THE UNIVERSITY
of EDINBURGH

ASSOCIATION OF BRITISH NEUROLOGISTS

Investigation of Endoplasmic Reticulum Stress and Protein Aggregates in *C9ORF72* Mutant Induced Pluripotent Stem Cell-Derived Spinal Motor Neurons

Angus Buchanan Gane

Declaration

I have composed this thesis myself acknowledging clearly in the text where material from others has been contributed with permission.

This work has not been submitted for any other degree or professional qualification.

Angus Gane

Acknowledgements

I started this project with no prior wet lab experience and am extremely grateful to Professor Siddharthan Chandran, my primary supervisor, for the opportunity to undertake this MScR which has allowed me to learn many techniques in a supportive environment.

Dr Bhuvaneish Thangaraj Selvaraj was my day-to-day supervisor and main teacher for molecular techniques, cell culture and imaging. The platedown and maintenance of neuronal cultures from which he too was conducting experiments was a joint effort. He has provided critical feedback on all aspects of the MScR including the completed thesis and has been patient and supportive through success and failure.

Dr Samantha Barton, Dr Navneet Vasistha and Mr Owen James provided 'benchside' advice and taught me techniques in molecular biology, cell culture and imaging. I am particularly grateful for the contributions of Dr Barton who, having trained elsewhere, has been able to provide alternative protocols for many of the techniques used which have enhanced my understanding and technical repertoire. Dr Barton has provided feedback on this manuscript.

Dr Dario Magnani has provided training on the ImageXpress along with advice on immunocytochemistry protocols. The majority of this data is not presented and will be published elsewhere.

Professor Giles Hardingham analysed the RNAseq data before the MScR commenced.

Dr Peter Connick guided statistical analysis of survival data. Initially, I approached Dr Connick with a list of statistical considerations which he expanded on before suggesting an appropriate model. The analysis presented in this thesis is my own, following some of the principles outlined by Dr Connick.

Mr Pratap Harbham and Dr Arpan Mehta joined the lab as I was approaching the end of my MScR. They were involved in the maintenance of neuronal cultures. Through conversation, our knowledge of neuroscience and statistics has grown.

Dr Jenna Gregory was always happy to talk about the molecular pathology of motor neuron disease and her experience in protein chaperones and neuropathology has guided my interpretation of the results obtained.

Dr Karen Burr, Mrs Karen Gladstone and Mrs Nicola Clements maintain the induced pluripotent stem cell lines used by the lab without which none of the following would have been possible.

David Story provided invaluable support as the lab manager by maintaining stocks, obtaining supplies and providing financial support.

I would also like to thank Dr Steve Morley, my personal tutor, for his support and encouragement in allowing me a leave of absence from the MBChB to complete this MScR.

I am grateful to the Association of British Neurologists who provided a £5000.00 bursary for me to complete this intercalated degree. This work was presented at their annual conference in Liverpool on the 5th of May, 2017.

Finally, it would be remiss for me not to thank family and friends, in particular, my parents. I need say no more than none of this would have been possible without them.

Abstract

Introduction: Protein aggregations are pathological hallmarks of motor neuron disease (MND) and have been shown to cause activation of endoplasmic reticulum (ER) stress sensors. ER stress-induced pathways can lead to multiple outcomes including degradation of protein aggregates or cell death. Using RNAseq data from *C9ORF72* mutants as a starting point we sought to investigate the relationship between ER stress and protein aggregation in induced pluripotent stem cell (iPSC) derived spinal motor neurons (sMNs).

Methods: iPSCs were differentiated into sMNs from subjects with the *C9ORF72* repeat expansion, $TARDBP^{G298S}$ and $TARDBP^{M337V}$ mutations with CRISPR/Cas9 gene-corrected controls. Previously obtained RNAseq data from *C9ORF72* mutants was subject to GO analysis with candidates validated by qPCR. ER stress pathways were assessed by PCR for XBP1 splicing and Western blotting for ATF6 and PERK in untreated and tunicamycin-treated sMNs at three weeks post platedown. A survival assay following tunicamycin treatment was performed on three-week-old sMNs. TDP-43 pathology has been characterised by Western blotting and immunocytochemistry (ICC) and dipeptide repeat (DPR) proteins by ICC.

Results: RNAseq data from *C9ORF72* mutant sMNs was enriched for genes involved in protein homeostasis. HSPB8, a protein implicated in clearing TDP-43 aggregations, was confirmed to be upregulated in *C9ORF72* mutant sMNs. TDP-43 pathology, however, was not present in these cells. DPR protein aggregates are present in *C9ORF72* mutants at this time but there was no evidence of ER stress activation. Treatment of *C9ORF72* mutants with tunicamycin, however, activates an ER stress response determined by increased PERK and ATF6 splicing to an extent greater than that seen in controls. Treatment with tunicamycin was not associated with a survival deficit. For comparison, *TARDBP* mutants were investigated which display TDP-43 pathology and ER stress evidenced by ATF6 splicing.

Conclusion: There is early evidence for differences in ER stress in both *C9ORF72* and *TARDBP* mutants. DPR proteins may disrupt cellular protein homeostasis in a manner distinct from that of TDP-43 pathology and are an early feature of disease in *C9ORF72* mutants. These findings have implications for understanding the pathogenesis of MND and drug screening.

Additional Chapter: In the absence of a gene-corrected control for $TARDBP^{M337V}$ a strategy for CRISPR/Cas9 mediated gene-correction of the M337V mutation is presented.

Lay Summary

Motor neuron disease (MND) is a disorder in which the cells that allow us to move muscle (motor neurons) die. Understanding why these motor neurons die may help us to treat MND.

One of the common features of MND is the presence of clumps of protein within motor neurons. In this study, we look at MND caused by two errors in DNA which can lead to the formation of such protein clumps. These two errors affect two different bits of DNA – one called *C9ORF72* and the other *TARDBP*. Both errors lead to clumping of a protein called TDP-43. The *C9ORF72* error also results in another type of protein clump containing dipeptide repeat (DPR) proteins. Our cells can only work if they contain the right amount of protein and this balance is maintained by a part of the cell called the endoplasmic reticulum (ER). Too much protein results in ER stress pathway activation.

To study the effect of these DNA errors on protein clumping and ER stress, cells from patients with the DNA errors were converted to motor neurons grown in dishes in the laboratory.

When levels of ER stress in three-week-old motor neurons from patients with MND due to the *C9ORF72* error were compared to those of healthy people no difference was observed. These cells contained the DPR protein clumps but not the TDP-43 clumps. When these cells were treated with the drug tunicamycin to cause ER stress they reacted more strongly than normal cells. This suggests that although they do not show ER stress under normal conditions they do have a vulnerability to stress. To see if such stress might impact the lifespan of motor neurons we treated cells with the *C9ORF72* error with tunicamycin but there was no difference in survival.

In comparison, the cells with the *TARDBP* DNA error did show TDP-43 aggregation at three weeks and evidence of an active ER stress response. This is in keeping with the idea that TDP-43 protein clumps can activate ER stress pathways that may lead to motor neuron death in MND.

Further investigation is required, but these findings are important because they help us to link the protein clumps observed in patients with MND to a possible mechanism by which motor neurons may die. It also suggests that even though TDP-43 protein clumps are present in most cases of MND they may play different roles in patients depending on the DNA error that caused the MND.

Table of Contents

List of Figures

Supplemental Figures

Supplemental Figure 1: HSP90B1 transcript levels in three-week sMNs are increased in C9-3 but consistent between other lines examined.

Supplemental Figure 2: Without outlier exclusion, significant upregulation of HSPB8 was observed only in C9-3.

Supplemental Figure 3: XPO1 transcript is not differentially expressed among control, *C9ORF72* mutant and gene corrected lines.

Supplemental Figure 4: Full y-axis plot of Figure 21C.

Supplemental Figure 5: Plot showing that mean normalised integrated density (mnID) does not decrease significantly with increasing doses of DMSO up to 10gL/mL.

