ASO) that targets superoxide dismutase 1 (*SOD1*) ALS. Administered through a lumbar puncture, the ASO can facilitate the degradation of SOD1 messenger RNA and thereby reduces the synthesis of the SOD1 protein (Miller *et al.*, 2022). In a phase 1/2 clinical trial, it was shown that Tofersen reduced SOD1 levels in the cerebrospinal fluid and was generally safe (Miller *et al.*, 2020). This was further validated in a later phase 3 trial (Miller *et al.*, 2022). In addition to Riluzole and Edaravone, the FDA has recently approved a new experimental drug for ALS called Relyvrio (Brown, 2022). It is a co-formulation of sodium phenylbutyrate and taurursodiol and has been found to reduce neuronal death in ALS models and slow down functional decline (Paganoni *et al.*, 2020). Furthermore, a long-term survival analysis of the participants from the trial has shown that Relyvrio has a long-term survival benefit with early initiation of the drug in ALS, highlighting that the drug has functional as well as survival benefits in ALS (Paganoni *et al.*, 2021). Currently, the mechanism of action of Relyvrio is unknown and is speculated from the processes that are affected by sodium phenylbutyrate and taurursodiol. Though pharmacological options are available, they either have modest effects where survival is prolonged by a few months, or further clinical trials are still required post-approval of the drug. Therefore, supportive care provided by multi-disciplinary teams of healthcare professionals remains the foundation of ALS management. Consequently, the identification of targets and the development of novel therapeutic strategies are of great importance for ALS.

Intracellular cytoplasmic protein aggregates are a known pathological hallmark feature of neurodegenerative disorders including ALS and can disrupt normal protein homeostasis and cause cellular stress (Blokhuis *et al.*, 2013). Several molecular constituents of aggregates in ALS have been identified over the years (Blokhuis *et al.*, 2013). The molecular makeup of protein aggregates deepens the understanding of disease pathogenesis involved in ALS and also provides novel targets for therapeutic interventions. Previously, two independent studies have shown that Arfaptin-2

(*ARFIP2*) is involved in the formation of protein aggregates in Huntington's disease (HD) and its C-terminal (HC-*ARFIP2*) can confer neuroprotective effects (Peters *et al.*, 2002; Rangone *et al.*, 2005). Since neurodegenerative disorders share some common features of pathogenesis, therapeutic strategies may be transferrable. Previous preliminary work in our lab implied the involvement of ARFIP2 in ALS (Mohammedeid, 2015). To our knowledge, ARFIP2's involvement in ALS has not been previously investigated. Accordingly, we sought to investigate the therapeutic potential of targeting ARFIP2 in ALS. In particular, with the C-terminal of ARFIP2 shown to confer neuroprotective effects, we explored whether HC-ARFIP2 could ameliorate the disease characteristics of ALS using *in vitro* and *in vivo* models.

6.1 Project Outcomes

The overall goal of the project was to identify whether ARFIP2 might represent a novel therapeutic target for ALS. ARFIP2 has previously been implicated in Huntington's disease (HD) where expression of ARFIP2 was significantly increased in the brain of HD mouse models and patients and has been shown to be involved in the modulation of protein aggregation in HD (Peters *et al.*, 2002; Rangone *et al.*, 2005). Various pathways are common to neurodegenerative disorders and therefore, we looked into ARFIP2 in ALS. In this project, we used *in vitro* and *in vivo* models that carried the *C9ORF72* mutation, the most common ALS-associated mutation. The *C9ORF72* mutation is similar to the disease-causing mutation in HD. HD is inherited in an autosomal dominant manner due to a mutation in the Huntingtin (*HTT*) gene (Ajitkumar and De Jesus, 2022). This mutation is an elongation of CAG repeats and is classed as a microsatellite repeat expansion mutation, similar to the type of mutation in the *C9ORF72* gene in ALS. This categorizes both, the most common genetic cause of ALS (*C9ORF72*-linked ALS) and HD, as repeat expansion diseases (Paulson, 2019). As in *C9ORF72*-linked ALS, RAN translation has been reported in HD (Bañez-Coronel *et al.*, 2015). Similar to RAN proteins translated in *C9ORF72*-ALS, Bañez-Coronel et al. (2015), reported that the HD expansion can be bidirectionally translated into four novel homopolymeric RAN proteins – polyAla, polySer, polyLeu and polyCys. Furthermore, there is evidence that the sense and antisense RAN proteins in HD can also accumulate and correlate with disease pathology (Bañez-Coronel *et al.*, 2015; Banez-Coronel and Ranum, 2019). Therefore, we chose to explore the effects of ARFIP2 in the *C9ORF72*-ALS background. Furthermore, with the *C9ORF72* mutation

being the most common in ALS, accounting for ~40% of the cases, any therapeutic benefit would apply to a larger proportion of ALS patients. Additionally, aspects of the pathophysiology of *C9ORF72*-ALS are known and recapitulated in models. Established and well-characterized *in vitro* (Sareen *et al.*, 2013) and *in vivo* (Shaw *et al.*, 2018) models of *C9ORF72*-ALS exist which were readily available for use in the project. Both the *in vitro* and *in vivo* models were able to display the DPR and RNA foci pathology associated with *C9ORF72*, and the *in vivo* model successfully displayed the motor deficit phenotype. All these phenotypes were used to assess the effects of ARFIP2 in ALS.

To investigate the effects of HC-ARFIP2 on *C9ORF72*-ALS disease characteristics, we first aimed to express HC-ARFIP2 in iPSC-derived motor neurons and zebrafish models of ALS used in this project. In order to achieve this, essential plasmids and appropriate delivery vectors were made and validated in less complex, cheaper and faster model systems (HEK293T cells). Through restriction enzyme digestion and Sanger sequencing, it was demonstrated that the cloning was successful and the *HC-ARFIP2* transgene was successfully inserted and expressed in the various viral vectors and tagged constructs. Lentiviruses were produced in-house with high functional titres and validated to show delivery of the transgene for expression in iPSC-derived MNs, while the cells still looked healthy post-transduction.

After the generation of the tools to be used in the project, we aimed to investigate whether endogenous ARFIP2 levels are altered in disease. We assessed the expression and distribution of endogenous ARFIP2 in control and ALS human brain tissue sections and *C9ORF72*-ALS iPSC-derived MNs. Immunohistochemistry staining of the human brain cerebellum and frontal cortex sections showed that endogenous levels of ARFIP2 appeared to be more extensively expressed in the cerebellum of *C9ORF72*-ALS cases and the frontal cortex of both *C9ORF72*-ALS and sALS cases, specifically in the Purkinje cells and the pyramidal neurons. The increase in ARFIP2 levels in the *C9ORF72*-ALS cases compared to controls offered support for the choice of the ALS genetic background used in this project. Moreover, with an increase of ARFIP2 seen in sALS cases, it warrants further exploration of ARFIP2 in other genotypes in ALS. In ALS, it is expected to see neuropathology in the motor cortex and spinal cord (Saberi *et al.*, 2015). Since the only other previous work on

ARFIP2 in neurodegenerative disorders was done in rodent models of HD and highlighted that ARFIP2 levels were increased in the cerebellum and cortex (Peters *et al.*, 2002), the cerebellum and frontal cortex sections were chosen for pilot investigation. It was planned to extend the investigation to include the more relevant regions where we expect to see ALS pathology, however, due to the delays in obtaining the human brain tissue because of the COVID-19 pandemic and the subsequent time constraints of the Ph.D., this could not be done and would be ideal to investigate in the future. Nonetheless, research has implicated the loss and degeneration of both the Purkinje cells and pyramidal neurons in ALS (Afshar *et al.*, 2017; Tan *et al.*, 2016; Mochizuki *et al.*, 2011). This increased immunoreactivity of ARFIP2 in previously associated neuronal types in ALS offered support to the hypothesis of targeting this protein in ALS, and also offered support for further investigation of ARFIP2 in ALS. Using cultured iPSC-derived motor neurons (MNs), we showed that in patient cell lines with the *C9ORF72*-ALS mutation, there was a significant increase in ARFIP2 expression compared to the control and isogenic cell lines used. An increase in endogenous ARFIP2 levels has also been reported in another neurodegenerative disease background where the levels of ARFIP2 were increased in patients with HD and in the striatum, cortex, and cerebellum of transgenic mice which were models of HD (Peters *et al.*, 2002; Rangone *et al.*, 2005). The finding that ARFIP2 expression is increased in a neurodegenerative background is consistent with the previous reports, however, its increase in expression in ALS specifically is a novel finding and implicated its possible involvement in the pathophysiology of ALS. This was encouraging and warranted further assessment into the effects of targeting *ARFIP2* on the pathogenesis of ALS in *C9ORF72*-ALS iPSC-derived MNs. iPSC-derived MNs are an expensive and time-consuming *in vitro* model. With the ubiquitous expression of ARFIP2, it can be argued that simpler models like fibroblasts from *C9ORF72*-ALS patients could be used initially to assess the alterations in ARFIP2 levels in ALS. It has been reported that in healthy individuals, ARFIP2 is moderately expressed in human fibroblasts (Karlsson *et al.*, 2021; *Human Protein Atlas*). While using fibroblasts would be inexpensive and quick, iPSC-derived MNs are a more relevant disease-specific model.

With the previously published finding that HC-ARFIP2 confers neuroprotective effects in HD by increasing survival of neurons and decreasing aggregation due to a potential

role in proteasome activity (Rangone *et al.*, 2005; Peters *et al.*, 2002), ARFIP2 was targeted in *C9ORF72*-ALS by expressing HC-ARFIP2 in iPSC-derived MNs. Metabolic abnormalities in ALS have been correlated with disease development (Vandoorne, De Bock and Van Den Bosch, 2018; Hor *et al.*, 2021). Using the resazurin reduction assay to conduct a longitudinal experiment, where the metabolic activity measured was proportional to cell viability, we showed that the *C9ORF72*-ALS iPSC-derived MN cell lines showed a steep decline in metabolic activity and therefore cell viability over time as the cells age. However, with HC-ARFIP2 expression the cell viability in the ALS cell lines was stable over time and comparable to the control cell lines. Therefore, it can be inferred that HC-ARFIP2 may improve cell viability in ALS and showed that HC-ARFIP2 may exert neuroprotective effects in ALS. Furthermore, these experiments also showed that HC-ARFIP2 expression post viral transduction at the selected dose was well tolerated and did not show any toxicity over time.

We then assessed whether the effect on cell viability was due to the effects of the transgene on *C9ORF72*-ALS-associated disease characteristics. Using a poly-GP MSD ELISA, we assessed the effects on poly-GP DPRs generated from both the sense and antisense transcripts, which serves as a proxy for sense and antisense DPRs. Results showed that expressing HC-ARFIP2 significantly decreased the levels of the poly-GP DPR in *C9ORF72*-ALS iPSC-derived MNs, returning it to the background level seen in healthy control. This shows a potential therapeutic effect of HC-ARFIP2 in ALS, however, further assessment of DPRs is required. While poly-GP DPRs were a good representative of the DPRs generated from both sense and antisense strands, it is not possible to predict and infer the effects of HC-ARFIP2 on other DPR species in *C9ORF72*-ALS. Research has shown that poly-GP, along with poly-PA, is unable to aggregate by itself due to a flexible coil structure. These DPRs also have fewer intracellular protein interactions compared to other DPRs, suggesting that they may have lower associated toxicity compared to other DPR species (Lee *et al.*, 2016; Schmitz *et al.*, 2021). The arginine-containing DPRs, poly-GR and poly-PR, are shown to have high associated toxicity in many *in vitro* models (Lee *et al.*, 2016; Wen *et al.*, 2014; Tao *et al.*, 2015) and are considered to be the most toxic DPRs (Schmitz *et al.*, 2021). Furthermore, each DPR has a different translation efficiency with varying expression patterns (Lee *et al.*, 2016) and the abundance of a particular DPR doesn't necessarily correlate with high associated toxicity (Schmitz *et al.*, 2021).

Therefore, to determine the effects of HC-ARFIP2 on DPRs, an assessment of all other DPRs, especially the more toxic arginine-containing DPRs is necessary. This would provide further insight into the potential neuroprotective effects of HC-ARFIP2 in ALS. However, this was not possible to do due to the time constraints of the Ph.D. The poly-GP MSD ELISA was the first in-house DPR assay to be developed. While the assays for the remaining four DPRs are now developed, they were not ready to use during the time period of this project.

Effects of HC-ARFIP2 expression on other disease characteristics including RNA foci and TDP-43 mislocalisation were also assessed. It was reported that the increase in transcription resulted in the accumulation of repeat-containing RNA foci in the *C9ORF72*-ALS iPSC-derived MN cultures (Sareen *et al.*, 2013). We showed that the antisense nuclear RNA foci were detected and measurable in our motor neuron cultures. This was consistent with the previous work focusing on RNA foci in the *in vitro* model used (Sareen *et al.*, 2013) as they also only reported the presence of antisense RNA foci as well. The previous study showing the presence of RNA foci in the cell model used showed that they detected foci in ~20% of the cells while we were able to detect foci in only ~5% of the cells. While the probes used and the fluorescence in situ hybridisation (FISH) and FISH/immunocytochemistry staining methods were the same, there were differences in the iPSC differentiation protocol and the detection methods used. The differentiation protocol used in this project follows a well-established protocol (Du *et al.*, 2015) which has a protocol length of 28 days. On the other hand, Sareen et al. (2013) use the protocol described in the study which spans 75 days. While the motor neurons were validated and cultured until day 40 for use in experiments shown in this thesis, the length of culture may not be enough to develop disease characteristics and might explain the low percentage of cells with RNA foci in our MN cultures. We detected the RNA foci in MN cultures using the Opera Phenix High Content Imaging System, whereas Sareen et al. (2013) used confocal microscopy for the detection method. Using the Opera Phoenix High Content Imaging system provides high resolution, unsupervised image acquisition, high throughput, allows for analysis of more cells and avoids spectral crosstalk. However, the exclusion of background in the analysis was not possible and any minimisation of noise/background is critical when assessing for small objects such as RNA foci. Therefore due to background and a low signal-to-noise ratio, the levels we report are

different from those previously reported. Using the confocal microscope for the detection of RNA foci could be a more sensitive method. In the future, it would be interesting to see whether using confocal microscopy can detect antisense foci in a larger proportion of cells as well as the detection of sense RNA foci in the MN cultures. With HC-ARFIP2 expression, the results showed no measurable effect on antisense RNA foci in *C9ORF72*-ALS iPSC-derived MNs. This can be attributed to the sensitivity of the assay and detection method used to assess RNA foci as well as the high background levels seen in the experiment.

A previous study reported that the antisense foci in motor neurons were associated with TDP-43 mislocalisation (Cooper-Knock *et al.*, 2015). It has been suggested that prolonged TDP-43 mislocalisation promotes aggregation which can be overwhelming for the degradation pathways, leading to eventual cell death (Winton *et al.*, 2008; Walker *et al.*, 2015). Targeting the aggregation may be too late and therefore early therapeutic strategies that rescue TDP-43 mislocalisation are important (Suk and Rousseaux, 2020). Since the presence of antisense foci was observed in our motor neuron cultures and developing therapeutic strategies targeting TDP-43 mislocalisation is important, we looked into the TDP-43 mislocalisation pathology in our model and assessed the effects of HC-ARFIP2 on this ALS phenotype. The analysis did not show a statistically significant TDP-43 mislocalisation phenotype in the *C9ORF72*-ALS iPSC-derived MNs but showed that there was a possible trend towards this phenotype, which may be attributed to the presence of antisense RNA foci. Even though TDP-43 mislocalisation and aggregation is a key component of C9ORF72-ALS at post-mortem, very few models to date have robustly shown this in hiPSC-derived MNs. It is unknown if this means that the TDP-43 pathology in *C9ORF72*-ALS can be disassociated from other characteristics of disease pathology, or if further cell maturation is needed for the TDP-43 pathology to develop (Giacomelli *et al.*, 2022). The iPSC-derived MNs used in the experiments were cultured for longer for further maturation and since it has been suggested that TDP-43 mislocalisation may be upstream of aggregation, it is likely that the cells may exhibit the TDP-43 aggregation pathology. If time permitted, it would be interesting to assess TDP-43 aggregation and whether HC-ARFIP2 has any effect on TDP-43 aggregation. Additionally, it would be interesting to assess the TDP-43 mislocalisation phenotype

at an earlier stage of the differentiation. Furthermore, with HC-ARFIP2 expression, there was no effect on TDP-43 localisation and distribution.

Previously, increasing the concentration of ARFIP2 has been shown to downregulate proteasome activity and expression of HC-ARFIP2 has been shown to dominantly rescue proteasomal activity (Peters *et al.*, 2002). Phosphorylation of ARFIP2 at a site on the C-terminal has also been shown to inhibit the blockage of the proteasome (Rangone *et al.*, 2005). Furthermore, a study reported that overexpression of poly-GP interferes with the ubiquitin-proteasome system (UPS) and demonstrated its associated cytotoxicity via UPS dysfunction (Yamakawa *et al.*, 2015). Therefore, to investigate the possible mechanism involved in the reduction of poly-GP DPRs due to HC-ARFIP2 expression, the effect of HC-ARFIP2 on proteasome activity was evaluated. However, due to assay variance, we were unable to draw any definite conclusions regarding the effect of HC-ARFIP2 on proteasome activity. While the results for the assessment of proteasome activity did not provide significant findings in the duration of the Ph.D., the literature suggests that the C-terminal (either through its expression or phosphorylation) dominantly rescues proteasomal activity (Peters *et al.*, 2002; Rangone *et al.*, 2005). Therefore, further investigation using alternative ways to measure proteasome activity is warranted.

ARFIP2 is also known as a partner of Rac1 (POR1) and binds to Rac1 and ADP-ribosylation factors (Arfs) in a GTP-dependent manner (Shin and Exton, 2001; Shin and Exton, 2005; D'Souza-Schorey *et al.*, 1997; Van Aelst, Joneson and Bar-Sagi, 1996). Rac1 plays a role in regulating multiple pathways including those involved in membrane ruffling, cytoskeleton organisation, cell proliferation and transcription (Bosco, Mulloy and Zheng, 2009). In the nervous system, it has been shown to regulate microtubule cytoskeletal dynamics and therefore it is central to the growth and stability of axons and dendrites. Furthermore, dysregulation and a decrease in active Rac1 (Rac1-GTP) have been implicated in the pathogenesis of ALS (D'Ambrosi *et al.*, 2014; Stankiewicz *et al.*, 2020). The small GTPase Arf1, 5 and 6 (in their GTP forms) have been shown to interact with ARIFP2 (Shin and Exton, 2005). Arf6 has many important functions and plays a role in protein trafficking and direct transport along microtubules, autophagy and cytoskeletal remodelling (Van Acker, Tavernier and Peelman, 2019). The exact function of ARFIP2 is unknown but the function of the

proteins that interact with ARFIP2 and its subcellular location indicates that it may be involved in regulating cytoskeletal remodelling via the microtubule network (D'Souza-Schorey *et al.*, 1997; Van Aelst, Joneson and Bar-Sagi, 1996; Shin and Exton, 2001; Peters *et al.*, 2002; Rangone *et al.*, 2005).

Rangone et al., (2005) showed that phosphorylating ARFIP2 at the C-terminus rescues proteasome impairment in HD and they also showed that loss of this phosphorylation at the C-terminus of ARFIP2 disrupts the microtubule network. The loss of proteostasis (protein homeostasis) has been implicated in the pathogenesis of ALS due to the dysfunction of the protein degradation systems (Webster *et al.*, 2017). The UPS catalyses most of the protein degradation in the cell and when the UPS is overwhelmed in disease, aggresomes comprising misfolded proteins form to organise the aberrant protein into a single location (Corboy, Thomas and Wigley, 2005). Aggresomes are juxtanuclear inclusions that form as an active and highly regulated cellular response to the expression of misfolded or aberrant proteins (Taylor *et al.*, 2003). Their formation involves the trafficking of aberrant proteins along the microtubules and therefore an intact network of microtubules is required in the sequestration of aggregates and subsequent degradation by the proteasome (Johnston, Ward and Kopito, 1998; Muchowski *et al.*, 2002). With the interaction of ARFIP2 with proteins that are involved in regulating microtubule cytoskeletal dynamics and direct transport along microtubules and evidence showing that targeting ARFIP2 disrupts the microtubule network (Rangone *et al.*, 2005), it provides an interesting link that ARFIP2 may act as a positive regulator in this process, highlighting a pivotal point in the pathway. In this project, we have shown that ARFIP2 expression is upregulated in ALS. It is possible that this increase in ARFIP2 may be positively regulating the formation of aggresomes which in turn overwhelms and negatively affects the UPS and may impair the degradation pathway. This would be consistent with the previously reported finding that increasing the concentration of ARFIP2 downregulates proteasome activity (Peters *et al.*, 2002). On the other hand, impairment of the proteasome pathway is a known pathological mechanism in ALS which could result in the accumulation of misfolded or aberrant proteins. This could increase the formation of aggresomes and therefore lead to an increase in ARFIP2 levels. Since impairment of the proteasome is common in many neurodegenerative disorders, it will be

interesting to assess this mechanism in the future and also evaluate the effects of HC-ARFIP2 expression.

Looking at the protein degradation system of DPRs in *C9ORF72*-ALS, it has been shown that poly-GP, along with poly-GR and poly-PR, are degraded by both UPS and autophagy pathways, while the UPS pathway does not contribute to degrading poly-GA (Yamakawa *et al.*, 2015; Cristofani *et al.*, 2017). Yamakawa et al (2015) assessed four out of five DPRs and poly-PA was not analysed in the study as it was difficult to detect quantitatively on blot analysis. On the other hand, it has been shown that the degradation of DPRs (including poly-PA) predominantly relies on functional autophagic clearance, while poly-GP is the only DPR degraded by both degradation pathways (Cristofani *et al.*, 2018). Dysregulation of the autophagy pathway has been widely implicated in *C9ORF72*-ALS, leading to impairment in DPR clearance and their accumulation (Sullivan *et al.*, 2016; Boivin *et al.*, 2020; Beckers, Tharkeshwar and Van Damme, 2021). Recently, ARFIP2 has been shown to be involved in autophagy by regulating the amino acid starvation-dependent distribution of ATG9A vesicles which then deliver the PI4-kinase, PI4KIIIβ, to the autophagosome initiation site (Judith *et al.*, 2019). Arf6, a small GTPase protein that ARFIP2 interacts with, has been shown to play a role and assist in autophagy (Van Acker, Tavernier and Peelman, 2019). With this knowledge, it would be interesting to assess the effects of expressing HC-ARFIP2 on autophagy in the future which could help explain the pathway by which HC-ARFIP2 expression may reduce the levels of the poly-GP DPR. Furthermore, this would also determine whether ARFIP2 is involved in regulating the degradation of other DPR species.

With an improvement of cell viability over time and reduced poly-GP DPRs seen when HC-ARFIP2 was expressed in *C9ORF72*-ALS iPSC-derived MNs, these findings warranted further investigation in a more complex and relevant *in vivo* disease model. We aimed to assess whether the expression of HC-ARFIP2 affects disease pathogenesis in a zebrafish model of *C9ORF72*-ALS (Shaw *et al.*, 2018). *ARFIP2* is highly conserved between humans and Zebrafish, with a 69% identity between the zebrafish and human amino acid sequence. Most of the homology seen is in the c-terminus conserved region. To assess the effects of HC-ARFIP2 in *in vivo*, we chose to take a more rigorous approach of generating stable *HC-ARFIP2* transgenic

zebrafish lines, rather than the quick and invasive approach of injecting an artificially synthesised HC-ARFIP2 peptide. An established method by two independent groups was followed to successfully generate the *HC-ARFIP2* transgenic lines (Higashijima *et al.*, 1997; Long *et al.*, 1997). We successfully found and established two stable lines that showed ubiquitous expression of the transgene and transgene inheritance through generations. This showed that HC-ARFIP2 expression was safe and had no associated toxicity in zebrafish. Both the established heterozygous transgenic lines were crossed with the heterozygous sense and antisense C9ORF72-ALS zebrafish lines to generate experimental populations of zebrafish with distinct genotypes.

The *C9ORF72*-ALS zebrafish models used in the study have been previously reported to show clear adult-onset ALS-like motor performance deficits (Shaw *et al.*, 2018). We assessed the swimming endurance of the fish using a swim tunnel assessment and our results were consistent with those previously reported. It was shown that at 6 months old, the C9 zebrafish in the sense and antisense background (C9-S and C9-A) show a significant phenotype of swimming endurance deficit. A significant rescue of this phenotype with HC-ARFIP2 expression was seen, however, this was only seen in the antisense background. It is possible that the effect of HC-ARFIP2 observed in the antisense background can be due to the more extreme phenotype seen in the antisense background (C9-A) compared to the sense background (C9-S). Some swim-tunnelled antisense zebrafish displayed signs of muscle atrophy as they became incapable to continue swimming, necessitating culling. The remaining swim-tunnelled zebrafish appeared healthy, and this indicated heterogeneity in the phenotype progression. This suggests that genetic, epigenetic, and cardiovascular involvement or other factors like gender and weight may modulate the motor performance phenotype. The effects of gender and weight on motor performance were assessed. Results showed that in both the sense and antisense background, the females had lower swimming endurance than their male clutch mates. Therefore the samples for each genotype included equal numbers for each gender to ensure that the results were not biased. It was reported that there was a significant decrease in the total weight of *C9ORF72*-ALS zebrafish compared to NTGs and the *C9ORF72*-ALS fish appeared to be smaller during the early stages of development (Shaw *et al.*, 2018). With this knowledge, the zebrafish in each sample population were size-matched and weighed to reduce the effects of external factors. Results showed that there was no

effect of weight on swimming endurance in the antisense background and a very weak effect in the sense background. While the gender and weight of individual fish were assessed, it would be interesting to assess the other stated factors. The motor performance analysis was conducted at 6 months of age, and it has been reported that the swimming ability of zebrafish at maximum speed continued to decrease as the fish aged to 12 months of age (Shaw *et al.*, 2018). With the development of a possible stronger motor performance deficit phenotype as the fish age, a significant rescue of the swimming endurance deficit may be achieved in the sense (C9-S) background at 12 months of age. Therefore conducting the swim tunnel analysis with more aged zebrafish might add further to the therapeutic potential of HC-ARFIP2 in ALS. Though the results showing the effects of HC-ARFIP2 expression on motor performance in zebrafish are promising, further validation of the experiment is required.

It is likely that the more severe motor performance phenotype observed in the antisense background is due to a model-specific effect or it may be due to genotype-specific effects. The arginine-rich poly-GR (from the sense strand) and poly-PR (from the antisense strand) DPRs are considered to be the more toxic DPRs due to their highly interactive nature and have been shown to be toxic in many disease models (Wen *et al.*, 2014; Tao *et al.*, 2015; Lee *et al.*, 2016). Their toxicity can be explained by their positive charge and high polarity due to arginine. In addition to cytoplasmic toxicity, it has also been reported that both DPRs can be easily transported into the nucleus and can disrupt RNA biogenesis and decrease cell viability (Kwon *et al.*, 2014; Schmitz *et al.*, 2021). Despite the studies showing the toxicity of arginine-rich DPRs, particularly in the nucleus, the relative contribution to the toxicity of poly-GR and poly-PR in ALS is still unclear. Furthermore, it is possible that the DPRs from sense and antisense strands are not translated at comparable levels (Lee *et al.*, 2016; Schmitz *et al.*, 2021). It is likely that, in addition to a model effect, the more severe motor performance phenotype in the antisense background zebrafish may be due to the toxicity elicited by the arginine-rich poly-PR DPRs. While poly-GR concentrates mainly in the cytoplasm, poly-PR mainly localises in the nucleolus (Schmitz *et al.*, 2021; Liu *et al.*, 2022). White et al., (2019) showed the involvement of poly-PR in the nucleolus and highlighted that it binds, sequesters and changes the biophysical properties of nucleophosmin1 (NPM1) (White *et al.*, 2019; Lee *et al.*, 2016). The transcription, translation and diffusion of NPM1 were also reported to be impaired by poly-PR (Chen

et al., 2021). The NPM1 protein is involved in many cellular processes including protein chaperoning, centrosome duplication, ribosome biogenesis and cell proliferation. NPM1 has been shown to play a key role in the protein quality control machinery of the nucleolus and poly-PR expression compromised this function (Frottin *et al.*, 2019). The downstream effects of disrupting NPM1 have not yet been clearly understood, however, evidence suggests that the binding, displacement and disruption of its functions associated with poly-PR could significantly further DPR-associated toxicity. It has also been reported that poly-PR expression can inhibit DNA double-strand break repair by binding to NPM1 (Andrade *et al.*, 2020). Furthermore, it was reported that a mouse model expressing the *C9ORF72* repeat expansion via a bacterial artificial chromosome (BAC) exhibited motor impairments and the affected lines of mice had upregulated antisense transcripts (Liu *et al.*, 2016). It may be likely that the impairment is associated with measurable poly-PR expression. Another study investigated the associated toxicity with individual DPR expression and identified poly-PR as the most toxic DPRs in cellular models (Darling *et al.*, 2019). Poly-PR has a strong tendency to interact with low complexity domain (LCD) containing proteins (Lee *et al.*, 2016; Lin *et al.*, 2016). LCDs are composed of repeats of a few amino acids and are very abundant in eukaryotic proteins. The binding of poly-PR to LCD-containing proteins (including FUS, hnRNPA1 and hnRNPA2B1) may disrupt many cellular functions leading to disrupted homeostasis and cytotoxicity (Lin *et al.*, 2016). This indicates that poly-PR-associated toxicity may explain the more severe motor performance phenotype observed in the antisense background zebrafish model, however, definite conclusions can only be made after assessing the relative expressions and toxicity of each of the DPR species in the zebrafish sense and antisense models.

In *C9ORF72*-ALS zebrafish models, the mutation-associated characteristic of DPRs has been previously reported (Ohki *et al.*, 2017; Shaw *et al.*, 2018). Similar to the assessment of the poly-GP DPR in the cellular model, a poly-GP MSD ELISA was used. We focused on the poly-GP DPR as it served as a proxy for sense and antisense DPRs as it is translated from both transcripts. Our results showed that the *C9ORF72*-ALS model used showed significant levels of the poly-GP DPR in both the sense and antisense background, where the DPRs were more abundant in the sense background compared to the antisense background. It has been reported that in the sense strand, the poly-GP repeat has a unique C-terminal sequence, whereas, in the antisense

strand, there is a stop codon immediately following the repeat (Zu *et al.*, 2013). This suggests that the poly-GP translated from the sense and antisense strands are not identical. It has also been reported that the C-terminal sequence of proteins from RAN translation can affect their relative toxicity and cellular distribution (He *et al.*, 2020), which suggests that the sense and antisense resultant poly-GP DPRs may have different molecular interactions and biochemical properties. Therefore, it may be possible that the in-house developed anti-poly-GP antibody used in the assay is more efficient against the sense-translated poly-GP DPR than the antisense-translated poly-GP DPR. This may explain the difference in the poly-GP DPR burden seen between the sense and antisense background. The observed difference in abundance could also be attributed to the amount of soluble and insoluble poly-GP DPRs present in the samples. The DPR MSD ELISA is able to measure the amount of soluble and insoluble poly-GP DPRs in the samples, however this is dependent on the extraction method used (Quaegebeur *et al.*, 2020). Here, the samples were solubilised using urea before running the assay, however, the concentration of urea may be limiting, and it might not be able to solubilise the samples completely. Increasing the concentration of urea was not possible because it would interfere with the binding of the antibody. Therefore, there could be high amounts of insoluble poly-GP DPRs in aggregates in the antisense background that were not measured. Previously, Zu *et al.* (2013) have also shown that most of the poly-GP inclusions were generated from the antisense strand. This offers an explanation for the lower abundance of poly-GP DPR observed in the antisense background as insoluble poly-GP was unmeasurable. In the future, a thorough investigation assessing the aggregation of poly-GP in the samples and further solubilisation of the samples before running the MSD ELISA would offer more insights. The possible presence of high insoluble poly-GP inclusions in the antisense background might also explain the observation of a more severe motor performance phenotype in the antisense background. Previously, a study showed that the presence of poly-GP DPR inclusions in the muscle is associated with significant muscle atrophy in *C9ORF72*-ALS patients (Cykowski *et al.*, 2019). This suggests that poly-GP DPR aggregates may contribute to muscle pathology in ALS. Furthermore, this is also consistent with the observation that swim-tunnelled antisense zebrafish displayed clear signs of muscle atrophy. With HC-ARFIP2 expression showing a rescue of the motor performance phenotype in the antisense background, it would be interesting to see the effect of HC-ARFIP2 expression on DPR inclusions in the muscle. Further

investigation of DPR inclusions in muscle samples of zebrafish will provide more insight into the cause of the more severe motor performance phenotype and the mechanism by which HC-ARFIP2 expression rescues this. Additionally, it may be possible that poly-GP DPRs do not affect motor performance in zebrafish and therefore the differential expression of DPRs warrants further investigation of all DPR species in ALS. Assays are now available in the lab to further this work.

The established *HC-ARFIP2* transgenic lines (2 and 4) were crossed with the sense and antisense *C9ORF72*-ALS lines. Though a strong effect of HC-ARFIP2 expression on poly-GP DPRs was seen *in vitro*, in the embryos of the line 2 genotypes, there was no effect of HC-ARFIP2 on poly-GP DPRs. On the other hand, in the line 4 genotypes, an effect of HC-ARFIP2 expression on poly-GP DPRs were observed using embryos, however, only in the sense background. This offered some support to the therapeutic potential of HC-ARFIP2 in ALS, however, further validation is needed. It was observed that the staining intensity of HC-ARFIP2 expression in line 4 F1 transgenic embryos appeared to be more intense than in line 2 F1 transgenic embryos. Therefore it is likely that the effect of HC-ARFIP2 on poly-GP DPRs in line 4 may be due to a higher expression of HC-ARFIP2. Furthermore, we cannot be sure if the effect of HC-ARFIP2 on poly-GP DPR is due to unmeasurable events like the locus of integration of the transgene or homologous recombination. Nonetheless, the preliminary results showing the decrease in poly-GP levels correlating to HC-ARFIP2 expression in *in vivo* warrant further investigation on the effects of HC-ARFIP2 on all other DPR species expressed in *C9ORF72*-ALS. Due to the interest of time, an assessment of the effect of HC-ARFIP2 using line 4 was only conducted for DPRs. Analysis of the effects of HC-ARFIP2 on all other read-outs using line 4 would be interesting and provide further insights into the effects of HC-ARFIP2 on disease phenotypes as well as the level of expression needed to see neuroprotective effects in *in vivo*.

Several different zebrafish *C9ORF72*-ALS models have displayed the RNA foci phenotype (Lee *et al.*, 2013; Ohki *et al.*, 2017; Swinnen *et al.*, 2018; Shaw *et al.*, 2018). Consistent with this, results showed that both sense and antisense RNA foci were observed in zebrafish embryos with C9-S and C9-A genotypes, respectively. Though the embryos displayed the RNA foci phenotype, embryos are not a good representative of the adult-onset disorder. Therefore, assessing the effect of HC-

ARFIP2 on RNA foci in adult zebrafish would provide better insight, as well as comparing the foci load in various regions of the brain and spinal cord in adult zebrafish. It was also observed that the RNA foci were more abundant in the sense background compared to the antisense background. A reason for this difference in abundance may be due to the detection of the foci. It is possible that the sense foci are more readily detectable than the antisense foci. It has been widely reported that sense RNA, due to the G-rich regions, have the propensity to form structures called G-quadruplexes (Fratta *et al.*, 2012; Reddy *et al.*, 2013; Conlon *et al.*, 2016; Schludi and Edbauer, 2018). The presence of these structures may confer a different conformation to the sense RNA foci which may make them more detectable than antisense RNA foci. Furthermore, it may be possible that sense RNA foci are more stable than antisense RNA foci as G-quadruplexes are highly stable structures. Recently, a more sensitive and enhanced detection method for RNA foci has been developed that, in comparison to the traditional FISH method, dramatically increases the signal of a single molecule, making it simpler to detect RNAs with a low abundance and eliminate background signal due to autofluorescence and off-target binding (Glineburg *et al.*, 2021). Further analysis of RNA foci using this method may provide more insights. Furthermore, contrasting results were seen in *in vitro* where antisense RNA foci were observed to be more abundant than sense RNA foci. In zebrafish, the RNA foci were assessed in the muscle. It may be possible that the difference in RNA foci levels may be due to a cell-specific effect and therefore it would be interesting to assess sense and antisense RNA foci in zebrafish MNs. As seen in *in vitro*, HC-ARFIP2 did not affect RNA foci levels in zebrafish. However further investigation and validation are required to draw definite conclusions.

6.2 Conclusions

The data presented in this thesis identifies ARFIP2 as a novel protein of interest in ALS and highlights that expressing its C-terminal affects disease characteristics. In this project, it was shown that HC-ARFIP2 improves cell viability in *C9ORF72*-ALS iPSC-derived MNs, affects the levels of the poly-GP DPRs, and significantly rescues motor performance deficits seen in the *C9ORF72*-ALS zebrafish model. Taken together, the data presented in this thesis identify the therapeutic potential of targeting ARFIP2 in ALS and show that expressing HC-ARFIP2 may ameliorate the disease phenotypes. Though the data presented is promising and encouraging, it is preliminary

and further work is required to completely comprehend the involvement and effects of HC-ARFIP2 in ALS. Nonetheless, it is important to highlight that the current project shows findings of great importance that may contribute to the identification and development of a novel therapeutic target for ALS.

ARFIP2 has been shown to be involved in autophagy by regulating the distribution of ATG9A vesicles which deliver the PI4-kinase (phosphatidylinositol-4IIIβ) to the autophagosome initiation site (Judith *et al.*, 2019). PI4KIIIβ interacts with ATG9A and ATG13 to control phosphatidylinositol-4-phosphate (PI4P) production at the initiation membrane site and the autophagic response. Therefore it can be inferred that ATG9A shapes the forming autophagosome through ARFIP2 and ARFIP2 modulates the composition of the ATG9A membranes. The regulation of the ATG9A vesicles distributions is amino acid starvation dependent. In *C9ORF72*-ALS, as DPRs are produced, an amino acid starvation condition may occur in the cell, upregulating autophagy. This may cause the ARFIP2 to be increased in C9-ALS as ARFIP2 is shown to be required to assist in autophagic mechanisms. Even if autophagy is impaired, ARFIP2 may be working upstream and may be upregulated during the amino acid starvation dependent induced autophagy. The data from this thesis offers support to this. Furthermore, when the UPS is overwhelmed due to aberrant/misfolded proteins, it triggers the aggresome formation initiation which needs an intact microtubule network to form aggresomes. It was shown that loss of phosphorylation of ARFIP2 at Ser260 (C-terminal on ARFIP2) disrupts the microtubule network and that phosphorylation inhibits the blockade of UPS. This highlights that ARFIP2 is a positive regulator in aggresome formation and therefore in the UPS pathway (Rangone *et al.*, 2005). With the above in mind and the previous evidence for the dysregulation of UPS and autophagy in ALS, it can be inferred that ARFIP2, particularly HC-ARFIP2, may present as a viable therapeutic target for use in ALS. However, further work to support the data in this thesis is required.

In addition to HD and ALS, protein aggregates are a hallmark for other neurodegenerative diseases and proteostasis disruption has been implicated in their disease development. ARFIP2 has been previously implicated in HD and the results from this project show the possible involvement of ARFIP2 in *C9ORF72*-ALS. Moreover, ARFIP2 may be involved in protein degradation pathways. Therefore the

importance of this study lies in the prospect that this concept may apply to other neurodegenerative disorders as well as proteinopathies. Additionally, *C9ORF72*-ALS and HD can be categorized as repeat expansion disorders. This further highlights that this concept may also be applied to other repeat expansion diseases.