Supplemental Figure 6: GA and GP DPR protein aggregates are present at higher levels in a *C9ORF72* mutant line relative to control lines in three-week-old sMN cultures.

List of Tables

Supplemental Tables

Abbreviations

3D – Three dimensional
A – Adenine
ALS – Amyotrophic lateral sclerosis
AM – Acetoxymethyl
ASO – Antisense oligonucleotide
ATF4 – Activating transcription factor 4
ATF6 – Activating transcription factor 6
BAC – Bacterial artificial chromosome
BCA – Bicinchoninic acid
BDNF – Brain-derived neurotrophic factor
bp – Base pair
BSA – Bovine serum albumin
C – Cytosine
C9ORF72 – Chromosome 9 open reading frame 72
cDNA – Complementary DNA
ChAT – Choline acetyltransferase
CHOP – CCAAT-enhancer-binding protein homologous protein
CI – Confidence interval
CNS – Central nervous system
CNTF – Ciliary-derived neurotrophic factor
CRISPR – Clustered regularly interspaced short palindromic repeats
DAPI – 4′,6-diamidino-2-phenylindole dichloride
DDR – DNA damage repair
DDT - Dichlorodiphenyltrichloroethylene
DF – Dilution factor
DMSO – Dimethyl sulfoxide
DNA – Deoxyribonucleic acid
DNAse – Deoxyribonuclease
dNTP – Deoxynucleotide
DPR – Dipeptide repeat
DSB – Double stranded break
DSHB – Developmental studies hybridoma bank
EDTA – Ethylenediaminetetraacetic acid
eIF2α – Eukaryotic transcription factor 2 alpha
ER – Endoplasmic reticulum
ERAD – Endoplasmic reticulum-associated degradation
ERSE – Endoplasmic reticulum stress response element
ESC – Embryonic stem cell
FACS – Fluorescence-activated cell sorting
FDU – Fluorodeoxyuridine
FGF2 – Fibroblast growth factor 2
fMND – Familial motor neuron disease
FTD – Frontotemporal dementia
FUS – Fused in sarcoma

G – Guanine
G_4C_2 – GGGGCC [nucleic acid sequence]
GA – Glycine-alanine
GADD34 – Growth arrest and DNA damage gene 34
GAPDH – Glyceraldehyde-3-phosphate dehydrogenase
GDNF – Glial-derived neurotrophic factor
GFAP – Glial fibrillary acidic protein
GO – Gene ontology
GP – Glycine-proline
GR – Glycine-arginine
gRNA – Guide RNA
GRP78 – Glucose related protein 78kDa
GWAS – Genome-wide association study
h – Hour(s)
HCl – Hydrochloric acid
HDR – Homology-directed repair
HRE – Hexanucleotide repeat expansion
HRT – Homology repair template
HSP – Hereditary spastic paraplegia
ICC – Immunocytochemistry
ID – Integrated density
IGF-1 – Insulin-like growth factor 1
IgG – Immunoglobulin G
iPSC – Induced pluripotent stem cell
IRE1 – Inositol-requiring enzyme 1
kDa – KiloDaltons
LMF – Laminin, MatriGel, Fibronectin
LMN – Lower motor neuron
MAP2 – Microtubule-associated protein 2
MG – MatriGel
MN – Motor neuron
MND – Motor neuron disease
MN-FM – Motor neuron feeding media
mnID – Mean normalised integrated density
mRNA – Messenger ribonucleic acid
NEB – New England Biolabs
NHEJ – Non-homologous end joining
NRT – No reverse transcriptase
NTC – No template control
PA – Proline-alanine
PAGE – Polyacrylamide gel electrophoresis
PANTHER – Protein analysis through evolutionary relationships
PBP – Progressive bulbar palsy
PBS – Phosphate buffered saline
PCR – Polymerase chain reaction
PDI – Protein disulphide isomerase
PERK – Protein kinase RNA-like endoplasmic reticulum kinase

PG – Proline-glycine
PH – Proportional hazards
PI – Propidium iodide
PLS – Primary lateral sclerosis
PMA – Progressive muscular atrophy
PR – Proline-arginine
PVDF – Polyvinylidene fluoride
qPCR – Quantitative PCR
RA – Retinoic acid
RAN – Repeat-associated non-ATG
RFLP – Restriction fragment length polymorphism
RI – Rock inhibitor
RIDD – IRE1 dependent decay
RIPA – Radioimmunoprecipitation assay
RNA – Ribonucleic acid
RNase – Ribonuclease
RNAseq – RNA sequencing
RNP – Ribonucleoprotein
ROS – Reactive oxygen species
RT – Reverse transcriptase
SAG – Smoothened agonist
SDS – Sodium dodecyl sulfate
SHH – Sonic hedgehog
SEM – Standard error of the mean
sgRNA – Single guide RNA
SMA – Spinal muscular atrophy
sMN – Spinal motor neuron
SMN – Survival motor neuron
SOD1 – Superoxide dismutase 1
ssODN – Single stranded oligodeoxynucleotide
T – Thymine
TALEN – Transcription activator-like effector nuclease
TAR – Transactive response element
TARDBP – TAR DNA binding protein
TDP-43 – TAR DNA binding protein 43
TE – Tris-EDTA
UMN – Upper motor neuron
UPR – Unfolded protein response
UPRE – Unfolded protein response element
WB – Western blot
WGS – Whole genome sequencing
WT – Wildtype
XBP1 – X-box binding protein
ZFN – Zinc finger nuclease

Chapter 1: Background

What is motor neuron disease?

The nomenclature surrounding motor neuron disease is inconsistent and complicated. The term motor neuron diseases (plural) is used to describe a family of conditions in which the clinical features of disease arise primarily from the degeneration of motor neurons (MNs) These disorders include spinal muscular atrophy (SMA), hereditary spastic paraplegia (HSP), post-irradiation lumbosacral radiculopathy, motor neuron diseases of a known infective or metabolic aetiology and motor neuron disease (MND) itself. MND is further divided into different clinical syndromes depending on the location of affected motor neurons (**Table 1**). Upper motor neurons (UMNs) are neurons situated in the brain or brainstem which project axons to lower motor neurons (LMNs). The cell bodies of LMNs lie within the brainstem or anterior horn of the spinal cord and project axons to muscle.

Motor Neurons Affected	**Syndrome Name**
LMNs	Progressive muscular atrophy (PMA)
UMNs	Primary lateral sclerosis (PLS)
LMNs and UMNs throughout the body	Amyotrophic lateral sclerosis (ALS)
LMNs and UMNs of the bulbar region	Progressive bulbar palsy (PBP)

Table 1: MND subtypes by MN group affected

The most common MND syndrome is amyotrophic lateral sclerosis (ALS). In the United Kingdom, these terms are often used interchangeably though a difference exists. In the United States, ALS and MND are truly synonymous. Whether these different clinical syndromes reflect different underlying disease processes remains a question for debate. The clinical syndrome of individuals within families with familial MND (fMND) often varies though suggesting a common predisposition to MND which is modified by factors within the affected individual to determine the specific syndrome that develops (Mulder et al. 1986; Siddique and Pericak-Vance 1989).

The aetiology of MND is complex and, as for most other diseases, it is often stated that it is likely both genetic and environmental factors interact to cause disease. Various environmental

factors have been suggested over the years including viral and heavy metal based aetiologies but the evidence has never given prominence to any one theory (Ingre et al. 2015). Arguably more progress has been made in dissecting the genetic contributors to MND and an ever-growing number of genes are now associated with the disease. Some mutations identified appear able to cause MND with Mendelian inheritance and high penetrance, others act as risk factors or disease-modifiers (Leblond et al. 2014). Evidence of oligogenic MND (in which more than one MND-associated gene has been found to carry mutations in an individual with MND) is also increasing with a recent systematic review and meta-analysis suggesting as many as 0.4% of cases are oligogenic – far higher than would be expected by chance (Zou et al. 2017). This percentage will only increase as more mutations are associated with MND. This is in keeping with a population level study that used statistical modelling techniques to suggest the development of MND is a multistep process (Al-Chalabi et al. 2014). MND-associated with a hexanucleotide repeat expansion (HRE) in *C9ORF72* (termed C9-MND) is likely to provide an excellent insight into the multistep nature of this process given that the *C9ORF72* HRE has been associated with a number of neurodegenerative diseases including MND (DeJesus-Hernandez et al. 2011; Majounie et al. 2011), frontotemporal dementia (FTD) and less frequently in Huntington disease phenocopies (Hensman et al. 2014; Kostic et al. 2014); Alzheimer's disease (Majounie et al. 2012; Kwon et al. 2013) and others.

The increasing number of genetic mutations associated with MND has given rise to a plethora of model systems which have been used to study neurodegenerative mechanisms. Two affected genes have been studied by the Chandran lab – *C9ORF72* and *TARDBP*. The *C9ORF72* HRE is the most common cause of MND in European populations where it accounts for around 33.7% of familial and 5.1% of sporadic cases at the latest count (Zou et al. 2017). Mutations in *TARDBP* also potentially allow inference about a large proportion of MND cases because, although mutations in *TARDBP* are much rarer, the protein product, TDP-43, is a major component of inclusion bodies in most MND cases (Neumann et al. 2006). The notable exceptions are MND caused by mutations in *SOD1* (Mackenzie et al. 2007) and *FUS* (Vance et al. 2009).

C9-MND pathology is the focus of this work and endoplasmic reticulum stress and the unfolded protein response have been investigated in an attempt to understand the contribution of proteinopathy to MND with data from *TARDBP* mutant lines included for comparison at points.

Since association of the *C9ORF72* HRE with MND and FTD (DeJesus-Hernandez et al. 2011; Majounie et al. 2011) three main hypotheses have arisen seeking to explain how the HRE causes neurodegeneration: loss of function of endogenous C9ORF72; neurodegeneration as a result of toxic RNA species and toxicity resulting from dipeptide repeat (DPR) proteins. These processes are complex and have been the subject of recent reviews but some of the key findings and issues are described here.

A loss of function neurodegenerative mechanism relies on the observation of decreased expression or function of the C9ORF72 protein. The function of C9ORF72 is still poorly understood. In *C9ORF72* orthologs in *Caenorhabditis elegans* and zebrafish, *F18A1.6* and *zC9orf72* respectively, reducing ortholog expression resulted in motor phenotypes (Ciura et al. 2013; Therrien et al. 2013). In zebrafish, this deficit was rescued by expression of human C9ORF72 (Ciura et al. 2013). Neurodegeneration following C9orf72 knockdown has not, however, been replicated in rodent models (Lagier-Tourenne et al. 2013; Wen et al. 2014; Koppers et al. 2015). Work on human tissue has lent little clarity to the debate. The transcription of *C9ORF72* results in three different pre-mRNAs. In variants 1 and 3 the HRE lies within an intronic region whereas in variant 2 the HRE exists within the promoter. Different studies on post-mortem tissue and stem cell models have suggested that all transcript levels are decreased or that transcripts are differentially affected – reviewed by Gitler et al. (Gitler and Tsuiji 2016). A key argument against a loss of function mechanism of neurodegeneration comes from the observation that individuals homozygous for the mutant allele do not develop as severe a disease as might be expected given the impairment of endogenous C9ORF72 function (Fratta et al. 2013). Furthermore, mutations have not yet been identified with exons of *C9ORF72* which might be expected to cause pathology if C9ORF72 is required for normal functioning of the nervous system.

The second key hypothesis suggests that RNA foci which accumulate in the nuclei and cytoplasm of cells from patients with the *C9ORF72* HRE cause neurodegenerative disease (DeJesus-Hernandez et al. 2011; Majounie et al. 2011). In other diseases such as myotonic dystrophy type 1 RNA foci have been shown to result in sequestration of vital RNA processing machinery leading to toxicity (Echeverria and Cooper 2012). Similarly, many independent studies have shown that RNA foci in C9ORF72 sequester a number of RNA binding proteins such as Pur-alpha and hnRNPA1. In a *C9ORF72* HRE iPSC model treatment with antisense oligonucleotides (ASOs) targeting both within and out with the HRE sequence rescued a glutamate-induced excitotoxicity phenotype which it was claimed demonstrated an RNA foci-mediated toxic gain of function mechanism (Donnelly et al. 2013). The effect of ASO treatment on the production of DPRs though was not assessed and could confound interpretation of those results. Evidence from other sources argues against the role of a toxic RNA as a disease driver. The presence of RNA foci throughout a number of cell types in human MND and FTD not prone to neurodegeneration (Dejesus-hernandez et al. 2017) along with their presence in animal models such as the BAC transgenic mouse described earlier (Peters et al. 2015) makes it unlikely they alone are sufficient to cause neurodegeneration despite their close association with the *C9ORF72* HRE.

Finally, DPR proteinopathy is considered. Unconventional non-ATG mediated translation of the *C9ORF72* HRE in all six reading frames leads to the production of six different DPR proteins: glycine-proline (GP), glycine-alanine (GA), glycine-arginine (GR), proline-glycine (PG), proline-arginine (PR) and proline-alanine (PA). In the absence of RNA foci, evidence has been obtained that DPR proteins alone are capable of inducing neurodegenerative phenotypes. In one study poly-GA expression in mice caused neurodegeneration and behavioural abnormalities along with deficits in nucleocytoplasmic transport (Zhang et al. 2016). While this finding supports the idea that DPR proteins may drive neurodegeneration it leaves a number of questions unanswered relating to the distribution of DPR proteins in human disease and also in the lack of co-existing TDP-43 pathology which is present in almost all cases of MND and FTD associated with the *C9ORF72* HRE. Again, DPR pathology in human subjects correlates poorly with neurodegeneration (MacKenzie et al. 2013) and DPRs are rarely seen in sMNs which are predominantly affected in ALS (Gomez-Deza et al. 2015).

Motor neuron disease as a proteinopathy

The association of *TARDBP* mutations with the major component of protein aggregates in MND has supported the hypothesis of MND as a proteinopathy. The mechanisms by which cells maintain proteostasis are briefly reviewed focusing on those aspects relevant to the experimental data. Evidence suggesting endoplasmic reticulum (ER) stress and the unfolded protein response (UPR) is important in MND is then discussed.

Protein homeostasis

Protein folding describes the process by which proteins acquire a three-dimensional structure. It is noteworthy that not all proteins are folded. Alpha-synuclein, best known for its role in Parkinson's disease, is an intrinsically disordered protein which, in the central nervous system (CNS) of humans and other species, exists as a monomeric form with no consistent 3D structure (Fauvet et al. 2012).

The manner in which proteins fold is complex, described in a review article as a 'stochastic search' rather than 'mandatory steps' (Dobson 2003). Protein folding is subject to both intrinsic and extrinsic factors. Intrinsic factors include the primary, secondary and tertiary structures of a protein. In addition to quaternary protein structure are other environmental factors that affect protein folding. One such physiologically relevant environmental factor is a crowded molecular environment (Zhou 2013). Within the cell, different proteins are folded within different compartments including the cytoplasm, mitochondria and endoplasmic reticulum (Dobson and Ellis 1998).

A group of proteins called chaperones also serve important roles in the protein folding pathway. Chaperones improve the efficiency of protein folding by reducing the likelihood of competing reactions rather than acting as catalysts (Dobson 2003). In addition to chaperones, proteins with a catalytic role in protein folding have been described. Of particular note here with respect to MND are the protein disulphide isomerases (PDIs) which are involved in the

establishment of disulphide bonds (Wilkinson and Gilbert 2004). Dysfunction of both these classes of molecule has been implicated in the pathophysiology of MND.

If normal protein folding cannot be achieved, for example, because a protein is destabilised through mutation, an unfolded protein response (UPR) is launched by three stress sensors that reside in the membrane of the endoplasmic reticulum (ER) – IRE1, PERK and ATF6 (**Figure 1**).