6.3   Future perspectives

While the results from the project are promising, further validation is required through future work which has been discussed throughout the thesis. It would be interesting to assess the expression and distribution of ARFIP2 in the more relevant regions of the human brain where we expect to see ALS pathology. In addition, it would be interesting to assess whether the expression levels of ARFIP2 are increased in the zebrafish model used at the embryo stage when the DPR pathology and RNA foci were detected in the model, 6-months of age when the motor deficit pathology was assessed and at a more advanced age when more severe ALS-associated phenotypes are expected. Furthermore, it has been shown that the adult survival rates of the *C9ORF72*-ALS zebrafish model, which was used in this project, significantly reduced as they aged, compared to their NTG clutch mates (Shaw *et al.*, 2018). Therefore, it would be interesting to assess the effects of HC-ARFIP2 on survival in *in vivo*. While autophagy impairment has been widely implicated in ALS (Sullivan *et al.*, 2016; Boivin *et al.*, 2020; Beckers, Tharkeshwar and Van Damme, 2021; Chua *et al.*, 2022; Vicencio *et al.*, 2020), it has also been recently highlighted in HD (Oh *et al.*, 2022; Dewan *et al.*, 2021). With evidence showing the involvement of ARFIP2 in autophagy (Judith *et al.*, 2019) as well its implication in both ALS and HD, ARFIP2 may play a role in the autophagy pathway in neurodegenerative disorders. Assessing this in the future may further elucidate the involvement of ARFIP2 in autophagy and could help explain the pathway by which HC-ARFIP2 expression may reduce the levels of DPRs. Additionally, it was shown that the neuroprotective effects of ARFIP2 are due to the phosphorylation at a site on the C-terminal (Rangone *et al.*, 2005). Therefore, it would be interesting to assess the phosphorylation state of HC-ARFIP2.

While zebrafish have many advantages as a model, they serve as an intermediary step between assessing the effects in cells and then in mouse models. Zebrafish are small, prolific breeders and easy to house but still biologically complex, therefore, they can be used for faster analysis than in mice as well as providing valuable *in vivo* data.

The zebrafish model used in this project was able to recapitulate the ALS disease phenotypes seen in humans such as DPRs and RNA foci as well as reduced survival and motor deficits (Shaw *et al.*, 2018). While mouse models are able to display the DPR and RNA foci phenotypes, the current models fail to accurately mimic the survival and motor deficits (Mordes *et al.*, 2020). This supports the use of zebrafish in the study and highlights the importance of the zebrafish model for ALS. Nonetheless, using mouse models for assessment is still essential as they are larger and more biologically similar to humans. With further validation of the therapeutic potential of HC-ARFIP2 in zebrafish, it would then be interesting to further assess whether the results can be reproduced in a mouse model of *C9ORF72*-ALS and identify ARFIP2 as a novel therapeutic target for ALS. Furthermore, it would be interesting to assess the therapeutic potential of HC-ARFIP2 in models of ALS with other genetic backgrounds that have been implicated in the pathogenesis of the disease.

7. REFERENCES

Abe, K., Itoyama, Y., Sobue, G., Tsuji, S., Aoki, M., Doyu, M., Hamada, C., Kondo, K., Yoneoka, T., Akimoto, M., Yoshino, H. and Group, E. A. S. (2014) 'Confirmatory double-blind, parallel-group, placebo-controlled study of efficacy and safety of edaravone (MCI-186) in amyotrophic lateral sclerosis patients', *Amyotroph Lateral Scler Frontotemporal Degener,* 15(7-8), pp. 610-7.

Afshar, P., Ashtari, N., Jiao, X., Rahimi-Balaei, M., Zhang, X., Yaganeh, B., Del Bigio, M. R., Kong, J. and Marzban, H. (2017) 'Overexpression of Human SOD1 Leads to Discrete Defects in the Cerebellar Architecture in the Mouse', *Front Neuroanat,* 11, pp. 22.

Ahmed, S. A., Gogal, R. M. and Walsh, J. E. (1994) 'A new rapid and simple non-radioactive assay to monitor and determine the proliferation of lymphocytes: an alternative to [3H]thymidine incorporation assay', *J Immunol Methods,* 170(2), pp. 211-24.

Ajitkumar, A. and De Jesus, O. (2022) 'Huntington Disease', *StatPearls [Internet]*. Treasure Island (FL): StatPearls Publishing.

Al-Chalabi, A., Andersen, P. M., Chioza, B., Shaw, C., Sham, P. C., Robberecht, W., Matthijs, G., Camu, W., Marklund, S. L., Forsgren, L., Rouleau, G., Laing, N. G., Hurse, P. V., Siddique, T., Leigh, P. N. and Powell, J. F. (1998) 'Recessive amyotrophic lateral sclerosis families with the D90A SOD1 mutation share a common founder: evidence for a linked protective factor', *Hum Mol Genet,* 7(13), pp. 2045-50.

Al-Chalabi, A., Fang, F., Hanby, M. F., Leigh, P. N., Shaw, C. E., Ye, W. and Rijsdijk, F. (2010) 'An estimate of amyotrophic lateral sclerosis heritability using twin data', *J Neurol Neurosurg Psychiatry,* 81(12), pp. 1324-6.

Al-Chalabi, A. and Hardiman, O. (2013) 'The epidemiology of ALS: a conspiracy of genes, environment and time', *Nature Reviews Neurology,* 9(11), pp. 617-628.

Al-Saif, A., Al-Mohanna, F. and Bohlega, S. (2011) 'A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis', *Ann Neurol,* 70(6), pp. 913-9.

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(6), pp. 691-702.

Ambroggio, E. E., Sillibourne, J., Antonny, B., Manneville, J. B. and Goud, B. (2013) 'Arf1 and Membrane Curvature Cooperate to Recruit Arfaptin2 to Liposomes', *PLoS One: Vol. 4*.

Andersen, P. M., Forsgren, L., Binzer, M., Nilsson, P., Ala-Hurula, V., Keränen, M. L., Bergmark, L., Saarinen, A., Haltia, T., Tarvainen, I., Kinnunen, E., Udd, B. and Marklund, S. L. (1996) 'Autosomal recessive adult-onset amyotrophic lateral sclerosis associated with homozygosity

for Asp90Ala CuZn-superoxide dismutase mutation. A clinical and genealogical study of 36 patients', *Brain,* 119 (Pt 4), pp. 1153-72.

Andersen, P. M., Nilsson, P., Ala-Hurula, V., Keränen, M. L., Tarvainen, I., Haltia, T., Nilsson, L., Binzer, M., Forsgren, L. and Marklund, S. L. (1995) 'Amyotrophic lateral sclerosis associated with homozygosity for an Asp90Ala mutation in CuZn-superoxide dismutase', *Nat Genet,* 10(1), pp. 61-6.

Andersen, P. M., Sims, K. B., Xin, W. W., Kiely, R., O'Neill, G., Ravits, J., Pioro, E., Harati, Y., Brower, R. D., Levine, J. S., Heinicke, H. U., Seltzer, W., Boss, M. and Brown, R. H., Jr. (2003) 'Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes', *Amyotroph Lateral Scler Other Motor Neuron Disord,* 4(2), pp. 62-73.

Andrade, N. S., Ramic, M., Esanov, R., Liu, W., Rybin, M. J., Gaidosh, G., Abdallah, A., Del'Olio, S., Huff, T. C., Chee, N. T., Anatha, S., Gendron, T. F., Wahlestedt, C., Zhang, Y., Benatar, M., Mueller, C. and Zeier, Z. (2020) 'Dipeptide repeat proteins inhibit homology-directed DNA double strand break repair in C9ORF72 ALS/FTD', *Mol Neurodegener,* 15(1), pp. 13.

Andrews, J. A., Jackson, C. E., Heiman-Patterson, T. D., Bettica, P., Brooks, B. R. and Pioro, E. P. (2020) 'Real-world evidence of riluzole effectiveness in treating amyotrophic lateral sclerosis', *Amyotroph Lateral Scler Frontotemporal Degener,* 21(7-8), pp. 509-518.

Ansari, A. M., Ahmed, A. K., Matsangos, A. E., Lay, F., Born, L. J., Marti, G., Harmon, J. W. and Sun, Z. (2016) 'Cellular GFP Toxicity and Immunogenicity: Potential Confounders in in Vivo Cell Tracking Experiments', *Stem Cell Rev Rep,* 12(5), pp. 553-559.

Arai, T., Hasegawa, M., Akiyama, H., Ikeda, K., Nonaka, T., Mori, H., Mann, D., Tsuchiya, K., Yoshida, M., Hashizume, Y. and Oda, T. (2006) 'TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis', *Biochem Biophys Res Commun,* 351(3), pp. 602-11.

Armada-Moreira, A., Gomes, J. I., Pina, C. C., Savchak, O. K., Gonçalves-Ribeiro, J., Rei, N., Pinto, S., Morais, T. P., Martins, R. S., Ribeiro, F. F., Sebastião, A. M., Crunelli, V. and Vaz, S. H. (2020) 'Going the Extra (Synaptic) Mile: Excitotoxicity as the Road Toward Neurodegenerative Diseases', *Front Cell Neurosci,* 14, pp. 90.

Aschenbrenner, D. S. (2022) 'New Oral Form for ALS Drug', *Am J Nurs,* 122(9), pp. 24-25.

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., 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(4), pp. 639-46.

Auewarakul, P. (2005) 'Composition bias and genome polarity of RNA viruses', *Virus Res,* 109(1), pp. 33-7.

Ayala, Y. M., Zago, P., D'Ambrogio, A., Xu, Y. F., Petrucelli, L., Buratti, E. and Baralle, F. E. (2008) 'Structural determinants of the cellular localization and shuttling of TDP-43', *J Cell Sci,* 121(Pt 22), pp. 3778-85.

Babin, P. J., Goizet, C. and Raldúa, D. (2014) 'Zebrafish models of human motor neuron diseases: advantages and limitations', *Prog Neurobiol,* 118, pp. 36-58.

Babić Leko, M., Župunski, V., Kirincich, J., Smilović, D., Hortobágyi, T., Hof, P. R. and Šimić, G. (2019) 'Molecular Mechanisms of Neurodegeneration Related to C9orf72 Hexanucleotide Repeat Expansion', *Behav Neurol,* 2019, pp. 2909168.

Bacman, S. R., Bradley, W. G. and Moraes, C. T. (2006) 'Mitochondrial involvement in amyotrophic lateral sclerosis: trigger or target?', *Mol Neurobiol,* 33(2), pp. 113-31.

Baens, M., Noels, H., Broeckx, V., Hagens, S., Fevery, S., Billiau, A. D., Vankelecom, H. and Marynen, P. (2006) 'The dark side of EGFP: defective polyubiquitination', *PLoS One,* 1, pp. e54.

Baloh, R. H. (2011) 'TDP-43: The relationship between protein aggregation and neurodegeneration in ALS and FTLD', *FEBS J,* 278(19), pp. 3539-49.

Banez-Coronel, M. and Ranum, L. P. W. (2019) 'Repeat-associated non-AUG (RAN) translation: insights from pathology', *Lab Invest,* 99(7), pp. 929-942.

Bankhead, P. (2014) 'Analyzing fluorescence microscopy images with ImageJ'.

Bannwarth, S., Ait-El-Mkadem, S., Chaussenot, A., Genin, E. C., Lacas-Gervais, S., Fragaki, K., Berg-Alonso, L., Kageyama, Y., Serre, V., Moore, D. G., Verschueren, A., Rouzier, C., Le Ber, I., Augé, G., Cochaud, C., Lespinasse, F., N'Guyen, K., de Septenville, A., Brice, A., Yu-Wai-Man, P., Sesaki, H., Pouget, J. and Paquis-Flucklinger, V. (2014) 'A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement', *Brain: Vol. 8*, pp. 2329-45.

Barmada, S. J., Skibinski, G., Korb, E., Rao, E. J., Wu, J. Y. and Finkbeiner, S. (2010) 'Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial amyotrophic lateral sclerosis', *J Neurosci,* 30(2), pp. 639-49.

Bastian, F. B., Roux, J., Niknejad, A., Comte, A., Fonseca Costa, S. S., de Farias, T. M., Moretti, S., Parmentier, G., de Laval, V. R., Rosikiewicz, M., Wollbrett, J., Echchiki, A., Escoriza, A., Gharib, W. H., Gonzales-Porta, M., Jarosz, Y., Laurenczy, B., Moret, P., Person, E., Roelli, P., Sanjeev, K., Seppey, M. and Robinson-Rechavi, M. (2021) 'The Bgee suite: integrated curated

expression atlas and comparative transcriptomics in animals', *Nucleic Acids Res,* 49(D1), pp. D831-D847.

Bauer, A. P., Leikam, D., Krinner, S., Notka, F., Ludwig, C., Längst, G. and Wagner, R. (2010) 'The impact of intragenic CpG content on gene expression', *Nucleic Acids Res,* 38(12), pp. 3891-908.

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(4), pp. 667-77.

Beal, M. F. (2002) 'Oxidatively modified proteins in aging and disease', *Free Radic Biol Med,* 32(9), pp. 797-803.

Beal, M. F., Ferrante, R. J., Browne, S. E., Matthews, R. T., Kowall, N. W. and Brown, R. H. (1997) 'Increased 3-nitrotyrosine in both sporadic and familial amyotrophic lateral sclerosis', *Ann Neurol,* 42(4), pp. 644-54.

Beard, J. D. and Kamel, F. (2015) 'Military Service, Deployments, and Exposures in Relation to Amyotrophic Lateral Sclerosis Etiology and Survival', *Epidemiologic Reviews,* 37(1), pp. 55-70.

Beckers, J., Tharkeshwar, A. K. and Van Damme, P. (2021) 'C9orf72 ALS-FTD: recent evidence for dysregulation of the autophagy-lysosome pathway at multiple levels', *Autophagy,* 17(11), pp. 3306-3322.

Bendotti, C. and Carrì, M. T. (2004) 'Lessons from models of SOD1-linked familial ALS', *Trends Mol Med,* 10(8), pp. 393-400.

Bendotti, C., Marino, M., Cheroni, C., Fontana, E., Crippa, V., Poletti, A. and De Biasi, S. (2012) 'Dysfunction of constitutive and inducible ubiquitin-proteasome system in amyotrophic lateral sclerosis: implication for protein aggregation and immune response', *Prog Neurobiol,* 97(2), pp. 101-26.

Bensimon, G., Lacomblez, L. and Meininger, V. (1994) 'A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group', *N Engl J Med,* 330(9), pp. 585-91.

Birsa, N., Bentham, M. P. and Fratta, P. (2020) 'Cytoplasmic functions of TDP-43 and FUS and their role in ALS', *Semin Cell Dev Biol,* 99, pp. 193-201.

Bissaro, M., Federico, S., Salmaso, V., Sturlese, M., Spalluto, G. and Moro, S. (2018) 'Targeting Protein Kinase CK1δ with Riluzole: Could It Be One of the Possible Missing Bricks to Interpret Its Effect in the Treatment of ALS from a Molecular Point of View?', *ChemMedChem,* 13(24), pp. 2601-2605.

Blokhuis, A. M., Groen, E. J. N., Koppers, M., van den Berg, L. H. and Pasterkamp, R. J. (2013) 'Protein aggregation in amyotrophic lateral sclerosis', *Acta Neuropathol: Vol. 6*, pp. 777-94.

Boeynaems, S., Bogaert, E., Van Damme, P. and Van Den Bosch, L. (2016) 'Inside out: the role of nucleocytoplasmic transport in ALS and FTLD', *Acta Neuropathol,* 132(2), pp. 159-173.

Boillée, S., Vande Velde, C. and Cleveland, D. W. (2006) 'ALS: a disease of motor neurons and their nonneuronal neighbors', *Neuron,* 52(1), pp. 39-59.

Boivin, M., Pfister, V., Gaucherot, A., Ruffenach, F., Negroni, L., Sellier, C. and Charlet-Berguerand, N. (2020) 'Reduced autophagy upon C9ORF72 loss synergizes with dipeptide repeat protein toxicity in G4C2 repeat expansion disorders', *EMBO J,* 39(4), pp. e100574.

Bonafede, R. and Mariotti, R. (2017) 'ALS Pathogenesis and Therapeutic Approaches: The Role of Mesenchymal Stem Cells and Extracellular Vesicles', *Front Cell Neurosci,* 11, pp. 80.

Bonifacino, T., Zerbo, R. A., Balbi, M., Torazza, C., Frumento, G., Fedele, E., Bonanno, G. and Milanese, M. (2021) 'Nearly 30 Years of Animal Models to Study Amyotrophic Lateral Sclerosis: A Historical Overview and Future Perspectives', *Int J Mol Sci,* 22(22).

Boon, K. L., Xiao, S., McWhorter, M. L., Donn, T., Wolf-Saxon, E., Bohnsack, M. T., Moens, C. B. and Beattie, C. E. (2009) 'Zebrafish survival motor neuron mutants exhibit presynaptic neuromuscular junction defects', *Hum Mol Genet,* 18(19), pp. 3615-25.

Borawski, J., Lindeman, A., Buxton, F., Labow, M. and Gaither, L. A. (2007) 'Optimization procedure for small interfering RNA transfection in a 384-well format', *J Biomol Screen,* 12(4), pp. 546-59.

Bosco, D. A., Morfini, G., Karabacak, N. M., Song, Y., Gros-Louis, F., Pasinelli, P., Goolsby, H., Fontaine, B. A., Lemay, N., McKenna-Yasek, D., Frosch, M. P., Agar, J. N., Julien, J. P., Brady, S. T. and Brown, R. H. (2010) 'Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS', *Nat Neurosci,* 13(11), pp. 1396-403.

Bosco, E. E., Mulloy, J. C. and Zheng, Y. (2009) 'Rac1 GTPase: a "Rac" of all trades', *Cell Mol Life Sci,* 66(3), pp. 370-4.

Boston-Howes, W., Gibb, S. L., Williams, E. O., Pasinelli, P., Brown, R. H. and Trotti, D. (2006) 'Caspase-3 cleaves and inactivates the glutamate transporter EAAT2', *J Biol Chem,* 281(20), pp. 14076-84.

Bourke, S. C., Tomlinson, M., Williams, T. L., Bullock, R. E., Shaw, P. J. and Gibson, G. J. (2006) 'Effects of non-invasive ventilation on survival and quality of life in patients with amyotrophic lateral sclerosis: a randomised controlled trial', *Lancet Neurol,* 5(2), pp. 140-7.

Boyer, C., Teo, J., Phillips, P., Erlich, R. B., Sagnella, S., Sharbeen, G., Dwarte, T., Duong, H. T., Goldstein, D., Davis, T. P., Kavallaris, M. and McCarroll, J. (2015) 'Correction to "effective delivery of siRNA into cancer cells and tumors using well-defined biodegradable cationic star polymers"', *Mol Pharm,* 12(3), pp. 1004.

Braems, E., Tziortzouda, P. and Van Den Bosch, L. (2021) 'Exploring the alternative: Fish, flies and worms as preclinical models for ALS', *Neurosci Lett,* 759, pp. 136041.

Brettschneider, J., Van Deerlin, V. M., Robinson, J. L., Kwong, L., Lee, E. B., Ali, Y. O., Safren, N., Monteiro, M. J., Toledo, J. B., Elman, L., McCluskey, L., Irwin, D. J., Grossman, M., Molina-Porcel, L., Lee, V. M. and Trojanowski, J. Q. (2012) 'Pattern of ubiquilin pathology in ALS and FTLD indicates presence of C9ORF72 hexanucleotide expansion', *Acta Neuropathol,* 123(6), pp. 825-39.

Brooks, B. R. (1991) 'The role of axonal transport in neurodegenerative disease spread: a meta-analysis of experimental and clinical poliomyelitis compares with amyotrophic lateral sclerosis', *Can J Neurol Sci,* 18(3 Suppl), pp. 435-8.

Brooks, B. R., Miller, R. G., Swash, M., Munsat, T. L. and Diseases, W. F. o. N. R. G. o. M. N. (2000) 'El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis', *Amyotroph Lateral Scler Other Motor Neuron Disord,* 1(5), pp. 293-9.

Brown, A. (2022) 'FDA new drug approvals in Q3 2022', *Nat Rev Drug Discov,* 21(11), pp. 788.

Brown, R. H. and Al-Chalabi, A. (2017) 'Amyotrophic Lateral Sclerosis', *N Engl J Med,* 377(16), pp. 1602.

Bukau, B., Weissman, J. and Horwich, A. (2006) 'Molecular chaperones and protein quality control', *Cell,* 125(3), pp. 443-51.

Bulcha, J. T., Wang, Y., Ma, H., Tai, P. W. L. and Gao, G. (2021) 'Viral vector platforms within the gene therapy landscape', *Signal Transduct Target Ther,* 6(1), pp. 53.

Buratti, E. and Baralle, F. E. (2010) 'The multiple roles of TDP-43 in pre-mRNA processing and gene expression regulation', *RNA Biol,* 7(4), pp. 420-9.

Buratti, E., Brindisi, A., Giombi, M., Tisminetzky, S., Ayala, Y. M. and Baralle, F. E. (2005) 'TDP-43 binds heterogeneous nuclear ribonucleoprotein A/B through its C-terminal tail: an important region for the inhibition of cystic fibrosis transmembrane conductance regulator exon 9 splicing', *J Biol Chem,* 280(45), pp. 37572-84.

Buratti, E., De Conti, L., Stuani, C., Romano, M., Baralle, M. and Baralle, F. (2010) 'Nuclear factor TDP-43 can affect selected microRNA levels', *FEBS J,* 277(10), pp. 2268-81.

Burkhardt, M. F., Martinez, F. J., Wright, S., Ramos, C., Volfson, D., Mason, M., Garnes, J., Dang, V., Lievers, J., Shoukat-Mumtaz, U., Martinez, R., Gai, H., Blake, R., Vaisberg, E., Grskovic, M., Johnson, C., Irion, S., Bright, J., Cooper, B., Nguyen, L., Griswold-Prenner, I. and Javaherian, A. (2013) 'A cellular model for sporadic ALS using patient-derived induced pluripotent stem cells', *Mol Cell Neurosci,* 56, pp. 355-64.

Byrne, S., Heverin, M., Elamin, M., Walsh, C. and Hardiman, O. (2014) 'Intermediate repeat expansion length in C9orf72 may be pathological in amyotrophic lateral sclerosis', *Amyotroph Lateral Scler Frontotemporal Degener,* 15(1-2), pp. 148-50.

Cabezas de la Fuente, D. (2021) *Considerations for expression of fluorescent proteins and imaging in mammalian cells*. Focal Plane: Focal Plane. Available at: https://focalplane.biologists.com/2021/07/07/considerations-for-expression-of-fluorescent-proteins-and-imaging-in-mammalian-cells/#ref10.

Callaghan, B., Feldman, D., Gruis, K. and Feldman, E. (2011) 'The association of exposure to lead, mercury, and selenium and the development of amyotrophic lateral sclerosis and the epigenetic implications', *Neurodegener Dis,* 8(1-2), pp. 1-8.

Calvo, A., Canosa, A., Bertuzzo, D., Cugnasco, P., Solero, L., Clerico, M., De Mercanti, S., Bersano, E., Cammarosano, S., Ilardi, A., Manera, U., Moglia, C., Marinou, K., Bottacchi, E., Pisano, F., Mora, G., Mazzini, L. and Chiò, A. (2016) 'Influence of cigarette smoking on ALS outcome: a population-based study', *J Neurol Neurosurg Psychiatry,* 87(11), pp. 1229-1233.

Carmany-Rampey, A. and Moens, C. B. (2006) 'Modern mosaic analysis in the zebrafish', *Methods,* 39(3), pp. 228-38.

Cheah, B. C., Vucic, S., Krishnan, A. V. and Kiernan, M. C. (2010) 'Riluzole, neuroprotection and amyotrophic lateral sclerosis', *Curr Med Chem,* 17(18), pp. 1942-199.

Chen, C., Yamanaka, Y., Ueda, K., Li, P., Miyagi, T., Harada, Y., Tezuka, S., Narumi, S., Sugimoto, M., Kuroda, M., Hayamizu, Y. and Kanekura, K. (2021) 'Phase separation and toxicity of C9orf72 poly(PR) depends on alternate distribution of arginine', *J Cell Biol,* 220(11).

Chen, S., Sayana, P., Zhang, X. and Le, W. (2013) 'Genetics of amyotrophic lateral sclerosis: an update', *Mol Neurodegener*, pp. 28.

Chen, S., Zhang, X., Song, L. and Le, W. (2012) 'Autophagy dysregulation in amyotrophic lateral sclerosis', *Brain Pathol,* 22(1), pp. 110-6.

Chen, Y. Z., Bennett, C. L., Huynh, H. M., Blair, I. P., Puls, I., Irobi, J., Dierick, I., Abel, A., Kennerson, M. L., Rabin, B. A., Nicholson, G. A., Auer-Grumbach, M., Wagner, K., De Jonghe, P., Griffin, J. W., Fischbeck, K. H., Timmerman, V., Cornblath, D. R. and Chance, P. F. (2004) 'DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4)', *Am J Hum Genet,* 74(6), pp. 1128-35.

Chew, J., Cook, C., Gendron, T. F., Jansen-West, K., Del Rosso, G., Daughrity, L. M., Castanedes-Casey, M., Kurti, A., Stankowski, J. N., Disney, M. D., Rothstein, J. D., Dickson, D. W., Fryer, J. D., Zhang, Y. J. and Petrucelli, L. (2019) 'Aberrant deposition of stress granule-resident proteins linked to C9orf72-associated TDP-43 proteinopathy', *Mol Neurodegener,* 14(1), pp. 9.

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. (2015a) 'C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits', *Science,* 348(6239), pp. 1151-4.

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. (2015b) 'Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits', *Science,* 348(6239), pp. 1151-4.

Chia, R., Chiò, A. and Traynor, B. J. (2018) 'Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications', *Lancet Neurol,* 17(1), pp. 94-102.

Chio, 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(2), pp. 118-30.

Chiò, A., Calvo, A., Moglia, C., Mazzini, L., Mora, G. and group, P. s. (2011) 'Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based study', *J Neurol Neurosurg Psychiatry,* 82(7), pp. 740-6.

Chow, C. Y., Landers, J. E., Bergren, S. K., Sapp, P. C., Grant, A. E., Jones, J. M., Everett, L., Lenk, G. M., McKenna-Yasek, D. M., Weisman, L. S., Figlewicz, D., Brown, R. H. and Meisler, M. H. (2009) 'Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS', *Am J Hum Genet,* 84(1), pp. 85-8.

Chowdhury, G. M., Banasr, M., de Graaf, R. A., Rothman, D. L., Behar, K. L. and Sanacora, G. (2008) 'Chronic riluzole treatment increases glucose metabolism in rat prefrontal cortex and hippocampus', *J Cereb Blood Flow Metab,* 28(12), pp. 1892-7.

Chua, J. P., De Calbiac, H., Kabashi, E. and Barmada, S. J. (2022) 'Autophagy and ALS: mechanistic insights and therapeutic implications', *Autophagy,* 18(2), pp. 254-282.

Cirulli, E. T., Lasseigne, B. N., Petrovski, S., Sapp, P. C., Dion, P. A., Leblond, C. S., Couthouis, J., Lu, Y. F., Wang, Q., Krueger, B. J., Ren, Z., Keebler, J., Han, Y., Levy, S. E., Boone, B. E., Wimbish, J. R., Waite, L. L., Jones, A. L., Carulli, J. P., Day-Williams, A. G., Staropoli, J. F., Xin, W. W., Chesi, A., Raphael, A. R., McKenna-Yasek, D., Cady, J., Vianney de Jong, J. M., Kenna, K. P., Smith, B. N., Topp, S., Miller, J., Gkazi, A., Al-Chalabi, A., van den Berg, L. H., Veldink, J., Silani, V., Ticozzi,

N., Shaw, C. E., Baloh, R. H., Appel, S., Simpson, E., Lagier-Tourenne, C., Pulst, S. M., Gibson, S., Trojanowski, J. Q., Elman, L., McCluskey, L., Grossman, M., Shneider, N. A., Chung, W. K., Ravits, J. M., Glass, J. D., Sims, K. B., Van Deerlin, V. M., Maniatis, T., Hayes, S. D., Ordureau, A., Swarup, S., Landers, J., Baas, F., Allen, A. S., Bedlack, R. S., Harper, J. W., Gitler, A. D., Rouleau, G. A., Brown, R., Harms, M. B., Cooper, G. M., Harris, T., Myers, R. M. and Goldstein, D. B. (2015) 'Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways', *Science,* 347(6229), pp. 1436-41.

Cleveland, D. W., Laing, N., Hurse, P. V. and Brown, R. H. (1995) 'Toxic mutants in Charcot's sclerosis', *Nature,* 378(6555), pp. 342-3.

Cohen, T. J., Hwang, A. W., Restrepo, C. R., Yuan, C. X., Trojanowski, J. Q. and Lee, V. M. (2015) 'An acetylation switch controls TDP-43 function and aggregation propensity', *Nat Commun,* 6, pp. 5845.

Collaborators, G. M. N. D. (2018) 'Global, regional, and national burden of motor neuron diseases 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016', *Lancet Neurol,* 17(12), pp. 1083-1097.

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.

Cook, C. N., Wu, Y., Odeh, H. M., Gendron, T. F., Jansen-West, K., Del Rosso, G., Yue, M., Jiang, P., Gomes, E., Tong, J., Daughrity, L. M., Avendano, N. M., Castanedes-Casey, M., Shao, W., Oskarsson, B., Tomassy, G. S., McCampbell, A., Rigo, F., Dickson, D. W., Shorter, J., Zhang, Y. J. and Petrucelli, L. (2020) 'poly(GR) aggregation induces TDP-43 proteinopathy', *Sci Transl Med,* 12(559).

Cooper-Knock, J., Higginbottom, A., Stopford, M. J., Highley, J. R., Ince, P. G., Wharton, S. B., Pickering-Brown, S., Kirby, J., Hautbergue, G. M. and Shaw, P. J. (2015) 'Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy', *Acta Neuropathol,* 130(1), pp. 63-75.

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(Pt 7), pp. 2040-51.

Corboy, M. J., Thomas, P. J. and Wigley, W. C. (2005) 'Aggresome formation', *Methods Mol Biol,* 301, pp. 305-27.

Corcia, P., Beltran, S., Bakkouche, S. E. and Couratier, P. (2021) 'Therapeutic news in ALS', *Rev Neurol (Paris),* 177(5), pp. 544-549.

Corcia, P., Couratier, P., Blasco, H., Andres, C. R., Beltran, S., Meininger, V. and Vourc'h, P. (2017) 'Genetics of amyotrophic lateral sclerosis', *Revue Neurologique (Paris),* 173(5), pp. 254-262.

Costa, J., Swash, M. and de Carvalho, M. (2012) 'Awaji criteria for the diagnosis of amyotrophic lateral sclerosis:a systematic review', *Arch Neurol,* 69(11), pp. 1410-6.

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(8), pp. 1280-1303.

Cristofani, R., Crippa, V., Vezzoli, G., Rusmini, P., Galbiati, M., Cicardi, M. E., Meroni, M., Ferrari, V., Tedesco, B., Piccolella, M., Messi, E., Carra, S. and Poletti, A. (2018) 'The small heat shock protein B8 (HSPB8) efficiently removes aggregating species of dipeptides produced in C9ORF72-related neurodegenerative diseases', *Cell Stress Chaperones,* 23(1), pp. 1-12.

Cruz, M. P. (2018) 'Edaravone (Radicava): A Novel Neuroprotective Agent for the Treatment of Amyotrophic Lateral Sclerosis', *P T,* 43(1), pp. 25-28.

Cruz-Garcia, D., Ortega-Bellido, M., Scarpa, M., Villeneuve, J., Jovic, M., Porzner, M., Balla, T., Seufferlein, T. and Malhotra, V. (2013) 'Recruitment of arfaptins to the trans-Golgi network by PI(4)P and their involvement in cargo export', *EMBO J: Vol. 12*, pp. 1717-29.

Cui, Z., Gilda, J. E. and Gomes, A. V. (2014) 'Crude and purified proteasome activity assays are affected by type of microplate', *Anal Biochem,* 446, pp. 44-52.

Cunha-Oliveira, T., Montezinho, L., Mendes, C., Firuzi, O., Saso, L., Oliveira, P. J. and Silva, F. S. G. (2020) 'Oxidative Stress in Amyotrophic Lateral Sclerosis: Pathophysiology and Opportunities for Pharmacological Intervention', *Oxid Med Cell Longev,* 2020, pp. 5021694.

Cykowski, M. D., Dickson, D. W., Powell, S. Z., Arumanayagam, A. S., Rivera, A. L. and Appel, S. H. (2019) 'Dipeptide repeat (DPR) pathology in the skeletal muscle of ALS patients with C9ORF72 repeat expansion', *Acta Neuropathol,* 138(4), pp. 667-670.

D'Ambrosi, N., Rossi, S., Gerbino, V. and Cozzolino, M. (2014) 'Rac1 at the crossroad of actin dynamics and neuroinflammation in Amyotrophic Lateral Sclerosis', *Front Cell Neurosci,* 8, pp. 279.

D'Souza-Schorey, C., Boshans, R. L., McDonough, M., Stahl, P. D. and Van Aelst, L. (1997) 'A role for POR1, a Rac1-interacting protein, in ARF6-mediated cytoskeletal rearrangements', *Embo j,* 16(17), pp. 5445-54.

Darling, A. L., Breydo, L., Rivas, E. G., Gebru, N. T., Zheng, D., Baker, J. D., Blair, L. J., Dickey, C. A., Koren, J. and Uversky, V. N. (2019) 'Repeated repeat problems: Combinatorial effect of C9orf72-derived dipeptide repeat proteins', *Int J Biol Macromol,* 127, pp. 136-145.

Davis, E. (2014) 'Lentivirus: That's My MOI, and I'm Sticking To It'.

de Jong, S. W., Huisman, M. H. B., Sutedja, N. A., van der Kooi, A. J., de Visser, M., Schelhaas, H. J., Fischer, K., Veldink, J. H. and van den Berg, L. H. (2012) 'Smoking, Alcohol Consumption, and the Risk of Amyotrophic Lateral Sclerosis: A Population-based Study', *American Journal of Epidemiology,* 176(3), pp. 233-239.

Dedmon, M. M., Christodoulou, J., Wilson, M. R. and Dobson, C. M. (2005) 'Heat shock protein 70 inhibits alpha-synuclein fibril formation via preferential binding to prefibrillar species', *J Biol Chem,* 280(15), pp. 14733-40.

Deflorio, C., Onesti, E., Lauro, C., Tartaglia, G., Giovannelli, A., Limatola, C., Inghilleri, M. and Grassi, F. (2014) 'Partial block by riluzole of muscle sodium channels in myotubes from amyotrophic lateral sclerosis patients', *Neurol Res Int,* 2014, pp. 946073.

DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., Nicholson, A. M., Finch, N. A., Gilmer, H. F., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G. Y. R., Karydas, A., Seeley, W. W., Josephs, K. A., Coppola, G., Geschwind, D. H., Wszolek, Z. K., Feldman, H., Knopman, D., Petersen, R., Miller, B. L., Dickson, D., Boylan, K., Graff-Radford, N. and Rademakers, R. (2011) 'Expanded GGGGCC hexanucleotide repeat in non-coding region of C9ORF72 causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis', *Neuron,* 72(2), pp. 245-56.

Deng, H. X., Bigio, E. H., Zhai, H., Fecto, F., Ajroud, K., Shi, Y., Yan, J., Mishra, M., Ajroud-Driss, S., Heller, S., Sufit, R., Siddique, N., Mugnaini, E. and Siddique, T. (2011a) 'Differential Involvement of Optineurin in Amyotrophic Lateral Sclerosis With or Without SOD1 Mutations', *Arch Neurol,* 68(8), pp. 1057-61.

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. (2011b) 'Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia', *Nature,* 477(7363), pp. 211-5.

Deng, H. X., Zhai, H., Bigio, E. H., Yan, J., Fecto, F., Ajroud, K., Mishra, M., Ajroud-Driss, S., Heller, S., Sufit, R., Siddique, N., Mugnaini, E. and Siddique, T. (2010) 'FUS-immunoreactive inclusions are a common feature in sporadic and non-SOD1 familial amyotrophic lateral sclerosis', *Ann Neurol,* 67(6), pp. 739-48.

Dewan, R., Chia, R., Ding, J., Hickman, R. A., Stein, T. D., Abramzon, Y., Ahmed, S., Sabir, M. S., Portley, M. K., Tucci, A., Ibáñez, K., Shankaracharya, F. N. U., Keagle, P., Rossi, G., Caroppo, P., Tagliavini, F., Waldo, M. L., Johansson, P. M., Nilsson, C. F., Rowe, J. B., Benussi, L., Binetti, G., Ghidoni, R., Jabbari, E., Viollet, C., Glass, J. D., Singleton, A. B., Silani, V., Ross, O. A., Ryten, M., Torkamani, A., Tanaka, T., Ferrucci, L., Resnick, S. M., Pickering-Brown, S., Brady, C. B., Kowal, N., Hardy, J. A., Van Deerlin, V., Vonsattel, J. P., Harms, M. B., Morris, H. R., Ferrari, R., Landers, J. E., Chiò, A., Gibbs, J. R., Dalgard, C. L., Scholz, S. W., Traynor, B. J., (TAGC), A. G. C., Consortium, F. S., Consortium, G. E. R., (iAFGC), I. A. F. G. C., (IFGC), I. F. G. C., (iLBDGC), I. L. G. C., Consortium, N. A. and Consortium, P. (2021) 'Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis', *Neuron,* 109(3), pp. 448-460.e4.

Dewey, C. M., Cenik, B., Sephton, C. F., Johnson, B. A., Herz, J. and Yu, G. (2012) 'TDP-43 aggregation in neurodegeneration: are stress granules the key?', *Brain Res,* 1462, pp. 16-25.

Doble, A. (1996) 'The pharmacology and mechanism of action of riluzole', *Neurology,* 47(6 Suppl 4), pp. S233-41.

Doi, Y., Atsuta, N., Sobue, G., Morita, M. and Nakano, I. (2014) 'Prevalence and Incidence of Amyotrophic Lateral Sclerosis in Japan', *J Epidemiol*.

Dormann, D. and Haass, C. (2011) 'TDP-43 and FUS: a nuclear affair', *Trends Neurosci,* 34(7), pp. 339-48.

Dormann, D., Rodde, R., Edbauer, D., Bentmann, E., Fischer, I., Hruscha, A., Than, M. E., Mackenzie, I. R., Capell, A., Schmid, B., Neumann, M. and Haass, C. (2010) 'ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import', *Embo j,* 29(16), pp. 2841-57.

Dorst, J. and Genge, A. (2022) 'Clinical studies in amyotrophic lateral sclerosis', *Curr Opin Neurol,* 35(5), pp. 686-692.

Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., Amora, R., Hocking, T. D., Zhang, L., Rebar, E. J., Gregory, P. D., Urnov, F. D. and Amacher, S. L. (2008) 'Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases', *Nat Biotechnol,* 26(6), pp. 702-8.

Du, Z. W., Chen, H., Liu, H., Lu, J., Qian, K., Huang, C. L., Zhong, X., Fan, F. and Zhang, S. C. (2015) 'Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells', *Nat Commun,* 6, pp. 6626.

Dupuis, L., Pradat, P. F., Ludolph, A. C. and Loeffler, J. P. (2011) 'Energy metabolism in amyotrophic lateral sclerosis', *Lancet Neurol,* 10(1), pp. 75-82.

Edaravone (MCI-186) ALS 16 Study Group, T. (2017) 'A post-hoc subgroup analysis of outcomes in the first phase III clinical study of edaravone (MCI-186) in amyotrophic lateral sclerosis', *Amyotroph Lateral Scler Frontotemporal Degener,* 18(sup1), pp. 11-19.

Ederle, H., Funk, C., Abou-Ajram, C., Hutten, S., Funk, E. B. E., Kehlenbach, R. H., Bailer, S. M. and Dormann, D. (2018) 'Nuclear egress of TDP-43 and FUS occurs independently of Exportin-1/CRM1', *Sci Rep,* 8(1), pp. 7084.