Inositol Requiring Enzyme 1 (IRE1)

On ER stress, dissociation of glucose-related protein 78kDa (GRP78, also commonly referred to as BiP) from IRE1 activates RNase activity that results in excision of a 26bp unconventional intron leading to a truncated X-box binding protein 1 (XBP1) transcript and XBP1 protein. The truncated XBP1 transcript encodes a stable transcription factor which is capable of binding to the ER stress response element (ERSE) (Yoshida et al. 2001). The consequences of activation of the ERSE are described below.

Protein kinase RNA-like endoplasmic reticulum kinase

On ER stress, protein kinase RNA-like endoplasmic reticulum kinase (PERK) undergoes oligomerization and autophosphorylation (Bertolotti et al. 2000). Phosphorylation of PERK leads to phosphorylation of its downstream target eukaryotic translation initiation factor 2 alpha (eIF2α) (Harding et al. 1999). Phosphorylation of eIF2α causes almost global translational inhibition (Harding et al. 2000). One notably spared transcript is that of activating transcription factor 4 (ATF4) which is itself capable of inducing the transcription factor CCAAT-enhancer-binding protein homologous protein (CHOP) (Harding et al. 2000). The downstream effects of CHOP upregulation are broad and include induction of apoptosis and increasing reactive oxygen species (ROS) generation (Nishitoh 2012).

The response of PERK to ER stress is subject to a negative feedback loop involving growth arrest and DNA damage gene 34 (GADD34) (Novoa et al. 2001).

ATF6

ATF6 is inserted into the ER membrane following its synthesis. Under ER stress GRP78 dissociates from ATF6 exposing two Golgi localisation sequences resulting in translocation of ATF6 to the Golgi apparatus (Shen et al. 2002). In the Golgi apparatus, ATF6 is cleaved and this cleaved fraction translocates to the nucleus where it functions as a transcription factor. Promotors containing a cis-acting ERSE are activated leading to increased expression of a number of chaperone proteins (Yoshida et al. 1998). Upregulated targets include GRP78, CHOP and XBP1 (Yoshida et al. 2000).

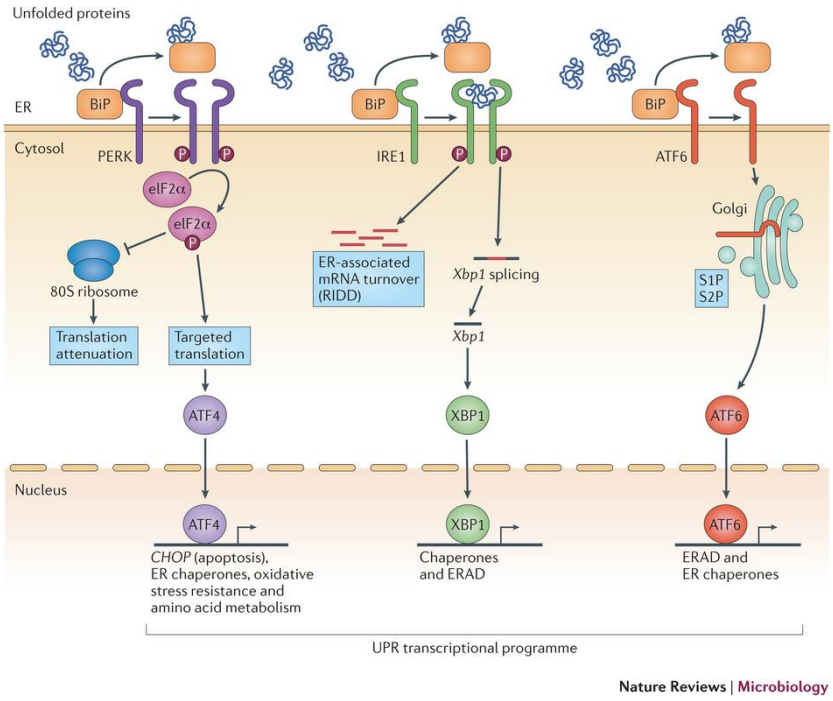


Figure 1: Simplified pathways activated by ER stress sensors PERK, IRE1 and ATF6 from (Celli and Tsolis 2014). Abbreviations: RIDD – IRE1 dependent decay, S1P – site 1 protease, S2P – site 2 protease, ERAD – endoplasmic reticulum-associated degradation. Processes following the activation of the three ER stress sensors are described in the text.

The unfolded protein response and endoplasmic reticulum stress in MND

Post-mortem examination the anterior horn cells of patients with MND has revealed abnormalities in the ER (Oyanagi et al. 2008). Pathology has also been reported to affect the Golgi apparatus – another key organelle involved in protein homeostasis (Gonatas et al. 1992). In *SOD1* mutant mouse models of MND Golgi fragmentation has been shown to occur prior to muscle denervation supporting a dying-forwards hypothesis in which soma degeneration precedes that of the axon (van Dis et al. 2014).

Of the two major groups of protein aggregates relevant to C9-MND, TDP-43 and the DPR proteins, the evidence favours a TDP-43 based pathology in humans. Evidence for and against a DPR proteinopathy has been presented in *C9-MND*. Here, it is noted that TDP-43 pathology is better spatially correlated with neurodegeneration in the human nervous system as opposed to DPR proteins which occur throughout the CNS in a variety of unaffected cells (MacKenzie et al. 2013; Davidson et al. 2014).

In cases of MND and FTD TDP-43 pathology has been described in several forms: cytoplasmic aggregation, nuclear clearance, truncated C-terminal fragments and hyperphosphorylated (Neumann et al. 2006). Evidence has recently been provided that TDP-43 pathology may be made worse by ER stress through inhibition of non-sense mediated RNA decay (Li et al. 2016). PERK appears the main ER stress sensor required for non-sense mediated RNA decay and has been shown to be upregulated in MND post-mortem tissue.

It should also be noted despite the common nature of inclusion bodies in MND there is evidence that *TARDBP* mutations can cause MND in the absence of TDP-43 aggregation or nuclear clearance though such cases are rare (Arnold et al. 2013)

The role of stem cell models

The ability to generate cellular models from human subjects with MND has granted new opportunities to investigate neurodegeneration in human biochemical systems. Several

methods now exist by which somatic human cells can be biopsied and converted to a cell type of interest. The two main strategies are:

- iPSC generation followed by differentiation
- direct reprogramming

Human cellular models of MND provide a mechanism by which genetic configurations that cause human disease can be studied in their entirety. This is in contrast to animal models in which gene mutations or mutant proteins are introduced recapitulating only specific parts of pathology though this of course in cases is capable of producing applicable phenotypes. The ability to study the interaction between many risk loci is likely to be particularly important in complex genetic diseases like MND. The same phenomenon allows dissection of pathophysiological processes underlying sporadic disease. Reprogramming technologies have proved particularly valuable in MND because of the practical issues obtaining neural tissue biopsies in life. Studies were, as a result, previously confined to end-stage pathology save in rare occurrences where the subjects death occurred early in the disease course in a manner not directly dependent on the disease. For this reason, the ability to examine the progression of pathology, as is possible in animal models, in human disease is invaluable. Components of these pathways can then be manipulated increasing our mechanistic understanding of the disease process.

Cellular models are not without limitations. While it is desirable that the base sequence be conserved though out the reprogramming process and subsequent passaging, sporadic mutations can occur (Ji et al. 2012). Concerns also arise regarding the conservation of epigenetic memory (Vaskova et al. 2013). For example, protective and degenerative roles for hypermethylation of the mutant *C9ORF72* allele have been proposed. Evidence of a protective effect of hypermethylation includes the observation that the number of RNA foci decrease on hypermethylation and increased vulnerability to autophagy and oxidative stress is abated (Liu et al. 2014). These potentially neuroprotective effects have also been studied in human patients (McMillan et al. 2015; Russ et al. 2015). It has also been suggested that hypermethylation in this area might lead to transcriptional silencing and loss-of-function neurodegeneration but strong evidence for this is currently lacking (Haeusler et al. 2016).

There is evidence that the specific methylation status of patients with *C9ORF7*2 HRE-associated disease is preserved following the reprogramming of C9-MND patient fibroblasts (Esanov et al. 2016). It is also important to remember that MND usually occurs after 50 years of age and the relevance of studying cells which are usually just weeks (even days) old is to be considered. There is excellent evidence though now that with longitudinal neuronal culture more mature cellular phenotypes are observed. A clear example of this comes from the field of Alzheimer's disease in which the abundance of different tau isoforms has been observed to vary in cellular models from developmental signatures to that seen in adult humans over a year in culture (Sposito et al. 2015).

Variability among different cell lines is considerable and has been widely reported. A review (Soldner and Jaenisch 2012) suggests three possible levels at which is variability could arise:

1. Reprogramming
2. Differentiation
3. Genetic background

In vivo differentiation protocols often produce a mixed cell population. In the spinal motor neuron (sMN) differentiation protocol used here approximately 60% of cells express motor neuron markers. The neuronal population is said to account for 90% of cells though we believe this closer to 100% at one week with a slowly increasing non-neuronal population as time in culture increases. A protocol generating a completely pure culture of sMNs would be valuable particularly for biochemical studies seeking to study this cell type specifically. Increasing the efficiency of differentiation protocols is complex and costly but other alternatives exist to enrich cultures of cell types of interest. These usually require additional genomic manipulation though such as the integration of antibiotic resistance genes for selection or reporters for processes like fluorescence-activated cell sorting (FACS). We have chosen not to take such measures in an effort to preserve the genome as much as possible. Processes like FACS are also more amenable to some cell types than others. FACS of MNs would axotomise cells and would be associated with a low survival though it is possible.

Differences in genetic background confound analysis in many disease models. Advances in gene–editing have allowed the creation of isogenic control lines in which mutations can be removed using the CRISPR/Cas9 system. This allows pairwise comparisons in which the only variables are the mutation (either present or corrected) and the gene-editing process. At a recent meeting of the Dementias Platform UK Stem Cell Network, it was suggested that mutant lines should be subject to sham correction. This would involve mock transfection, low-density platedown, clonal selection, screening and expansion. Such controls have not been used here though they may prove valuable in the future particularly following identification of seemingly pathological phenotypes in gene-corrected lines which are described later.

iPSC model systems have been able to recapitulate a number of aspects of *C9ORF72* pathology. RNA foci have been demonstrated in iPSC-derived motor neurons from C9-MND patients allowing mechanistic dissection of their protein binding partners (O'Rourke et al. 2013). The contribution of RNA foci to nucleocytoplasmic shuttling deficits has been reported and rescued with ASO treatment (Donnelly et al. 2014). Nucleolar and ER stress have been identified in a paper detailing the effects the DNA and RNA structures formed as a result of the *C9ORF72* HRE (Haeusler et al. 2014). Of particular relevance to this work, ER function has been investigated with evidence of calcium dysregulation in the ER and stress granule formation (Dafinca et al. 2016). The power of these models to dissect human biochemical pathways leading to a mechanistic understanding of C9-MND is immense and the use of such systems is only likely to increase with time.

Aims

The aim of this project was to validate RNA sequencing candidates involved in protein homeostasis and investigate the nature of the unfolded protein response in *C9ORF72* mutant iPSC-derived spinal motor neurons.

Hypotheses

Prediction 1) Motor neurons derived from subjects with motor neuron disease will exhibit activation of the unfolded protein response which will decrease survival.

Prediction 2) The presence of an activated unfolded protein response will correlate with TDP-43 and DPR protein pathology.

Chapter 2: Materials and Methods

Materials

Chemicals and Kits

Material	Company	Catalogue Number	Details
Accutase® Solution	SIGMA	A6964-100MG	
Advanced DMEM/F12	ThermoFisher Scientific	12634010	
Agarose, Molecular Grade	Bioline	BIO-41025	
Alexa Fluor® Dyes	ThermoFisher Scientific	Multiple products	
Amersham ECL Prime Western Blotting Detection Reagent	GE	RPN2109	Used for highly expressed proteins
Amersham ECL Select Western Blotting Detection Reagent	GE	RPN2235	
Anti-ATF6	Novus Biologicals	70B1413	Mouse, IgG1, Unconjugated, WB @ 1:1,000
Anti-Calnexin	Cell Signal	C5C9	Rabbit, Polyclonal, Unconjugated, WB @ 1:10,000
Anti-ChAT	Merck Millipore	Ab144P	Goat, Polyclonal, Unconjugated, ICC @1:100
Anti-GAPDH	Merck	CB1001	Mouse, IgG1, Unconjugated, WB @ 1:10,000
Anti-GFAP	Sigma	C9205	Mouse, IgG1, Cy3 conjugated ICC @ 1:1,000
Anti-ISLET1/2	DSHB	39.4D5	Mouse, IgG2b, Unconjugated, ICC @ 1:50
Anti-MAP2	Sigma	M9942	Mouse, IgG1, Unconjugated, ICC @ 1:500
Anti-TARDBP	ProteinTech	12893-1-AP	Rabbit, Polyclonal, Unconjugated, WB @ 1:1,000
Anti-TDP43	Abnova	H00023435	Mouse, IgG1, Unconjugated, ICC @ 1:200
Anti-Tau	Dako	A0024	Rabbit,

			Polyclonal, Unconjugated, ICC @ 1:1000
Ascorbic Acid	Sigma	A4403	
B27	Life Technologies	17504044	
BDNF	Tocris	248-BD	10µg/mL in 0.1% BSA
Bovine Pancreatic DNase	Sigma	DN25	
Bromophenol blue	Sigma	B0126	
Bovine Albumin Fraction V	ThermoFisher Scientific	15260037	7.5% Abbreviated to BSA
Calcein AM	ThermoFisher Scientific	C1430	
CHIR99021	Tocris	4423	3mM in DMSO 10mg in 7.16mL DMSO
CNTF	Tocris	257-NF	10µg/mL in 0.1% BSA
DAPI (4′-6-Diamidino-2-phenylindole dihydrochloride	SIGMA	D9542-10MG	10mg/mL in water
DAPT	Tocris	2634	10mM stock
DMSO	Sigma	D2438-50ML	
Earl's Balanced Salt Solution	Life Technologies	24010043	
Ethanol, Absolute	ThermoFisher Scientific	BP2818500	
FDU	Sigma	F0503	See * below
FGF2	Peprotech	450-33	20µg/mL in 0.1% BSA in DPBS
Fibronectin	Sigma	F2006	1mg/mL in cold PBS Leave in the incubator for 30 minutes before aliquoting to reconstitute
FluroSave™ Reagent	Calbiochem	345789-20ML	
GDNF	Tocris	212-GD	10µg/mL in 0.1% BSA
GelCode™ Blue Safe Protein Stain	ThermoFisher Scientific	24594	Referred to colloquially as Coomassie
Orange G	Sigma	O3756	PCR loading gel
GelRed™ Nucleic Acid Gel Stain	Biotium	41003	

Gibco™ 2-Mercaptoethanol	ThermoFisher Scientific	21985023	
Glucose	Sigma	G8898	
GlutaMax™ Supplement	ThermoFisher Scientific	35050038	
Glycine	Sigma	G8898-1KG	
IGF-1	Peprotech	AF100-11	20µg/mL in 0.1% BSA in DPBS 1mg in 50mL 0.1% BSA in DPBS (aliquot in 5mL and store at -80°C)
Immersion Oil Immersol 518 F fluorescence free, oiler 20 ml	Zeiss	444960-0000-000	
Laminin	Sigma	L2020	
LDN193189	Stratech Scientific	S2618	100µM Aliquots in DMSO 50µL of 10mM LDN in 5mL DMSO
L-Glutamic Acid (GluE)	Sigma	G5888	25mM
Matrigel®	Corning	354230	1:2 in cold Advanced DMEM/F12 Thaw in fridge for 1 hour 10mL MG in 10mL cold Advanced DMEM/F12 1mL aliquots (will make 30 of 1:60)
MEM Non-Essential Amino Acids Solution (100X)	ThermoFisher Scientific	11140035	
Methanol	ThermoFisher Scientific	M/3900/17	
$MgCl_2$	Sigma	M8266	
N-2 Supplement	ThermoFisher Scientific	17502048	
Non-flat dry milk powder	NEB	99995	Made to 5% in TBST for membrane blocking during WB
Normal Goat Serum	Vector Laboratories	S100	
Nuclease-Free Water	Ambion	AM9937	