Elden, A. C., Kim, H. J., Hart, M. P., Chen-Plotkin, A. S., Johnson, B. S., Fang, X., Armakola, M., Geser, F., Greene, R., Lu, M. M., Padmanabhan, A., Clay, D., McCluskey, L., Elman, L., Juhr, D., Gruber, P. J., Rüb, U., Auburger, G., Trojanowski, J. Q., Lee, V. M. Y., Van Deerlin, V. M., Bonini, N. M. and Gitler, A. D. (2010) 'Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS', *Nature,* 466(7310), pp. 1069-75.

Fang, T., Al Khleifat, A., Meurgey, J. H., Jones, A., Leigh, P. N., Bensimon, G. and Al-Chalabi, A. (2018) 'Stage at which riluzole treatment prolongs survival in patients with amyotrophic lateral sclerosis: a retrospective analysis of data from a dose-ranging study', *Lancet Neurol,* 17(5), pp. 416-422.

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(13), pp. 3579-3595.

Fecto, F., Yan, J., Vemula, S. P., Liu, E., Yang, Y., Chen, W., Zheng, J. G., Shi, Y., Siddique, N., Arrat, H., Donkervoort, S., Ajroud-Driss, S., Sufit, R. L., Heller, S. L., Deng, H. X. and Siddique, T. (2011) 'SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis', *Arch Neurol,* 68(11), pp. 1440-6.

Feldman, E. L., Goutman, S. A., Petri, S., Mazzini, L., Savelieff, M. G., Shaw, P. J. and Sobue, G. (2022) 'Amyotrophic lateral sclerosis', *The Lancet*.

Ferraiuolo, L., Kirby, J., Grierson, A. J., Sendtner, M. and Shaw, P. J. (2011) 'Molecular pathways of motor neuron injury in amyotrophic lateral sclerosis', *Nature Reviews Neurology,* 7(11), pp. 616.

Feuillette, S., Delarue, M., Riou, G., Gaffuri, A. L., Wu, J., Lenkei, Z., Boyer, O., Frébourg, T., Campion, D. and Lecourtois, M. (2017) 'Neuron-to-Neuron Transfer of FUS in Drosophila Primary Neuronal Culture Is Enhanced by ALS-Associated Mutations', *J Mol Neurosci,* 62(1), pp. 114-122.

Finley, D. (2009) 'Recognition and processing of ubiquitin-protein conjugates by the proteasome', *Annu Rev Biochem,* 78, pp. 477-513.

Fiorito, V., Chiabrando, D. and Tolosano, E. (2018) 'Mitochondrial Targeting in Neurodegeneration: A Heme Perspective', *Pharmaceuticals (Basel),* 11(3).

Foerster, B. R., Pomper, M. G., Callaghan, B. C., Petrou, M., Edden, R. A., Mohamed, M. A., Welsh, R. C., Carlos, R. C., Barker, P. B. and Feldman, E. L. (2013) 'An imbalance between excitatory and inhibitory neurotransmitters in amyotrophic lateral sclerosis revealed by use of 3-T proton magnetic resonance spectroscopy', *JAMA Neurol,* 70(8), pp. 1009-16.

Fortier, G., Butti, Z. and Patten, S. A. (2020) 'Modelling C9orf72-Related Amyotrophic Lateral Sclerosis in Zebrafish', *Biomedicines,* 8(10).

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

Fredi, M., Cavazzana, I., Biasiotto, G., Filosto, M., Padovani, A., Monti, E., Tincani, A., Franceschini, F. and Zanella, I. (2019) 'C9orf72 Intermediate Alleles in Patients with Amyotrophic Lateral Sclerosis, Systemic Lupus Erythematosus, and Rheumatoid Arthritis', *Neuromolecular Med,* 21(2), pp. 150-159.

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(7567), pp. 129-33.

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Muller, K., Marroquin, N., Nordin, F., Hubers, A., Weydt, P., Pinto, S., Press, R., Millecamps, S., Molko, N., Bernard, E., Desnuelle, C., Soriani, M. H., Dorst, J., Graf, E., Nordstrom, U., Feiler, M. S., Putz, S., Boeckers, T. M., Meyer, T., Winkler, A. S., Winkelman, J., de Carvalho, M., Thal, D. R., Otto, M., Brannstrom, 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', *Nat Neurosci,* 18(5), pp. 631-6.

Frottin, F., Schueder, F., Tiwary, S., Gupta, R., Körner, R., Schlichthaerle, T., Cox, J., Jungmann, R., Hartl, F. U. and Hipp, M. S. (2019) 'The nucleolus functions as a phase-separated protein quality control compartment', *Science,* 365(6451), pp. 342-347.

Fus-Kujawa, A., Prus, P., Bajdak-Rusinek, K., Teper, P., Gawron, K., Kowalczuk, A. and Sieron, A. L. (2021a) 'An Overview of Methods and Tools for Transfection of Eukaryotic Cells', *Front Bioeng Biotechnol,* 9, pp. 701031.

Fus-Kujawa, A., Teper, P., Botor, M., Klarzyńska, K., Sieroń, Ł., Verbelen, B., Smet, M., Sieroń, A. L., Mendrek, B. and Kowalczuk, A. (2021b) 'Functional star polymers as reagents for efficient

nucleic acids delivery into HT-1080 cells', *International Journal of Polymeric Materials and Polymeric Biomaterials,* 70(5), pp. 356-370.

Fávero, F. M., Voos, M. C., Castro, I., Caromano, F. A. and Oliveira, A. S. B. (2017) 'Epidemiological and clinical factors impact on the benefit of riluzole in the survival rates of patients with ALS', *Arq Neuropsiquiatr,* 75(8), pp. 515-522.

Gatto, N., Dos Santos Souza, C., Shaw, A. C., Bell, S. M., Myszczynska, M. A., Powers, S., Meyer, K., Castelli, L. M., Karyka, E., Mortiboys, H., Azzouz, M., Hautbergue, G. M., Márkus, N. M., Shaw, P. J. and Ferraiuolo, L. (2021) 'Directly converted astrocytes retain the ageing features of the donor fibroblasts and elucidate the astrocytic contribution to human CNS health and disease', *Aging Cell,* 20(1), pp. e13281.

Geevasinga, N., Menon, P., Ng, K., Van Den Bos, M., Byth, K., Kiernan, M. C. and Vucic, S. (2016) 'Riluzole exerts transient modulating effects on cortical and axonal hyperexcitability in ALS', *Amyotroph Lateral Scler Frontotemporal Degener,* 17(7-8), pp. 580-588.

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(6), pp. 829-44.

Gendron, T. F. and Petrucelli, L. (2018) 'Disease mechanisms of C9ORF72 repeat expansions', *Cold Spring Harb Perspect Med,* 8(4).

Giacomelli, E., Vahsen, B. F., Calder, E. L., Xu, Y., Scaber, J., Gray, E., Dafinca, R., Talbot, K. and Studer, L. (2022) 'Human stem cell models of neurodegeneration: From basic science of amyotrophic lateral sclerosis to clinical translation', *Cell Stem Cell,* 29(1), pp. 11-35.

Glineburg, M. R., Zhang, Y., Krans, A., Tank, E. M., Barmada, S. J. and Todd, P. K. (2021) 'Enhanced detection of expanded repeat mRNA foci with hybridization chain reaction', *Acta Neuropathol Commun,* 9(1), pp. 73.

Gois, A. M., Mendonça, D. M. F., Freire, M. A. M. and Santos, J. R. (2020) 'IN VITRO AND IN VIVO MODELS OF AMYOTROPHIC LATERAL SCLEROSIS: AN UPDATED OVERVIEW', *Brain Res Bull,* 159, pp. 32-43.

Grabher, C., Joly, J. S. and Wittbrodt, J. (2004) 'Highly efficient zebrafish transgenesis mediated by the meganuclease I-SceI', *Methods Cell Biol,* 77, pp. 381-401.

Grad, L. I., Guest, W. C., Yanai, A., Pokrishevsky, E., O'Neill, M. A., Gibbs, E., Semenchenko, V., Yousefi, M., Wishart, D. S., Plotkin, S. S. and Cashman, N. R. (2011) 'Intermolecular transmission of

superoxide dismutase 1 misfolding in living cells', *Proc Natl Acad Sci U S A,* 108(39), pp. 16398-403.

Graham, A. J., Macdonald, A. M. and Hawkes, C. H. (1997) 'British motor neuron disease twin study', *J Neurol Neurosurg Psychiatry,* 62(6), pp. 562-9.

Green, M. R. and Sambrook, J. (2018) 'Touchdown Polymerase Chain Reaction (PCR)', *Cold Spring Harb Protoc,* 2018(5).

Greenway, M. J., Andersen, P. M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., Patterson, V., Swingler, R., Kieran, D., Prehn, J., Morrison, K. E., Green, A., Acharya, K. R., Brown, R. H. and Hardiman, O. (2006) 'ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis', *Nature Genetics,* 38(4), pp. 411.

Gromicho, M., Oliveira Santos, M., Pinto, A., Pronto-Laborinho, A. and De Carvalho, M. (2017) 'Young-onset rapidly progressive ALS associated with heterozygous FUS mutation', *Amyotroph Lateral Scler Frontotemporal Degener,* 18(5-6), pp. 451-453.

Gros-Louis, F., Gaspar, C. and Rouleau, G. A. (2006) 'Genetics of familial and sporadic amyotrophic lateral sclerosis', *Biochim Biophys Acta,* 1762(11-12), pp. 956-972.

Group, W. and Group, E. M.-A. S. (2017) 'Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial', *Lancet Neurol,* 16(7), pp. 505-512.

Guerrero, E. N., Mitra, J., Wang, H., Rangaswamy, S., Hegde, P. M., Basu, P., Rao, K. S. and Hegde, M. L. (2019) 'Amyotrophic lateral sclerosis-associated TDP-43 mutation Q331K prevents nuclear translocation of XRCC4-DNA ligase 4 complex and is linked to genome damage-mediated neuronal apoptosis', *Hum Mol Genet,* 28(15), pp. 2459-2476.

Guo, H., Lai, L., Butchbach, M. E., Stockinger, M. P., Shan, X., Bishop, G. A. and Lin, C. L. (2003) 'Increased expression of the glial glutamate transporter EAAT2 modulates excitotoxicity and delays the onset but not the outcome of ALS in mice', *Hum Mol Genet,* 12(19), pp. 2519-32.

Gurney, M. E., Pu, H., Chiu, A. Y., Dal Canto, M. C., Polchow, C. Y., Alexander, D. D., Caliendo, J., Hentati, A., Kwon, Y. W. and Deng, H. X. (1994) 'Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation', *Science,* 264(5166), pp. 1772-5.

Hanby, M. F., Scott, K. M., Scotton, W., Wijesekera, L., Mole, T., Ellis, C. E., Nigel Leigh, P., Shaw, C. E. and Al-Chalabi, A. (2011) 'The risk to relatives of patients with sporadic amyotrophic lateral sclerosis', *Brain: A Journal of Neurology,* 134(12), pp. 3451-3454.

Hannaford, A., Pavey, N., van den Bos, M., Geevasinga, N., Menon, P., Shefner, J. M., Kiernan, M. C. and Vucic, S. (2021) 'Diagnostic Utility of Gold Coast Criteria in Amyotrophic Lateral Sclerosis', *Ann Neurol,* 89(5), pp. 979-986.

Hannon, G. J. and Rossi, J. J. (2004) 'Unlocking the potential of the human genome with RNA interference', *Nature,* 431(7006), pp. 371-8.

Hardiman, O., Al-Chalabi, A., Chio, A., Corr, E. M., Logroscino, G., Robberecht, W., Shaw, P. J., Simmons, Z. and Berg, L. H. v. d. (2017) 'Amyotrophic lateral sclerosis', *Nature Reviews Disease Primers,* 3, pp. 17071.

Hayashi, Y., Homma, K. and Ichijo, H. (2016) 'SOD1 in neurotoxicity and its controversial roles in SOD1 mutation-negative ALS', *Adv Biol Regul,* 60, pp. 95-104.

He, F., Flores, B. N., Krans, A., Frazer, M., Natla, S., Niraula, S., Adefioye, O., Barmada, S. J. and Todd, P. K. (2020) 'The carboxyl termini of RAN translated GGGGCC nucleotide repeat expansions modulate toxicity in models of ALS/FTD', *Acta Neuropathol Commun,* 8(1), pp. 122.

He, L., Liang, J., Chen, C., Chen, J., Shen, Y., Sun, S. and Li, L. (2022) 'C9orf72 functions in the nucleus to regulate DNA damage repair', *Cell Death Differ*.

Heinemeyer, W., Fischer, M., Krimmer, T., Stachon, U. and Wolf, D. H. (1997) 'The active sites of the eukaryotic 20 S proteasome and their involvement in subunit precursor processing', *J Biol Chem,* 272(40), pp. 25200-9.

Hennig, S., Kong, G., Mannen, T., Sadowska, A., Kobelke, S., Blythe, A., Knott, G. J., Iyer, K. S., Ho, D., Newcombe, E. A., Hosoki, K., Goshima, N., Kawaguchi, T., Hatters, D., Trinkle-Mulcahy, L., Hirose, T., Bond, C. S. and Fox, A. H. (2015) 'Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles', *J Cell Biol,* 210(4), pp. 529-39.

Herranz-Martin, S., Chandran, J., Lewis, K., Mulcahy, P., Higginbottom, A., Walker, C., Valenzuela, I. M. Y., Jones, R. A., Coldicott, I., Iannitti, T., Akaaboune, M., El-Khamisy, S. F., Gillingwater, T. H., Shaw, P. J. and Azzouz, M. (2017) 'Viral delivery of C9orf72 hexanucleotide repeat expansions in mice leads to repeat-length-dependent neuropathology and behavioural deficits', *Dis Model Mech,* 10(7), pp. 859-868.

Hewitt, C., Kirby, J., Highley, J. R., Hartley, J. A., Hibberd, R., Hollinger, H. C., Williams, T. L., Ince, P. G., McDermott, C. J. and Shaw, P. J. (2010) 'Novel FUS/TLS mutations and pathology in familial and sporadic amyotrophic lateral sclerosis', *Arch Neurol,* 67(4), pp. 455-61.

Higashijima, S. (2008) 'Transgenic zebrafish expressing fluorescent proteins in central nervous system neurons', *Dev Growth Differ,* 50(6), pp. 407-13.

Higashijima, S., Okamoto, H., Ueno, N., Hotta, Y. and Eguchi, G. (1997) 'High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin', *Dev Biol,* 192(2), pp. 289-99.

Higgins, C. M., Jung, C. and Xu, Z. (2003) 'ALS-associated mutant SOD1G93A causes mitochondrial vacuolation by expansion of the intermembrane space and by involvement of SOD1 aggregation and peroxisomes', *BMC Neurosci,* 4, pp. 16.

Highley, J. R., Kirby, J., Jansweijer, J. A., Webb, P. S., Hewamadduma, C. A., Heath, P. R., Higginbottom, A., Raman, R., Ferraiuolo, L., Cooper-Knock, J., McDermott, C. J., Wharton, S. B., Shaw, P. J. and Ince, P. G. (2014) 'Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones', *Neuropathol Appl Neurobiol,* 40(6), pp. 670-85.

Hor, J. H., Santosa, M. M., Lim, V. J. W., Ho, B. X., Taylor, A., Khong, Z. J., Ravits, J., Fan, Y., Liou, Y. C., Soh, B. S. and Ng, S. Y. (2021) 'ALS motor neurons exhibit hallmark metabolic defects that are rescued by SIRT3 activation', *Cell Death Differ,* 28(4), pp. 1379-1397.

Howe, K. and Clark, M. D. and Torroja, C. F. and Torrance, J. and Berthelot, C. and Muffato, M. and Collins, J. E. and Humphray, S. and McLaren, K. and Matthews, L. and McLaren, S. and Sealy, I. and Caccamo, M. and Churcher, C. and Scott, C. and Barrett, J. C. and Koch, R. and Rauch, G. J. and White, S. and Chow, W. and Kilian, B. and Quintais, L. T. and Guerra-Assunção, J. A. and Zhou, Y. and Gu, Y. and Yen, J. and Vogel, J. H. and Eyre, T. and Redmond, S. and Banerjee, R. and Chi, J. and Fu, B. and Langley, E. and Maguire, S. F. and Laird, G. K. and Lloyd, D. and Kenyon, E. and Donaldson, S. and Sehra, H. and Almeida-King, J. and Loveland, J. and Trevanion, S. and Jones, M. and Quail, M. and Willey, D. and Hunt, A. and Burton, J. and Sims, S. and McLay, K. and Plumb, B. and Davis, J. and Clee, C. and Oliver, K. and Clark, R. and Riddle, C. and Elliot, D. and Eliott, D. and Threadgold, G. and Harden, G. and Ware, D. and Begum, S. and Mortimore, B. and Mortimer, B. and Kerry, G. and Heath, P. and Phillimore, B. and Tracey, A. and Corby, N. and Dunn, M. and Johnson, C. and Wood, J. and Clark, S. and Pelan, S. and Griffiths, G. and Smith, M. and Glithero, R. and Howden, P. and Barker, N. and Lloyd, C. and Stevens, C. and Harley, J. and Holt, K. and Panagiotidis, G. and Lovell, J. and Beasley, H. and Henderson, C. and Gordon, D. and Auger, K. and Wright, D. and Collins, J. and Raisen, C. and Dyer, L. and Leung, K. and Robertson, L. and Ambridge, K. and Leongamornlert, D. and McGuire, S. and Gilderthorp, R. and Griffiths, C. and Manthravadi, D. and Nichol, S. and Barker, G. and Whitehead, S. and Kay, M. and Brown, J. and Murnane, C. and Gray, E. and Humphries, M. and Sycamore, N. and Barker, D. and Saunders, D. and Wallis, J. and Babbage, A. and Hammond, S. and Mashreghi-Mohammadi, M. and Barr, L. and Martin, S. and Wray, P. and Ellington, A. and Matthews, N. and Ellwood, M. and Woodmansey, R. and Clark, G. and Cooper, J. and Tromans, A. and Grafham, D. and Skuce, C. and Pandian, R. and Andrews, R. and Harrison, E. and Kimberley, A. and Garnett, J. and Fosker, N. and Hall, R. and Garner, P.

and Kelly, D. and Bird, C. and Palmer, S. and Gehring, I. and Berger, A. and Dooley, C. M. and Ersan-Ürün, Z. and Eser, C. and Geiger, H. and Geisler, M. and Karotki, L. and Kirn, A. and Konantz, J. and Konantz, M. and Oberländer, M. and Rudolph-Geiger, S. and Teucke, M. and Lanz, C. and Raddatz, G. and Osoegawa, K. and Zhu, B. and Rapp, A. and Widaa, S. and Langford, C. and Yang, F. and Schuster, S. C. and Carter, N. P. and Harrow, J. and Ning, Z. and Herrero, J. and Searle, S. M. and Enright, A. and Geisler, R. and Plasterk, R. H. and Lee, C. and Westerfield, M. and de Jong, P. J. and Zon, L. I. and Postlethwait, J. H. and Nüsslein-Volhard, C. and Hubbard, T. J. and Roest Crollius, H. and Rogers, J. and Stemple, D. L. (2013) 'The zebrafish reference genome sequence and its relationship to the human genome', *Nature,* 496(7446), pp. 498-503.

Human Protein Atlas Available at: proteinatlas.org.

Iacoangeli, A., Al Khleifat, A., Jones, A. R., Sproviero, W., Shatunov, A., Opie-Martin, S., Morrison, K. E., Shaw, P. J., Shaw, C. E., Fogh, I., Dobson, R. J., Newhouse, S. J., Al-Chalabi, A. and Initiative, A. s. D. N. (2019) 'C9orf72 intermediate expansions of 24-30 repeats are associated with ALS', *Acta Neuropathol Commun,* 7(1), pp. 115.

Igaz, L. M., Kwong, L. K., Chen-Plotkin, A., Winton, M. J., Unger, T. L., Xu, Y., Neumann, M., Trojanowski, J. Q. and Lee, V. M. (2009) 'Expression of TDP-43 C-terminal Fragments in Vitro Recapitulates Pathological Features of TDP-43 Proteinopathies', *J Biol Chem,* 284(13), pp. 8516-24.

Inukai, Y., Nonaka, T., Arai, T., Yoshida, M., Hashizume, Y., Beach, T. G., Buratti, E., Baralle, F. E., Akiyama, H., Hisanaga, S. and Hasegawa, M. (2008) 'Abnormal phosphorylation of Ser409/410 of TDP-43 in FTLD-U and ALS', *FEBS Lett,* 582(19), pp. 2899-904.

Ito, D. and Suzuki, N. (2011) 'Conjoint pathologic cascades mediated by ALS/FTLD-U linked RNA-binding proteins TDP-43 and FUS', *Neurology,* 77(17), pp. 1636-43.

Ito, H., Wate, R., Zhang, J., Ohnishi, S., Kaneko, S., Nakano, S. and Kusaka, H. (2008) 'Treatment with edaravone, initiated at symptom onset, slows motor decline and decreases SOD1 deposition in ALS mice', *Exp Neurol,* 213(2), pp. 448-55.

Jaiswal, M. K. (2016) 'Riluzole But Not Melatonin Ameliorates Acute Motor Neuron Degeneration and Moderately Inhibits SOD1-Mediated Excitotoxicity Induced Disrupted Mitochondrial Ca', *Front Cell Neurosci,* 10, pp. 295.

Jakobsson, J., Ericson, C., Jansson, M., Björk, E. and Lundberg, C. (2003) 'Targeted transgene expression in rat brain using lentiviral vectors', *J Neurosci Res,* 73(6), pp. 876-85.

Jana, N. R., Zemskov, E. A., Wang, G. and Nukina, N. (2001) 'Altered proteasomal function due to the expression of polyglutamine-expanded truncated N-terminal huntingtin induces apoptosis by

caspase activation through mitochondrial cytochrome c release', *Hum Mol Genet,* 10(10), pp. 1049-59.

Johnson, B. S., McCaffery, J. M., Lindquist, S. and Gitler, A. D. (2008) 'A yeast TDP-43 proteinopathy model: Exploring the molecular determinants of TDP-43 aggregation and cellular toxicity', *Proc Natl Acad Sci U S A,* 105(17), pp. 6439-44.

Johnson, B. S., Snead, D., Lee, J. J., McCaffery, J. M., Shorter, J. and Gitler, A. D. (2009) 'TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity', *J Biol Chem,* 284(30), pp. 20329-39.

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., Monsurro, M. R., Battistini, S., Salvi, F., Spataro, R., Sola, P., Borghero, G., Galassi, G., Scholz, S. W., Taylor, J. P., Restagno, G., Chio, A. and Traynor, B. J. (2010) 'Exome sequencing reveals VCP mutations as a cause of familial ALS', *Neuron,* 68(5), pp. 857-64.

Johnson, J. O., Pioro, E. P., Boehringer, A., Chia, R., Feit, H., Renton, A. E., Pliner, H. A., Abramzon, Y., Marangi, G., Winborn, B. J., Gibbs, J. R., Nalls, M. A., Morgan, S., Shoai, M., Hardy, J., Pittman, A., Orrell, R. W., Malaspina, A., Sidle, K. C., Fratta, P., Harms, M. B., Baloh, R. H., Pestronk, A., Weihl, C. C., Rogaeva, E., Zinman, L., Drory, V. E., Borghero, G., Mora, G., Calvo, A., Rothstein, J. D., Drepper, C., Sendtner, M., Singleton, A. B., Taylor, J. P., Cookson, M. R., Restagno, G., Sabatelli, M., Bowser, R., Chiò, A. and Traynor, B. J. (2014a) 'Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis', *Nat Neurosci,* 17(5), pp. 664-6.

Johnson, J. O., Pioro, E. P., Boehringer, A., Chia, R., Feit, H., Renton, A. E., Pliner, H. A., Abramzon, Y., Marangi, G., Winborn, B. J., Gibbs, J. R., Nalls, M. A., Morgan, S., Shoai, M., Hardy, J., Pittman, A., Orrell, R. W., Malaspina, A., Sidle, K. C., Fratta, P., Harms, M. B., Baloh, R. H., Pestronk, A., Weihl, C. C., Rogaeva, E., Zinman, L., Drory, V. E., Borghero, G., Mora, G., Calvo, A., Rothstein, J. D., Drepper, C., Sendtner, M., Singleton, A. B., Taylor, J. P., Cookson, M. R., Restagno, G., Sabatelli, M., Bowser, R., Chiò, A., Traynor, B. J. and ITALSGEN (2014b) 'Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis', *Nat Neurosci,* 17(5), pp. 664-666.

Johnston, C. A., Stanton, B. R., Turner, M. R., Gray, R., Blunt, A. H., Butt, D., Ampong, M. A., Shaw, C. E., Leigh, P. N. and Al-Chalabi, A. (2006) 'Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London', *J Neurol,* 253(12), pp. 1642-3.

Johnston, J. A., Ward, C. L. and Kopito, R. R. (1998) 'Aggresomes: a cellular response to misfolded proteins', *J Cell Biol,* 143(7), pp. 1883-98.

JR., B. (1964) 'The respiratory metabolism and swimming performance of young sockeye salmon', *J Fish Res Bd Can,* 21, pp. 1183–1226.

Judith, D., Jefferies, H. B. J., Boeing, S., Frith, D., Snijders, A. P. and Tooze, S. A. (2019) 'ATG9A shapes the forming autophagosome through Arfaptin 2 and phosphatidylinositol 4-kinase IIIβ', *J Cell Biol,* 218(5), pp. 1634-1652.

Kabashi, E., Lin, L., Tradewell, M. L., Dion, P. A., Bercier, V., Bourgouin, P., Rochefort, D., Bel Hadj, S., Durham, H. D., Vande Velde, C., Rouleau, G. A. and Drapeau, P. (2010) 'Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo', *Hum Mol Genet,* 19(4), pp. 671-83.

Kaneb, H. M., Folkmann, A. W., Belzil, V. V., Jao, L. E., Leblond, C. S., Girard, S. L., Daoud, H., Noreau, A., Rochefort, D., Hince, P., Szuto, A., Levert, A., Vidal, S., André-Guimont, C., Camu, W., Bouchard, J. P., Dupré, N., Rouleau, G. A., Wente, S. R. and Dion, P. A. (2015) 'Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis', *Hum Mol Genet: Vol. 5*, pp. 1363-73.

Kanoh, H., Williger, B. T. and Exton, J. H. (1997) 'Arfaptin 1, a putative cytosolic target protein of ADP-ribosylation factor, is recruited to Golgi membranes', *J Biol Chem,* 272(9), pp. 5421-9.

Karbowski, M. and Neutzner, A. (2012) 'Neurodegeneration as a consequence of failed mitochondrial maintenance', *Acta Neuropathol,* 123(2), pp. 157-71.

Karlsson, M., Zhang, C., Méar, L., Zhong, W., Digre, A., Katona, B., Sjöstedt, E., Butler, L., Odeberg, J., Dusart, P., Edfors, F., Oksvold, P., von Feilitzen, K., Zwahlen, M., Arif, M., Altay, O., Li, X., Ozcan, M., Mardinoglu, A., Fagerberg, L., Mulder, J., Luo, Y., Ponten, F., Uhlén, M. and Lindskog, C. (2021) 'A single-cell type transcriptomics map of human tissues', *Sci Adv,* 7(31).

Karra, D. and Dahm, R. (2010) 'Transfection techniques for neuronal cells', *J Neurosci,* 30(18), pp. 6171-7.

Kelekar, A. (2005) 'Autophagy', *Ann N Y Acad Sci,* 1066, pp. 259-71.

Khosravi, B., LaClair, K. D., Riemenschneider, H., Zhou, Q., Frottin, F., Mareljic, N., Czuppa, M., Farny, D., Hartmann, H., Michaelsen, M., Arzberger, T., Hartl, F. U., Hipp, M. S. and Edbauer, D. (2020) 'Cell-to-cell transmission of C9orf72 poly-(Gly-Ala) triggers key features of ALS/FTD', *EMBO J,* 39(8), pp. e102811.

Kim, D. and Rossi, J. (2008) 'RNAi mechanisms and applications', *Biotechniques,* 44(5), pp. 613-6.

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(7442), pp. 467-73.

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) 'Prion-like domain mutations in hnRNPs cause multisystem proteinopathy and ALS', *Nature,* 495(7442), pp. 467-73.

Kim, T. K. and Eberwine, J. H. (2010) 'Mammalian cell transfection: the present and the future', *Anal Bioanal Chem,* 397(8), pp. 3173-8.

King, A. E., Woodhouse, A., Kirkcaldie, M. T. and Vickers, J. C. (2016) 'Excitotoxicity in ALS: Overstimulation, or overreaction?', *Exp Neurol,* 275 Pt 1, pp. 162-71.

Kino, Y., Washizu, C., Kurosawa, M., Yamada, M., Miyazaki, H., Akagi, T., Hashikawa, T., Doi, H., Takumi, T., Hicks, G. G., Hattori, N., Shimogori, T. and Nukina, N. (2015) 'FUS/TLS deficiency causes behavioral and pathological abnormalities distinct from amyotrophic lateral sclerosis', *Acta Neuropathol Commun,* 3, pp. 24.

Kiskinis, E., Sandoe, J., Williams, L. A., Boulting, G. L., Moccia, R., Wainger, B. J., Han, S., Peng, T., Thams, S., Mikkilineni, S., Mellin, C., Merkle, F. T., Davis-Dusenbery, B. N., Ziller, M., Oakley, D., Ichida, J., Di Costanzo, S., Atwater, N., Maeder, M. L., Goodwin, M. J., Nemesh, J., Handsaker, R. E., Paull, D., Noggle, S., McCarroll, S. A., Joung, J. K., Woolf, C. J., Brown, R. H. and Eggan, K. (2014) 'Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1', *Cell Stem Cell,* 14(6), pp. 781-95.

Klim, J. R., Williams, L. A., Limone, F., Guerra San Juan, I., Davis-Dusenbery, B. N., Mordes, D. A., Burberry, A., Steinbaugh, M. J., Gamage, K. K., Kirchner, R., Moccia, R., Cassel, S. H., Chen, K., Wainger, B. J., Woolf, C. J. and Eggan, K. (2019) 'ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair', *Nat Neurosci,* 22(2), pp. 167-179.

Kraemer, B. C., Schuck, T., Wheeler, J. M., Robinson, L. C., Trojanowski, J. Q., Lee, V. M. and Schellenberg, G. D. (2010) 'Loss of murine TDP-43 disrupts motor function and plays an essential role in embryogenesis', *Acta Neuropathol,* 119(4), pp. 409-19.

Kudla, G., Lipinski, L., Caffin, F., Helwak, A. and Zylicz, M. (2006) 'High guanine and cytosine content increases mRNA levels in mammalian cells', *PLoS Biol,* 4(6), pp. e180.

Kwiatkowski, T. J., Jr., Bosco, D. A., Leclerc, A. L., Tamrazian, E., Vanderburg, C. R., Russ, C., Davis, A., Gilchrist, J., Kasarskis, E. J., Munsat, T., Valdmanis, P., Rouleau, G. A., Hosler, B. A., Cortelli, P., de Jong, P. J., Yoshinaga, Y., Haines, J. L., Pericak-Vance, M. A., Yan, J., Ticozzi, N., Siddique, T., McKenna-Yasek, D., Sapp, P. C., Horvitz, H. R., Landers, J. E. and Brown, R. H., Jr. (2009) 'Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis', *Science,* 323(5918), pp. 1205-8.

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(6201), pp. 1139-45.

Lacomblez, L., Bensimon, G., Leigh, P. N., Guillet, P. and Meininger, V. (1996) 'Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II', *Lancet,* 347(9013), pp. 1425-31.

Lagier-Tourenne, C., Polymenidou, M. and Cleveland, D. W. (2010) 'TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration', *Hum Mol Genet,* 19(R1), pp. R46-64.

Lattante, S., Rouleau, G. A. and Kabashi, E. (2013) 'TARDBP and FUS mutations associated with amyotrophic lateral sclerosis: summary and update', *Hum Mutat,* 34(6), pp. 812-26.

Leblond, C. S., Kaneb, H. M., Dion, P. A. and Rouleau, G. A. (2014) 'Dissection of genetic factors associated with amyotrophic lateral sclerosis', *Exp Neurol,* 262 Pt B, pp. 91-101.

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. (2016) 'C9orf72 Dipeptide Repeats Impair the Assembly, Dynamics, and Function of Membrane-Less Organelles', *Cell,* 167(3), pp. 774-788.e17.

Lee, S. M., Asress, S., Hales, C. M., Gearing, M., Vizcarra, J. C., Fournier, C. N., Gutman, D. A., Chin, L. S., Li, L. and Glass, J. D. (2019) 'TDP-43 cytoplasmic inclusion formation is disrupted in C9orf72-associated amyotrophic lateral sclerosis/frontotemporal lobar degeneration', *Brain Commun,* 1(1), pp. fcz014.

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(5), pp. 1178-86.

Lemmens, R., Van Hoecke, A., Hersmus, N., Geelen, V., D'Hollander, I., Thijs, V., Van Den Bosch, L., Carmeliet, P. and Robberecht, W. (2007) 'Overexpression of mutant superoxide dismutase 1 causes a motor axonopathy in the zebrafish', *Hum Mol Genet,* 16(19), pp. 2359-65.

Li, H. Y., Yeh, P. A., Chiu, H. C., Tang, C. Y. and Tu, B. P. (2011) 'Hyperphosphorylation as a defense mechanism to reduce TDP-43 aggregation', *PLoS One,* 6(8), pp. e23075.

Li, M., Husic, N., Lin, Y., Christensen, H., Malik, I., McIver, S., LaPash Daniels, C. M., Harris, D. A., Kotzbauer, P. T., Goldberg, M. P. and Snider, B. J. (2010) 'Optimal promoter usage for lentiviral vector-mediated transduction of cultured central nervous system cells', *J Neurosci Methods,* 189(1), pp. 56-64.

Liachko, N. F., Guthrie, C. R. and Kraemer, B. C. (2010) 'Phosphorylation promotes neurotoxicity in a Caenorhabditis elegans model of TDP-43 proteinopathy', *J Neurosci,* 30(48), pp. 16208-19.

Lieschke, G. J. and Currie, P. D. (2007) 'Animal models of human disease: zebrafish swim into view', *Nat Rev Genet,* 8(5), pp. 353-67.

Lim, S. T., Airavaara, M. and Harvey, B. K. (2010) 'Viral vectors for neurotrophic factor delivery: a gene therapy approach for neurodegenerative diseases of the CNS', *Pharmacol Res,* 61(1), pp. 14-26.

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(3), pp. 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(3), pp. 789-802.e12.

Linden-Junior, E., Becker, J., Schestatsky, P., Rotta, F. T., Marrone, C. D. and Gomes, I. (2013) 'Prevalence of amyotrophic lateral sclerosis in the city of Porto Alegre, in Southern Brazil', *Arq Neuropsiquiatr,* 71(12), pp. 959-62.

Ling, J. P., Pletnikova, O., Troncoso, J. C. and Wong, P. C. (2015) 'TDP-43 repression of nonconserved cryptic exons is compromised in ALS-FTD', *Science,* 349(6248), pp. 650-5.

Link, V., Shevchenko, A. and Heisenberg, C. P. (2006) 'Proteomics of early zebrafish embryos', *BMC Dev Biol,* 6, pp. 1.

Liu, F., Morderer, D., Wren, M. C., Vettleson-Trutza, S. A., Wang, Y., Rabichow, B. E., Salemi, M. R., Phinney, B. S., Oskarsson, B., Dickson, D. W. and Rossoll, W. (2022) 'Proximity proteomics of C9orf72 dipeptide repeat proteins identifies molecular chaperones as modifiers of poly-GA aggregation', *Acta Neuropathol Commun,* 10(1), pp. 22.

Liu, H. S., Jan, M. S., Chou, C. K., Chen, P. H. and Ke, N. J. (1999) 'Is green fluorescent protein toxic to the living cells?', *Biochem Biophys Res Commun,* 260(3), pp. 712-7.

Liu, P., Jenkins, N. A. and Copeland, N. G. (2003) 'A highly efficient recombineering-based method for generating conditional knockout mutations', *Genome Res,* 13(3), pp. 476-84.

Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O., Borchelt, D. R., Yachnis, A. T. and Ranum, L. P. (2016) 'C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD', *Neuron,* 90(3), pp. 521-34.

Liu, Y. P. and Berkhout, B. (2014) 'HIV-1-based lentiviral vectors', *Methods Mol Biol,* 1087, pp. 273-84.

Logan, R., Dubel-Haag, J., Schcolnicov, N. and Miller, S. J. (2022) 'Novel Genetic Signatures Associated With Sporadic Amyotrophic Lateral Sclerosis', *Front Genet,* 13, pp. 851496.

Logroscino, G., Traynor, B. J., Hardiman, O., Chiò, A., Mitchell, D., Swingler, R. J., Millul, A., Benn, E., Beghi, E. and EURALS (2010) 'Incidence of amyotrophic lateral sclerosis in Europe', *J Neurol Neurosurg Psychiatry,* 81(4), pp. 385-90.

Long, Q., Meng, A., Wang, H., Jessen, J. R., Farrell, M. J. and Lin, S. (1997) 'GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene', *Development,* 124(20), pp. 4105-11.

Lundstrom, K. (2019) 'RNA Viruses as Tools in Gene Therapy and Vaccine Development', *Genes (Basel),* 10(3).

Luty, A. A., Kwok, J. B., Dobson-Stone, C., Loy, C. T., Coupland, K. G., Karlstrom, H., Sobow, T., Tchorzewska, J., Maruszak, A., Barcikowska, M., Panegyres, P. K., Zekanowski, C., Brooks, W. S., Williams, K. L., Blair, I. P., Mather, K. A., Sachdev, P. S., Halliday, G. M. and Schofield, P. R. (2010) 'Sigma nonopioid intracellular receptor 1 mutations cause frontotemporal lobar degeneration-motor neuron disease', *Ann Neurol,* 68(5), pp. 639-49.

Lutz, C. (2018) 'Mouse models of ALS: Past, present and future', *Brain Res,* 1693(Pt A), pp. 1-10.

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. and Kroemer, G. (2013) 'The hallmarks of aging', *Cell,* 153(6), pp. 1194-217.

Ma, X. R., Prudencio, M., Koike, Y., Vatsavayai, S. C., Kim, G., Harbinski, F., Briner, A., Rodriguez, C. M., Guo, C., Akiyama, T., Schmidt, H. B., Cummings, B. B., Wyatt, D. W., Kurylo, K., Miller, G., Mekhoubad, S., Sallee, N., Mekonnen, G., Ganser, L., Rubien, J. D., Jansen-West, K., Cook, C. N., Pickles, S., Oskarsson, B., Graff-Radford, N. R., Boeve, B. F., Knopman, D. S., Petersen, R. C., Dickson, D. W., Shorter, J., Myong, S., Green, E. M., Seeley, W. W., Petrucelli, L. and Gitler, A. D. (2022) 'TDP-43 represses cryptic exon inclusion in the FTD-ALS gene UNC13A', *Nature,* 603(7899), pp. 124-130.

Mackenzie, I. R., Bigio, E. H., Ince, P. G., Geser, F., Neumann, M., Cairns, N. J., Kwong, L. K., Forman, M. S., Ravits, J., Stewart, H., Eisen, A., McClusky, L., Kretzschmar, H. A., Monoranu, C. M., Highley, J. R., Kirby, J., Siddique, T., Shaw, P. J., Lee, V. M. and Trojanowski, J. Q. (2007) 'Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations', *Ann Neurol,* 61(5), pp. 427-34.

Mackenzie, I. R., Frick, P., Grässer, F. A., Gendron, T. F., Petrucelli, L., Cashman, N. R., Edbauer, D., Kremmer, E., Prudlo, J., Troost, D. and Neumann, M. (2015) 'Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers', *Acta Neuropathol,* 130(6), pp. 845-61.