PBS, pH 7.4	ThermoFisher Scientific	10010023	
PFA (16%)	Agar Scientific	R1026	
Phosphatase Inhibitor	Roche	04906837	
Phosphate Buffered Saline (PBS) Tablets	ThermoFisher Scientific	18912014	
Pierce™ BCA Protein Assay Kit	ThermoScientific	23225	Reaction volume reduced from 2ml to 1ml. Sample volume reduced from 100µL and 50µL.
Polyornithine	Sigma	P4957	
Protease Inhibitor	Roche	05892970001	
PSF (Antibiotic-Antimycotic)	ThermoFisher Scientific	15240062	
Q5® High-Fidelity DNA Polymerase	NEB	M0491S	
Q5® High-Fidelity 2x Master Mix	NEB	M0492S	
Retinoic acid (RA) all-trans	Sigma	R2625	1mM in ethanol Stock at 100mM - dilute 100µL in 9.9mL ethanol for 1mM
RevertAid RT Reverse Transcription Kit	ThermoFisher Scientific	K1691	
RNeasy Mini Kit	QIAGEN	74106	
SAG	Merck	566660	1mM in H_2O 1mg in 1.67mL embryo transfer water
SB431542	Tocris	1614	10mM in DMSO 10mg in 2.4mL DMSO
SDS (Sodium dodecyl sulfate)	SIGMA	L377-100MG	
Sodium azide	Merck	S2002	
Sodium chloride	Thermofisher Scientific	S/3160/60	
Sodium deoxycholate	Sigma	D6750	
TAE Buffer	In-house	NA	
TBS Ultra Pure	VWR Life Sciences	K859-100	
Thiourea	Sigma	T8656-100MG	
TriDye™ 100bp DNA Ladder	NEB	N32715	
Triton™ X100	Sigma	9002931	
Trizma® base	Sigma	T1503-1KG	

Trypan Blue	Sigma	T8154-100ML	
Trypsin (0.25%, phenol red)	Life Technologies	12634028	
Tunicamycin from *Streptomyces* sp.	Sigma	T7765	Made to 1mg/mL in DMSO
TWEEN20	Aldrich	27,434-8	
UltraPure™ 0.5M EDTA, pH 8.0	Life Technologies	155775020	
Urea	Sigma	U0631	
Uridine	Sigma	U3003	See * below
Y-27632 (Rock Inhibitor)	R&D Systems	1254	10mM in DMSO 10mg in 3.125mL DMSO
Wizard® SV Genomic DNA Purification System	Promega	A2361	

Table 2: Table detailing chemicals and kits used for experiments.

*Preparation of Uridine/FDU aliquots. Uridine and FDU were resuspended separately in H_2O at 20mM then mixed at a ratio of 1:1 such that each component had a final concentration of 10mM. Small aliquots were prepared to avoid repeated freeze-thaw.

Software

- ImageJ, Version 1.51h
- RStudio, Version 1.0.136
- GraphPad Prism 7
- SnapGene® Viewer, Version 3.14
- Primer3, http://primer3.ut.ee/
- GitHub, https://github.com/
- CRISPR Design, http://crispr.mit.edu/

Methods

The generation of gene-corrected induced pluripotent stem cells for the study of motor neuron disease

Previously generated induced-pluripotent stem cell (iPSC) lines (**Table 3**) derived from patients with MND harbouring *C9ORF72* hexanucleotide repeat expansion (HRE) (C9-MND) and *TARDBP*^{G298S} mutations have been gene-corrected by Dr B. T. Selvaraj.

Line	Reprogramming Method	Reprogramming Site	Sex	Age at Biopsy	Diagnosis
Con 1	SeV (non-integrating)	Edinburgh	-	-	None (Family history of psychiatric disease)
Con 2	Retroviral (integrating)	Edinburgh	F	40	None
C9-1	SeV (non-integrating)	Cambridge	F	39	ALS
C9-2	Retroviral (integrating)	Cambridge	M	67	ALS
C9-3	Retroviral (integrating)	Cambridge	M	58	ALS
G298S	Retroviral (integrating)	Cambridge	-	-	ALS
M337V	Episomal (non-integrating)	Edinburgh	-	-	ALS

Table 3: Table showing iPSC lines used with reprogramming method, site and selected patient details. Unknown details are denoted '-'.

The *C9ORF72* HRE was excised using two guide RNAs (gRNAs) flanking either side of the G_4C_2 repeats. Double-stranded DNA breaks caused by the gRNAs in the absence of a repair template undergo non-homologous end joining (NHEJ). Guide RNAs can bind to both wildtype (WT) (containing 2-30 hexanucleotide repeats) and mutant (>30 hexanucleotide repeats) alleles.

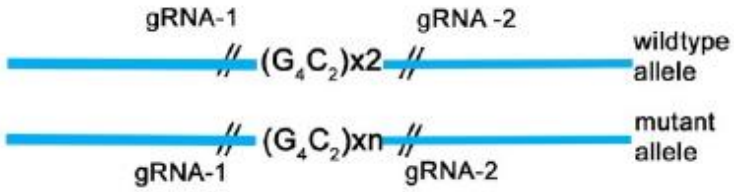


Figure 2: Diagram showing gRNA targeting sites flanking G_4C_2 repeats on WT and mutant alleles. Diagram created by Dr B. T Selvaraj.

If both gRNAs successfully target a single allele, NHEJ may occur excluding the G_4C_2 expansion. This process can result in three outcomes:

1. Cells in which the G_4C_2 repeats are not removed because double-stranded DNA breaks were not introduced by both gRNAs.
2. Cells heterozygous for the G_4C_2 repeats in which one allele only is successfully targeted by both gRNAs with NHEJ.
3. Homozygous deletion of G_4C_2 repeats in which both alleles have been successfully targeted by both gRNAs with NHEJ.

Following transfection of the gRNAs and Cas9 plasmids, clones were picked and screened for the absence of G_4C_2 repeats by repeat-primed PCR. One corrected clone was expanded for each of the three patient-derived lines available. The correction process is described in a manuscript by Dr B. T. Selvaraj (in review). The correction of the *TARDBP*^{G298S} was also achieved using a CRISPR-Cas9 system but is more complex featuring homology-directed repair. This process is described in detail in *Chapter 5* in which a strategy to correct the *TARDBP*^{M337V} mutation is presented. The pairing of mutant and gene corrected lines is shown in **Table 4**.

Disease Line	Gene Corrected Partner
C9-1	C9-Δ1
C9-2	C9-Δ2
C9-3	C9-Δ3
G298S-1	G298S-Δ1, G298S-Δ2
G298S-2	Not corrected
M337V	Not corrected

Table 4: Table showing disease lines paired with their gene corrected controls. Both G298S-1 and G298S-2 were derived from the same patient. G298S-1 has two gene corrected controls: G298S-Δ1 (biallelic correction) and G298S-Δ2 (monoallelic correction).

Generation of spinal motor neurons from iPSCs

Spinal motor neurons (sMNs) were generated from iPSCs using a previously published protocol (Maury et al. 2014). This work used a combinatorial approach to identify chemicals, doses and temporal factors capable of increasing the speed and efficiency of sMN conversions. Generation of sMNs from human iPSCs is shown in **Figure 3**. In brief, following neural induction with dual-SMAD inhibition by SB431542 (20μM) and LDN193189 (0.1μM), and potentiation with the Wnt-agonist CHIR99021 (3μM), rostrocaudal patterning is achieved with smoothened agonist (SAG) (500nM) and retinoic acid (RA) (1μM), respectively. DAPT (10μM) is then added to the sMN progenitors which markedly improves the efficiency of post-mitotic sMN generation in a mechanism that involves inhibition of Notch signalling which plays a role in maintaining pluripotency (Ben-Shushan et al. 2015).