Madgwick, A., Fort, P., Hanson, P. S., Thibault, P., Gaudreau, M. C., Lutfalla, G., Möröy, T., Abou Elela, S., Chaudhry, B., Elliott, D. J., Morris, C. M. and Venables, J. P. (2015) 'Neural differentiation modulates the vertebrate brain specific splicing program', *PLoS One,* 10(5), pp. e0125998.

Maekawa, S., Al-Sarraj, S., Kibble, M., Landau, S., Parnavelas, J., Cotter, D., Everall, I. and Leigh, P. N. (2004) 'Cortical selective vulnerability in motor neuron disease: a morphometric study', *Brain,* 127(Pt 6), pp. 1237-51.

Magrané, J. and Manfredi, G. (2009) 'Mitochondrial function, morphology, and axonal transport in amyotrophic lateral sclerosis', *Antioxid Redox Signal,* 11(7), pp. 1615-26.

Maher, P. (2014) 'Proteasome Assay in Cell Lysates', *Bio Protoc,* 4(2).

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., 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(4), pp. 323-30.

Man, Z., Kondo, Y., Koga, H., Umino, H., Nakayama, K. and Shin, H. W. (2011) 'Arfaptins are localized to the trans-Golgi by interaction with Arl1, but not Arfs', *J Biol Chem,* 286(13), pp. 11569-78.

Manfredsson, F. P. and Mandel, R. J. (2010) 'Development of gene therapy for neurological disorders', *Discov Med,* 9(46), pp. 204-11.

Maniecka, Z. and Polymenidou, M. (2015) 'From nucleation to widespread propagation: A prion-like concept for ALS', *Virus Res,* 207, pp. 94-105.

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, pp. 68.

Mansilla, M. J., Comabella, M., Río, J., Castilló, J., Castillo, M., Martin, R., Montalban, X. and Espejo, C. (2014) 'Up-regulation of inducible heat shock protein-70 expression in multiple sclerosis patients', *Autoimmunity,* 47(2), pp. 127-33.

Martin, L. J., Liu, Z., Chen, K., Price, A. C., Pan, Y., Swaby, J. A. and Golden, W. C. (2007) 'Motor neuron degeneration in amyotrophic lateral sclerosis mutant superoxide dismutase-1 transgenic mice: mechanisms of mitochondriopathy and cell death', *J Comp Neurol,* 500(1), pp. 20-46.

Martin, S., Al Khleifat, A. and Al-Chalabi, A. (2017) 'What causes amyotrophic lateral sclerosis?', *F1000Res,* 6.

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., Kinoshita, Y., Kamada, M., Nodera, H., Suzuki, H., Komure, O., Matsuura, S., Kobatake, K., Morimoto, N., Abe, K., Suzuki, N., Aoki, M., Kawata, A., Hirai, T., Kato, T., Ogasawara, K., Hirano, A., Takumi, T., Kusaka, H., Hagiwara, K., Kaji, R. and Kawakami, H. (2010) 'Mutations of optineurin in amyotrophic lateral sclerosis', *Nature,* 465(7295), pp. 223-6.

Masrori, P. and Van Damme, P. (2020) 'Amyotrophic lateral sclerosis: a clinical review', *Eur J Neurol,* 27(10), pp. 1918-1929.

Mayer, M. P. and Bukau, B. (2005) 'Hsp70 chaperones: cellular functions and molecular mechanism', *Cell Mol Life Sci,* 62(6), pp. 670-84.

McGown, A., McDearmid, J. R., Panagiotaki, N., Tong, H., Al Mashhadi, S., Redhead, N., Lyon, A. N., Beattie, C. E., Shaw, P. J. and Ramesh, T. M. (2013) 'Early interneuron dysfunction in ALS: insights from a mutant sod1 zebrafish model', *Ann Neurol,* 73(2), pp. 246-58.

McWhorter, M. L., Monani, U. R., Burghes, A. H. and Beattie, C. E. (2003) 'Knockdown of the survival motor neuron (Smn) protein in zebrafish causes defects in motor axon outgrowth and pathfinding', *J Cell Biol,* 162(5), pp. 919-31.

Mehta, P., Antao, V., Kaye, W., Sanchez, M., Williamson, D., Bryan, L., Muravov, O., Horton, K., Division of Toxicology and Human Health Sciences, A. f. T. S. a. D. R., Atlanta, G.orgia and (CDC), C. f. D. C. a. P. (2014) 'Prevalence of amyotrophic lateral sclerosis - United States, 2010-2011', *MMWR Suppl,* 63(7), pp. 1-14.

Meininger, V. (2011) 'ALS, what new 144 years after Charcot?', *Arch Ital Biol,* 149(1), pp. 29-37.

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?', *Front Neurosci,* 13, pp. 1310.

Meng, X., Noyes, M. B., Zhu, L. J., Lawson, N. D. and Wolfe, S. A. (2008) 'Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases', *Nat Biotechnol,* 26(6), pp. 695-701.

Mertens, J., Paquola, A. C. M., Ku, M., Hatch, E., Böhnke, L., Ladjevardi, S., McGrath, S., Campbell, B., Lee, H., Herdy, J. R., Gonçalves, J. T., Toda, T., Kim, Y., Winkler, J., Yao, J., Hetzer, M. W. and Gage, F. H. (2015) 'Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects', *Cell Stem Cell,* 17(6), pp. 705-718.

Meyer, K., Ferraiuolo, L., Miranda, C. J., Likhite, S., McElroy, S., Renusch, S., Ditsworth, D., Lagier-Tourenne, C., Smith, R. A., Ravits, J., Burghes, A. H., Shaw, P. J., Cleveland, D. W., Kolb, S. J. and Kaspar, B. K. (2014) 'Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS', *Proc Natl Acad Sci U S A,* 111(2), pp. 829-32.

Miletic, H., Fischer, Y. H., Neumann, H., Hans, V., Stenzel, W., Giroglou, T., Hermann, M., Deckert, M. and Von Laer, D. (2004) 'Selective transduction of malignant glioma by lentiviral vectors pseudotyped with lymphocytic choriomeningitis virus glycoproteins', *Hum Gene Ther,* 15(11), pp. 1091-100.

Miller, T., Cudkowicz, M., Shaw, P. J., Andersen, P. M., Atassi, N., Bucelli, R. C., Genge, A., Glass, J., Ladha, S., Ludolph, A. L., Maragakis, N. J., McDermott, C. J., Pestronk, A., Ravits, J., Salachas, F., Trudell, R., Van Damme, P., Zinman, L., Bennett, C. F., Lane, R., Sandrock, A., Runz, H., Graham, D., Houshyar, H., McCampbell, A., Nestorov, I., Chang, I., McNeill, M., Fanning, L., Fradette, S. and Ferguson, T. A. (2020) 'Phase 1-2 Trial of Antisense Oligonucleotide Tofersen for', *N Engl J Med,* 383(2), pp. 109-119.

Miller, T. M., Cudkowicz, M. E., Genge, A., Shaw, P. J., Sobue, G., Bucelli, R. C., Chiò, A., Van Damme, P., Ludolph, A. C., Glass, J. D., Andrews, J. A., Babu, S., Benatar, M., McDermott, C. J., Cochrane, T., Chary, S., Chew, S., Zhu, H., Wu, F., Nestorov, I., Graham, D., Sun, P., McNeill, M., Fanning, L., Ferguson, T. A., Fradette, S. and Group, V. a. O. W. (2022) 'Trial of Antisense Oligonucleotide Tofersen for SOD1 ALS', *N Engl J Med,* 387(12), pp. 1099-1110.

Minciacchi, D., Kassa, R. M., Del Tongo, C., Mariotti, R. and Bentivoglio, M. (2009) 'Voronoi-based spatial analysis reveals selective interneuron changes in the cortex of FALS mice', *Exp Neurol,* 215(1), pp. 77-86.

Mirska, D., Schirmer, K., Funari, S. S., Langner, A., Dobner, B. and Brezesinski, G. (2005) 'Biophysical and biochemical properties of a binary lipid mixture for DNA transfection', *Colloids Surf B Biointerfaces,* 40(1), pp. 51-9.

Mitchell, J., Paul, P., Chen, H. J., Morris, A., Payling, M., Falchi, M., Habgood, J., Panoutsou, S., Winkler, S., Tisato, V., Hajitou, A., Smith, B., Vance, C., Shaw, C., Mazarakis, N. D. and de Belleroche, J. (2010) 'Familial amyotrophic lateral sclerosis is associated with a mutation in D-amino acid oxidase', *Proc Natl Acad Sci U S A,* 107(16), pp. 7556-61.

Mitchell, J. C., McGoldrick, P., Vance, C., Hortobagyi, T., Sreedharan, J., Rogelj, B., Tudor, E. L., Smith, B. N., Klasen, C., Miller, C. C., Cooper, J. D., Greensmith, L. and Shaw, C. E. (2013) 'Overexpression of human wild-type FUS causes progressive motor neuron degeneration in an age- and dose-dependent fashion', *Acta Neuropathol,* 125(2), pp. 273-88.

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. (2014) 'C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins', *Science,* 345(6201), pp. 1192-1194.

Mizielinska, S., Lashley, T., Norona, F. E., Clayton, E. L., Ridler, C. E., Fratta, P. and Isaacs, A. M. (2013) 'C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci', *Acta Neuropathol,* 126(6), pp. 845-57.

Mizuno, Y., Amari, M., Takatama, M., Aizawa, H., Mihara, B. and Okamoto, K. (2006) 'Immunoreactivities of p62, an ubiqutin-binding protein, in the spinal anterior horn cells of patients with amyotrophic lateral sclerosis', *J Neurol Sci,* 249(1), pp. 13-8.

Mochizuki, Y., Mizutani, T., Shimizu, T. and Kawata, A. (2011) 'Proportional neuronal loss between the primary motor and sensory cortex in amyotrophic lateral sclerosis', *Neurosci Lett,* 503(1), pp. 73-5.

Mohammedeid, A. (2015) *Regulation of Protein Aggregation by Arfaptin2 in Amyotrophic Lateral Sclerosis.* PhD Thesis, University of Sheffield

Mohammedeid, A. M., Lukashchuk, V. and Ning, K. (2014) 'Protein aggregation and Arfaptin2: A novel therapeutic target against neurodegenerative diseases', *New Horizons in Translational Medicine,* 2(1), pp. 12-15.

Mordes, D. A., Morrison, B. M., Ament, X. H., Cantrell, C., Mok, J., Eggan, P., Xue, C., Wang, J. Y., Eggan, K. and Rothstein, J. D. (2020) 'Absence of Survival and Motor Deficits in 500 Repeat C9ORF72 BAC Mice', *Neuron,* 108(4), pp. 775-783.e4.

Mordes, D. A., Prudencio, M., Goodman, L. D., Klim, J. R., Moccia, R., Limone, F., Pietilainen, O., Chowdhary, K., Dickson, D. W., Rademakers, R., Bonini, N. M., Petrucelli, L. and Eggan, K. (2018) 'Dipeptide repeat proteins activate a heat shock response found in C9ORF72-ALS/FTLD patients', *Acta Neuropathol Commun,* 6(1), pp. 55.

Morgan, S. and Orrell, R. W. (2016) 'Pathogenesis of amyotrophic lateral sclerosis', *British Medical Bulletin,* 119(1), pp. 87-98.

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 Neuropathol,* 126(6), pp. 881-93.

Mori, K., Lammich, S., Mackenzie, I. R., Forné, 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(3), pp. 413-23.

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(6125), pp. 1335-8.

Mosimann, C., Kaufman, C. K., Li, P., Pugach, E. K., Tamplin, O. J. and Zon, L. I. (2011) 'Ubiquitous transgene expression and Cre-based recombination driven by the ubiquitin promoter in zebrafish', *Development,* 138(1), pp. 169-77.

Muchowski, P. J., Ning, K., D'Souza-Schorey, C. and Fields, S. (2002) 'Requirement of an intact microtubule cytoskeleton for aggregation and inclusion body formation by a mutant huntingtin fragment', *Proc Natl Acad Sci U S A,* 99(2), pp. 727-32.

Mulder, D. W., Kurland, L. T., Offord, K. P. and Beard, C. M. (1986) 'Familial adult motor neuron disease: amyotrophic lateral sclerosis', *Neurology,* 36(4), pp. 511-7.

Murphy, M. P., Holmgren, A., Larsson, N. G., Halliwell, B., Chang, C. J., Kalyanaraman, B., Rhee, S. G., Thornalley, P. J., Partridge, L., Gems, D., Nyström, T., Belousov, V., Schumacker, P. T. and Winterbourn, C. C. (2011) 'Unraveling the biological roles of reactive oxygen species', *Cell Metab,* 13(4), pp. 361-366.

Muyderman, H. and Chen, T. (2014) 'Mitochondrial dysfunction in amyotrophic lateral sclerosis - a valid pharmacological target?', *Br J Pharmacol,* 171(8), pp. 2191-205.

Mórotz, G. M., De Vos, K. J., Vagnoni, A., Ackerley, S., Shaw, C. E. and Miller, C. C. (2012) 'Amyotrophic lateral sclerosis-associated mutant VAPBP56S perturbs calcium homeostasis to disrupt axonal transport of mitochondria', *Hum Mol Genet,* 21(9), pp. 1979-88.

Münch, C., O'Brien, J. and Bertolotti, A. (2011) 'Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells', *Proc Natl Acad Sci U S A,* 108(9), pp. 3548-53.

Nassif, M., Woehlbier, U. and Manque, P. A. (2017) 'The Enigmatic Role of C9ORF72 in Autophagy', *Front Neurosci,* 11, pp. 442.

National Institute of Neurological Disorders and Stroke (NINDS) (2014) *Motor Neuron Diseases Fact Sheet*. Available at: https://web.archive.org/web/20140413093035/http://www.ninds.nih.gov/disorders/motor _neuron_diseases/detail_motor_neuron_diseases.htm.

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. (2006) 'Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis', *Science,* 314(5796), pp. 130-3.

Ng, A. S., Rademakers, R. and Miller, B. L. (2015) 'Frontotemporal dementia: a bridge between dementia and neuromuscular disease', *Ann N Y Acad Sci,* 1338(1), pp. 71-93.

Nguyen, H. P., Van Broeckhoven, C. and van der Zee, J. (2018) 'ALS Genes in the Genomic Era and their Implications for FTD', *Trends Genet,* 34(6), pp. 404-423.

Nihei, K., McKee, A. C. and Kowall, N. W. (1993) 'Patterns of neuronal degeneration in the motor cortex of amyotrophic lateral sclerosis patients', *Acta Neuropathol,* 86(1), pp. 55-64.

Nishimura, A. L., Mitne-Neto, M., Silva, H. C., Richieri-Costa, A., Middleton, S., Cascio, D., Kok, F., Oliveira, J. R., Gillingwater, T., Webb, J., Skehel, P. and Zatz, M. (2004) 'A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis', *Am J Hum Genet,* 75(5), pp. 822-31.

Nomura, T., Watanabe, S., Kaneko, K., Yamanaka, K., Nukina, N. and Furukawa, Y. (2014) 'Intranuclear aggregation of mutant FUS/TLS as a molecular pathomechanism of amyotrophic lateral sclerosis', *J Biol Chem,* 289(2), pp. 1192-202.

Nonaka, T., Masuda-Suzukake, M., Arai, T., Hasegawa, Y., Akatsu, H., Obi, T., Yoshida, M., Murayama, S., Mann, D. M., Akiyama, H. and Hasegawa, M. (2013) 'Prion-like properties of pathological TDP-43 aggregates from diseased brains', *Cell Rep,* 4(1), pp. 124-34.

Nonaka, T., Suzuki, G., Tanaka, Y., Kametani, F., Hirai, S., Okado, H., Miyashita, T., Saitoe, M., Akiyama, H., Masai, H. and Hasegawa, M. (2016) 'Phosphorylation of TAR DNA-binding Protein of 43

kDa (TDP-43) by Truncated Casein Kinase 1δ Triggers Mislocalization and Accumulation of TDP-43', *J Biol Chem,* 291(11), pp. 5473-5483.

O'Brien, J., Wilson, I., Orton, T. and Pognan, F. (2000) 'Investigation of the Alamar Blue (resazurin) fluorescent dye for the assessment of mammalian cell cytotoxicity', *Eur J Biochem,* 267(17), pp. 5421-6.

Oakes, J. A., Davies, M. C. and Collins, M. O. (2017) 'TBK1: a new player in ALS linking autophagy and neuroinflammation', *Molecular Brain,* 10(1), pp. 5.

Oh, Y. M., Lee, S. W., Kim, W. K., Chen, S., Church, V. A., Cates, K., Li, T., Zhang, B., Dolle, R. E., Dahiya, S., Pak, S. C., Silverman, G. A., Perlmutter, D. H. and Yoo, A. S. (2022) 'Age-related Huntington's disease progression modeled in directly reprogrammed patient-derived striatal neurons highlights impaired autophagy', *Nat Neurosci,* 25(11), pp. 1420-1433.

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', *Mol Neurodegener,* 12(1), pp. 6.

Opie-Martin, S., Jones, A., Iacoangeli, A., Al-Khleifat, A., Oumar, M., Shaw, P. J., Shaw, C. E., Morrison, K. E., Wootton, R. E., Davey-Smith, G., Pearce, N. and Al-Chalabi, A. (2020) 'UK case control study of smoking and risk of amyotrophic lateral sclerosis', *Amyotroph Lateral Scler Frontotemporal Degener,* 21(3-4), pp. 222-227.

Orlacchio, A., Babalini, C., Borreca, A., Patrono, C., Massa, R., Basaran, S., Munhoz, R. P., Rogaeva, E. A., St George-Hyslop, P. H., Bernardi, G. and Kawarai, T. (2010) 'SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis', *Brain: Vol. 2*, pp. 591-8.

Oskarsson, B., Horton, D. K. and Mitsumoto, H. (2015) 'Potential Environmental Factors in Amyotrophic Lateral Sclerosis', *Neurol Clin,* 33(4), pp. 877-888.

Paganoni, S., Hendrix, S., Dickson, S. P., Knowlton, N., Macklin, E. A., Berry, J. D., Elliott, M. A., Maiser, S., Karam, C., Caress, J. B., Owegi, M. A., Quick, A., Wymer, J., Goutman, S. A., Heitzman, D., Heiman-Patterson, T. D., Jackson, C. E., Quinn, C., Rothstein, J. D., Kasarskis, E. J., Katz, J., Jenkins, L., Ladha, S., Miller, T. M., Scelsa, S. N., Vu, T. H., Fournier, C. N., Glass, J. D., Johnson, K. M., Swenson, A., Goyal, N. A., Pattee, G. L., Andres, P. L., Babu, S., Chase, M., Dagostino, D., Hall, M., Kittle, G., Eydinov, M., McGovern, M., Ostrow, J., Pothier, L., Randall, R., Shefner, J. M., Sherman, A. V., St Pierre, M. E., Tustison, E., Vigneswaran, P., Walker, J., Yu, H., Chan, J., Wittes, J., Yu, Z. F., Cohen, J., Klee, J., Leslie, K., Tanzi, R. E., Gilbert, W., Yeramian, P. D., Schoenfeld, D. and Cudkowicz, M. E. (2021) 'Long-term survival of participants in the CENTAUR trial of sodium phenylbutyrate-taurursodiol in amyotrophic lateral sclerosis', *Muscle Nerve,* 63(1), pp. 31-39.

Paganoni, S., Macklin, E. A., Hendrix, S., Berry, J. D., Elliott, M. A., Maiser, S., Karam, C., Caress, J. B., Owegi, M. A., Quick, A., Wymer, J., Goutman, S. A., Heitzman, D., Heiman-Patterson, T., Jackson, C. E., Quinn, C., Rothstein, J. D., Kasarskis, E. J., Katz, J., Jenkins, L., Ladha, S., Miller, T. M., Scelsa, S. N., Vu, T. H., Fournier, C. N., Glass, J. D., Johnson, K. M., Swenson, A., Goyal, N. A., Pattee, G. L., Andres, P. L., Babu, S., Chase, M., Dagostino, D., Dickson, S. P., Ellison, N., Hall, M., Hendrix, K., Kittle, G., McGovern, M., Ostrow, J., Pothier, L., Randall, R., Shefner, J. M., Sherman, A. V., Tustison, E., Vigneswaran, P., Walker, J., Yu, H., Chan, J., Wittes, J., Cohen, J., Klee, J., Leslie, K., Tanzi, R. E., Gilbert, W., Yeramian, P. D., Schoenfeld, D. and Cudkowicz, M. E. (2020) 'Trial of Sodium Phenylbutyrate-Taurursodiol for Amyotrophic Lateral Sclerosis', *N Engl J Med,* 383(10), pp. 919-930.

Page, B., Page, M. and Noel, C. (1993) 'A new fluorometric assay for cytotoxicity measurements in-vitro', *Int J Oncol,* 3(3), pp. 473-6.

Pansarasa, O., Bordoni, M., Diamanti, L., Sproviero, D., Gagliardi, S. and Cereda, C. (2018) 'SOD1 in Amyotrophic Lateral Sclerosis: "Ambivalent" Behavior Connected to the Disease', *Int J Mol Sci,* 19(5).

Panula, P., Sallinen, V., Sundvik, M., Kolehmainen, J., Torkko, V., Tiittula, A., Moshnyakov, M. and Podlasz, P. (2006) 'Modulatory neurotransmitter systems and behavior: towards zebrafish models of neurodegenerative diseases', *Zebrafish,* 3(2), pp. 235-47.

Papale, A., Cerovic, M. and Brambilla, R. (2009) 'Viral vector approaches to modify gene expression in the brain', *J Neurosci Methods,* 185(1), pp. 1-14.

Parkinson, N., Ince, P. G., Smith, M. O., Highley, R., Skibinski, G., Andersen, P. M., Morrison, K. E., Pall, H. S., Hardiman, O., Collinge, J., Shaw, P. J. and Fisher, E. M. (2006) 'ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B)', *Neurology,* 67(6), pp. 1074-7.

Patel, R., Rinker, L., Peng, J. and Chilian, W. M. (2017) 'Reactive Oxygen Species: The Good and the Bad', in C. Filip, E.A. (ed.) *Reactive Oxygen Species (ROS) in Living Cells*.

Paul, M. S. and Limaiem, F. (2022) 'Histology, Purkinje Cells', *StatPearls [Internet]*. Treasure Island (FL): tatPearls Publishing.

Paulson, H. (2019) 'Repeat expansion diseases', *Handb Clin Neurol,* 147, pp. 105-123.

Perez, N., Sugar, J., Charya, S., Johnson, G., Merril, C., Bierer, L., Perl, D., Haroutunian, V. and Wallace, W. (1991) 'Increased synthesis and accumulation of heat shock 70 proteins in Alzheimer's disease', *Brain Res Mol Brain Res,* 11(3-4), pp. 249-54.

Peter, B. J., Kent, H. M., Mills, I. G., Vallis, Y., Butler, P. J., Evans, P. R. and McMahon, H. T. (2004) 'BAR domains as sensors of membrane curvature: the amphiphysin BAR structure', *Science,* 303(5657), pp. 495-9.

Peters, P. J., Ning, K., Palacios, F., Boshans, R. L., Kazantsev, A., Thompson, L. M., Woodman, B., Bates, G. P. and D'Souza-Schorey, C. (2002) 'Arfaptin 2 regulates the aggregation of mutant huntingtin protein', *Nature Cell Biology,* 4(3), pp. 240.

Philips, T. and Rothstein, J. D. (2015) 'Rodent Models of Amyotrophic Lateral Sclerosis', *Curr Protoc Pharmacol,* 69, pp. 5.67.1-5.67.21.

Phillips, J. B. and Westerfield, M. (2014) 'Zebrafish models in translational research: tipping the scales toward advancements in human health', *Dis Model Mech,* 7(7), pp. 739-43.

Picher-Martel, V., Valdmanis, P. N., Gould, P. V., Julien, J. P. and Dupré, N. (2016) 'From animal models to human disease: a genetic approach for personalized medicine in ALS', *Acta Neuropathol Commun,* 4(1), pp. 70.

Pinarbasi, E. S., Cağatay, T., Fung, H. Y. J., Li, Y. C., Chook, Y. M. and Thomas, P. J. (2018) 'Active nuclear import and passive nuclear export are the primary determinants of TDP-43 localization', *Sci Rep,* 8(1), pp. 7083.

Pizzino, G., Irrera, N., Cucinotta, M., Pallio, G., Mannino, F., Arcoraci, V., Squadrito, F., Altavilla, D. and Bitto, A. (2017) 'Oxidative Stress: Harms and Benefits for Human Health', *Oxid Med Cell Longev,* 2017, pp. 8416763.

Plaut, I. (2000) 'Effects of fin size on swimming performance, swimming behaviour and routine activity of zebrafish Danio rerio', *J Exp Biol,* 203(Pt 4), pp. 813-20.

Powell, S. K., Rivera-Soto, R. and Gray, S. J. (2015) 'Viral expression cassette elements to enhance transgene target specificity and expression in gene therapy', *Discov Med,* 19(102), pp. 49-57.

Pramatarova, A., Figlewicz, D. A., Krizus, A., Han, F. Y., Ceballos-Picot, I., Nicole, A., Dib, M., Meininger, V., Brown, R. H. and Rouleau, G. A. (1995) 'Identification of new mutations in the Cu/Zn superoxide dismutase gene of patients with familial amyotrophic lateral sclerosis', *Am J Hum Genet,* 56(3), pp. 592-596.

Protter, D. S. W. and Parker, R. (2016) 'Principles and Properties of Stress Granules', *Trends Cell Biol,* 26(9), pp. 668-679.

Pugdahl, K., Camdessanché, J. P., Cengiz, B., de Carvalho, M., Liguori, R., Rossatto, C., Oliveira Santos, M., Vacchiano, V. and Johnsen, B. (2021) 'Gold Coast diagnostic criteria increase sensitivity in amyotrophic lateral sclerosis', *Clin Neurophysiol,* 132(12), pp. 3183-3189.

Qiu, H., Lee, S., Shang, Y., Wang, W. Y., Au, K. F., Kamiya, S., Barmada, S. J., Finkbeiner, S., Lui, H., Carlton, C. E., Tang, A. A., Oldham, M. C., Wang, H., Shorter, J., Filiano, A. J., Roberson, E. D.,

Tourtellotte, W. G., Chen, B., Tsai, L. H. and Huang, E. J. (2014) 'ALS-associated mutation FUS-R521C causes DNA damage and RNA splicing defects', *J Clin Invest,* 124(3), pp. 981-99.

Quaegebeur, A., Glaria, I., Lashley, T. and Isaacs, A. M. (2020) 'Soluble and insoluble dipeptide repeat protein measurements in C9orf72-frontotemporal dementia brains show regional differential solubility and correlation of poly-GR with clinical severity', *Acta Neuropathol Commun,* 8(1), pp. 184.

Rafferty, S. A. and Quinn, T. A. (2018) 'A beginner's guide to understanding and implementing the genetic modification of zebrafish', *Prog Biophys Mol Biol,* 138, pp. 3-19.

Ramesh, T., Lyon, A. N., Pineda, R. H., Wand, C., Janssen, P. M. L., Canan, B. D., Burghes, A. H. M. and Beattie, C. E. (2010a) 'A genetic model of amyotrophic lateral sclerosis in zebrafish displays phenotypic hallmarks of motoneuron disease', *Disease Models & Mechanisms,* 3(9-10), pp. 652-662.

Ramesh, T., Lyon, A. N., Pineda, R. H., Wang, C., Janssen, P. M., Canan, B. D., Burghes, A. H. and Beattie, C. E. (2010b) 'A genetic model of amyotrophic lateral sclerosis in zebrafish displays phenotypic hallmarks of motoneuron disease', *Dis Model Mech,* 3(9-10), pp. 652-62.

Rangone, H., Pardo, R., Colin, E., Girault, J. A., Saudou, F. and Humbert, S. (2005) 'Phosphorylation of arfaptin 2 at Ser260 by Akt Inhibits PolyQ-huntingtin-induced toxicity by rescuing proteasome impairment', *J Biol Chem,* 280(23), pp. 22021-8.

Ratti, A. and Buratti, E. (2016) 'Physiological functions and pathobiology of TDP-43 and FUS/TLS proteins', *J Neurochem,* 138 Suppl 1, pp. 95-111.

Ray, P. D., Huang, B. W. and Tsuji, Y. (2012) 'Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling', *Cell Signal,* 24(5), pp. 981-90.

Reaume, A. G., Elliott, J. L., Hoffman, E. K., Kowall, N. W., Ferrante, R. J., Siwek, D. F., Wilcox, H. M., Flood, D. G., Beal, M. F., Brown, R. H., Scott, R. W. and Snider, W. D. (1996) 'Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury', *Nat Genet,* 13(1), pp. 43-7.

Reddy, K., Zamiri, B., Stanley, S. Y. R., Macgregor, R. B. 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(14), pp. 9860-9866.

Renton, A. E., Chiò, A. and Traynor, B. J. (2014) 'State of play in amyotrophic lateral sclerosis genetics', *Nat Neurosci,* 17(1), pp. 17-23.

Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., Schymick, J. C., Laaksovirta, H., van Swieten, J. C., Myllykangas, L., 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., 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., Hölttä-Vuori, M., Ikonen, E., Sulkava, R., Benatar, M., Wuu, J., Chiò, A., Restagno, G., Borghero, G., Sabatelli, M., Heckerman, D., Rogaeva, E., Zinman, L., Rothstein, J., 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(2), pp. 257-68.

Riviere, M., Meininger, V., Zeisser, P. and Munsat, T. (1998) 'An analysis of extended survival in patients with amyotrophic lateral sclerosis treated with riluzole', *Arch Neurol,* 55(4), pp. 526-8.

Rizzu, P., Blauwendraat, C., Heetveld, S., Lynes, E. M., Castillo-Lizardo, M., Dhingra, A., Pyz, E., Hobert, M., Synofzik, M., Simón-Sánchez, J., Francescatto, M. and Heutink, P. (2016) 'C9orf72 is differentially expressed in the central nervous system and myeloid cells and consistently reduced in C9orf72, MAPT and GRN mutation carriers', *Acta Neuropathol Commun,* 4(1), pp. 37.

Robberecht, W., Aguirre, T., Van den Bosch, L., Tilkin, P., Cassiman, J. J. and Matthijs, G. (1996) 'D90A heterozygosity in the SOD1 gene is associated with familial and apparently sporadic amyotrophic lateral sclerosis', *Neurology,* 47(5), pp. 1336-9.

Robberecht, W. and Philips, T. (2013) 'The changing scene of amyotrophic lateral sclerosis', *Nature Reviews Neuroscience,* 14(4), pp. 248-264.

Robinson, H. K., Deykin, A. V., Bronovitsky, E. V., Ovchinnikov, R. K., Ustyugov, A. A., Shelkovnikova, T. A., Kukharsky, M. S., Ermolkevich, T. G., Goldman, I. L., Sadchikova, E. R., Kovrazhkina, E. A., Bachurin, S. O., Buchman, V. L. and Ninkina, N. N. (2015) 'Early lethality and neuronal proteinopathy in mice expressing cytoplasm-targeted FUS that lacks the RNA recognition motif', *Amyotroph Lateral Scler Frontotemporal Degener,* 16(5-6), pp. 402-9.

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. and et al. (1993) 'Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis', *Nature,* 362(6415), pp. 59-62.

Rowland, L. P. and Shneider, N. A. (2001) 'Amyotrophic lateral sclerosis', *N Engl J Med,* 344(22), pp. 1688-700.

Ryan, M., Heverin, M., McLaughlin, R. L. and Hardiman, O. (2019) 'Lifetime Risk and Heritability of Amyotrophic Lateral Sclerosis', *JAMA Neurol,* 76(11), pp. 1367-1374.

Ryan, S., Rollinson, S., Hobbs, E. and Pickering-Brown, S. (2022) 'C9orf72 dipeptides disrupt the nucleocytoplasmic transport machinery and cause TDP-43 mislocalisation to the cytoplasm', *Sci Rep,* 12(1), pp. 4799.

Saberi, S., Stauffer, J. E., Schulte, D. J. and Ravits, J. (2015) 'Neuropathology of Amyotrophic Lateral Sclerosis and Its Variants', *Neurol Clin,* 33(4), pp. 855-76.

Sadelain, M. (2017) 'CD19 CAR T Cells', *Cell,* 171(7), pp. 1471.

Safran, M., Rosen, N., Twik, M., BarShir, R., Iny Stein, T., Dahary, D., Fishilevich, S. and Lancet, D. (2022) '*The GeneCards Suite* Chapter', *Practical Guide to Life Science Databases*, pp. 27-56.

Saitoh, Y. and Takahashi, Y. (2020) 'Riluzole for the treatment of amyotrophic lateral sclerosis', *Neurodegener Dis Manag,* 10(6), pp. 343-355.

Sakowski, S. A., Lunn, J. S., Busta, A. S., Oh, S. S., Zamora-Berridi, G., Palmer, M., Rosenberg, A. A., Philip, S. G., Dowling, J. J. and Feldman, E. L. (2012) 'Neuromuscular effects of G93A-SOD1 expression in zebrafish', *Mol Neurodegener,* 7, pp. 44.

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(208), pp. 208ra149.

Sasaki, S., Horie, Y. and Iwata, M. (2007) 'Mitochondrial alterations in dorsal root ganglion cells in sporadic amyotrophic lateral sclerosis', *Acta Neuropathol,* 114(6), pp. 633-9.

Sasaki, S. and Iwata, M. (2007) 'Mitochondrial alterations in the spinal cord of patients with sporadic amyotrophic lateral sclerosis', *J Neuropathol Exp Neurol,* 66(1), pp. 10-6.

Sasayama, H., Shimamura, M., Tokuda, T., Azuma, Y., Yoshida, T., Mizuno, T., Nakagawa, M., Fujikake, N., Nagai, Y. and Yamaguchi, M. (2012) 'Knockdown of the Drosophila fused in sarcoma (FUS) homologue causes deficient locomotive behavior and shortening of motoneuron terminal branches', *PLoS One,* 7(6), pp. e39483.

Sawada, Y., Kajiwara, G., Iizuka, A., Takayama, K., Shuvaev, A. N., Koyama, C. and Hirai, H. (2010) 'High transgene expression by lentiviral vectors causes maldevelopment of Purkinje cells in vivo', *Cerebellum,* 9(3), pp. 291-302.

Scekic-Zahirovic, J., Sendscheid, O., El Oussini, H., Jambeau, M., Sun, Y., Mersmann, S., Wagner, M., Dieterlé, S., Sinniger, J., Dirrig-Grosch, S., Drenner, K., Birling, M. C., Qiu, J., Zhou, Y., Li, H., Fu, X. D., Rouaux, C., Shelkovnikova, T., Witting, A., Ludolph, A. C., Kiefer, F., Storkebaum, E.,

Lagier-Tourenne, C. and Dupuis, L. (2016) 'Toxic gain of function from mutant FUS protein is crucial to trigger cell autonomous motor neuron loss', *EMBO J,* 35(10), pp. 1077-97.

Scherer, L. J. and Rossi, J. J. (2003) 'Approaches for the sequence-specific knockdown of mRNA', *Nat Biotechnol,* 21(12), pp. 1457-65.

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 Neuropathol,* 134(2), pp. 241-254.

Schludi, M. H. and Edbauer, D. (2018) 'Targeting RNA G-quadruplexes as new treatment strategy for', *EMBO Mol Med,* 10(1), pp. 4-6.

Schmitz, A., Pinheiro Marques, J., Oertig, I., Maharjan, N. and Saxena, S. (2021) 'Emerging Perspectives on Dipeptide Repeat Proteins in C9ORF72 ALS/FTD', *Front Cell Neurosci,* 15, pp. 637548.

Schneider, B., Zufferey, R. and Aebischer, P. (2008) 'Viral vectors, animal models and new therapies for Parkinson's disease', *Parkinsonism Relat Disord,* 14 Suppl 2, pp. S169-71.

Sephton, C. F., Good, S. K., Atkin, S., Dewey, C. M., Mayer, P., Herz, J. and Yu, G. (2010) 'TDP-43 is a developmentally regulated protein essential for early embryonic development', *J Biol Chem,* 285(9), pp. 6826-34.

Sharma, A., Lyashchenko, A. K., Lu, L., Nasrabady, S. E., Elmaleh, M., Mendelsohn, M., Nemes, A., Tapia, J. C., Mentis, G. Z. and Shneider, N. A. (2016) 'ALS-associated mutant FUS induces selective motor neuron degeneration through toxic gain of function', *Nat Commun,* 7, pp. 10465.

Shaw, B. F. and Valentine, J. S. (2007) 'How do ALS-associated mutations in superoxide dismutase 1 promote aggregation of the protein?', *Trends Biochem Sci,* 32(2), pp. 78-85.

Shaw, M. P., Higginbottom, A., McGown, A., Castelli, L. M., James, E., Hautbergue, G. M., Shaw, P. J. and Ramesh, T. M. (2018) 'Stable transgenic C9orf72 zebrafish model key aspects of the ALS/FTD phenotype and reveal novel pathological features', *Acta Neuropathol Commun,* 6(1), pp. 125.

Shaw, P. (2005) 'Molecular and cellular pathways of neurodegeneration in motor neurone disease', *J Neurol Neurosurg Psychiatry,* 76(8), pp. 1046-57.

Shaw, P. J. and Eggett, C. J. (2000) 'Molecular factors underlying selective vulnerability of motor neurons to neurodegeneration in amyotrophic lateral sclerosis', *J Neurol,* 247 Suppl 1, pp. I17-27.

Shaw, P. J. and Wood-Allum, C. (2010) 'Motor neurone disease: a practical update on diagnosis and management', *Clin Med (Lond): Vol. 3*, pp. 252-8.

Shefner, J. M., Al-Chalabi, A., Baker, M. R., Cui, L. Y., de Carvalho, M., Eisen, A., Grosskreutz, J., Hardiman, O., Henderson, R., Matamala, J. M., Mitsumoto, H., Paulus, W., Simon, N., Swash, M., Talbot, K., Turner, M. R., Ugawa, Y., van den Berg, L. H., Verdugo, R., Vucic, S., Kaji, R., Burke, D. and Kiernan, M. C. (2020) 'A proposal for new diagnostic criteria for ALS', *Clin Neurophysiol,* 131(8), pp. 1975-1978.

Shelkovnikova, T. A., Peters, O. M., Deykin, A. V., Connor-Robson, N., Robinson, H., Ustyugov, A. A., Bachurin, S. O., Ermolkevich, T. G., Goldman, I. L., Sadchikova, E. R., Kovrazhkina, E. A., Skvortsova, V. I., Ling, S. C., Da Cruz, S., Parone, P. A., Buchman, V. L. and Ninkina, N. N. (2013) 'Fused in sarcoma (FUS) protein lacking nuclear localization signal (NLS) and major RNA binding motifs triggers proteinopathy and severe motor phenotype in transgenic mice', *J Biol Chem,* 288(35), pp. 25266-25274.

Shi, B., Xue, M., Wang, Y., Li, D., Zhao, X. and Li, X. (2018) 'An improved method for increasing the efficiency of gene transfection and transduction', *Int J Physiol Pathophysiol Pharmacol,* 10(2), pp. 95-104.

Shibuya, K., Park, S. B., Geevasinga, N., Menon, P., Howells, J., Simon, N. G., Huynh, W., Noto, Y., Götz, J., Kril, J. J., Ittner, L. M., Hodges, J., Halliday, G., Vucic, S. and Kiernan, M. C. (2016) 'Motor cortical function determines prognosis in sporadic ALS', *Neurology,* 87(5), pp. 513-20.

Shin, O. H. and Exton, J. H. (2001) 'Differential binding of arfaptin 2/POR1 to ADP-ribosylation factors and Rac1', *Biochem Biophys Res Commun,* 285(5), pp. 1267-73.

Shin, O. H. and Exton, J. H. (2005) 'Assays and properties of arfaptin 2 binding to Rac1 and ADP-ribosylation factors (Arfs)', *Methods Enzymol,* 404, pp. 359-67.

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(1), pp. 22-31.

Smith, B. N., Ticozzi, N., Fallini, C., Gkazi, A. S., Topp, S., Kenna, K. P., Scotter, E. L., Kost, J., Keagle, P., Miller, J. W., Calini, D., Vance, C., Danielson, E. W., Troakes, C., Tiloca, C., Al-Sarraj, S., Lewis, E. A., King, A., Colombrita, C., Pensato, V., Castellotti, B., de Belleroche, J., Baas, F., ten Asbroek, A. L., Sapp, P. C., McKenna-Yasek, D., McLaughlin, R. L., Polak, M., Asress, S., Esteban-

Pérez, J., oz-Blanco, J. L. M., Simpson, M., Consortium, S., van Rheenen, W., Diekstra, F. P., Lauria, G., Duga, S., Corti, S., Cereda, C., Corrado, L., Sorarù, G., Morrison, K. E., Williams, K. L., Nicholson, G. A., Blair, I. P., Dion, P. A., Leblond, C. S., Rouleau, G. A., Hardiman, O., Veldink, J. H., van den Berg, L. H., Al-Chalabi, A., Pall, H., Shaw, P. J., Turner, M. R., Talbot, K., Taroni, F., García-Redondo, A., Wu, Z., Glass, J. D., Gellera, C., Ratti, A., Brown, R. H., Silani, V., Shaw, C. E. and Landers, J. E. (2014) 'Exome-wide Rare Variant Analysis Identifies TUBA4A Mutations Associated with Familial ALS', *Neuron,* 84(2), pp. 324-31.

Son, E. Y., Ichida, J. K., Wainger, B. J., Toma, J. S., Rafuse, V. F., Woolf, C. J. and Eggan, K. (2011) 'Conversion of mouse and human fibroblasts into functional spinal motor neurons', *Cell Stem Cell,* 9(3), pp. 205-18.

Sork, H., Nordin, J. Z., Turunen, J. J., Wiklander, O. P., Bestas, B., Zaghloul, E. M., Margus, H., Padari, K., Duru, A. D., Corso, G., Bost, J., Vader, P., Pooga, M., Smith, C. E., Wood, M. J., Schiffelers, R. M., Hällbrink, M. and Andaloussi, S. E. (2016) 'Lipid-based Transfection Reagents Exhibit Cryo-induced Increase in Transfection Efficiency', *Mol Ther Nucleic Acids,* 5, pp. e290.

Soto, C. and Pritzkow, S. (2018) 'Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases', *Nat Neurosci,* 21(10), pp. 1332-1340.

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(5870), pp. 1668-72.

Stankiewicz, T. R., Pena, C., Bouchard, R. J. and Linseman, D. A. (2020) 'Dysregulation of Rac or Rho elicits death of motor neurons and activation of these GTPases is altered in the G93A mutant hSOD1 mouse model of amyotrophic lateral sclerosis', *Neurobiol Dis,* 136, pp. 104743.

Statland, J. M., Barohn, R. J., McVey, A. L., Katz, J. S. and Dimachkie, M. M. (2015) 'Patterns of Weakness, Classification of Motor Neuron Disease, and Clinical Diagnosis of Sporadic Amyotrophic Lateral Sclerosis', *Neurol Clin,* 33(4), pp. 735-48.

Stelzer, G., Rosen, N., Plaschkes, I., Zimmerman, S., Twik, M., Fishilevich, S., Stein, T. I., Nudel, R., Lieder, I., Mazor, Y., Kaplan, S., Dahary, D., Warshawsky, D., Guan-Golan, Y., Kohn, A., Rappaport, N., Safran, M. and Lancet, D. (2016) 'The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses', *Curr Protoc Bioinformatics,* 54, pp. 1.30.1-1.30.33.

Stephens, B., Guiloff, R. J., Navarrete, R., Newman, P., Nikhar, N. and Lewis, P. (2006) 'Widespread loss of neuronal populations in the spinal ventral horn in sporadic motor neuron disease. A morphometric study', *J Neurol Sci,* 244(1-2), pp. 41-58.

Stil, A. and Drapeau, P. (2016) 'Neuronal labeling patterns in the spinal cord of adult transgenic Zebrafish', *Dev Neurobiol,* 76(6), pp. 642-60.

Studer, L., Vera, E. and Cornacchia, D. (2015) 'Programming and Reprogramming Cellular Age in the Era of Induced Pluripotency', *Cell Stem Cell,* 16(6), pp. 591-600.

Suk, T. R. and Rousseaux, M. W. C. (2020) 'The role of TDP-43 mislocalization in amyotrophic lateral sclerosis', *Mol Neurodegener,* 15(1), pp. 45.

Sullivan, P. M., Zhou, X., Robins, A. M., Paushter, D. H., Kim, D., Smolka, M. B. and Hu, F. (2016) 'The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway', *Acta Neuropathol Commun,* 4(1), pp. 51.

Sun, Z., Diaz, Z., Fang, X., Hart, M. P., Chesi, A., Shorter, J. and Gitler, A. D. (2011) 'Molecular determinants and genetic modifiers of aggregation and toxicity for the ALS disease protein FUS/TLS', *PLoS Biol,* 9(4), pp. e1000614.

Sundaram, S. R., Gowtham, L. and Nayak, B. S. (2012) 'The role of excitatory neurotransmitter glutamate in brain physiology and pathology', *Asian Journal of Pharmaceutical and Clinical Research,* 5(2).

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(3), pp. 427-443.

Swinnen, B. and Robberecht, W. (2014) 'The phenotypic variability of amyotrophic lateral sclerosis', *Nat Rev Neurol,* 10(11), pp. 661-70.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007) 'Induction of pluripotent stem cells from adult human fibroblasts by defined factors', *Cell,* 131(5), pp. 861-72.

Takahashi, K. and Yamanaka, S. (2006) 'Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors', *Cell,* 126(4), pp. 663-76.

Takahashi, Y., Fukuda, Y., Yoshimura, J., Toyoda, A., Kurppa, K., Moritoyo, H., Belzil, V. V., Dion, P. A., Higasa, K., Doi, K., Ishiura, H., Mitsui, J., Date, H., Ahsan, B., Matsukawa, T., Ichikawa, Y., Moritoyo, T., Ikoma, M., Hashimoto, T., Kimura, F., Murayama, S., Onodera, O., Nishizawa, M., Yoshida, M., Atsuta, N., Sobue, G., Fifita, J. A., Williams, K. L., Blair, I. P., Nicholson, G. A., Gonzalez-Perez, P., Brown, R. H., Jr., Nomoto, M., Elenius, K., Rouleau, G. A., Fujiyama, A., Morishita, S., Goto, J. and Tsuji, S. (2013) 'ERBB4 mutations that disrupt the neuregulin-ErbB4 pathway cause amyotrophic lateral sclerosis type 19', *Am J Hum Genet,* 93(5), pp. 900-5.

Takalo, M., Salminen, A., Soininen, H., Hiltunen, M. and Haapasalo, A. (2013) 'Protein aggregation and degradation mechanisms in neurodegenerative diseases', *Am J Neurodegener Dis,* 2(1), pp. 1-14.

Tan, C. F., Eguchi, H., Tagawa, A., Onodera, O., Iwasaki, T., Tsujino, A., Nishizawa, M., Kakita, A. and Takahashi, H. (2007) 'TDP-43 immunoreactivity in neuronal inclusions in familial amyotrophic lateral sclerosis with or without SOD1 gene mutation', *Acta Neuropathol,* 113(5), pp. 535-42.

Tan, R. H., Kril, J. J., McGinley, C., Hassani, M., Masuda-Suzukake, M., Hasegawa, M., Mito, R., Kiernan, M. C. and Halliday, G. M. (2016) 'Cerebellar neuronal loss in amyotrophic lateral sclerosis cases with ATXN2 intermediate repeat expansions', *Ann Neurol,* 79(2), pp. 295-305.

Tao, Q. Q. and Wu, Z. Y. (2017) 'Amyotrophic Lateral Sclerosis: Precise Diagnosis and Individualized Treatment', *Chin Med J (Engl): Vol. 19*, pp. 2269-72.

Tao, Z., Wang, H., Xia, Q., Li, K., Jiang, X., Xu, G., Wang, G. and Ying, Z. (2015) 'Nucleolar stress and impaired stress granule formation contribute to C9orf72 RAN translation-induced cytotoxicity', *Hum Mol Genet,* 24(9), pp. 2426-41.

Tashiro, Y., Urushitani, M., Inoue, H., Koike, M., Uchiyama, Y., Komatsu, M., Tanaka, K., Yamazaki, M., Abe, M., Misawa, H., Sakimura, K., Ito, H. and Takahashi, R. (2012) 'Motor neuron-specific disruption of proteasomes, but not autophagy, replicates amyotrophic lateral sclerosis', *J Biol Chem,* 287(51), pp. 42984-94.

Taylor, J. P., Tanaka, F., Robitschek, J., Sandoval, C. M., Taye, A., Markovic-Plese, S. and Fischbeck, K. H. (2003) 'Aggresomes protect cells by enhancing the degradation of toxic polyglutamine-containing protein', *Hum Mol Genet,* 12(7), pp. 749-57.

Tefera, T. W., Steyn, F. J., Ngo, S. T. and Borges, K. (2021) 'CNS glucose metabolism in Amyotrophic Lateral Sclerosis: a therapeutic target?', *Cell Biosci,* 11(1), pp. 14.

Teyssou, E., Takeda, T., Lebon, V., Boillée, S., Doukouré, B., Bataillon, G., Sazdovitch, V., Cazeneuve, C., Meininger, V., LeGuern, E., Salachas, F., Seilhean, D. and Millecamps, S. (2013) 'Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology', *Acta Neuropathol,* 125(4), pp. 511-22.

Thakor, K., Naud, S., Howard, D., Tandan, R. and Waheed, W. (2021) 'Effect of riluzole on weight in short-term and long-term survivors of amyotrophic lateral sclerosis', *Amyotroph Lateral Scler Frontotemporal Degener,* 22(5-6), pp. 360-367.

The Writing Group, E. M.-A. S. G. (2017) 'Safety and efficacy of edaravone in well defined patients with amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial', *Lancet Neurol,* 16(7), pp. 505-512.

Thielsen, K. D., Moser, J. M., Schmitt-John, T., Jensen, M. S., Jensen, K. and Holm, M. M. (2013) 'The Wobbler mouse model of amyotrophic lateral sclerosis (ALS) displays hippocampal hyperexcitability, and reduced number of interneurons, but no presynaptic vesicle release impairments', *PLoS One,* 8(12), pp. e82767.

Thompson, J. M., Yakhnitsa, V., Ji, G. and Neugebauer, V. (2018) 'Small conductance calcium activated potassium (SK) channel dependent and independent effects of riluzole on neuropathic pain-related amygdala activity and behaviors in rats', *Neuropharmacology,* 138, pp. 219-231.

Tohgi, H., Abe, T., Yamazaki, K., Murata, T., Ishizaki, E. and Isobe, C. (1999) 'Remarkable increase in cerebrospinal fluid 3-nitrotyrosine in patients with sporadic amyotrophic lateral sclerosis', *Ann Neurol,* 46(1), pp. 129-31.

Tollervey, J. R., Curk, T., Rogelj, B., Briese, M., Cereda, M., Kayikci, M., König, J., Hortobágyi, T., Nishimura, A. L., Zupunski, V., Patani, R., Chandran, S., Rot, G., Zupan, B., Shaw, C. E. and Ule, J. (2011) 'Characterizing the RNA targets and position-dependent splicing regulation by TDP-43', *Nat Neurosci,* 14(4), pp. 452-8.

Traynor, B. J., Alexander, M., Corr, B., Frost, E. and Hardiman, O. (2003) 'An outcome study of riluzole in amyotrophic lateral sclerosis--a population-based study in Ireland, 1996-2000', *J Neurol,* 250(4), pp. 473-9.

Troakes, C., Maekawa, S., Wijesekera, L., Rogelj, B., Siklós, L., Bell, C., Smith, B., Newhouse, S., Vance, C., Johnson, L., Hortobágyi, T., Shatunov, A., Al-Chalabi, A., Leigh, N., Shaw, C. E., King, A. and Al-Sarraj, S. (2012) 'An MND/ALS phenotype associated with C9orf72 repeat expansion: abundant p62-positive, TDP-43-negative inclusions in cerebral cortex, hippocampus and cerebellum but without associated cognitive decline', *Neuropathology,* 32(5), pp. 505-14.

Tropepe, V. and Sive, H. L. (2003) 'Can zebrafish be used as a model to study the neurodevelopmental causes of autism?', *Genes Brain Behav,* 2(5), pp. 268-81.

Trotti, D., Rolfs, A., Danbolt, N. C., Brown, R. H. and Hediger, M. A. (1999) 'SOD1 mutants linked to amyotrophic lateral sclerosis selectively inactivate a glial glutamate transporter', *Nat Neurosci,* 2(9), pp. 848.

Turner, M. R., Hardiman, O., Benatar, M., Brooks, B. R., Chio, A., de Carvalho, M., Ince, P. G., Lin, C., Miller, R. G., Mitsumoto, H., Nicholson, G., Ravits, J., Shaw, P. J., Swash, M., Talbot, K., Traynor, B. J., Van den Berg, L. H., Veldink, J. H., Vucic, S. and Kiernan, M. C. (2013) 'Controversies and priorities in amyotrophic lateral sclerosis', *Lancet Neurol,* 12(3), pp. 310-322.

Turner, M. R. and Kiernan, M. C. (2012) 'Does interneuronal dysfunction contribute to neurodegeneration in amyotrophic lateral sclerosis?', *Amyotroph Lateral Scler,* 13(3), pp. 245-50.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., Olsson, I., Edlund, K., Lundberg, E., Navani, S., Szigyarto, C. A., Odeberg, J., Djureinovic, D., Takanen, J. O., Hober, S., Alm, T., Edqvist, P. H., Berling, H., Tegel, H., Mulder, J., Rockberg, J., Nilsson, P., Schwenk, J. M., Hamsten, M., von Feilitzen, K., Forsberg, M., Persson, L., Johansson, F., Zwahlen, M., von Heijne, G., Nielsen, J. and Pontén, F. (2015) 'Proteomics. Tissue-based map of the human proteome', *Science,* 347(6220), pp. 1260419.

Van Acker, T., Tavernier, J. and Peelman, F. (2019) 'The Small GTPase Arf6: An Overview of Its Mechanisms of Action and of Its Role in Host‑Pathogen Interactions and Innate Immunity', *Int J Mol Sci,* 20(9).

Van Aelst, L., Joneson, T. and Bar-Sagi, D. (1996) 'Identification of a novel Rac1-interacting protein involved in membrane ruffling', *EMBO J,* 15(15), pp. 3778-86.

van Blitterswijk, M., DeJesus-Hernandez, M. and Rademakers, R. (2012) 'How do C9ORF72 repeat expansions cause ALS and FTD: can we learn from other non-coding repeat expansion disorders?', *Curr Opin Neurol,* 25(6), pp. 689-700.

Van Damme, P., Robberecht, W. and Van Den Bosch, L. (2017) 'Modelling amyotrophic lateral sclerosis: progress and possibilities', *Dis Model Mech,* 10(5), pp. 537-549.

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', *Lancet,* 390(10107), pp. 2084-2098.

Van Langenhove, T., van der Zee, J. and Van Broeckhoven, C. (2012) 'The molecular basis of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum', *Ann Med,* 44(8), pp. 817-28.

Van Mossevelde, S., van der Zee, J., Cruts, M. and Van Broeckhoven, C. (2017) 'Relationship between C9orf72 repeat size and clinical phenotype', *Curr Opin Genet Dev,* 44, pp. 117-124.

Vance, C., Rogelj, B., Hortobagyi, T., De Vos, K. J., Nishimura, A. L., Sreedharan, J., Hu, X., Smith, B., Ruddy, D., Wright, P., Ganesalingam, J., Williams, K. L., Tripathi, V., Al-Saraj, S., Al-Chalabi, A., Leigh, P. N., Blair, I. P., Nicholson, G., de Belleroche, J., Gallo, J. M., Miller, C. C. and Shaw, C. E. (2009) 'Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6', *Science,* 323(5918), pp. 1208-1211.

Vance, C., Scotter, E. L., Nishimura, A. L., Troakes, C., Mitchell, J. C., Kathe, C., Urwin, H., Manser, C., Miller, C. C., Hortobágyi, T., Dragunow, M., Rogelj, B. and Shaw, C. E. (2013) 'ALS mutant FUS disrupts nuclear localization and sequesters wild-type FUS within cytoplasmic stress granules', *Hum Mol Genet,* 22(13), pp. 2676-88.

Vandoorne, T., De Bock, K. and Van Den Bosch, L. (2018) 'Energy metabolism in ALS: an underappreciated opportunity?', *Acta Neuropathol,* 135(4), pp. 489-509.

Vicencio, E., Beltrán, S., Labrador, L., Manque, P., Nassif, M. and Woehlbier, U. (2020) 'Implications of Selective Autophagy Dysfunction for ALS Pathology', *Cells,* 9(2).

Visser, J., van den Berg-Vos, R. M., Franssen, H., van den Berg, L. H., Wokke, J. H., de Jong, J. M., Holman, R., de Haan, R. J. and de Visser, M. (2007) 'Disease course and prognostic factors of progressive muscular atrophy', *Arch Neurol,* 64(4), pp. 522-8.

Vucic, S., Rothstein, J. D. and Kiernan, M. C. (2014) 'Advances in treating amyotrophic lateral sclerosis: insights from pathophysiological studies', *Trends Neurosci,* 37(8), pp. 433-42.

Waelter, S., Boeddrich, A., Lurz, R., Scherzinger, E., Lueder, G., Lehrach, H. and Wanker, E. E. (2001) 'Accumulation of mutant huntingtin fragments in aggresome-like inclusion bodies as a result of insufficient protein degradation', *Mol Biol Cell,* 12(5), pp. 1393-407.

Wagle-Shukla, A., Ni, Z., Gunraj, C. A., Bahl, N. and Chen, R. (2009) 'Effects of short interval intracortical inhibition and intracortical facilitation on short interval intracortical facilitation in human primary motor cortex', *J Physiol,* 587(Pt 23), pp. 5665-78.

Waite, A. J., Bäumer, D., East, S., Neal, J., Morris, H. R., Ansorge, O. and Blake, D. J. (2014) 'Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion', *Neurobiol Aging,* 35(7), pp. 1779.e5-1779.e13.

Walhout, R., Verstraete, E., van den Heuvel, M. P., Veldink, J. H. and van den Berg, L. H. (2017) 'Patterns of symptom development in patients with motor neuron disease', *Amyotroph Lateral Scler Frontotemporal Degener,* 19(1-2), pp. 21-28.

Walker, A. K., Spiller, K. J., Ge, G., Zheng, A., Xu, Y., Zhou, M., Tripathy, K., Kwong, L. K., Trojanowski, J. Q. and Lee, V. M. (2015) 'Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43', *Acta Neuropathol,* 130(5), pp. 643-60.

Wang, H., O'Reilly É, J., Weisskopf, M. G., Logroscino, G., McCullough, M. L., Thun, M., Schatzkin, A., Kolonel, L. N. and Ascherio, A. (2011) 'Smoking and risk of amyotrophic lateral sclerosis: a pooled analysis of five prospective cohorts', *Arch Neurol,* 68(2), pp. 207-213.

Wang, M. D., Gomes, J., Cashman, N. R., Little, J. and Krewski, D. (2014) 'A Meta-Analysis of Observational Studies of the Association Between Chronic Occupational Exposure to Lead and Amyotrophic Lateral Sclerosis', *Journal of Occupational and Environmental Medicine,* 56(12), pp. 1235-1242.

Wang, W. Y., Pan, L., Su, S. C., Quinn, E. J., Sasaki, M., Jimenez, J. C., Mackenzie, I. R., Huang, E. J. and Tsai, L. H. (2013) 'Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons', *Nat Neurosci,* 16(10), pp. 1383-91.

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(15), pp. 1656-76.

Webster, C. P., Smith, E. F., Grierson, A. J. and De Vos, K. J. (2018) 'C9orf72 plays a central role in Rab GTPase-dependent regulation of autophagy', *Small GTPases,* 9(5), pp. 399-408.

Webster, C. P., Smith, E. F., Shaw, P. J. and De Vos, K. J. (2017) 'Protein Homeostasis in Amyotrophic Lateral Sclerosis: Therapeutic Opportunities?', *Front Mol Neurosci,* 10, pp. 123.

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(6), pp. 1213-25.

Westerfield, M. (2000) *A guide for the laboratory use of zebrafish (Danio rerio) 4. The Zebrafish Book*: Eugene: Univ. of Oregon Press.

White, M. R., Mitrea, D. M., Zhang, P., Stanley, C. B., Cassidy, D. E., Nourse, A., Phillips, A. H., Tolbert, M., Taylor, J. P. and Kriwacki, R. W. (2019) 'C9orf72 Poly(PR) Dipeptide Repeats Disturb Biomolecular Phase Separation and Disrupt Nucleolar Function', *Mol Cell,* 74(4), pp. 713-728.e6.

Wiedemann, F. R., Manfredi, G., Mawrin, C., Beal, M. F. and Schon, E. A. (2002) 'Mitochondrial DNA and respiratory chain function in spinal cords of ALS patients', *J Neurochem,* 80(4), pp. 616-25.

Wijesekera, L. C. and Leigh, P. N. (2009) 'Amyotrophic lateral sclerosis', *Orphanet J Rare Dis*, pp. 3.

Williams, K. L., Topp, S., Yang, S., Smith, B., Fifita, J. A., Warraich, S. T., Zhang, K. Y., Farrawell, N., Vance, C., Hu, X., Chesi, A., Leblond, C. S., Lee, A., Rayner, S. L., Sundaramoorthy, V., Dobson-Stone, C., Molloy, M. P., van Blitterswijk, M., Dickson, D. W., Petersen, R. C., Graff-Radford, N. R., Boeve, B. F., Murray, M. E., Pottier, C., Don, E., Winnick, C., McCann, E. P., Hogan, A., Daoud, H., Levert, A., Dion, P. A., Mitsui, J., Ishiura, H., Takahashi, Y., Goto, J., Kost, J., Gellera, C., Gkazi, A. S., Miller, J., Stockton, J., Brooks, W. S., Boundy, K., Polak, M., Muñoz-Blanco, J. L., Esteban-Pérez, J., Rábano, A., Hardiman, O., Morrison, K. E., Ticozzi, N., Silani, V., de Belleroche, J., Glass, J. D., Kwok, J. B. J., Guillemin, G. J., Chung, R. S., Tsuji, S., Brown, R. H., García-Redondo, A., Rademakers, R., Landers, J. E., Gitler, A. D., Rouleau, G. A., Cole, N. J.,

Yerbury, J. J., Atkin, J. D., Shaw, C. E., Nicholson, G. A. and Blair, I. P. (2016) 'CCNF mutations in amyotrophic lateral sclerosis and frontotemporal dementia', *Nat Commun*.

Wils, H., Kleinberger, G., Janssens, J., Pereson, S., Joris, G., Cuijt, I., Smits, V., Ceuterick-de Groote, C., Van Broeckhoven, C. and Kumar-Singh, S. (2010) 'TDP-43 transgenic mice develop spastic paralysis and neuronal inclusions characteristic of ALS and frontotemporal lobar degeneration', *Proc Natl Acad Sci U S A,* 107(8), pp. 3858-63.

Winton, M. J., Igaz, L. M., Wong, M. M., Kwong, L. K., Trojanowski, J. Q. and Lee, V. M. (2008) 'Disturbance of nuclear and cytoplasmic TAR DNA-binding protein (TDP-43) induces disease-like redistribution, sequestration, and aggregate formation', *J Biol Chem,* 283(19), pp. 13302-9.

Witzel, S., Maier, A., Steinbach, R., Grosskreutz, J., Koch, J. C., Sarikidi, A., Petri, S., Günther, R., Wolf, J., Hermann, A., Prudlo, J., Cordts, I., Lingor, P., Löscher, W. N., Kohl, Z., Hagenacker, T., Ruckes, C., Koch, B., Spittel, S., Günther, K., Michels, S., Dorst, J., Meyer, T., Ludolph, A. C. and (MND-NET), G. M. N. D. N. (2022) 'Safety and Effectiveness of Long-term Intravenous Administration of Edaravone for Treatment of Patients With Amyotrophic Lateral Sclerosis', *JAMA Neurol,* 79(2), pp. 121-130.

Wiznerowicz, M. and Trono, D. (2003) 'Conditional suppression of cellular genes: lentivirus vector-mediated drug-inducible RNA interference', *J Virol,* 77(16), pp. 8957-61.

Wolf, J., Safer, A., Wöhrle, J. C., Palm, F., Nix, W. A., Maschke, M. and Grau, A. J. (2014) 'Variability and prognostic relevance of different phenotypes in amyotrophic lateral sclerosis - data from a population-based registry', *J Neurol Sci,* 345(1-2), pp. 164-7.

Wood, J. D., Beaujeux, T. P. and Shaw, P. J. (2003) 'Protein aggregation in motor neurone disorders', *Neuropathol Appl Neurobiol,* 29(6), pp. 529-45.

Wood, J. D., Landers, J. A., Bingley, M., McDermott, C. J., Thomas-McArthur, V., Gleadall, L. J., Shaw, P. J. and Cunliffe, V. T. (2006) 'The microtubule-severing protein Spastin is essential for axon outgrowth in the zebrafish embryo', *Hum Mol Genet,* 15(18), pp. 2763-71.

Woods, N. B., Muessig, A., Schmidt, M., Flygare, J., Olsson, K., Salmon, P., Trono, D., von Kalle, C. and Karlsson, S. (2003) 'Lentiviral vector transduction of NOD/SCID repopulating cells results in multiple vector integrations per transduced cell: risk of insertional mutagenesis', *Blood,* 101(4), pp. 1284-9.

Wu, C. H., Fallini, C., Ticozzi, N., Keagle, P. J., Sapp, P. C., Piotrowska, K., Lowe, P., Koppers, M., McKenna-Yasek, D., Baron, D. M., 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, R. H., Jr. and Landers, J. E. (2012) 'Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis', *Nature,* 488(7412), pp. 499-503.

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(19), pp. 7778-83.

Yamakawa, M., Ito, D., Honda, T., Kubo, K., Noda, M., Nakajima, K. and Suzuki, N. (2015) 'Characterization of the dipeptide repeat protein in the molecular pathogenesis of c9FTD/ALS', *Hum Mol Genet,* 24(6), pp. 1630-45.

Yamashita, S. and Ando, Y. (2015) 'Genotype-phenotype relationship in hereditary amyotrophic lateral sclerosis', *Transl Neurodegener,* 4, pp. 13.

Yang, X., Ji, Y., Wang, W., Zhang, L., Chen, Z., Yu, M., Shen, Y., Ding, F., Gu, X. and Sun, H. (2021) 'Amyotrophic Lateral Sclerosis: Molecular Mechanisms, Biomarkers, and Therapeutic Strategies', *Antioxidants (Basel),* 10(7).

Yang, Y., Hentati, A., Deng, H. X., Dabbagh, O., Sasaki, T., Hirano, M., Hung, W. Y., Ouahchi, K., Yan, J., Azim, A. C., Cole, N., Gascon, G., Yagmour, A., Ben-Hamida, M., Pericak-Vance, M., Hentati, F. and Siddique, T. (2001) 'The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis', *Nat Genet,* 29(2), pp. 160-5.

Yunusova, Y., Plowman, E. K., Green, J. R., Barnett, C. and Bede, P. (2019) 'Clinical Measures of Bulbar Dysfunction in ALS', *Front Neurol,* 10, pp. 106.

Zamft, B., Bintu, L., Ishibashi, T. and Bustamante, C. (2012) 'Nascent RNA structure modulates the transcriptional dynamics of RNA polymerases', *Proc Natl Acad Sci U S A,* 109(23), pp. 8948-53.

Zarei, S., Carr, K., Reiley, L., Diaz, K., Guerra, O., Altamirano, P. F., Pagani, W., Lodin, D., Orozco, G. and Chinea, A. (2015) 'A comprehensive review of amyotrophic lateral sclerosis', *Surgical Neurology International*, pp. 171.

Zhang, B., Pan, C., Feng, C., Yan, C., Yu, Y., Chen, Z., Guo, C. and Wang, X. (2022) 'Role of mitochondrial reactive oxygen species in homeostasis regulation', *Redox Rep,* 27(1), pp. 45-52.

Zhang, F., Hackett, N. R., Lam, G., Cheng, J., Pergolizzi, R., Luo, L., Shmelkov, S. V., Edelberg, J., Crystal, R. G. and Rafii, S. (2003) 'Green fluorescent protein selectively induces HSP70-mediated up-regulation of COX-2 expression in endothelial cells', *Blood,* 102(6), pp. 2115-21.

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(7567), pp. 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', *Nat Med,* 24(8), pp. 1136-1142.

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., 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 Neuropathol,* 128(4), pp. 505-24.

Zhang, Y. J., Xu, Y. F., Cook, C., Gendron, T. F., Roettges, P., Link, C. D., Lin, W. L., Tong, J., Castanedes-Casey, M., Ash, P., Gass, J., Rangachari, V., Buratti, E., Baralle, F., Golde, T. E., Dickson, D. W. and Petrucelli, L. (2009) 'Aberrant cleavage of TDP-43 enhances aggregation and cellular toxicity', *Proc Natl Acad Sci U S A,* 106(18), pp. 7607-12.

Zhenfei, L., Shiru, D., Xiaomeng, Z., Cuifang, C. and Yaling, L. (2019) 'Discontiguous or Contiguous Spread Patterns Affect the Functional Staging in Patients With Sporadic Amyotrophic Lateral Sclerosis', *Front Neurol,* 10, pp. 523.

Zoing, M. C., Burke, D., Pamphlett, R. and Kiernan, M. C. (2006) 'Riluzole therapy for motor neurone disease: an early Australian experience (1996-2002)', *J Clin Neurosci,* 13(1), pp. 78-83.

Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O., Lewis, J., Miller, T. M., Harms, M. B., Falchook, A. E., Subramony, S. H., 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(51), pp. E4968-77.

8. APPENDICES

8.1 Appendix A

8.1.1 The use of genetically modified organisms and transgenic animals training record

Department of Neuroscience
Use of Genetically Modified Organisms and Transgenic Animals –Training Records

<u>NAME:</u> ANUSHKA BHARGAV <u>STATUS:</u> PHD STUDENT   <u>DATE:</u> 26/3/19

1. I have read and understood the Safety Services training notes on use of GMOs and transgenic organisms:
    Signed: *Anoushka Bhargava*

2. I have received training from my PI (or delegated person) on the GM project(s) I will be working on. GM training should include:
    Details about the biohazard.
    The findings of any risk assessments undertaken.
    The control measures required to prevent exposure to this material.
    Training in good microbiological practice
    Familiarisation with the local rules
    Training in emergency procedures

    Signed: *Anoushka Bhargava.*

    Signed by trainer: *[signature]*

    Name of Trainer: DR KE NING

3. I have read and signed the Risk Assessment(s) relating to the GM projects (s) I am working on.
    Signed: *Anoushka Bhargava.*

4. GM Licence numbers and titles for procedures I will be working on:
    GMO Project Number: GMO 2006_07B
    Lentiviral vector –based gene transfer to transduced terminally differentiated cells and neuronal cells using both in vitro and in vivo methods of neurological disorders.

5. The Licence holder(s) of the licence(s) I will be working on:
    Professor Mimoun Azzouz
    Dr. Adrian Higginbottom.

    <u>RETURN TO ANNE GREGORY BEFORE WORKING WITH GM ORGANISMS OR TRANSGENIC ANIMALS.</u>

    iPS cell culture : 2001_14aGMAG (Neurodegeneration Cell culture models Feb 2018 update)
    47
    GM Zebrafish : ~~7018058~~ GMO2012-13 and GMO2012-15

8.1.2 Ethical approval for the use of human brain tissue

The University Of Sheffield.

SITraN
Sheffield Institute for Translational Neuroscience

Department of Neuroscience
Faculty of Medicine, Dentistry & Health

4th June 2014

Alex Harris
International Strategy Manager
Medical Research Council
14th Floor
One Kemble Street
London
WC2B 4AN

Professor Paul Ince MBBS MD FRCPath

Professor of Neuropathology,
Head of Neuroscience Department
Consultant Neuropathologist
Director of the Sheffield Brain Tissue Bank

Sheffield Institute of Translational Neuroscience
University of Sheffield
Room B35
385a Glossop Road
Sheffield
S10 2HQ

Tel: +44 (0)114 22 22234
Fax: +44 (0)114 22 22290
Email:p.g.ince@sheffield.ac.uk

Dear Mr Harris,

Re: Investigating mechanistic causes of C9ORF72-related amyotrophic lateral sclerosis
(ALS) (Principal Investigator: Dr Ke Ning, Submitter's Reference: MR/M010864/1)

I am the Director of the Sheffield Brain Tissue Bank and pleased to give permission to the above project to access the Bank. The Bank has been approved by the Scotland A Reserch Ethics Committee (REC reference: 08/MRE00/103). Please see attachment for details.

Please do not hesitate to let me know if you need further information.

With best wishes,

Yours sincerely,

Professor Paul Ince

Professor of Neuropathology

2011
THE AWARDS
AWARD WINNER
UNIVERSITY OF THE YEAR

8.2   Appendix 1

8.2.1   Sequences

pLV-FL

TGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCT

ACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGG

GTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCA

GATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGT

GAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTT

GACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTT

CAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCC

AGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTG

CATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCT

GAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAA

AGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT

AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGC

GCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAG

GACTCGGCTTGCTGAAGCGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTG

GTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTG

CGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTC

GGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAA

GCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA

GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGA

AGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGG

ATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAAC

AAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAG

GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT

TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAG

AAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA

GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATT

ATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGC

AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGA

ATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG

TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG

TAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGA

GAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACC
AGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGT
GGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATA
GTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATA
GAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGA
GGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAG
AGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGGTTAACTTTTAA
AAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAA
TAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAA
TTTTATCGATGGTCGAGTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCT
GGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCC
TCTGGCCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGCTCCGTTC
TTTGGTGGCCCCTTCGCGCCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCC
CCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTC
TCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGC
CTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAG
GCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGGCGGGCTCAGGGGCG
GGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAA
GCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCT
CTAGGATCTGCCACCATGACGGACGGGATCCTAGGGAAGGCAGCCACAATGG
AGATCCCTATCCACGGGAACGGCGAAGCCAGGCAGCTTCCTGAAGATGATGGG
CTGGAGCAGGACCTCCAGCAGGTGATGGTGTCAGGACCCAACCTCAATGAAAC
CAGCATTGTGTCTGGTGGCTATGGGGGCTCTGGTGATGGACTCATCCCCACAG
GGTCTGGCCGCCATCCATCTCACAGCACCACTCCTTCTGGCCCTGGAGATGAG
GTGGCTCGGGGCATTGCTGGAGAAAGTTTGACATCGTCAAGAAATGGGGCAT
CAACACCTATAAGTGCACAAAGCAACTGTTATCAGAACGATTTGGTCGAGGCTC
ACGGACTGTGGACCTGGAGCTAGAGCTGCAGATTGAGTTGCTGCGTGAGACGA
AGCGCAAGTATGAGAGTGTCCTGCAGCTGGGCCGGGCACTGACAGCCCACCT
CTACAGCCTGCTGCAGACCCAGCATGCACTGGGTGATGCCTTTGCTGACCTCA
GCCAGAAGTCCCCAGAGCTTCAGGAGGAATTTGGCTACAATGCAGAGACACAG
AAACTACTATGCAAGAATGGGGAAACGCTGCTAGGAGCCGTGAACTTCTTTGTC
TCTAGCATCAACACATTGGTCACCAAGACCATGGAAGACACGCTCATGACTGTG
AAACAGTATGAGGCTGCCAGGCTGGAATATGATGCCTACCGAACAGACTTAGA
GGAGCTGAGTCTAGGCCCCCGGGATGCAGGGACACGTGGTCGACTTGAGAGT

GCCCAGGCCACTTTCCAGGCCCATCGGGACAAGTATGAGAAGCTGCGGGGAG
ATGTGGCCATCAAGCTCAAGTTCCTGGAAGAAAACAAGATCAAGGTGATGCACA
AGCAGCTGCTGCTCTTCCACAATGCTGTGTCCGCCTACTTTGCTGGGAACCAG
AAACAGCTGGAGCAGACCCTGCAGCAGTTCAACATCAAGCTGCGGCCTCCAGG
AGCTGAGAAACCCTCCTGGCTAGAGGAGCAGTACCCATACGATGTTCCAGATT
ACGCTTGACTCGAGGGAATTCCGATAATCAACCTCTGGATTACAAAATTTGTGA
AAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGGATACGCT
GCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTC
CTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAG
GCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGG
GCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTA
TTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGC
TCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCT
TTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTC
TGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCT
GCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGA
TCTCCCTTTGGGCCGCCTCCCCGCATCGGGAATTCGAGCTCGGTACCTTTAAA
ACCAATGACTTACAAGGCAGCTGTAAATCTTAGCCACTTTTTAAAAGAAAAGGG
GGGACTGGAAGGGCTAATTCACTCCCAACGAAAACAAAATCTGCTTTTTGCTTG
TACTGGGTCTCTCTGGTTAGACCAAATCTGAGCCTGGGAGCTCTCTGGCTAACT
AGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAAGTAGT
GTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAAACCCTTTTA
GTCAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCCGTGTATTCTATAGT
GTCACCTAAATCGTATGTGTATGATACATAAGGTTATGTATTAATTGTAGCCGCG
TTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTGGCTTACTGAAGCA
GACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTTTGGTTGGACGAAC
CTTCTGAGTTTCTGGTAACGCCGTCCCGCACCCGGAAATGGTCAGCGAACCAA
TCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACGCATCGTGGCCGG
CATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACC
GATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGT
GGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTG
CATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGGG
CTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGATATGGTGCACTCT
CAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAA

CACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA

CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATC

ACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAA

TGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGT

GCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTC

ATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGA

GTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCC

TGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTT

GGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGATCCTTG

AGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCT

ATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCC

GCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGC

ATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGA

GTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAG

CTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGG

GAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCC

TGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCT

AGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGAC

CACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAG

CCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA

GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATG

AACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAAC

TGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAA

TTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTT

AACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGAT

CTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACC

ACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCC

GAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTA

GCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCG

CTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTA

CCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTG

AACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAA

CTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAG

AAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACG

AGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCG
CCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCC
TATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGC
CTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTAT
TACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGC
AGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTC
TCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGTGTGTCAGTTAG
GGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCAT
CTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAG
AAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAAC
TCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATG
GCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGC
TATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGC
TTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCACTTTATCCCGCG
TCAGGGAGAGGCAGTGCGTAAAAAGACGCGGACTCATGTGAAATACTGGTTTT
TAGTGCGCCAGATCTCTATAATCTCGCGCAACCTATTTTCCCCTCGAACACTTTT
TAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAAAAATACATCGTC
ACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGCCGACTGATGCC
TTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGGTCTGTACCGGG
TGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATACCGTTTGTATTTC
CAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAAACGCGCAGAAG
GCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTACT
GCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCGCA
AAACCGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAGAT
ACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTCCCGCCAATCTC
CGGTCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGTTCTTTTTAACTT
CAGGCGGGTTACAATAGTTTCCAGTAAGTATTCTGGAGGCTGCATCCATGACAC
AGGCAAACCTGAGCGAAACCCTGTTCAAACCCCGCTTTAAACATCCTGAAACCT
CGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTGTGCAGTCGGCC
CTTGATGGTAAAACCATCCCTCACTGGTATCGCATGATTAACCGTCTGATGTGG
ATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAGGCACGTATTGT
GATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACGGTGATTGGCT
ACCGTGGCGGCAACTGGATTTATGAGTGGGCCCCGGATCTTTGTGAAGGAACC
TTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCT

AAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATTGT
TTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGGAA
TGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGAT
GAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTA
GAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGT
TTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAAGGAAAAAGCTGC
ACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAGGCAT
AACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGAGTGT
CTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTGTAAA
GGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAATCAG
CCATACCACATTTGTAGAGCTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCC
CTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAG
CTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTT
TTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATCATGT
GTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAACTTCCCACCCC
ATACCCTATTACCACTGCCAAATTACCTGTGGTTTCATTTACTCTAAACCTGTGA
TTCCTCTGAATTATTTTCATTTTAAAGAATTGTATTTGTTAAATATGTACTACAAA
CTTAGTAGT

pLV-HC

TGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCT
ACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGG
GTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCA
GATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGT
GAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTT
GACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTT
CAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCC
AGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTG
CATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCT
GAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAA
AGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT
AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGC
GCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAG
GACTCGGCTTGCTGAAGCGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTG

GTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTG
CGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTC
GGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCAA
GCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGA
AGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGG
ATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAAC
AAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAG
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAG
AAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATT
ATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGC
AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGA
ATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG
TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGA
GAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACC
AGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGT
GGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATA
GTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATA
GAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGA
GGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAG
AGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGGTTAACTTTTAA
AAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAA
TAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAA
TTTTATCGATGGTCGAGTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCT
GGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCC
TCTGGCCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGCTCCGTTC
TTTGGTGGCCCCTTCGCGCCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCC
CCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTC
TCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGC
CTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAG
GCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGGCGGGCTCAGGGGCG

GGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAA
GCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCT
CTAGGATCCGCCACCATGCTCAGCCAGAAGTCCCCAGAGCTTCAGGAGGAATT
TGGCTACAATGCAGAGACACAGAAACTACTATGCAAGAATGGGGAAACGCTGC
TAGGAGCCGTGAACTTCTTTGTCTCTAGCATCAACACATTGGTCACCAAGACCA
TGGAAGACACGCTCATGACTGTGAAACAGTATGAGGCTGCCAGGCTGGAATAT
GATGCCTACCGAACAGACTTAGAGGAGCTGAGTCTAGGCCCCCGGGATGCAG
GGACACGTGGTCGACTTGAGAGTGCCCAGGCCACTTTCCAGGCCCATCGGGA
CAAGTATGAGAAGCTGCGGGGAGATGTGGCCATCAAGCTCAAGTTCCTGGAAG
AAAACAAGATCAAGGTGATGCACAAGCAGCTGCTGCTCTTCCACAATGCTGTGT
CCGCCTACTTTGCTGGGAACCAGAAACAGCTGGAGCAGACCCTGCAGCAGTTC
AACATCAAGCTGCGGCCTCCAGGAGCTGAGAAACCCTCCTGGCTAGAGGAGCA
GTACCCATACGATGTTCCAGATTACGCTTGACTCGAGGGAATTCCGATAATCAA
CCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTC
CTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTC
CCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTAT
GAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGC
TGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCG
GGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGC
CTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGT
GTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCT
GGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCG
GACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCG
CCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCATCGGG
AATTCGAGCTCGGTACCTTTAAAACCAATGACTTACAAGGCAGCTGTAAATCTTA
GCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGA
AAACAAAATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAAATCTGAG
CCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCT
TGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACT
AGAGATCCCTCAAACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATCTAGA
ATTAATTCCGTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGATACATAAG
GTTATGTATTAATTGTAGCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTG
TAGCATCTGGCTTACTGAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCA
GAGTCGGTTTGGTTGGACGAACCTTCTGAGTTTCTGGTAACGCCGTCCCGCAC

CCGGAAATGGTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTA
CGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGC
GCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCT
CATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGA
CTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAA
CGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAG
AGCGTCGATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAG
CCAGCCCCGACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTG
CTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTG
TCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGA
TACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGG
TGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATA
CATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAAT
ATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTT
TTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGT
AAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATC
TCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGA
TGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCG
GGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGT
ACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTAT
GCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAA
CGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCAT
GTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGA
CGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTAT
TAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGG
AGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTG
GTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTG
CAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACG
GGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGC
CTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAG
ATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGA
TAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGA
CCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGAT

CAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATA
CCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCT
GTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCC
AGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGA
TAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTG
GAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGG
GTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATC
TTTATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGAT
GCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTT
ACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCC
CCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGC
CGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAG
CGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAG
CTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG
CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTC
CCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAG
CAACCATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGT
TCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCC
GAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGG
AGGCCTAGGCTTTTGCAAAAAGCTTGGACACAAGACAGGCTTGCGAGATATGTT
TGAGAATACCACTTTATCCCGCGTCAGGGAGAGGCAGTGCGTAAAAAGACGCG
GACTCATGTGAAATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCGCAA
CCTATTTTCCCCTCGAACACTTTTTAAGCCGTAGATAAACAGGCTGGGACACTT
CACATGAGCGAAAAATACATCGTCACCTGGGACATGTTGCAGATCCATGCACGT
AAACTCGCAAGCCGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTA
AGCCGTGGCGGTCTGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGT
CATGTCGATACCGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTT
AAAGTGCTGAAACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGA
CCTGGTGGATACCGGTGGTACTGCGGTTGCGATTCGTGAAATGTATCCAAAAG
CGCACTTTGTCACCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACT
ATGTTGTTGATATCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGC
GTCGTATTCGTCCCGCCAATCTCCGGTCGCTAATCTTTTCAACGCCTGGCACTG
CCGGGCGTTGTTCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCAGTAAGTATT

CTGGAGGCTGCATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAACC

CCGCTTTAAACATCCTGAAACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGC

ACAACCGCCTGTGCAGTCGGCCCTTGATGGTAAAACCATCCCTCACTGGTATC

GCATGATTAACCGTCTGATGTGGATCTGGCGCGGCATTGACCCACGCGAAATC

CTCGACGTCCAGGCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGA

TTTATACGATACGGTGATTGGCTACCGTGGCGGCAACTGGATTTATGAGTGGG

CCCCGGATCTTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAA

CTACCTACAGAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATG

TGTTAAACTACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTG

ATGAATGGGAGCAGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAG

AAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTC

CTCCAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGC

TAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATT

TACACCACAAAGGAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATT

CTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTT

ACTCCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTA

CCTTTAGCTTTTTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGC

CTTGACTAGAGATCATAATCAGCCATACCACATTTGTAGAGCTTTTACTTGCTTT

AAAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTT

GTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA

CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAA

CTCATCAATGTATCTTATCATGTGTGGATCAACTGGATAACTCAAGCTAACCAAA

ATCATCCCAAACTTCCCACCCCATACCCTATTACCACTGCCAAATTACCTGTGGT

TTCATTTACTCTAAACCTGTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGT

ATTTGTTAAATATGTACTACAAACTTAGTAGT

pLV-GFP

TGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCT

ACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGGG

GTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCCA

GATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGT

GAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTT

GACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTT

CAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCC

AGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTG
CATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCT
GAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAA
AGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT
AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGC
GCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAG
GACTCGGCTTGCTGAAGCGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTG
GTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTG
CGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATTC
GGTTAAGGCCAGGGGGAAAGAAAAATATAAATTAAAACATATAGTATGGGCAA
GCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGA
AGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGG
ATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAC
AAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAG
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAG
AAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACAGGCCAGACAATT
ATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGC
AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGA
ATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG
TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTGGAATCACACGACCTGGATGGAGTGGGACAGA
GAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAACC
AGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGT
GGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGATA
GTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAATA
GAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGA
GGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAG
AGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGGTTAACTTTTAA
AAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAA
TAGCAACAGACATACAAACTAAAGAATTACAAAAACAAATTACAAAAATTCAAAA
TTTTATCGATGGTCGAGTACCGGGTAGGGGAGGCGCTTTTCCCAAGGCAGTCT

GGAGCATGCGCTTTAGCAGCCCCGCTGGGCACTTGGCGCTACACAAGTGGCC

TCTGGCCTCGCACACATTCCACATCCACCGGTAGGCGCCAACCGGCTCCGTTC

TTTGGTGGCCCCTTCGCGCCACCTTCTACTCCTCCCCTAGTCAGGAAGTTCCCC

CCCGCCCCGCAGCTCGCGTCGTGCAGGACGTGACAAATGGAAGTAGCACGTC

TCACTAGTCTCGTGCAGATGGACAGCACCGCTGAGCAATGGAAGCGGGTAGGC

CTTTGGGGCAGCGGCCAATAGCAGCTTTGCTCCTTCGCTTTCTGGGCTCAGAG

GCTGGGAAGGGGTGGGTCCGGGGGCGGGCTCAGGGGCGGGCTCAGGGGCG

GGGCGGGCGCCCGAAGGTCCTCCGGAGGCCCGGCATTCTGCACGCTTCAAAA

GCGCACGTCTGCCGCGCTGTTCTCCTCTTCCTCATCTCCGGGCCTTTCGACCT

CTAGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC

GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGT

TCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT

GAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGA

CCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG

CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC

CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCG

AGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGA

GGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACG

TCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCC

GCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAA

CACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGC

ACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCT

GCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTAC

AAGTAAAGCGGCCGCCTCGAGGGAATTCCGATAATCAACCTCTGGATTACAAAA

TTTGTGAAAGATTGACTGGTATTCTTAACTATGTTGCTCCTTTTACGCTATGTGG

ATACGCTGCTTTAATGCCTTTGTATCATGCTATTGCTTCCCGTATGGCTTTCATT

TTCTCCTCCTTGTATAAATCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCG

TTGTCAGGCAACGTGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACT

GGTTGGGGCATTGCCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCC

CCTCCCTATTGCCACGGCGGAACTCATCGCCGCCTGCCTTGCCCGCTGCTGGA

CAGGGGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTTGTCGGGGAAGCT

GACGTCCTTTCCATGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGGA

CGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCTTCCCGC

GGCCTGCTGCCGGCTCTGCGGCCTCTTCCGCGTCTTCGCCTTCGCCCTCAGAC

GAGTCGGATCTCCCTTTGGGCCGCCTCCCCGCATCGGGAATTCGAGCTCGGTA
CCTTTAAAACCAATGACTTACAAGGCAGCTGTAAATCTTAGCCACTTTTTAAAAG
AAAAGGGGGGACTGGAAGGGCTAATTCACTCCCAACGAAAACAAAATCTGCTTT
TTGCTTGTACTGGGTCTCTCTGGTTAGACCAAATCTGAGCCTGGGAGCTCTCTG
GCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTC
AAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAAAC
CCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCCGTGTATT
CTATAGTGTCACCTAAATCGTATGTGTATGATACATAAGGTTATGTATTAATTGTA
GCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTGGCTTACT
GAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTTTGGTTGG
ACGAACCTTCTGAGTTTCTGGTAACGCCGTCCCGCACCCGGAAATGGTCAGCG
AACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACGCATCGTG
GCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACA
TCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTC
GGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATC
TCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACT
ACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGATATGGTG
CACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCC
GCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCT
TACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACC
GTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTATTTTTATA
GGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGG
AAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATC
CGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGA
GTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTG
CCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGA
TCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGA
TCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
TCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCG
GTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAG
AAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAA
CCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCG
AAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGAT
CGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCAC

GATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACT
TACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGC
AGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATC
TGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGAT
GGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTAT
GGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTG
GTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAAACTTCATT
TTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGACCAAAATC
CCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAA
GGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAA
AACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTT
TTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAG
TGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACC
TCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGT
CTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGG
CTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACC
GAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGG
GAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGC
ACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTT
TCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGA
GCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCT
GGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCG
TATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGC
GCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCC
TCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGTGTGTCAGTTA
GGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGC
ATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGC
AGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTA
ACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCA
TGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGA
GCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAA
GCTTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCACTTTATCCCG
CGTCAGGGAGAGGCAGTGCGTAAAAAGACGCGGACTCATGTGAAATACTGGTT
TTTAGTGCGCCAGATCTCTATAATCTCGCGCAACCTATTTTCCCCTCGAACACTT

TTTAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAAAAATACATCG

TCACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGCCGACTGATG

CCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGGTCTGTACCG

GGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATACCGTTTGTATT

TCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAAACGCGCAGA

AGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATACCGGTGGTA

CTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCACCATCTTCG

CAAAACCGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATATCCCGCAAG

ATACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTCCCGCCAATC

TCCGGTCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGTTCTTTTTAAC

TTCAGGCGGGTTACAATAGTTTCCAGTAAGTATTCTGGAGGCTGCATCCATGAC

ACAGGCAAACCTGAGCGAAACCCTGTTCAAACCCCGCTTTAAACATCCTGAAAC

CTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTGTGCAGTCGG

CCCTTGATGGTAAAACCATCCCTCACTGGTATCGCATGATTAACCGTCTGATGT

GGATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAGGCACGTATT

GTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACGGTGATTGG

CTACCGTGGCGGCAACTGGATTTATGAGTGGGCCCCGGATCTTTGTGAAGGAA

CCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCT

CTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTGATTCTAATT

GTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGCAGTGGTGG

AATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGAT

GATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAG

GTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGAGTCATGCT

GTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAAGGAAAAAG

CTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTTATAAGTAG

GCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAGGCATAGA

GTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTTTTAATTTG

TAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGATCATAAT

CAGCCATACCACATTTGTAGAGCTTTTACTTGCTTTAAAAAACCTCCCACACCTC

CCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTG

CAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGC

ATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTATC

ATGTGTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAACTTCCCA

CCCCATACCCTATTACCACTGCCAAATTACCTGTGGTTTCATTTACTCTAAACCT

GTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGTATTTGTTAAATATGTACTA
CAAACTTAGTAGT

pcDNA3.1-FL-ARFIP2

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGC
TCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTC
GCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGTAGGGATAACAGGGTAATAGCTTTTAATCTCAAAAAACATTAAATGAAA
TGCATACAAGGTTTTATCCTGCTTTAGAACTGTTTGTATTTAATTATCAAACTATA
AGACAGACAATCTAATGCCAGTACACGCTACTCAAAGTTGTAAAACCTCAGATTT
AACTTCAGTAGAAGCTGATTCTCAAAATTGTTAGTGTCAAGCCTAGCTCTTTTGG
GGCTGAAAAGCAATCCTGCAGTGCTGAAAAGCCTCTCACAGGCAGCCGATGCG
GGAAGAGGTGTATTAGTCTTGATAGAGAGGCTGCAAATAGCAGGAAACGTGAG
CAGAGACTCCCTGGTGTCTGAAACACAGGCCAGATGGGTTTAAACGAATTCGT
CGAGACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGA
AAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACG
GCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATA
GGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCC
TTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTG
TTCAAGCTCATGGAGGGATCAGTGACCTGATTCCACAATGACTAGGCCTAATCC
AGAAATTAGATGACTGTCAACATAAAAAGGCACAGCACTCACTAGCTGCCCTAT
ATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACT
TTCTGGTTAAAGAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTCAG
GGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTC
AACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACAT
GCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAA
TCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTC
AGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTT
CACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGC
CTAGGCTAAAACACAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAAT
GTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATA
TTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACA
CATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTAC
AGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGA

TTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGA

CTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGA

TAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAACATAAC

CATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGA

AAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAAT

ACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAG

GAGCCGTACTGTTCCGTTTCAACGCAAATATTAACGGTAAGAGCGAATTTCCTA

GTTTGTTTTCATGCCATTCTTTAAAACCATAGCGTATTACTTTAATTATAGTAAAC

TTTCGCTTTCTTTATTACAAGAGACGTTTTGTGTTGATTCTCCGCGGACATTTTC

GGTCAGACAATCAGAAAATGACCGCGGAGGACCAGTAACTTGCATTACACGTA

AGTTAAATCTTCGTGTATTAAAATGGTTAGGTTGTTAACGTCAAATAGGTTACCG

TGTTTGCGTGTGATCAGGTTGGTTTTGTTAGATTTTTGTCAGTATTTTTAATTTAT

TTGTTTTAGTTTATTTATTTTTTTTGCTGAATCATAGTTTGTGAACAAAGAACCCG

GATGTTACATACAGTACAGCCGCCATGTTACAGAGAGTTATAACTTAATCATTTT

AAAAATAATTTTGCCTTACTTTTAGTTTGTCATGTTGAGAAATGAGGAAATGTTAA

AATGAGGAAATATCCAATTAATTTAATATATCAAAATAATCCATGATTACAATGCA

CTGAACTGGAGAAAATTAAGATGTTTTCTAGTGTCATGAAACAAATGTAAGAGAT

GTACATTGTAGATGTTTTATGTCAAGAATTGGCTAGTTGATGCAGCATACTGGC

GATACTCAGTTGTAATAACAGTAACGTTACATGTTAATAGACTACTGAGTATGCT

GTTCTGTCTATGTATGCTCTGTAAGCTACGAGAAGGACTTTTTTAAACAGTAAAG

GGTGCAATATTTTTACAAATTGAATTAAATAAAGGCTGTCTATTAAGTAATATGCT

TGATATTTTTCTTACTTGATCGGAAATAAGAAAAAATATAAACGTTGTTGCTCTAA

AAATCCTAGTTCAGTTTAGCCAACCACAAATACCTTTTTGTTCCTCCAACAGTTT

TTTTTTCTTCTCTATAATATTTGGCAGTCTATAGTACTCCAAATGTTTCCCCACAG

TCTAACTAATTGGTACAGCCAAAATCATGACACTTATTGCAATAATAATTTTGGTT

CATTGGCATTGTTGATAGCCTGTGCCACTAATATGGTCGATTGATCATGCTTCA

GGAAGAAACTATATTGTTTGATGTAAGATTATTAAATCTTCACCTGCTTCCATTA

CAAACTATTCCCATCTTATTGAATTCTGGTATGTCTTAAAGGATTAGTTCACTTCC

CAAATCAAAATTTACTTAGCCTTTTTTCATCCATGATCCCTTTTTTTCATCATTAAT

GAAGAAATTGTTTTTGAAAAGTTTCAAGATTTTTTTCTCTATATTGTGGAGCTTG

TTAGTTTAAAATTCCAAAATGCAATATGTGGCTTCAAATGGTTCTAAATGATCCC

AGTCAAGGAATAACAGTCTTATCTAATGAAACCATTAGACCTTTAAAAATAAAAA

AAATAAAAGTATTTATTTTTAAATGACTGAGTGATTAAGTTGAATTTCAGCGTTTC

CTTACTGTGTAGAAGTCCTTCCTTACTGGCCCCACCCTTTGGTTCTCTGCCAAT

CTGCTACCTAATGTAATGTTGTGGAACATTATTATTCTTTATTTCTTAATTTTTTAT
TTTTTATTTTAAAACAATGTAAACTGCACAGATGTGCAGTTTGTTTAAAATGGCCA
ATGCTTTGGAAATGCATGACATAATTAGATTTCATGATGCACAAAGCCAAATCTC
AGAGCTTGTGCAAAATGAGCTATCATTTCACTAGGTAAGACCCTAAATTTTCATA
TAGGATCATTTGGACAATTTTGCTGCAGGTAAAATGCATTCTATAGTCCACTGTC
AGCCATTGTTTTGGATAGTATTTATTTTTCTCTACAAGTATAGTCAATAGTTTTCT
ATTATTTTAAAGGTTTGTAACATTTAAGGGTGACCAAATGCAAAGTAAAATTTCAT
TTTCGGGTGAACTATCTCGTTTAACATGGGAGAAGTGCAAAACATACATTATTG
GCTAGAACATTGTAGTATTTTTTAAATGGAAATGTGTGATTGCTAATCTTACTTTG
AATTTGTTTACAGGGATCTATAACTTCGTATAAGCTTGCCACCATGACGGACGG
GATCCTAGGGAAGGCAGCCACAATGGAGATCCCTATCCACGGGAACGGCGAA
GCCAGGCAGCTTCCTGAAGATGATGGGCTGGAGCAGGACCTCCAGCAGGTGA
TGGTGTCAGGACCCAACCTCAATGAAACCAGCATTGTGTCTGGTGGCTATGGG
GGCTCTGGTGATGGACTCATCCCCACAGGGTCTGGTCGCCATCCATCTCACAG
CACCACTCCTTCTGGCCCTGGAGATGAGGTGGCTCGGGGCATTGCTGGAGAAA
AGTTTGACATCGTCAAGAAATGGGGCATCAACACCTATAAGTGCACAAAGCAAC
TGTTATCAGAACGATTTGGTCGAGGCTCACGGACTGTGGACCTGGAGCTAGAG
CTGCAGATTGAGTTGCTGCGTGAGACGAAGCGCAAGTATGAGAGTGTCCTGCA
GCTGGGCCGGGCACTGACAGCCCACCTCTACAGCCTGCTGCAGACCCAGCAT
GCACTGGGTGATGCCTTTGCTGACCTCAGCCAGAAGTCCCCAGAGCTTCAGGA
GGAATTTGGCTACAATGCAGAGACACAGAAACTACTATGCAAGAATGGGGAAAC
GCTGCTAGGAGCCGTGAACTTCTTTGTCTCTAGCATCAACACATTGGTCACCAA
GACCATGGAAGACACGCTCATGACTGTGAAACAGTATGAGGCTGCCAGGCTGG
AATATGATGCCTACCGAACAGACTTAGAGGAGCTGAGTCTAGGCCCCCGGGAT
GCAGGGACACGTGGTCGACTTGAGAGTGCCCAGGCCACTTTCCAGGCCCATC
GGGACAAGTATGAGAAGCTGCGGGGAGATGTGGCCATCAAGCTCAAGTTCCTG
GAAGAAACAAGATCAAGGTGATGCACAAGCAGCTGCTGCTCTTCCACAATGCT
GTGTCCGCCTACTTTGCTGGGAACCAGAAACAGCTGGAGCAGACCCTGCAGCA
GTTCAACATCAAGCTGCGGCCTCCAGGAGCTGAGAAACCTCCTGGCTAGAGG
AGCAGTACCCATACGATGTTCCAGATTACGCTTGACTCGAGTCTAGAGGGCCCT
TCGAACAAAACTCATCTCAGAAGAGGATCTGAATATGCATACCGGTCATCATC
ACCATCACCATTGAGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTT
GCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTG

AGTAGGTGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGG
AGGATTGGGAAGACAATAGCCGGTACGGATCCTTCAGGGGTGTCGCTTGGTGA
TTTCCAAAAATCAAATTAATTTTATTAAACTATTAGAACGAGCATGTTTTGTCTAT
ATGCTACAGAAGATAAAAAATAATAGGAGTTAACAGTTATAAAACAACACACTTT
GTTTCTATTGATTGTTGACCACACTGGGGTCTCATTAAGTTAGATTAAAGACACA
CTAACTGGGTCAAATGCAGCAGATTGATTTCATGGCACCAGGGTAAACTTTCTA
ACACTTTTACGGCAATCATATACATCAAAATTAAATACAGGCCACGACTGAACAA
GGAGGATGATCTCCAAAATTAAACAAAGAGACTTGTGCCTATTTCTCTGAGGGT
AAACATGACCTCTCAAGTTAGCAAGTTGTTTTTAACACTACAAAAATAGTTAAGA
CTGCAATCCCAGAATAAAGTATTGGTTTTAACCAATCAATATAGTACAGTAAACA
TCCATTTGTTTTGTTGAAACGTTAAACGAATCTGACCAAGCTATTAGCTTATATAA
AACAGGTTTGCCTTTTATGTAGCTGAAAATACCACAGGCCCGATTTTGCTACTGT
GTAAACATTTCAGCAAGATTTTTTTTATATTGCATTTTTTTCTACTGAATCGTTC
AAACATTTTATCATTTTAGTTTGTTCATTCATTGCAACTGGAAAAACAACACACCA
CACAACCGCACATTTTTCAGCAATAAGTACAATAAAACACTCAAATAAAAAAAAA
CTTTTTAAATCTCTTTGTATTTTTGACCGCTGTTTCGCGTAATTTCACGGTAAAAC
TCTGGAAATCTCCACTACATTCCTCTCAGCGGCTCCTCTCAATGACAGCTGAAG
AAGTGACGCGGCTGCCTGCTGTGTTTTGATTGGTCGAATTCACTGGAGGCTTC
CAGAACAGTGTAGAGTCTGAACGGGTGCGCGCTCTGCTGTATTTAAAGGGCGA
AAGAGAGACCGCAGAGAAACTCAACCGAAGAGAAGCGACTTGACAAAGAAGAA
AAGAGCAGCCTGACAGGACTTTTCCCCGACGAGGTGTTTATTCGCTCTATTTAA
GAATCTACTGTAAGGTAAGTCTCAATATATTGTACTCTAGTGGCTAATCAAAATT
TTATAGAGATTATATGTACTTAATGTCAAAAAATCTACTTTGTATATGTAATCTTTT
TACATGTGGACTGCCTATGTTCATCTTATTTTAGGTCTACTAGAAAATTATATTTC
CCGTTTTCACAATAAGGATTTAAAAAAAAAGCAATGAACAGACTGGCATTTACTT
TATGTTGCTGACATTATTATATATGAGCATAATAACCATAAATACTAGCAAATGTC
CTAAATGAATTTGTGTTAATGTTGTCTACAAAGAAAGAAAATTAGCGTTTTACT
TGTACAACTAATAATAACTTAGTTATTAAGAGAATTTCACTTGTTGACTAGAAAAA
TCCTTTCATAATGAAACAATTGCACCATAAATTGTATAAATATAAAATTAATTCTA
ATTGTTTTTTTTTTTTCCTGCAGTCGAGCGCCACCATGGCCTACTCCGAGAAAGT
CATCACCGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCACCGTGAACGGC
CACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCCAC
AACAACGTGAAGCTGAAGGTGACCAAGGGCGGCCACCTGCCCTTCGCCTGGG
ACATCCTGTCCCACCAGTTCCAGTACGGCTCCAAGGTGTACGTGAAGCACCCC

GCCGACATCACCGACTACAAGAAGCTGTCATTCCCCGAGGGCTTCAAGTGGGA
GCGCGTGATGAACTTCGAGGACGGCGGCGTGGCGACAGTGACCCAGGACTCC
TCCCTGCAGGACGGCTGCTTCATCTACAAGGTGAAGTTCATCGGAGTGAACTTC
ACCTCCGACGGCCCCGTGATGCAGAAGAAGACCATGGGCTGGGAGGCCTCAA
CCGAGCGCATGTACCCCCGCGACGGCGTGCTGAAGGGCGAGACCCACAAGGC
CCTGAAGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGTCCATCTACA
TGGACAAGAAGCCAGTGCAGCTGACCGGCTACTACTACGTGGACGCCAAGCTG
GACATCACCTCCCACAACGAGGACTACACCATCGTGGAGCAGTACGAGCGCAC
CGAGGGCCGCCACCACCTGTTCTAAGCGGCCGCGACTCTAGATCATAATCAGC
CATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCC
TGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGC
TTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATAAAGCATTTT
TTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAGGTACC
TAGGGATAACAGGGTAATTAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGG
CTTCTGAGGCGGAAAGAACCAGCTGGGGCTCTAGGGGGTATCCCCACGCGCC
CTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACC
GCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTT
CTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTT
AGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGG
TGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGAC
GTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACT
CAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCC
TATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTAATTCTGTG
GAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAA
GTATGCAAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAG
GCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACC
ATAGTCCCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGC
CCATTCTCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGC
CGCCTCTGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCC
TAGGCTTTTGCAAAAAGCTCCCGGGAGCTTGTATATCCATTTTCGGATCTGATC
AAGAGACAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCA
GGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACA
GACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGC
CCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGA

CGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCT
GTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGT
GCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTATCCA
TCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCA
TTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGC
CGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCA
GCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCG
TCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGC
TTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGA
CATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTG
ACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCC
TTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGCGAAAT
GACCGACCAAGCGACGCCCAACCTGCCATCACGAGATTTCGATTCCACCGCCG
CCTTCTATGAAAGGTTGGGCTTCGGAATCGTTTTCCGGGACGCCGGCTGGATG
ATCCTCCAGCGCGGGGATCTCATGCTGGAGTTCTTCGCCCACCCCAACTTGTTT
ATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACAAATA
AAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCT
TATCATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTAATCATGGTCA
TAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAG
CCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACA
TTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA
GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG
CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCG
GCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAG
GGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACG
AGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTA
TAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCC
GACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG
CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCT
CCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTT
ATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGAACAGTATTT

GGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAGAGTTGGTAGCTCT
TGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACG
GGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAG
ATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT
CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAG
TGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACT
CCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTG
CTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAA
ACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGC
CTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGT
TAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTC
GTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGT
TGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCA
TAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTAC
TCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCG
GCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATC
ATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGA
TCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTT
TCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAG
GGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATT
ATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTAT
TTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC
TGACGTC

pcDNA3.1-HC-ARFIP2

GACGGATCGGGAGATCTCCCGATCCCCTATGGTGCACTCTCAGTACAATCTGC
TCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTC
GCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGA
CAATTGTAGGGATAACAGGGTAATAGCTTTTAATCTCAAAAAACATTAAATGAAA
TGCATACAAGGTTTTATCCTGCTTTAGAACTGTTTGTATTTAATTATCAAACTATA
AGACAGACAATCTAATGCCAGTACACGCTACTCAAAGTTGTAAAACCTCAGATTT
AACTTCAGTAGAAGCTGATTCTCAAAATTGTTAGTGTCAAGCCTAGCTCTTTTGG

GGCTGAAAAGCAATCCTGCAGTGCTGAAAAGCCTCTCACAGGCAGCCGATGCG
GGAAGAGGTGTATTAGTCTTGATAGAGAGGCTGCAAATAGCAGGAAACGTGAG
CAGAGACTCCCTGGTGTCTGAAACACAGGCCAGATGGGTTTAAACGAATTCGT
CGAGACCAGCAAAGTTCTAGAATTTGTCGAAACATTTATGTTATATATTTCCTGA
AAAAAATTCTGAGTAAGTTCTTAAGTGTTATTGCCAGCAACATAAACAACAGACG
GCAAAATGAATAAATGATAACAAAGCAGTAGGCTTAAATAAACCTAATTTTTATA
GGCTGTTCTCTACAACCCTCAAACAGTGATTAGTTTTGTACTTATAAACTTGCCC
TTTCATTCATATTTCAAGAAAATTGGTTCAGAAGATCTGGATATTCTAGCAGTTG
TTCAAGCTCATGGAGGGATCAGTGACCTGATTCCACAATGACTAGGCCTAATCC
AGAAATTAGATGACTGTCAACATAAAAGGCACAGCACTCACTAGCTGCCCTAT
ATATTTTATTATATTTTACATATATTATTTTATTTATTTAGCTCTGAGTGCTGTACT
TTCTGGTTAAAGAAACTGCTTACAACAGCTAACCTGTACTACCTCAGGCTCAG
GGAATTTGGAACAGGTTTGTCTGGTTTGTTTCTTTAACCATGCATGCTTGTTTTC
AACTATGGCAACACAGTCACATGGGACATTACAGAAATGATTTGTCGATGACAT
GCGACTTTTCTTTAATAAAGCGCAAAGATCCCAAAAAGCAAACTTTTAACAAAAA
TCATATAATTATATTTTCAATCCAGCTTTGTAGCAACTTTGTGCTGCTGTTCACTC
AGCAACAGATAGTCAGTATAAGGTCAGTGTGTCTCAAAGCAGTGCCATCTGTTT
CACACATTGCGTTCTATATATAAGTGTGCTGGTTGACACGACACTGTATAAGGC
CTAGGCTAAAACACAAACAATGTAGAATGACACTGTGTTTTTTTTGTAAACAAAT
GTTGTTTTTGGTTAAACATCTTTGTGAAAACATCCTCCTGTCATGTATTTGCTATA
TTCAAATGTTAAACCCGTGCAGAATAGAACATATACAAAAAAAAACAACACAACA
CATTTTTAAACATTATTAAATATCAAGTATTGCTGGCAGTTCTGTTTCTGTTTTAC
AGTACCCTTTGCCACAGTTCTCCGCTTTTCCTGGTCCAGATTCCACAAGTCTGA
TTCACCAATAGCAAAGCGAATAAACAACCAAAGCAGCCAATCACTGCTTGTAGA
CTGTCCTGCGAGACCGGCCCATTCCAGCACATTCTGGAAACTTCCTTTATATGA
TAATTATAAATACATTTAAATTATTGATACAAAACATGTAATTCCTAGAACATAAC
CATAGCAATCATTAGTTTTCAGGGTAATTATGTATTTTTAGGATTTGACTGCGGA
AAGATCTGGTCATGTGACGTCTCATGAACGTCACGGCCCTGGGTTTCTATAAAT
ACAGTAGGACTCTCGACCATCGGCAGATTTTTCGAAGAAGAAGATCAGTTTCAG
GAGCCGTACTGTTCCGTTTCAACGCAAATATTAACGGTAAGAGCGAATTTCCTA
GTTTGTTTTCATGCCATTCTTTAAAACCATAGCGTATTACTTTAATTATAGTAAAC
TTTCGCTTTCTTTATTACAAGAGACGTTTTGTGTTGATTCTCCGCGGACATTTTC
GGTCAGACAATCAGAAAATGACCGCGGAGGACCAGTAACTTGCATTACACGTA
AGTTAAATCTTCGTGTATTAAAATGGTTAGGTTGTTAACGTCAAATAGGTTACCG

TGTTTGCGTGTGATCAGGTTGGTTTTGTTAGATTTTTGTCAGTATTTTTAATTTAT
TTGTTTTAGTTTATTTATTTTTTTTGCTGAATCATAGTTTGTGAACAAAGAACCCG
GATGTTACATACAGTACAGCCGCCATGTTACAGAGAGTTATAACTTAATCATTTT
AAAAATAATTTTGCCTTACTTTTAGTTTGTCATGTTGAGAAATGAGGAAATGTTAA
AATGAGGAAATATCCAATTAATTTAATATATCAAAATAATCCATGATTACAATGCA
CTGAACTGGAGAAAATTAAGATGTTTTCTAGTGTCATGAAACAAATGTAAGAGAT
GTACATTGTAGATGTTTTATGTCAAGAATTGGCTAGTTGATGCAGCATACTGGC
GATACTCAGTTGTAATAACAGTAACGTTACATGTTAATAGACTACTGAGTATGCT
GTTCTGTCTATGTATGCTCTGTAAGCTACGAGAAGGACTTTTTAAACAGTAAAG
GGTGCAATATTTTTACAAATTGAATTAAATAAAGGCTGTCTATTAAGTAATATGCT
TGATATTTTTCTTACTTGATCGGAAATAAGAAAAAATATAAACGTTGTTGCTCTAA
AAATCCTAGTTCAGTTTAGCCAACCACAAATACCTTTTTGTTCCTCCAACAGTTT
TTTTTTCTTCTCTATAATATTTGGCAGTCTATAGTACTCCAAATGTTTCCCCACAG
TCTAACTAATTGGTACAGCCAAAATCATGACACTTATTGCAATAATAATTTTGGTT
CATTGGCATTGTTGATAGCCTGTGCCACTAATATGGTCGATTGATCATGCTTCA
GGAAGAAACTATATTGTTTGATGTAAGATTATTAAATCTTCACCTGCTTCCATTA
CAAACTATTCCCATCTTATTGAATTCTGGTATGTCTTAAAGGATTAGTTCACTTCC
CAAATCAAAATTTACTTAGCCTTTTTTCATCCATGATCCCTTTTTTTCATCATTAAT
GAAGAAATTGTTTTTGAAAAGTTTCAAGATTTTTTTCTCTATATTGTGGAGCTTG
TTAGTTTAAAATTCCAAAATGCAATATGTGGCTTCAAATGGTTCTAAATGATCCC
AGTCAAGGAATAACAGTCTTATCTAATGAAACCATTAGACCTTTAAAAATAAAAA
AAATAAAAGTATTTATTTTTAAATGACTGAGTGATTAAGTTGAATTTCAGCGTTTC
CTTACTGTGTAGAAGTCCTTCCTTACTGGCCCCACCCTTTGGTTCTCTGCCAAT
CTGCTACCTAATGTAATGTTGTGGAACATTATTATTCTTTATTTCTTAATTTTTTAT
TTTTTATTTTAAAACAATGTAAACTGCACAGATGTGCAGTTTGTTTAAAATGGCCA
ATGCTTTGGAAATGCATGACATAATTAGATTTCATGATGCACAAAGCCAAATCTC
AGAGCTTGTGCAAAATGAGCTATCATTTCACTAGGTAAGACCCTAAATTTTCATA
TAGGATCATTTGGACAATTTTGCTGCAGGTAAAATGCATTCTATAGTCCACTGTC
AGCCATTGTTTTGGATAGTATTTATTTTTCTCTACAAGTATAGTCAATAGTTTTCT
ATTATTTTAAAGGTTTGTAACATTTAAGGGTGACCAAATGCAAAGTAAAATTTCAT
TTTCGGGTGAACTATCTCGTTTAACATGGGAGAAGTGCAAAACATACATTATTG
GCTAGAACATTGTAGTATTTTTTAAATGGAAATGTGTGATTGCTAATCTTACTTTG
AATTTGTTTACAGGGATCTATAACTTCGTATAAGCTTGCCACCATGCTCAGCCAG
AAGTCCCCAGAGCTTCAGGAGGAATTTGGCTACAATGCAGAGACACAGAAACT

ACTATGCAAGAATGGGGAAACGCTGCTAGGAGCCGTGAACTTCTTTGTCTCTAG
CATCAACACATTGGTCACCAAGACCATGGAAGACACGCTCATGACTGTGAAACA
GTATGAGGCTGCCAGGCTGGAATATGATGCCTACCGAACAGACTTAGAGGAGC
TGAGTCTAGGCCCCCGGGATGCAGGGACACGTGGTCGACTTGAGAGTGCCCA
GGCCACTTTCCAGGCCCATCGGGACAAGTATGAGAAGCTGCGGGGAGATGTG
GCCATCAAGCTCAAGTTCCTGGAAGAAACAAGATCAAGGTGATGCACAAGCA
GCTGCTGCTCTTCCACAATGCTGTGTCCGCCTACTTTGCTGGGAACCAGAAACA
GCTGGAGCAGACCCTGCAGCAGTTCAACATCAAGCTGCGGCCTCCAGGAGCT
GAGAAACCCTCCTGGCTAGAGGAGCAGTACCCATACGATGTTCCAGATTACGC
TTGACTCGAGTCTAGAGGGCCCTTCGAACAAAAACTCATCTCAGAAGAGGATCT
GAATATGCATACCGGTCATCATCACCATCACCATTGAGTTTAAACCCGCTGATC
AGCCTCGACTGTGCCTTCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGT
GCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAATAAAATGA
GGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGG
TGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCCGGTACGGATC
CTTCAGGGGTGTCGCTTGGTGATTTCCAAAAATCAAATTAATTTTATTAAACTATT
AGAACGAGCATGTTTTGTCTATATGCTACAGAAGATAAAAAATAATAGGAGTTAA
CAGTTATAAACAACACACTTTGTTTCTATTGATTGTTGACCACACTGGGGTCTC
ATTAAGTTAGATTAAAGACACACTAACTGGGTCAAATGCAGCAGATTGATTTCAT
GGCACCAGGGTAAACTTTCTAACACTTTTACGGCAATCATATACATCAAAATTAA
ATACAGGCCACGACTGAACAAGGAGGATGATCTCCAAAATTAAACAAAGAGACT
TGTGCCTATTTCTCTGAGGGTAAACATGACCTCTCAAGTTAGCAAGTTGTTTTTA
ACACTACAAAAATAGTTAAGACTGCAATCCCAGAATAAAGTATTGGTTTTAACCA
ATCAATATAGTACAGTAAACATCCATTTGTTTTGTTGAAACGTTAAACGAATCTG
ACCAAGCTATTAGCTTATATAAACAGGTTTGCCTTTTATGTAGCTGAAAATACC
ACAGGCCCGATTTTGCTACTGTGTAAACATTTCAGCAAGATTTTTTTTATATTG
CATTTTTTTCTACTGAATCGTTCAAACATTTTATCATTTTAGTTTGTTCATTCATTG
CAACTGGAAAAACAACACACCACACAACCGCACATTTTTCAGCAATAAGTACAA
TAAAACACTCAAATAAAAAAAAACTTTTTAAATCTCTTTGTATTTTTGACCGCTGT
TTCGCGTAATTTCACGGTAAAACTCTGGAAATCTCCACTACATTCCTCTCAGCG
GCTCCTCTCAATGACAGCTGAAGAAGTGACGCGGCTGCCTGCTGTGTTTTGATT
GGTCGAATTCACTGGAGGCTTCCAGAACAGTGTAGAGTCTGAACGGGTGCGCG
CTCTGCTGTATTTAAAGGGCGAAAGAGAGACCGCAGAGAAACTCAACCGAAGA
GAAGCGACTTGACAAAGAAGAAAGAGCAGCCTGACAGGACTTTTCCCCGACG