Emily Rhodes Lowry modified the Maury protocol and this modified version is used by the Chandran lab. Rock inhibitor (10μM) has been added to the initial feed to improve iPSC survival following accutase dissociation. 1μM of RA was used throughout the conversion protocol varying from the Lowry protocol in which during the differentiation phase 100nM RA is used. The original analysis by Maury et al. showed no statistical difference between these doses however there is a trend towards an increased proportion of OLIG2 positive cells with the higher dose.

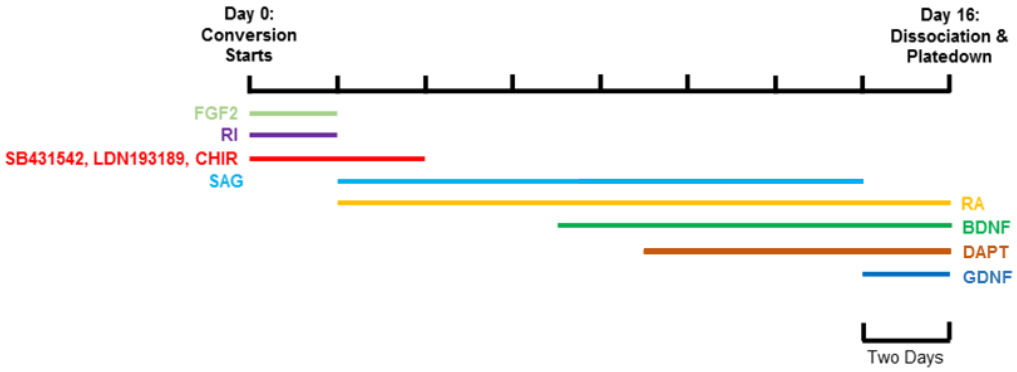


Figure 3: Diagram showing sMN generation. Abbreviations: FGF2 – fibroblast growth factor 2; RI – Rock inhibitor; SAG – smoothened agonist; RA – retinoic acid, BDNF – brain-derived neurotrophic factor; DAPT – N-[N-(3,5-Diflurophenacetyl-L-alanyl)]-(S)-phenylglycine t-butyl ester and GDNF – glial-derived neurotrophic factor. A scale bar is shown in at the bottom right.

At day 16, neurospheres were dissociated and plated down into a monolayer of neurons. Motor neurons were maintained in motor neuron feeding media (MN-FM) for three weeks at which point experiments were performed. The three-week timepoint was chosen for because of the relatively mature electrophysiological properties of MNs at this time point.

Immunocytochemistry (ICC)

Cells were washed once with phosphate buffered saline (PBS) and fixed in 4% paraformaldehyde (PFA) for 20 minutes. 3% Goat serum in 0.2% Triton-X100 PBS was used for blocking and membrane permeabilisation. For ICC involving ChAT, the only antibody not raised in goat, 3% donkey serum was used instead for blocking. Primary antibodies were applied in 0.2% Triton-X100 PBS solution with 3% goat serum and cells were incubated at room temperature for 2 hours. Primary antibodies were removed and cells washed 3 times with PBS before secondary antibodies were added in PBS for 1 hour. Cells were then incubated with 4',6-diamidino-2-phenylindole (DAPI) (1:10,000) for 5 minutes for nuclear staining. Cells were washed again with PBS 3 times before imaging. Concentrations used for each primary antibody are detailed with the relevant antibody in **Table 2**. Secondary antibodies were used at 1:1000. Cells were mounted on glass slides with FluroSave™ reagent.

Gene Ontology Analysis

Enrichment analysis was performed on RNA sequencing (RNAseq) data using the PANTHER Classification System (Mi et al. 2013).

First strand cDNA synthesis

Three RNAseq hits involved in protein homeostasis were selected for validation with qPCR – HSP90B1, HSPB8 and XPO1. cDNA was synthesised using the RevertAid RT Reverse Transcription Kit from the same RNA samples used to generate the RNAseq data. RNA concentration was determined by spectrophotometry (NanoDrop1000) with a TE buffer blank. The manufacturer's protocol for first strand cDNA synthesis was followed. Following addition of 250ng of template RNA to the random hexamers (300ng), samples were vortexed and incubated at 65°C for 5 minutes before the synthesis reaction was performed with 1X Buffer and 1μL of Revert Aid reverse transcriptase. No template control (NTC) and no RT (NRT) control were performed for one sample per reaction. PCR parameters were as follows: 25°C – 10 minutes; 37°C – 40 minutes; 85°C – 5 minutes; 12°C - ∞.

Quantitative PCR

Two sets of intron-spanning primers were designed for each of three target genes (HSP90B1, HSPB8, XPO1) using SnapGene and Primer3 software[A]. Primer annealing temperatures were optimised with gradient qPCR with the DyNAmo kit. The reaction mix was prepared as specified by the manufacturer with 1μL of 1:10 diluted control cDNA loaded per sample for optimisation reactions. For each pair of primers, the optimum annealing temperature and number of melt curve products were assessed by gradient qPCR. One primer from each pair was selected for efficiency analysis using a standard curve (**Table 5**). Three independent experiments were performed on the five lines each with three technical replicates, an NRT control and an NTC. HSP90B1 was run at 58°C with a β-actin housekeeper. HSPB8 and XPO1 were run with a GAPDH housekeeper at 64°C[B]. Housekeeper primers are detailed in **Table 6**. Outliers among the technical replicates were defined as results lying out with the 95% confidence interval (CI). These were excluded from further analysis (25 of 313 values) and the mean Ct value was used to calculate expression fold change in each experiment using the

ΔΔCt method (Livak and Schmittgen 2001). Ct values were normalised to Con 2 and the relevant housekeeper (β-actin for HSP90B1, GAPDH for others). Con 2 data was not used in statistical analysis as by definition there was no variance between experimental replicates. The Kruskal-Wallis test was performed using ΔΔCt values with Dunn's multiple comparisons correction in cases where the null hypothesis was refuted. Pairwise testing was not performed as removal of outliers in some cases lead to missing values.

Gene	Sequence
HSP90B1	FW:GCGAGACTCTTCAGCAACAT RV: TCCCCTTCAGCAGTAAAGTGA
HSPB8	FW: AAGCCAGAGGAGTTGATGGT RV: AGGAAGCTGGATTTTCTTTGTGA
XPO1	FW: AGCAATGCATGAAGAGGACG RV: CTTTGCCTCTTTTCTGTTCACA

Table 5: Sequences of experimental group primers for RNAseq validation by qPCR.

Gene	Sequence
B-Actin	FW: GTTACAGGAAGTCCCTTGCCATCC RV: CACCTCCCCTGTGTGGACTTGGG
GAPDH	FW: GAGTCCACTGGCGTCTTCAC RV: ATGACGAACATGGGGGCAT

Table 6: Sequences of housekeeper primers.

RNA Extraction

RNA was extracted with the RNeasy Mini Kit following the manufacturer's protocol with no modifications. RNA was stored at -80°C.

X-box Binding Protein 1 Splicing Polymerase Chain Reaction Assay

Three-week-old sMN cultures were treated with tunicamycin 5μg/mL in normal MN-FM media or a DMSO by volume (0.1%) control[c]. RNA was extracted 24 hours after treatment and quantified by spectrophotometry (NanoDrop1000). cDNA was synthesised as described in *Quantitative PCR* with 100-250ng of total RNA per reaction depending on the RNA yield. To amplify both the full length and spliced ATF6 transcripts, previously published primers were used (see **Table 7**) (Samali et al. 2010).

Primer	Sequence
XBP1 Forward	TTACGAGAGAAAACTCATGGCC
XBP1 Reverse	GGGTCCAAGTTGTCCAGAATGC

Table 7: XBP1 primer set used by Samali et al.