AGGTGTTTATTCGCTCTATTTAAGAATCTACTGTAAGGTAAGTCTCAATATATTGT
ACTCTAGTGGCTAATCAAAATTTTATAGAGATTATATGTACTTAATGTCAAAAAT
CTACTTTGTATATGTAATCTTTTTACATGTGGACTGCCTATGTTCATCTTATTTTA
GGTCTACTAGAAAATTATATTTCCCGTTTTCACAATAAGGATTTAAAAAAAAAGC
AATGAACAGACTGGCATTTACTTTATGTTGCTGACATTATTATATGAGCATAAT
AACCATAAATACTAGCAAATGTCCTAAATGAATTTGTGTTAATGTTGTCTACAAAA
GAAAAGAAATTAGCGTTTTACTTGTACAACTAATAATAACTTAGTTATTAAGAGA
ATTTCACTTGTTGACTAGAAAAATCCTTTCATAATGAAACAATTGCACCATAAATT
GTATAAATATAAAATTAATTCTAATTGTTTTTTTTTTTCCTGCAGTCGAGCGCCAC
CATGGCCTACTCCGAGAAAGTCATCACCGAGTTCATGCGCTTCAAGGTGCGCA
TGGAGGGCACCGTGAACGGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGG
GCCGCCCCTACGAGGGCCACAACAACGTGAAGCTGAAGGTGACCAAGGGCGG
CCACCTGCCCTTCGCCTGGGACATCCTGTCCCACCAGTTCCAGTACGGCTCCA
AGGTGTACGTGAAGCACCCCGCCGACATCACCGACTACAAGAAGCTGTCATTC
CCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGG
CGACAGTGACCCAGGACTCCTCCCTGCAGGACGGCTGCTTCATCTACAAGGTG
AAGTTCATCGGAGTGAACTTCACCTCCGACGGCCCCGTGATGCAGAAGAAGAC
CATGGGCTGGGAGGCCTCAACCGAGCGCATGTACCCCCGCGACGGCGTGCTG
AAGGGCGAGACCCACAAGGCCCTGAAGCTGAAGGACGGCGGCCACTACCTGG
TGGAGTTCAAGTCCATCTACATGGACAAGAAGCCAGTGCAGCTGACCGGCTAC
TACTACGTGGACGCCAAGCTGGACATCACCTCCCACAACGAGGACTACACCAT
CGTGGAGCAGTACGAGCGCACCGAGGGCCGCCACCACCTGTTCTAAGCGGCC
GCGACTCTAGATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTA
AAAAACCTCCCACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTT
GTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCA
CAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAA
CTCATCAATGTATCTTAGGTACCTAGGGATAACAGGGTAATTAGCAGGCATGCT
GGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCT
CTAGGGGGTATCCCCACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGT
GGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCT
CCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAA
GCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTC
GACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTG
ATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACT

CTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTA
TAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAA
AATTTAACGCGAATTAATTCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTC
CCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAG
CAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGC
ATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCCCATCCCG
CCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTTT
TTTATTTATGCAGAGGCCGAGGCCGCCTCTGCCTCTGAGCTATTCCAGAAGTAG
TGAGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAGCTCCCGGGAGCTTG
TATATCCATTTTCGGATCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGAT
TGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTAT
TCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC
CGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCG
GTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCAC
GACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGG
GACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCT
TGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATA
CGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAG
CGAGCACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGA
AGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGC
ATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTTGCCGA
ATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTG
GGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGA
AGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCG
CTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAG
CGGGACTCTGGGGTTCGCGAAATGACCGACCAAGCGACGCCCAACCTGCCAT
CACGAGATTTCGATTCCACCGCCGCCTTCTATGAAAGGTTGGGCTTCGGAATC
GTTTTCCGGGACGCCGGCTGGATGATCCTCCAGCGCGGGGATCTCATGCTGG
AGTTCTTCGCCCACCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAG
CAATAGCATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTG
GTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAG
CTAGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCC
GCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGG
GTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTT

TCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCG
GGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTC
GCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG
GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGC
AAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT
TCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAG
AGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAG
CTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCG
CCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC
TCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC
GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCC
GGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCA
GAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTAC
GGCTACACTAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCA
AGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTC
ACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTT
TTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTC
TGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTT
CGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACC
GGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAG
CTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTA
CAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT
CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTA
GCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCAC
TCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGAT
GCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGC
GGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACAT
AGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTC
TCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCC
AACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAG

GAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATA
CTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCAT
GAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCG
CACATTTCCCCGAAAAGTGCCACCTGACGTC

pLVTHM-471

TTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATC
TACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGG
GGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCC
AGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTT
GACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTT
CAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCC
AGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTG
CATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCT
GAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAA
AGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT
AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGC
GCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACG
CAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACT
GGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGT
GCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATT
CGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCA
AGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGA
AGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGG
ATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAAC
AAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAG
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAG
AAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATT
ATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGC
AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGA

ATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG
TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAG
AGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAAC
CAGCAAGAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTG
TGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGAT
AGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAAT
AGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCG
AGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACA
GAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGATCACGAGACT
AGCCTCGACCATCGCGATGTCGACGATAAGCTTTGCAAAGATGGATAAAGTTTT
AAACAGAGAGGAATCTTTGCAGCTAATGGACCTTCTAGGTCTTGAAAGGAGTGG
GAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCC
CGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGG
CGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGA
GGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTC
GCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGG
GCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTG
GCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGA
GTTCGAGGCCTTGCGCTTAAGGAGCCCCTTCGCCTCGTGCTTGAGTTGAGGCC
TGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCT
GTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGC
GACGCTTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACAC
TGGTATTTCGGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAG
CGCACATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAGAATCGGA
CGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGC
CGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGTTGC
GTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGG
AGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAA
GGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCG
CCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGT
TGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGA
CTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTT
TGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTT

CTTCCATTTCAGGTGTCGTGAGGAATTTCGACATTTAAATTTAATTAATCTCGAC
GGTATCGGTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGG
GAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAAC
AAATTACAAAAATTCAAAATTTTATCGATCACGAGACTAGCCTCGAGGTTTAAAC
TACGGGATCTTCGAAGGCCTAAGCTTACGCGCGCGTCCTAGCGCTACCGGTCG
CCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCT
GGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC
CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGC
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAA
GTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG
ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGA
CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGA
CGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG
AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGAC
CGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGAT
CTCGACTAGCTAGTAGCTAGCTAGCTAGTCGAGCTCAAGCTTCGGGGACTAGT
CATATGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCT
TAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTAT
CATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGT
TGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTG
TGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTG
TCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACT
CATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACT
GACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGC
CTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGG
CCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCT
CTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGC
CTCCCCGCATCGGTACGTATGGCCAGGTACCTTTAAGACCAATGACTTACAAGG
CAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAA
TTCACTCCCAACGAAGACAAGATGGGATCAATTCACCATGCTAGTGGATCCCCC

TCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCA
CTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGT
GATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAG
TGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGT
TTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCT
ATCAGTGATAGAGAAAGTGAAAGTCGGGGCTGCAGGAATTCGAACGCTGACG
TCATCAACCCGCTCCAAGGAATCGCGGGCCCAGTGTCACTAGGCGGGAACACC
CAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGAGGGACAGGGGAGTGG
CGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTG
AAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACGCGTCCCC
GGAGCTAGAGCTGCAGATTTTCAAGAGAAATCTGCAGCTCTAGCTCCTTTTTGG
AAATATACCGTCGCATGGGAATAACTTCGTATAGCATACATTATACGAAGTTATG
CTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAG
CTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGA
GTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCC
CTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCC
GTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGATACATAAGGTTATGTAT
TAATTGTAGCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTG
GCTTACTGAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTT
TGGTTGGACGAACCTTCTGAGTTTCTGGTAACGCCGTCCCGCACCCGGAAATG
GTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACG
CATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATC
GCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCG
CTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGG
CGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCA
ACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGA
ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCG
ACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCA
TCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTT
TTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT
TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTT
TCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGG

CATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATG
CTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGC
GGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACT
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG
CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCA
GTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCT
GCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGA
GGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCG
CCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG
ACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGC
GAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGAT
AAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGG
GGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATT
AAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAA
ACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGA
CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAA
AGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCA
AACAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTAC
CAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTG
TTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGC
CTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA
AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAG
CGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGA
CCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTT
CCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG
GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCT
GTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG
GGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGT
GGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA
CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACG
CAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGT

GTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGC
AAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCC
CCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC
CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTC
TCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTC
GGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCT
TTTGCAAAAGCTTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCA
CTTTATCCCGCGTCAGGGAGAGGCAGTGCGTAAAAGACGCGGACTCATGTGA
AATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCGCAACCTATTTTCCCC
TCGAACACTTTTTAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAA
AAATACATCGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGC
CGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGG
TCTGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATAC
CGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAA
ACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATA
CCGGTGGTACTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCA
CCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATA
TCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTC
CCGCCAATCTCCGGTCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGT
TCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCAGTAAGTATTCTGGAGGCTGC
ATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAACCCCGCTTTAAACA
TCCTGAAACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTG
TGCAGTCGGCCCTTGATGGTAAAACCATCCCTCACTGGTATCGCATGATTAACC
GTCTGATGTGGATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAG
GCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACG
GTGATTGGCTACCGTGGCGGCAACTGGATTTATGAGTGGGCCCCGGATCTTTG
TGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGA
TTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTG
ATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGC
AGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCAT
CTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGA
AGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGA
GTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAA
GGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTT

ATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAG
GCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTT
TTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAG
ATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCC
CACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTT
GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACA
AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGT
ATCTTATCATGTCTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAA
CTTCCCACCCCATACCCTATTACCACTGCCAATTACCTGTGGTTTCATTTACTCT
AAACCTGTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGTATTTGTTAAATAT
GTACTACAAACTTAGTAG

pLVTHM-480

TTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATC
TACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGG
GGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCC
AGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTT
GACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTT
CAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCC
AGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTG
CATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCT
GAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAA
AGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT
AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGC
GCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACG
CAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACT
GGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGT
GCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATT
CGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCA
AGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGA
AGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGG
ATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAAC

AAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAG
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAG
AAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATT
ATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGC
AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGA
ATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG
TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAG
AGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAAC
CAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTG
TGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGAT
AGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAAT
AGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCG
AGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACA
GAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGATCACGAGACT
AGCCTCGACCATCGCGATGTCGACGATAAGCTTTGCAAAGATGGATAAAGTTTT
AAACAGAGAGGAATCTTTGCAGCTAATGGACCTTCTAGGTCTTGAAAGGAGTGG
GAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCC
CGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGG
CGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGA
GGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTC
GCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGG
GCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTG
GCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGA
GTTCGAGGCCTTGCGCTTAAGGAGCCCCTTCGCCTCGTGCTTGAGTTGAGGCC
TGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCT
GTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGC
GACGCTTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACAC
TGGTATTTCGGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAG
CGCACATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAGAATCGGA
CGGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGC
CGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGTTGC

GTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGG
AGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAA
GGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCG
CCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGT
TGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGA
CTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTT
TGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTT
CTTCCATTTCAGGTGTCGTGAGGAATTTCGACATTTAAATTTAATTAATCTCGAC
GGTATCGGTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGG
GAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAAC
AAATTACAAAAATTCAAAATTTTATCGATCACGAGACTAGCCTCGAGGTTTAAAC
TACGGGATCTTCGAAGGCCTAAGCTTACGCGCGCGTCCTAGCGCTACCGGTCG
CCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCT
GGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC
CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGC
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAA
GTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG
ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGA
CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGA
CGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG
AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGAC
CGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGAT
CTCGACTAGCTAGTAGCTAGCTAGCTAGTCGAGCTCAAGCTTCGGGGACTAGT
CATATGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCT
TAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTAT
CATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGT
TGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTG
TGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTG
TCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACT
CATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACT

GACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGC
CTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGG
CCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCT
CTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGC
CTCCCCGCATCGGTACGTATGGCCAGGTACCTTTAAGACCAATGACTTACAAGG
CAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAA
TTCACTCCCAACGAAGACAAGATGGGATCAATTCACCATGCTAGTGGATCCCCC
TCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCA
CTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGT
GATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAG
TGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGT
TTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCT
ATCAGTGATAGAGAAAGTGAAAGTCGGGGCTGCAGGAATTCGAACGCTGACG
TCATCAACCCGCTCCAAGGAATCGCGGGCCCAGTGTCACTAGGCGGGAACACC
CAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGAGGGACAGGGGAGTGG
CGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTG
AAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACGCGTCCCC
GCTGCAGATTGAGTTGCTGTTCAAGAGACAGCAACTCAATCTGCAGCTTTTTGG
AAATATACCGTCGCATGGGAATAACTTCGTATAGCATACATTATACGAAGTTATG
CTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAG
CTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGA
GTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCC
CTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCC
GTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGATACATAAGGTTATGTAT
TAATTGTAGCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTG
GCTTACTGAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTT
TGGTTGGACGAACCTTCTGAGTTTCTGGTAACGCCGTCCCGCACCCGGAAATG
GTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACG
CATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATC
GCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCG
CTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGG
CGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCA
ACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGA
ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCG

ACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCA
TCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTT
TTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT
TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTT
TCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGG
CATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATG
CTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGC
GGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACT
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG
CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCA
GTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCT
GCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGA
GGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCG
CCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG
ACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGC
GAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGAT
AAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGG
GGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATT
AAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAA
ACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGA
CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAA
AGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCA
AACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTAC
CAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTG
TTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGC
CTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA
AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAG
CGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGA
CCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTT
CCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG

GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCT
GTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG
GGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGT
GGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA
CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACG
CAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGT
GTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGC
AAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCC
CCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC
CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTC
TCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTC
GGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCT
TTTGCAAAAAGCTTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCA
CTTTATCCCGCGTCAGGGAGAGGCAGTGCGTAAAAAGACGCGGACTCATGTGA
AATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCGCAACCTATTTTCCCC
TCGAACACTTTTTAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAA
AAATACATCGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGC
CGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGG
TCTGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATAC
CGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAA
ACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATA
CCGGTGGTACTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCA
CCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATA
TCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTC
CCGCCAATCTCCGGTCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGT
TCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCAGTAAGTATTCTGGAGGCTGC
ATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAACCCGCTTTAAACA
TCCTGAAACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTG
TGCAGTCGGCCCTTGATGGTAAAACCATCCCTCACTGGTATCGCATGATTAACC
GTCTGATGTGGATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAG
GCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACG
GTGATTGGCTACCGTGGCGGCAACTGGATTTATGAGTGGGCCCCGGATCTTTG
TGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGA

TTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTG
ATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGC
AGTGGTGGAATGCCTTTAATGAGGAAAACCTGTTTTGCTCAGAAGAAATGCCAT
CTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGA
AGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGA
GTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAA
GGAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTT
ATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAG
GCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTT
TTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAG
ATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCC
CACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTT
GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACA
AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGT
ATCTTATCATGTCTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAA
CTTCCCACCCCATACCCTATTACCACTGCCAATTACCTGTGGTTTCATTTACTCT
AAACCTGTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGTATTTGTTAAATAT
GTACTACAAACTTAGTAG

pLVTHM-749

TTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATC
TACCACACACAAGGCTACTTCCCTGATTAGCAGAACTACACACCAGGGCCAGG
GGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGTACCAGTTGAGCC
AGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTG
TGAGCCTGCATGGGATGGATGACCCGGAGAGAGAAGTGTTAGAGTGGAGGTTT
GACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAGTACTT
CAAGAACTGCTGATATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCC
AGGGAGGCGTGGCCTGGGCGGGACTGGGGAGTGGCGAGCCCTCAGATCCTG
CATATAAGCAGCTGCTTTTTGCCTGTACTGGGTCTCTCTGGTTAGACCAGATCT
GAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAA
AGCTTGCCTTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGT
AACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGC
GCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTCTCTCGACG
CAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACT

GGTGAGTACGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGT
GCGAGAGCGTCAGTATTAAGCGGGGGGAGAATTAGATCGCGATGGGAAAAATT
CGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTATGGGCA
AGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGAAACATCAGAA
GGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTTCAGACAGGATCAGA
AGAACTTAGATCATTATATAATACAGTAGCAACCCTCTATTGTGTGCATCAAAGG
ATAGAGATAAAAGACACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAAC
AAAAGTAAGACCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAG
GAGATATGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAGAGAG
AAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCTTGGGTTCTTGGGAGCAGCA
GGAAGCACTATGGGCGCAGCGTCAATGACGCTGACGGTACAGGCCAGACAATT
ATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGC
AACAGCATCTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGA
ATCCTGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATTTGGGG
TTGCTCTGGAAAACTCATTTGCACCACTGCTGTGCCTTGGAATGCTAGTTGGAG
TAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGATGGAGTGGGACAG
AGAAATTAACAATTACACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAAAC
CAGCAAGAAAAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTG
TGGAATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAATGAT
AGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCTATAGTGAAT
AGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCG
AGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACA
GAGACAGATCCATTCGATTAGTGAACGGATCTCGACGGTATCGATCACGAGACT
AGCCTCGACCATCGCGATGTCGACGATAAGCTTTGCAAAGATGGATAAAGTTTT
AAACAGAGAGGAATCTTTGCAGCTAATGGACCTTCTAGGTCTTGAAAGGAGTGG
GAATTGGCTCCGGTGCCCGTCAGTGGGCAGAGCGCACATCGCCCACAGTCCC
CGAGAAGTTGGGGGGAGGGGTCGGCAATTGAACCGGTGCCTAGAGAAGGTGG
CGCGGGGTAAACTGGGAAAGTGATGTCGTGTACTGGCTCCGCCTTTTTCCCGA
GGGTGGGGGAGAACCGTATATAAGTGCAGTAGTCGCCGTGAACGTTCTTTTTC
GCAACGGGTTTGCCGCCAGAACACAGGTAAGTGCCGTGTGTGGTTCCCGCGG
GCCTGGCCTCTTTACGGGTTATGGCCCTTGCGTGCCTTGAATTACTTCCACCTG
GCTGCAGTACGTGATTCTTGATCCCGAGCTTCGGGTTGGAAGTGGGTGGGAGA
GTTCGAGGCCTTGCGCTTAAGGAGCCCCTTCGCCTCGTGCTTGAGTTGAGGCC

TGGCCTGGGCGCTGGGGCCGCCGCGTGCGAATCTGGTGGCACCTTCGCGCCT
GTCTCGCTGCTTTCGATAAGTCTCTAGCCATTTAAAATTTTTGATGACCTGCTGC
GACGCTTTTTTTCTGGCAAGATAGTCTTGTAAATGCGGGCCAAGATCTGCACAC
TGGTATTTCGGTTTTTGGGGCCGCGGGCGGCGACGGGGCCCGTGCGTCCCAG
CGCACATGTTCGGCGAGGCGGGGCCTGCGAGCGCGGCCACCGAGAATCGGA
CGGGGGTAGTCTCAAGCTGGCCGGCCTGCTCTGGTGCCTGGCCTCGCGCCGC
CGTGTATCGCCCCGCCCTGGGCGGCAAGGCTGGCCCGGTCGGCACCAGTTGC
GTGAGCGGAAAGATGGCCGCTTCCCGGCCCTGCTGCAGGGAGCTCAAAATGG
AGGACGCGGCGCTCGGGAGAGCGGGCGGGTGAGTCACCCACACAAAGGAAAA
GGGCCTTTCCGTCCTCAGCCGTCGCTTCATGTGACTCCACGGAGTACCGGGCG
CCGTCCAGGCACCTCGATTAGTTCTCGAGCTTTTGGAGTACGTCGTCTTTAGGT
TGGGGGGAGGGGTTTTATGCGATGGAGTTTCCCCACACTGAGTGGGTGGAGA
CTGAAGTTAGGCCAGCTTGGCACTTGATGTAATTCTCCTTGGAATTTGCCCTTTT
TGAGTTTGGATCTTGGTTCATTCTCAAGCCTCAGACAGTGGTTCAAAGTTTTTTT
CTTCCATTTCAGGTGTCGTGAGGAATTTCGACATTTAAATTTAATTAATCTCGAC
GGTATCGGTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGG
GAAAGAATAGTAGACATAATAGCAACAGACATACAAACTAAAGAATTACAAAAAC
AAATTACAAAAATTCAAAATTTTATCGATCACGAGACTAGCCTCGAGGTTTAAAC
TACGGGATCTTCGAAGGCCTAAGCTTACGCGCGCGTCCTAGCGCTACCGGTCG
CCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCT
GGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG
GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC
CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGC
GTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAA
GTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACG
ACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGT
GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTG
GGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGA
CAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGG
ACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGA
CGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTG
AGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGAC
CGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTCCGGACTCAGAT
CTCGACTAGCTAGTAGCTAGCTAGCTAGTCGAGCTCAAGCTTCGGGGACTAGT

CATATGATAATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCT
TAACTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTAT
CATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAATCCTGGT
TGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGCGTGGTG
TGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTGCCACCACCTG
TCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTATTGCCACGGCGGAACT
CATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGGGCTCGGCTGTTGGGCACT
GACAATTCCGTGGTGTTGTCGGGGAAGCTGACGTCCTTTCCATGGCTGCTCGC
CTGTGTTGCCACCTGGATTCTGCGCGGGACGTCCTTCTGCTACGTCCCTTCGG
CCCTCAATCCAGCGGACCTTCCTTCCCGCGGCCTGCTGCCGGCTCTGCGGCCT
CTTCCGCGTCTTCGCCTTCGCCCTCAGACGAGTCGGATCTCCCTTTGGGCCGC
CTCCCCGCATCGGTACGTATGGCCAGGTACCTTTAAGACCAATGACTTACAAGG
CAGCTGTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAA
TTCACTCCCAACGAAGACAAGATGGGATCAATTCACCATGCTAGTGGATCCCCC
TCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCA
CTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGT
GATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAG
TGAAAGTCGAGTTTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGT
TTACCACTCCCTATCAGTGATAGAGAAAGTGAAAGTCGAGTTTACCACTCCCT
ATCAGTGATAGAGAAAGTGAAAGTCGGGGCTGCAGGAATTCGAACGCTGACG
TCATCAACCCGCTCCAAGGAATCGCGGGCCCAGTGTCACTAGGCGGGAACACC
CAGCGCGCGTGCGCCCTGGCAGGAAGATGGCTGTGAGGGACAGGGGAGTGG
CGCCCTGCAATATTTGCATGTCGCTATGTGTTCTGGGAAATCACCATAAACGTG
AAATGTCTTTGGATTTGGGAATCTTATAAGTTCTGTATGAGACCACGCGTCCCC
CCATGGAAGACACGCTCATTTCAAGAGAATGAGCGTGTCTTCCATGGTTTTTGG
AAATATACCGTCGCATGGGAATAACTTCGTATAGCATACATTATACGAAGTTATG
CTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAG
CTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGA
GTGCTTCAAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCC
CTCAGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGCATCTAGAATTAATTCC
GTGTATTCTATAGTGTCACCTAAATCGTATGTGTATGATACATAAGGTTATGTAT
TAATTGTAGCCGCGTTCTAACGACAATATGTACAAGCCTAATTGTGTAGCATCTG
GCTTACTGAAGCAGACCCTATCATCTCTCTCGTAAACTGCCGTCAGAGTCGGTT
TGGTTGGACGAACCTTCTGAGTTTCTGGTAACGCCGTCCCGCACCCGGAAATG

GTCAGCGAACCAATCAGCAGGGTCATCGCTAGCCAGATCCTCTACGCCGGACG
CATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATC
GCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCG
CTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGG
CGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCA
ACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGA
ATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCG
ACACCCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCA
TCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTT
TTCACCGTCATCACCGAAACGCGCGAGACGAAAGGGCCTCGTGATACGCCTAT
TTTTATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTT
TCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAA
GGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGG
CATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATG
CTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGC
GGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACT
TTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAG
CAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCA
GTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCT
GCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGA
GGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCG
CCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTG
ACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGC
GAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGAT
AAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGG
GGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAG
GCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATT
AAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGATTGATTTAAA
ACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTTGATAATCTCATGA
CCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAA
AGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCA
AACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTAC

CAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTG
TTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGC
CTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATA
AGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAG
CGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGA
CCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTT
CCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG
GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCT
GTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG
GGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGG
CCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGT
GGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAA
CGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACG
CAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGTGGAATGT
GTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGC
AAAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCC
CCAGCAGGCAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTC
CCGCCCCTAACTCCGCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTC
TCCGCCCCATGGCTGACTAATTTTTTTTATTTATGCAGAGGCCGAGGCCGCCTC
GGCCTCTGAGCTATTCCAGAAGTAGTGAGGAGGCTTTTTTGGAGGCCTAGGCT
TTTGCAAAAAGCTTGGACACAAGACAGGCTTGCGAGATATGTTTGAGAATACCA
CTTTATCCCGCGTCAGGGAGAGGCAGTGCGTAAAAAGACGCGGACTCATGTGA
AATACTGGTTTTTAGTGCGCCAGATCTCTATAATCTCGCGCAACCTATTTTCCCC
TCGAACACTTTTTAAGCCGTAGATAAACAGGCTGGGACACTTCACATGAGCGAA
AAATACATCGTCACCTGGGACATGTTGCAGATCCATGCACGTAAACTCGCAAGC
CGACTGATGCCTTCTGAACAATGGAAAGGCATTATTGCCGTAAGCCGTGGCGG
TCTGTACCGGGTGCGTTACTGGCGCGTGAACTGGGTATTCGTCATGTCGATAC
CGTTTGTATTTCCAGCTACGATCACGACAACCAGCGCGAGCTTAAAGTGCTGAA
ACGCGCAGAAGGCGATGGCGAAGGCTTCATCGTTATTGATGACCTGGTGGATA
CCGGTGGTACTGCGGTTGCGATTCGTGAAATGTATCCAAAAGCGCACTTTGTCA
CCATCTTCGCAAAACCGGCTGGTCGTCCGCTGGTTGATGACTATGTTGTTGATA
TCCCGCAAGATACCTGGATTGAACAGCCGTGGGATATGGGCGTCGTATTCGTC
CCGCCAATCTCCGGTCGCTAATCTTTTCAACGCCTGGCACTGCCGGGCGTTGT
TCTTTTTAACTTCAGGCGGGTTACAATAGTTTCCAGTAAGTATTCTGGAGGCTGC

ATCCATGACACAGGCAAACCTGAGCGAAACCCTGTTCAAACCCCGCTTTAAACA
TCCTGAAACCTCGACGCTAGTCCGCCGCTTTAATCACGGCGCACAACCGCCTG
TGCAGTCGGCCCTTGATGGTAAAACCATCCCTCACTGGTATCGCATGATTAACC
GTCTGATGTGGATCTGGCGCGGCATTGACCCACGCGAAATCCTCGACGTCCAG
GCACGTATTGTGATGAGCGATGCCGAACGTACCGACGATGATTTATACGATACG
GTGATTGGCTACCGTGGCGGCAACTGGATTTATGAGTGGGCCCCGGATCTTTG
TGAAGGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGA
TTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAAACTACTG
ATTCTAATTGTTTGTGTATTTTAGATTCCAACCTATGGAACTGATGAATGGGAGC
AGTGGTGGAATGCCTTTAATGAGGAAACCTGTTTTGCTCAGAAGAAATGCCAT
CTAGTGATGATGAGGCTACTGCTGACTCTCAACATTCTACTCCTCCAAAAAAGA
AGAGAAAGGTAGAAGACCCCAAGGACTTTCCTTCAGAATTGCTAAGTTTTTTGA
GTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTTGCTATTTACACCACAAA
GGAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATATTCTGTAACCTTT
ATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACTCCACACAG
GCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGCTTT
TTAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAG
ATCATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCC
CACACCTCCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTT
GTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTTCACA
AATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGT
ATCTTATCATGTCTGGATCAACTGGATAACTCAAGCTAACCAAAATCATCCCAAA
CTTCCCACCCCATACCCTATTACCACTGCCAATTACCTGTGGTTTCATTTACTCT
AAACCTGTGATTCCTCTGAATTATTTTCATTTTAAAGAAATTGTATTTGTTAAATAT
GTACTACAAACTTAGTAG

8.3 Appendix 2

8.3.1 Parameters used for imaging RNA foci on the Opera Phoenix High Content Screening system

The nuclei were detected using the M method with the diameter set to 23 μm and a 0.3 splitting coefficient and common threshold, which split up the nuclei that were clumped together. The threshold for nuclei was then set to remove any auto-fluorescent areas or clumps. The parameters were nuclear Hoechst intensity ≥ 225, nucleus roundness ≥ 0.55 and nucleus area ≥ 50 μm^2. RNA foci spots were then detected using the A method and the thresholds for the parameters were set at high to best remove auto-fluorescence. The thresholds were based on the following parameters:

- Corrected Spot Intensity – 10 ≤ x ≤ 50
- Spot to Region Intensity – 0.6 ≤ x ≤ 3 to 4
- Spot Area (px^2) – 15 ≤ x ≤ 35 to 45
- Spot Background Intensity – 250 to 300 ≤ x ≤ 450 to 600
- Spot Contrast – 0.1 < x ≤ 0.3
- Region Intensity – x ≥ 250 to 300

The number of selected nuclear spots (RNA foci) per cell was then calculated and the cells were categorised into cells with foci, cells without foci, cells with 1 focus, cells with 2-4 foci and cells with 5+ foci.

8.4 Appendix 3

8.4.1 Genotyping results for 3-month-old adult zebrafish by fin clipping
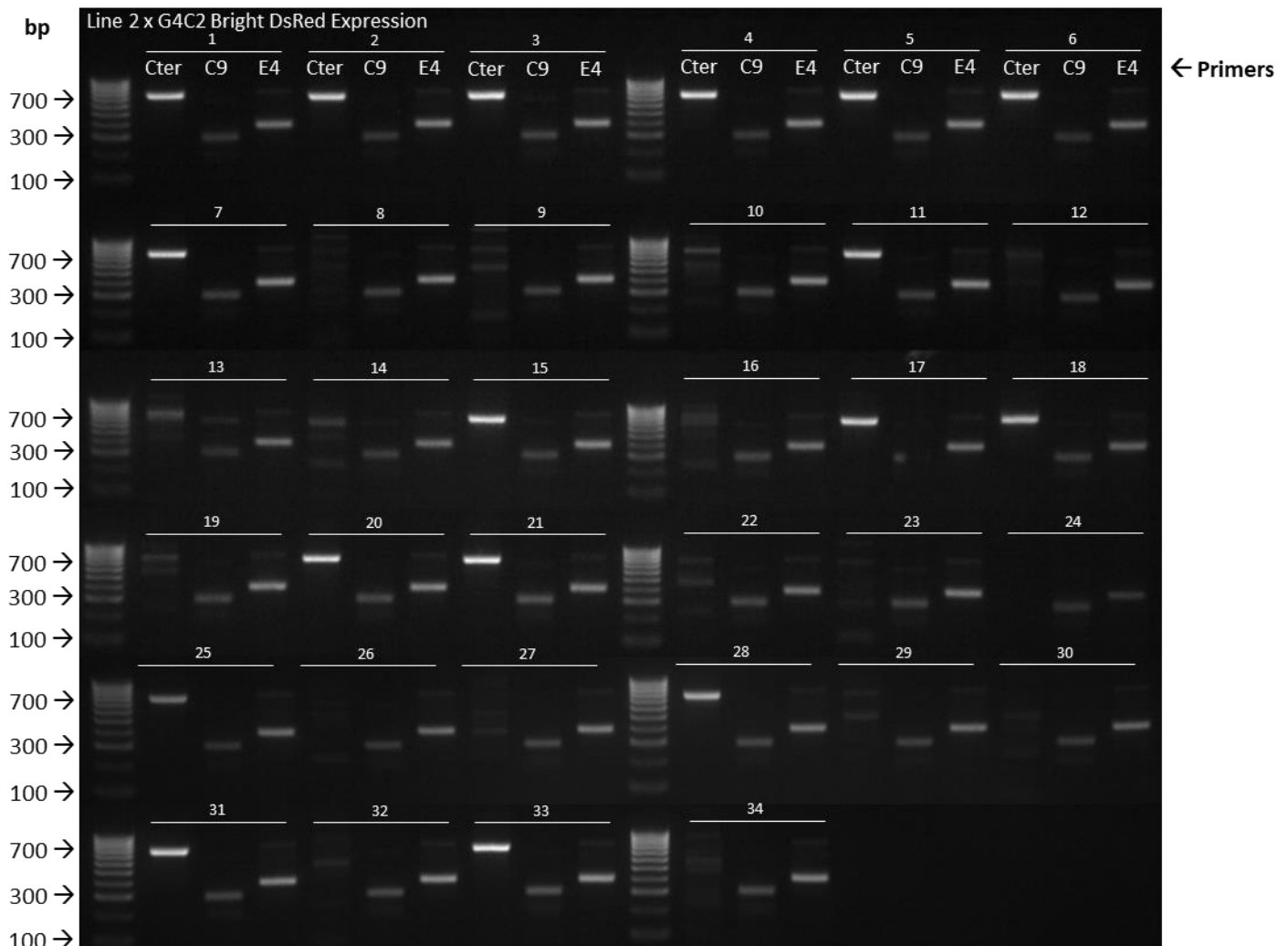

Figure A 1: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 2 transgenics with sense (G4C2) C9ORF72 line that showed Bright DsRed expression

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with sense (G4C2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and bright DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The bright DsRed expressors are a mixed population of double mutants with both the HC-ARFIP2 insert and C9 repeats (HC/C9); and fish with C9 repeats but no HC-ARFIP2 insert (C9). Genotyping showed that there were 18 HC/C9 and 16 C9 only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
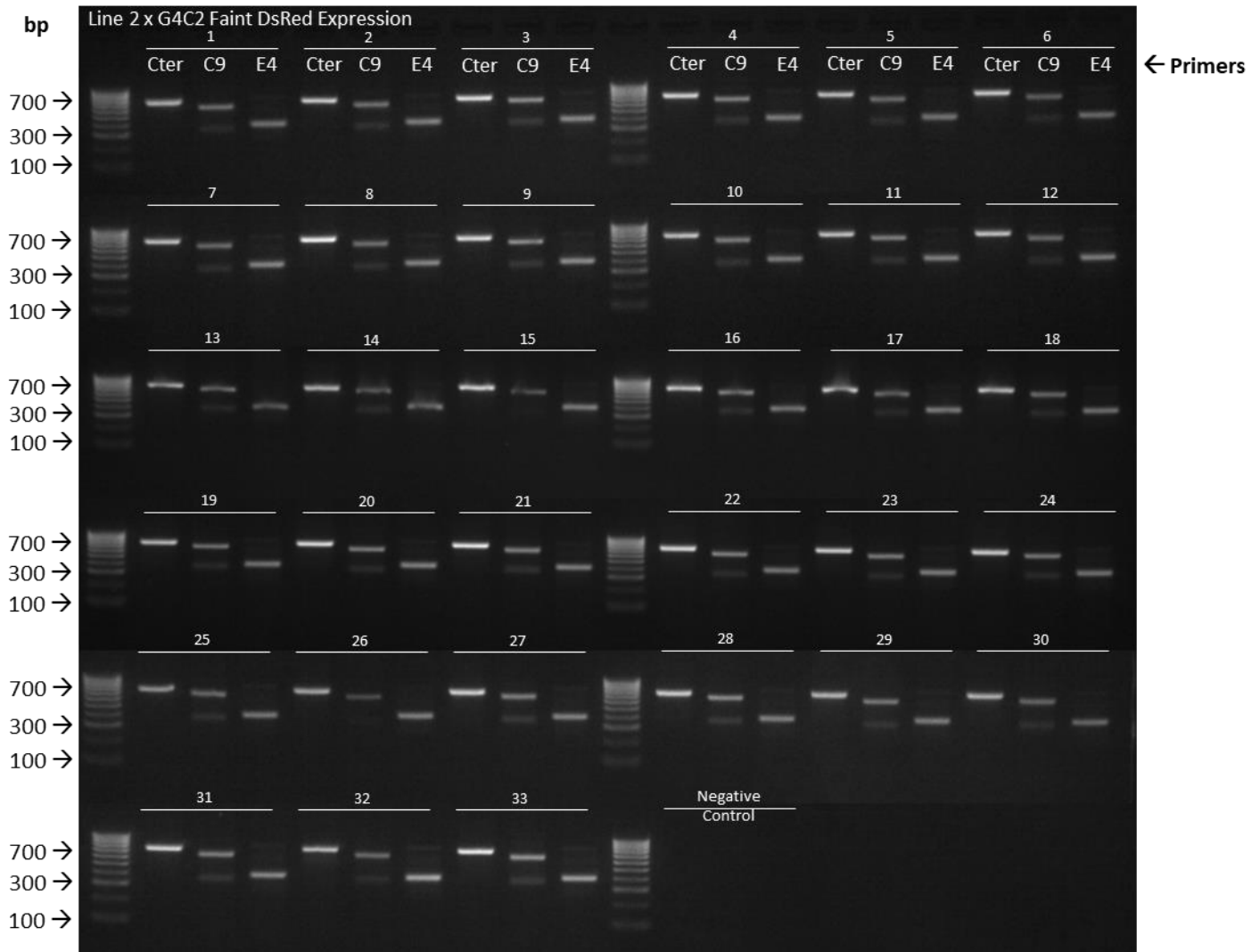
Figure A 2: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 2 transgenics with sense (G4C2) C9ORF72 line that showed faint DsRed expression

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with sense (G4C2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and faint DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The faint DsRed expressors are HC-ARFIP2 transgenics with no C9 repeats (HC). Genotyping showed that there were 33 HC only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
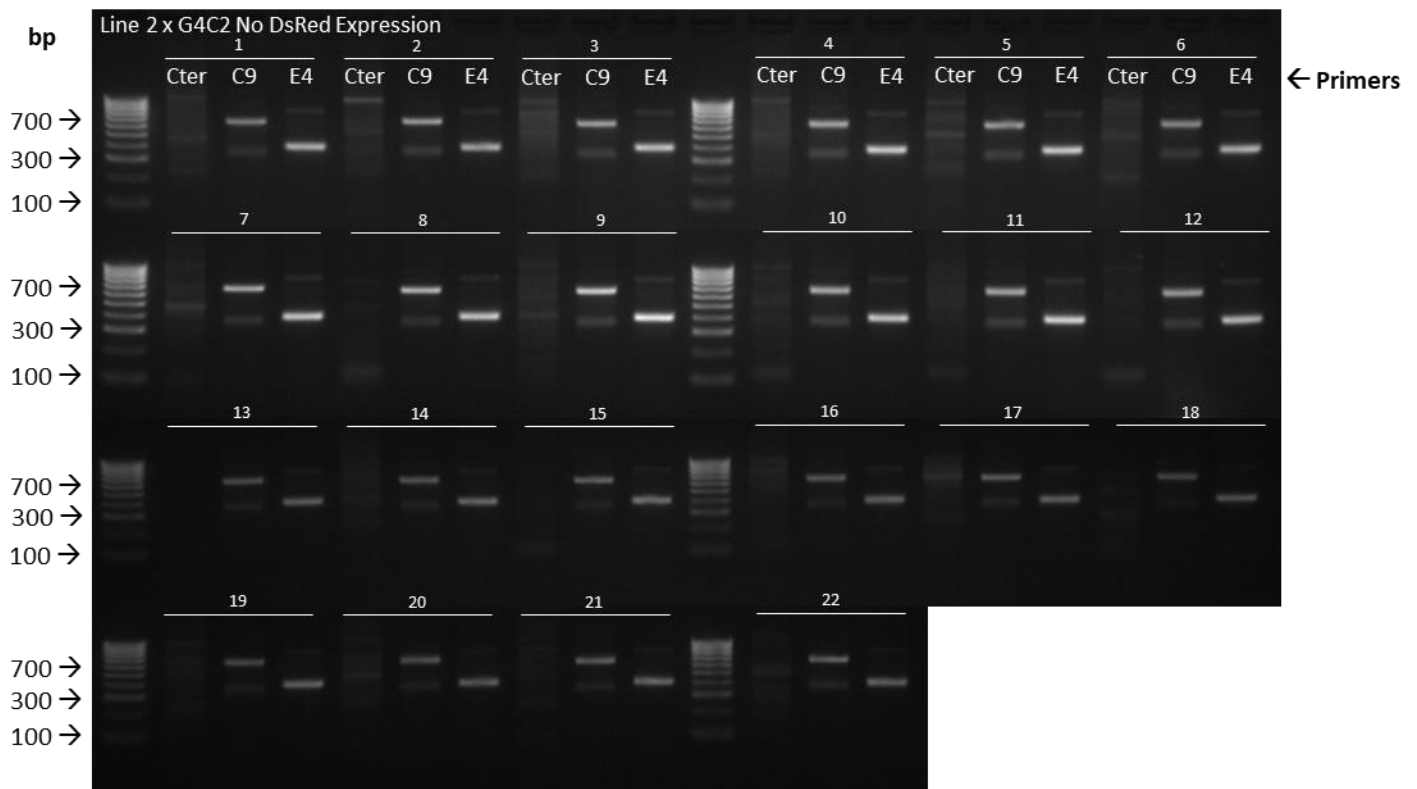
Figure A 3: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 2 transgenics with sense (G4C2) C9ORF72 line that showed no DsRed expression

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with sense (G4C2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and no DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The no DsRed expressors are non-transgenics with no HC-ARFIP2 or C9 repeats (NTG). Genotyping showed that there were 22 NTG adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
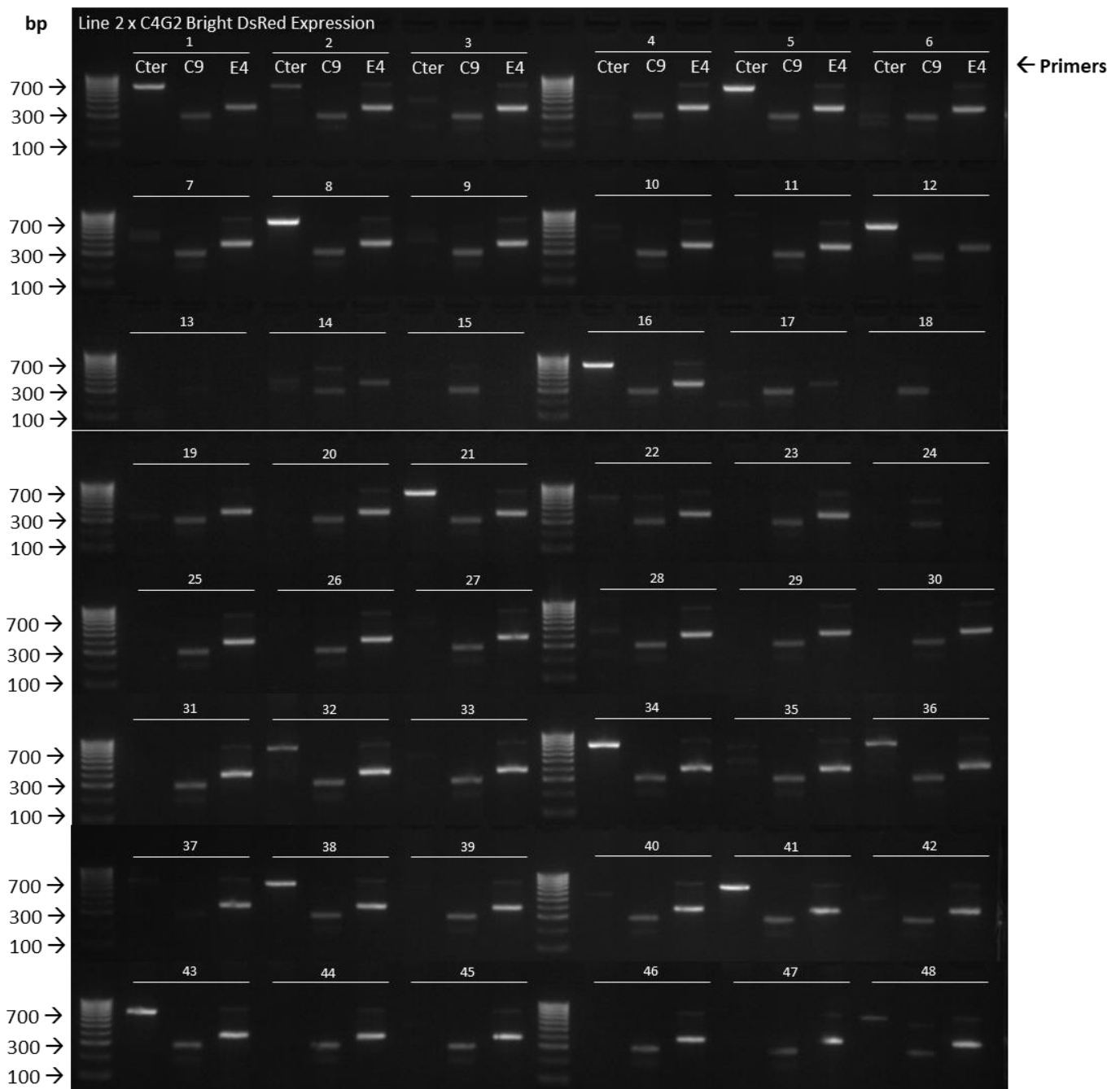
Figure A 4: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 2 transgenics with antisense (C4G2) C9ORF72 line that showed bright DsRed expression

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with antisense (C4G2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and bright DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The bright DsRed expressors are a mixed population of double mutants with both the HC-ARFIP2 insert and C9 repeats (HC/C9); and fish with C9 repeats but no HC-ARFIP2 insert (C9). Genotyping showed that there were 15 HC/C9 and 33 C9 only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
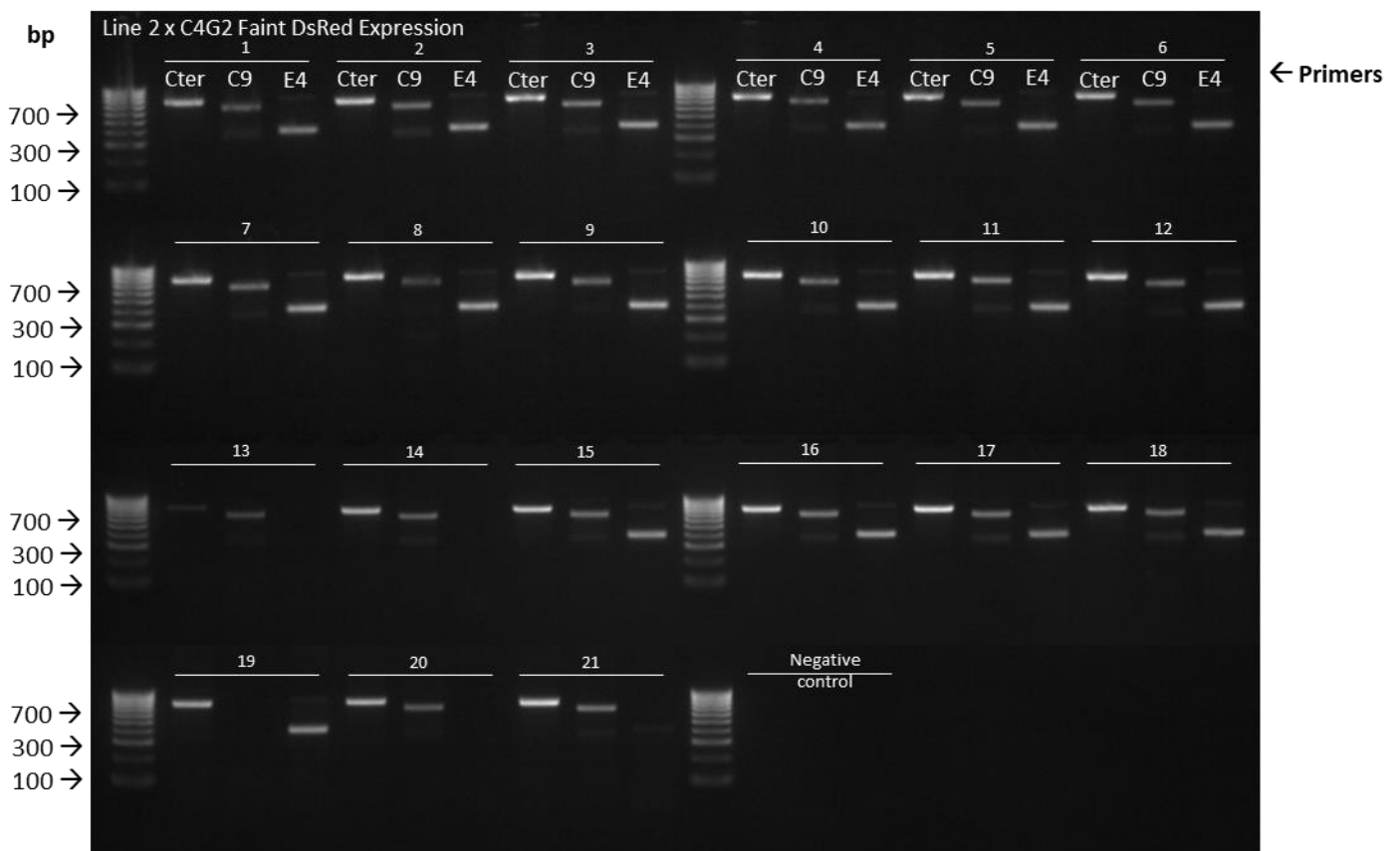
Figure A 5: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 2 transgenics with antisense (C4G2) C9ORF72 line that showed faint DsRed expression

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with antisense (C4G2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and faint DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The faint DsRed expressors are HC-ARFIP2 transgenics with no C9 repeats (HC). Genotyping showed that there were 21 HC only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
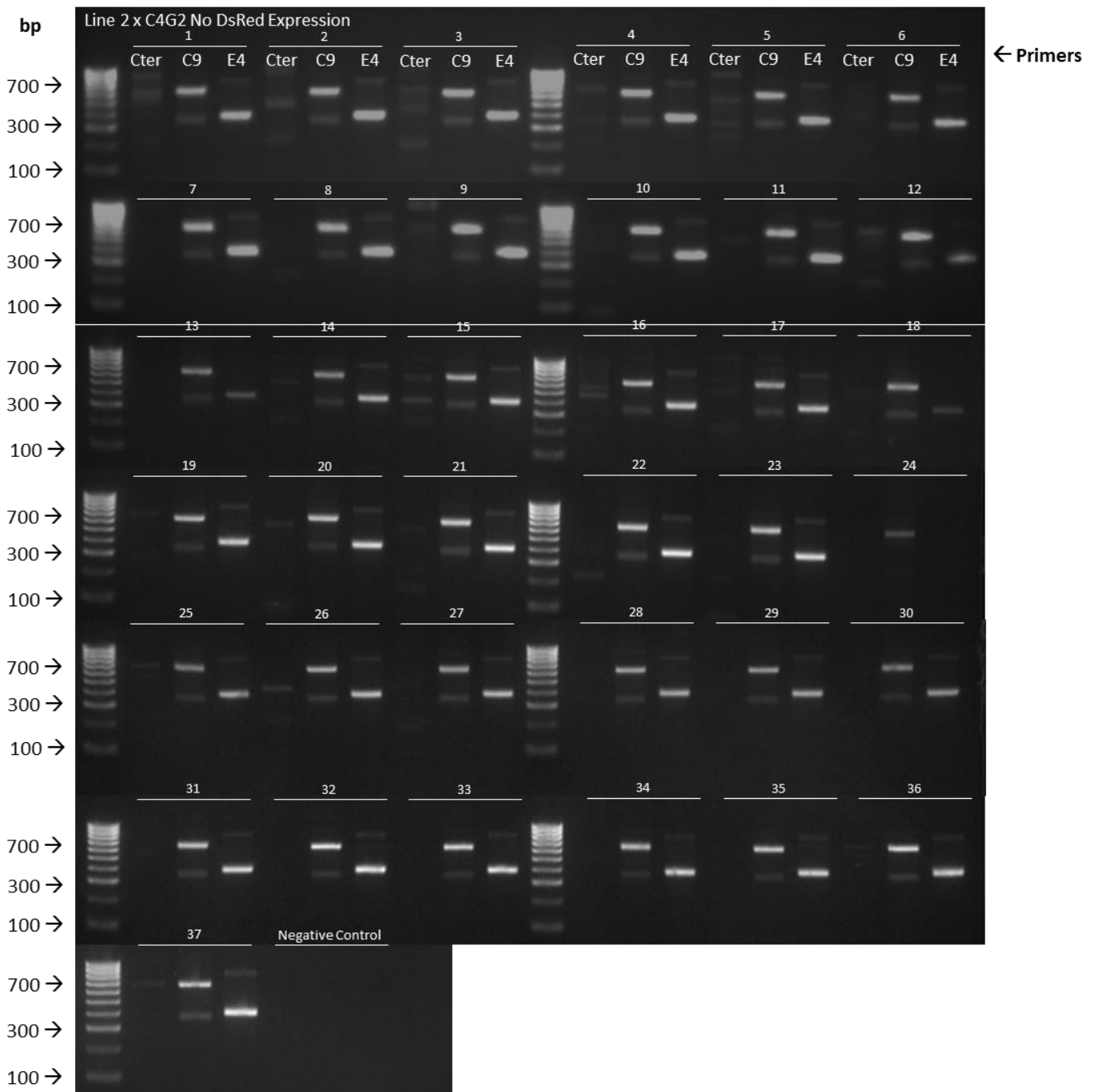
Figure A 6: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 2 transgenics with antisense (C4G2) C9ORF72 line that showed no DsRed expression

HC-ARFIP2 Line 2 heterozygous transgenics were crossed with antisense (C4G2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and no DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The no DsRed expressors are non-transgenics with no HC-ARFIP2 or C9 repeats (NTG). Genotyping showed that there were 37 NTG adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
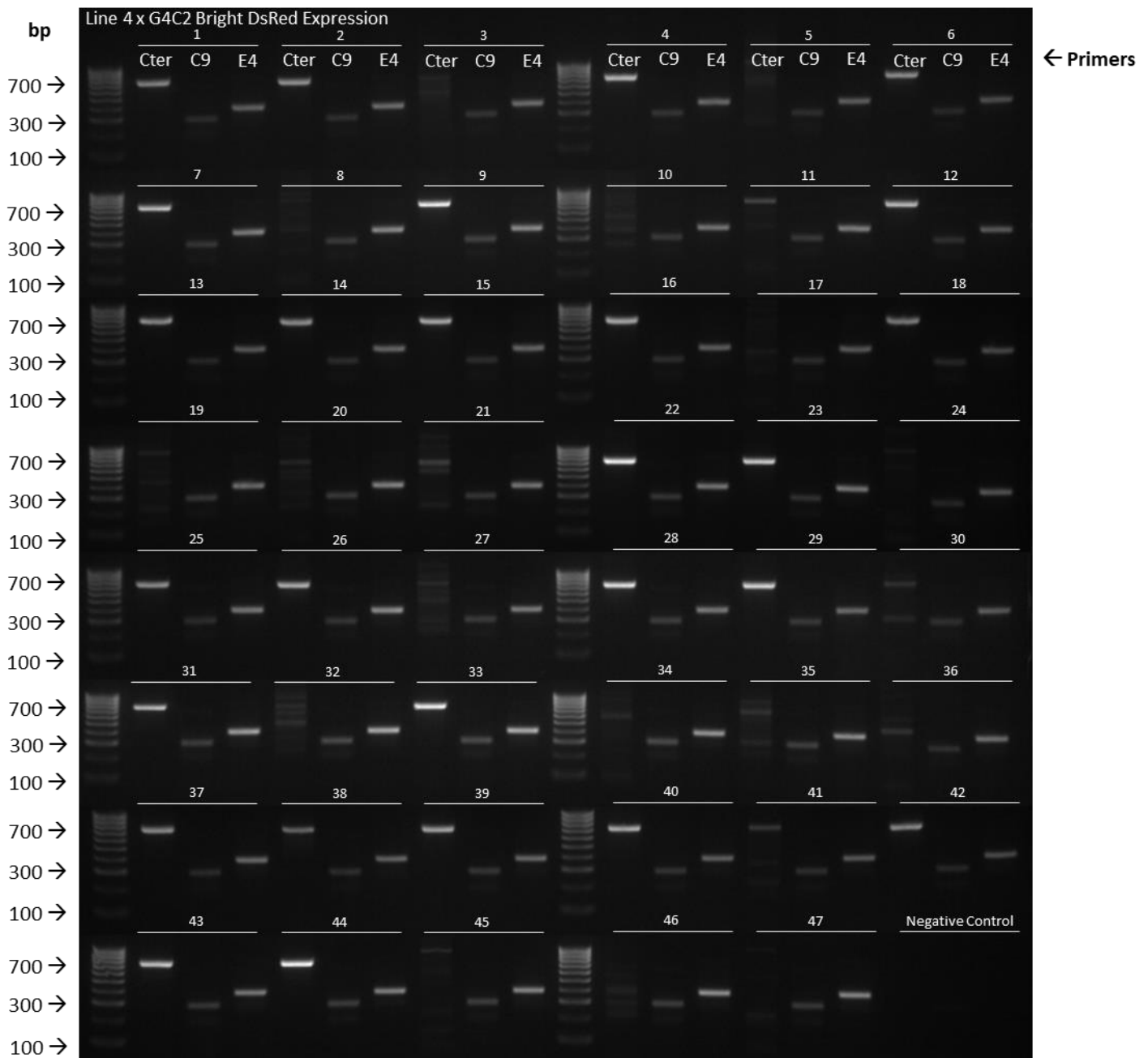
Figure A 7: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 4 transgenics with sense (G4C2) C9ORF72 line that showed Bright DsRed expression

HC-ARFIP2 Line 4 heterozygous transgenics were crossed with sense (G4C2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and bright DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The bright DsRed expressors are a mixed population of double mutants with both the HC-ARFIP2 insert and C9 repeats (HC/C9); and fish with C9 repeats but no HC-ARFIP2 insert (C9). Genotyping showed that there were 29 HC/C9 and 18 C9 only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).

Figure A 8: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 4 transgenics with sense (G4C2) C9ORF72 line that showed faint DsRed expression

HC-ARFIP2 Line 4 heterozygous transgenics were crossed with sense (G4C2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and faint DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The faint DsRed expressors are HC-ARFIP2 transgenics with no C9 repeats (HC). Genotyping showed that there were 41 HC only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
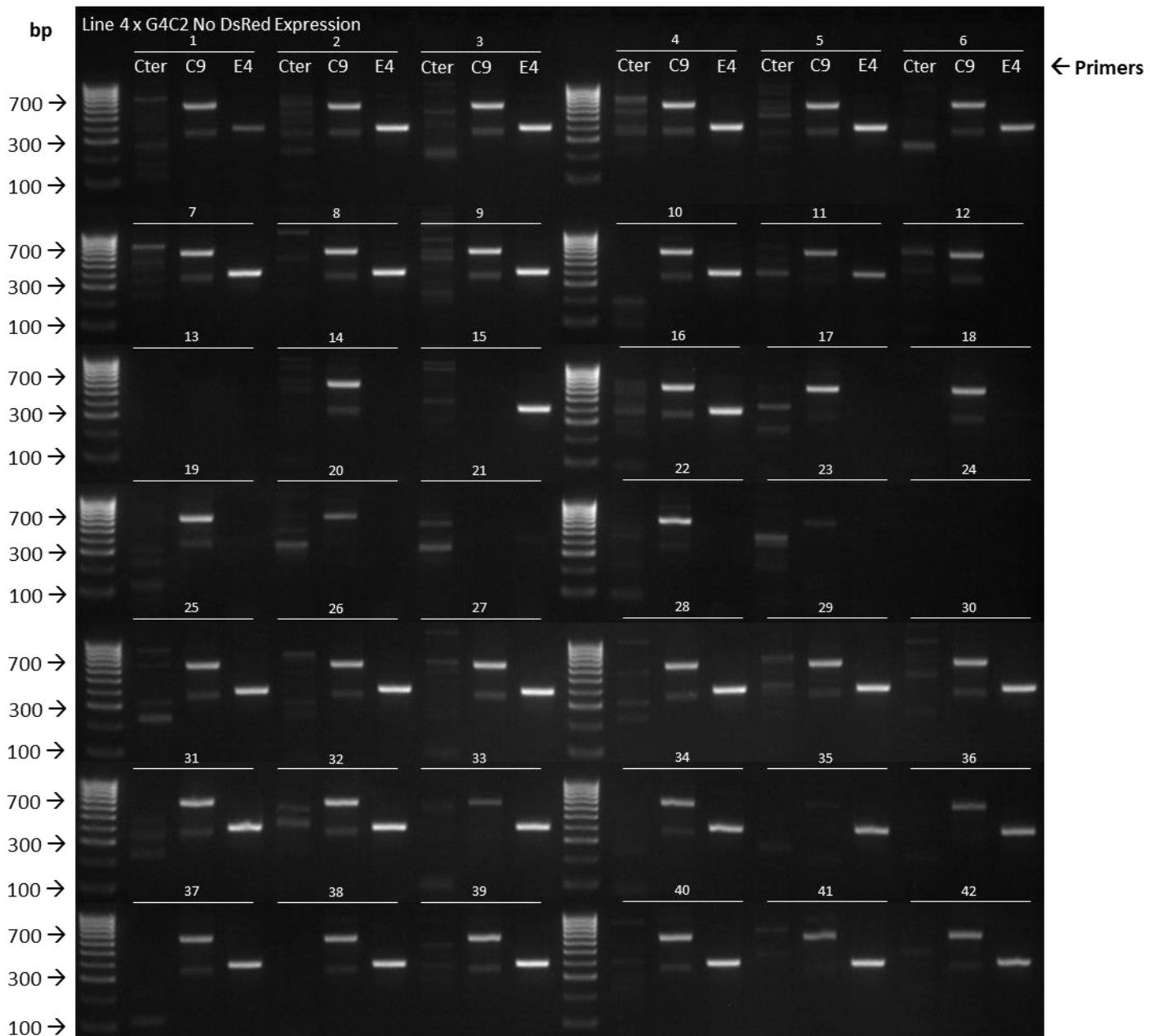
Figure A 9: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 4 transgenics with sense (G4C2) C9ORF72 line that showed no DsRed expression

HC-ARFIP2 Line 4 heterozygous transgenics were crossed with sense (G4C2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and no DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The no DsRed expressors are non-transgenics with no HC-ARFIP2 or C9 repeats (NTG). Genotyping showed that there were 42 NTG adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).

Line 4 x C4G2 Bright DsRed Expression

Figure A 10: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 4 transgenics with antisense (C4G2) C9ORF72 line that showed bright DsRed expression

HC-ARFIP2 Line 4 heterozygous transgenics were crossed with antisense (C4G2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and bright DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The bright DsRed expressors are a mixed population of double mutants with both the HC-ARFIP2 insert and C9 repeats (HC/C9); and fish with C9 repeats but no HC-ARFIP2 insert (C9). Genotyping showed that there were 54 HC/C9 and 17 C9 only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).
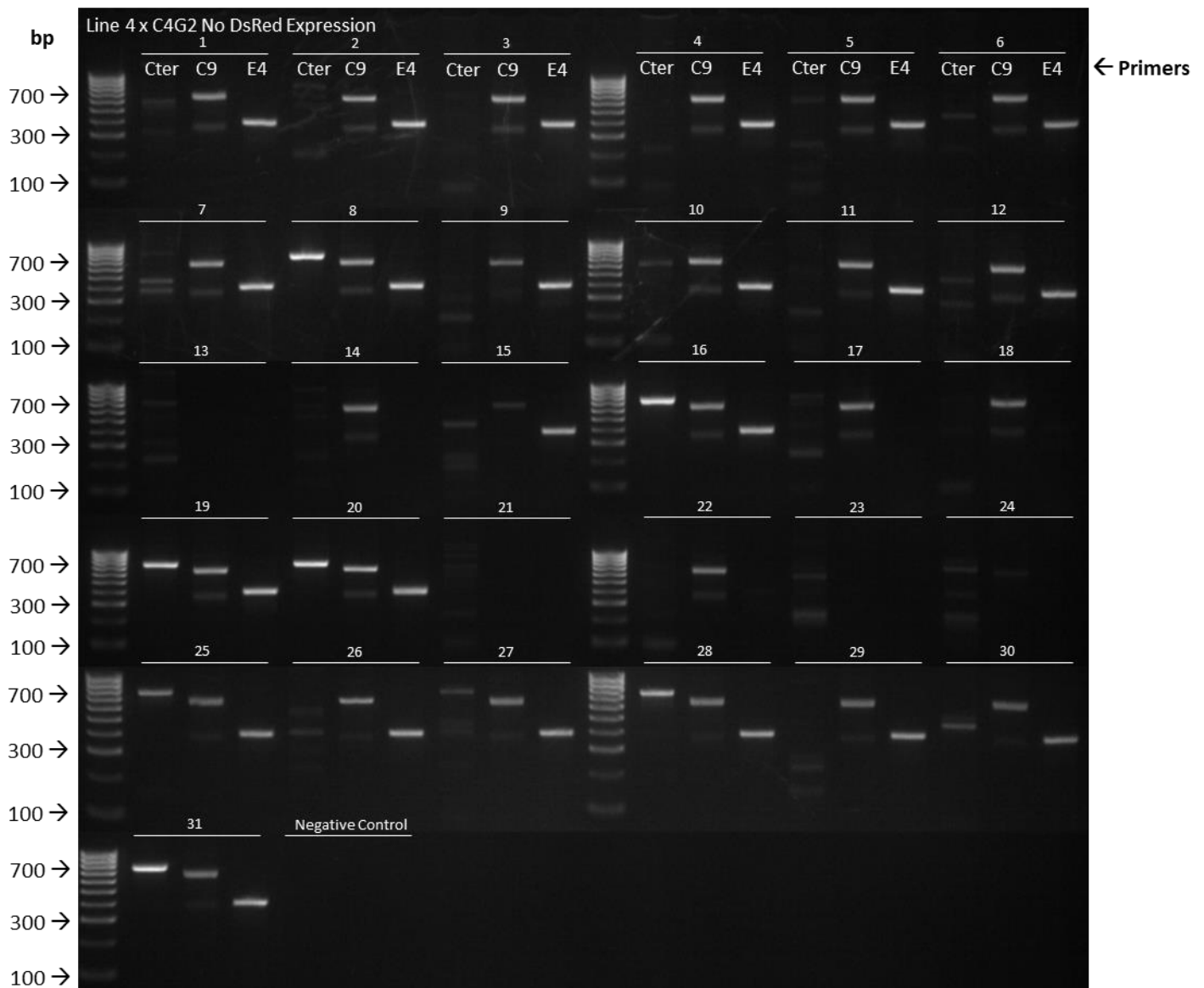
Figure A 11: Genotyping by fin clipping of adult zebrafish raised from crossing HC-ARFIP2 Line 4 transgenics with antisense (C4G2) C9ORF72 line that showed no DsRed expression

HC-ARFIP2 Line 4 heterozygous transgenics were crossed with antisense (C4G2) C9ORF72 heterozygous lines. The zebrafish fry was separated according to DsRed phenotypes and no DsRed expressors were raised to adulthood. Adults at 3 months were fin clipped and DNA was extracted. They were then genotyped using primers specific for HC-ARFIP2 (cter), C9ORF2 sense and antisense repeats (C9) and a reference gene (E4) as positive control. The no DsRed expressors are non-transgenics with no HC-ARFIP2 or C9 repeats (NTG). Genotyping showed that there were 24 NTG adults, but also 7 HC only adults. Expected bands (bp) – Cter (703), C9 (280), and E4 (353).

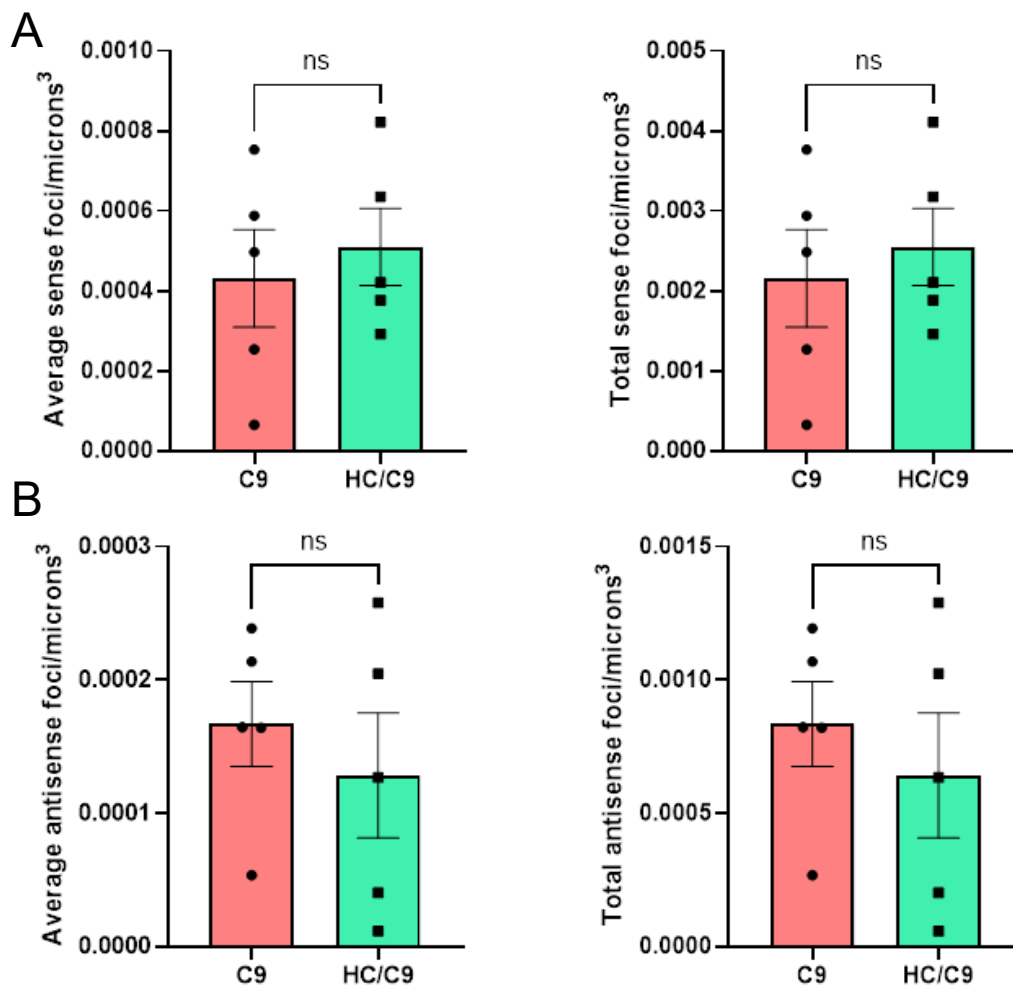### 8.4.2 Blind analysis of RNA foci in zebrafish embryos

Figure A 12: Additional quantification for effect of HC-ARFIP2 on RNA foci in zebrafish to show conservation of trends among individual quantifiers

Zebrafish from crossing HC-ARFIP2 Line 2 with sense and antisense C9ORF72 lines were stained with locked nucleic acid (LNA) probes specific for sense and antisense RNA foci at 5 days post fertilisation (dpf). **(A)** HC-ARFIP2 Line 2 was crossed with the sense C9ORF72 line. Average number of sense RNA foci observed in the C9-S population was 0.00043 and 0.00051 in the HC/C9-S double mutant population. Total number of sense RNA foci observed in the C9-S population was 0.0022 and 0.0025 in the HC/C9-S double mutant population. There was no significant difference in sense RNA foci between C9-S and HC/C9-S populations. **(B)** HC-ARFIP2 Line 2 was crossed with the antisense C9ORF72 line. Average number of antisense RNA foci observed in the C9-A population was 0.00017 and 0.00013 in the HC/C9-A double mutant population. Total number of antisense RNA foci observed in the C9-A population was 0.00083 and 0.00064 in the HC/C9-A double mutant population. There was no significant difference in antisense RNA foci between C9-A and HC/C9-A populations. Data were analysed using an unpaired t-test. Error bars represent ± SEM; $p < 0.05$. Ns = $P > 0.05$, * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$.

8.5 Outcomes of Ph.D. Programme

Publications

1. Sandoval-Castellanos, A. M.*, **Bhargava, A.***, Zhao, M., Xu, J. and Ning, K. (2023) 'Serine and arginine rich splicing factor 1: a potential target for neuroprotection and other diseases', Neural Regen Res, 18(7), pp. 1411-1416.

2. **Bhargava, A.***, Sandoval Castellanos, A. M.*, Shah, S. and Ning, K. (2022) 'An insight into the iPSCs-derived two-dimensional culture and three-dimensional organoid models for neurodegenerative disorders', Interface Focus, 12(5), pp. 20220040.

3. Boddy, S., Islam, M., Moll, T., Kurz, J., Burrows, D., McGown, A., **Bhargava, A.**, Julian, T. H., Harvey, C., Marshall, J. N., Hall, B. P., Allen, S. P., Kenna, K. P., Sanderson, E., Zhang, S., Ramesh, T., Snyder, M. P., Shaw, P. J., McDermott, C. and Cooper-Knock, J. (2022) 'Unbiased metabolome screen leads to personalized medicine strategy for amyotrophic lateral sclerosis', Brain Commun, 4(2), pp. fcac069.

4. Karyka, E., Berrueta Ramirez, N., Webster, C. P., Marchi, P. M., Graves, E. J., Godena, V. K., Marrone, L., **Bhargava, A.**, Ray, S., Ning, K., Crane, H., Hautbergue, G. M., El-Khamisy, S. F. and Azzouz, M. (2022) 'SMN-deficient cells exhibit increased ribosomal DNA damage', Life Sci Alliance, 5(8).

Equal contribution

Manuscripts in preparation

Anushka Bhargava, Cleide dos Santos de Souza, Allan C Shaw, David Burrows, Ergita Bali, Adrian Higginbottom, Ramesh Tennore, Pamela J Shaw, and Ke Ning. Identification of Arfaptin-2 as a therapeutic target for Amyotrophic Lateral Sclerosis (ALS), *in preparation*.

Conferences

- **Oral presentations**
  - Neurodegeneration Early-Career Researcher (ECR) Meeting, Virtual, 2021 (One of four selected speakers at the international conference by an external judging panel) – *Identification of Arfaptin-2 as a potential therapeutic target for Amyotrophic Lateral Sclerosis (ALS) using iPS-*

derived motor neurons and zebrafish as models of ALS. Anushka Bhargava, Tennore Ramesh, Pamela J Shaw, Ke Ning.

- o North-East Postgraduate Conference (NEPG), Virtual, 2020 – *Developing a gene therapy for the disease that inspired the 'Ice Bucket Challenge'.* Anushka Bhargava, Tennore Ramesh, Pamela J Shaw, Ke Ning.

- **Poster presentations**
    - o American Society for Gene and Cell Therapy 25th Annual Meeting (ASGCT), Washington D.C., USA 2022 – *Identification of Arfaptin-2 as a potential therapeutic target for Amyotrophic Lateral Sclerosis (ALS) using iPSC-derived motor neurons and zebrafish as models of ALS.* Anushka Bhargava, Tennore Ramesh, Pamela J Shaw and Ke Ning.
    - o Motor Neuron Disease Association (MNDA), International Symposium on ALS/MND, Virtual, 2021 – *Identification of Arfaptin-2 as a potential therapeutic target for Amyotrophic Lateral Sclerosis (ALS) using iPSC-derived motor neurons and zebrafish as models of ALS.* Anushka Bhargava, Tennore Ramesh, Pamela J Shaw and Ke Ning.
    - o Medical School Annual Research Meeting, Sheffield, UK, 2020 - *Identification of Arfaptin-2 as a potential therapeutic target for Amyotrophic Lateral Sclerosis (ALS) using iPSC-derived motor neurons and zebrafish as models of ALS.* Anushka Bhargava, Tennore Ramesh, Pamela J Shaw and Ke Ning.
    - o British Society for Gene and Cell Therapy Annual Conference (BSGCT), Sheffield, UK, 2019 (selected for a 1-minute poster pitch) – *Arfaptin 2 as a therapeutic target for amyotrophic lateral sclerosis (ALS).* Aida Mohammedeid, Anushka Bhargava, Vinay K Godena, Su-Chii Kong, Ramesh Tennore, Andrew J Grierson, Jun Xu, Pamela J Shaw, Mimoun Azzouz and Ke Ning.

Published Abstracts

1. Bhargava, T. Ramesh, P. Shaw and K. Ning (2021) Theme 07 - PRE-CLINICAL THERAPEUTIC STRATEGIES, Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 22:sup2, 121-134, DOI:

10.1080/21678421.2021.1985793 - From the 32nd International Symposium on ALS/MND.

2. Mohammedeid, A., A. Bhargava, V. K. Godena, S. Kong, R. Tennore, A. J. Grierson, J. Xu, P. J. Shaw, M. Azzouz, and K. Ning. "Arfaptin-2 as a therapeutic target for amyotrophic lateral sclerosis (ALS)." In HUMAN GENE THERAPY, vol. 30, no. 8, pp. A19-A20.

3. Bhargava, A., Ramesh, T., Shaw, P. J., & Ning, K. (2022, April). Identification of Arfaptin-2 as a Potential Therapeutic Target for Amyotrophic Lateral Sclerosis (ALS) Using iPSC-Derived Motor Neurons and Zebrafish as Models of ALS. In MOLECULAR THERAPY (Vol. 30, No. 4, pp. 298-299).

Awards/Grants

- Awarded a research grant from the Motor Neuron Disease Association (MNDA) worth £10,000.

- Awarded a travel grant from The Guarantors of Brain worth £1000 as a contribution to the cost of travel to Washington DC, USA for the ASGCT 2022 conference.

- Received funding worth £760.50 from the Think Ahead Research-led Activities fund (UoS) to organise events for researcher professional development in collaboration with the Research Staff Association.

- £250 prize for being selected as a Ph.D. ECR speaker at the Neurodegeneration Early-Career Researcher (ECR) Meeting, 2021.

Outreach

- Writer for the 'Hot Off the Press' lay summaries initiative.

Teaching and Training Opportunities

- Designed and managed 6-week placement projects for four undergraduate students (2019 and 2020).

- Supervised theses for two MSc students (2019 and 2021).