30µL PCR reactions with 4 units Q5® High-Fidelity DNA Polymerase, x1 Q5® High-Fidelity 2x Master Mix, dNTPs (200nM) nuclease-free water and 5µL cDNA. 5µL PCR product mixed with Gel Orange and run on 1% agarose gel (1g agarose in 100mL TAE buffer with 5µL Gel Red) with 50bp and 100bp reference ladders for 1 hour at 100V.

Protein Extraction and Quantification

Soluble and insoluble protein fractions were obtained following lysis with radioimmunoprecipitation assay (RIPA) buffer. RIPA buffer was prepared fresh with 50mM Tris-HCl (pH 7.4), 150mM NaCl, 1% Triton-X100, 0.5% sodium deoxycholate, 0.1% SDS, 2mM EDTA, 1x protease inhibitor and 1x phosphatase inhibitor in water. Following cell lysis, samples were centrifuged at 13,000rcf for 30 minutes at 4°C. The supernatant was collected as the soluble fraction. 200µL of PBS was added to the pellet followed by centrifugation again at 13,000rcf for 10 minutes at 4°C. The PBS wash was removed and 75µL of urea buffer (**Table 8**) was added and the sample incubated on ice for 10 minutes before sonication on ice for 5 seconds. Samples were centrifuged a final time at 13,000rcf for 30 minutes at 4°C and the supernatant was taken as the insoluble fraction. Proteins were stored at -20°C short term and -80°C long term.

Component	Final Concentration
Urea	7M
Thiourea	2M
CHAPS	4%
Trizma® Base	30mM

Table 8: Urea buffer composition.

Protein concentration was estimated using a bicinchoninic acid (BCA) assay carried out as described by the manufacturer with the following alterations. A 1mL reaction mix of Reagents A and B was used (instead of 2mL) and correspondingly 50µL of protein added (instead of 100µL). Absorbance at 562nm was determined using an Eppendorf BioPhotometer with a RIPA blank. A standard curve was produced using the 'Test Tube' protocol. Three technical replicates were performed and the mean calculated. Mean absorbance was plotted against the concentration **Figure 4**.

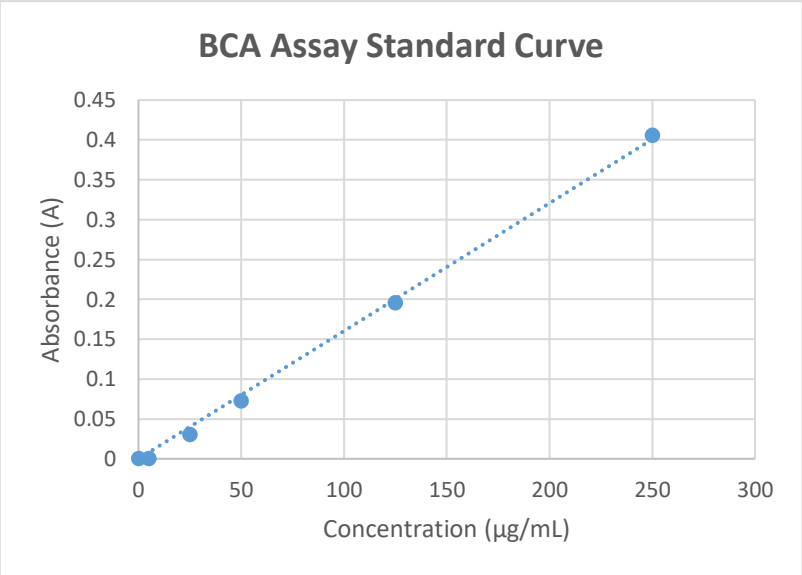


Figure 4: BCA standard curve produced using the Test Tube protocol. Each point represents the mean of three technical replicates. A linear model with intercept = 0 was fitted of the form y=0.0016x (R^2=0.9979).

Protein concentrations were subsequently determined as:

$$Concentration\ (\mu g/mL) = \frac{1}{m}(A.DF)$$

m is the gradient of the standard curve, A is the absorbance and DF is the dilution factor of the sample used in the BCA assay.

Western Blots

Following determination of protein concentration by BCA, Laemmli buffer (X5) (**Table 9**) was added to 10 – 30µg protein samples with volumes equalised with RIPA buffer[D].

Substance	Quantity
1M pH 6.8	3.5mL
Glycerol (100%)	5mL
SDS	1g
1% Bromophenol blue	0.125mL
β-ME	1mL
H_2O	0.875mL

Table 9: Laemmli buffer (X5) composition.

Samples and protein ladders were loaded into 4 – 20% polyacrylamide gels and run at 150V until the bromophenol blue running front had reached the bottom of the gel. Protein was then transferred to a methanol-activated PVDF membrane (Immobilon®-P, Merck, IPVH00010) and transferred using a MiniGel Tank (Life Technologies) at 20V for 1.5 hours. Following transfer, membranes were blocked for 1 hour with 5% milk in TBS-T. Primary antibodies against targets ATF6 (1:1000), PERK (1:1,000), calnexin (1:10,000) or GAPDH (1:10,000) were added in 5% milk in TSB-T and left overnight at 4°C. Following three TBS-T washes HRP-conjugated secondary antibodies were added in 5% milk in TBS-T for 1 hour. Three more TBS-T washes were performed before membranes were exposed to detection Amersham ECL detection solutions for 5 minutes. Amersham ECL Prime was used for highly expressed proteins while Select was used for less abundant proteins. Membranes were exposed on radiographic film (Amersham hyperfilm MP, GE, 28906842). Band intensity was quantified by densitometry (Image J, Gel Analysis Plugin).

For Western blots using samples from the insoluble protein fraction, BCA could not be performed accurately because of the composition of the urea buffer. In these cases, equal volumes of protein were subject to electrophoresis followed by Coomassie staining and densitometric band analysis to determine ratios that would allow equal loading of samples. Following Western blotting, the membrane was again stained with Coomassie to confirm equivalent loading.

Tunicamycin Survival Analysis

Survival assays determining the effect of tunicamycin on neurons derived from 8 iPSC lines (Con 1, Con 2, C9-1, C9-Δ1, C9-2, C9-Δ2, C9-3, C9-Δ3) were performed at three weeks post platedown. Calcein AM was used as a fluorescent reporter of cell viability. Esterases produce an acetoxymethyl (AM) derivative of calcein (calcein AM), which is trapped in cells with intact membranes hence the intensity of calcein AM is a measure of membrane integrity[E]. Three half media changes were used to dilute existing MN-FM media with a basal media (neurobasal (X1), B27 (X50), PSF (X100) and GlutaMax(X100)) to exacerbate any phenotype. Calcein was added at 4μM and 0h imaging performed at 488nm at x10 magnification with a Zeiss microscope with incubation at 37°C. Image parameters were kept constant over each 48h experiment but varied between replicates. Following imaging, cells were subjected to tunicamycin at 1μg/mL, 3μg/mL, 5μg/mL, 10μg/mL or a DMSO (matched by volume) control in basal media. Repeat imaging was undertaken at 24h and 48h. Prior to each subsequent imaging period was a half media change with basal media supplemented with calcein AM. Note that drugs were not re-added during these media changes thus the concentration of drug in the media reduced by approximately 50% every 24 hours in addition to the decay resulting from the half-life of the drug. There was insufficient literature to establish this in this culture system.

Stacks were prepared for each well imaged containing the baseline, 24h and 48h images. The resulting images were to be analysed by manual cell tracking through stacks. Concerns regarding the precision of these measurements lead to attempts to develop alternative analysis protocols one of which is described and is based on the determination of cellular mean integrated density.

Manual Cell Tracking

If cells could be accurately and precisely classified as alive or dead cell tracking would prove an excellent way to analyse survival. Biological systems are rarely so simple though and different markers of life and death have their advantages and disadvantages. Calcein was used because of prior experience in the lab of this technique. As with many other survival

techniques, the intensity of calcein AM fluorescence is a function of membrane integrity. Compromise of cell membranes, as occurs during cell death, results in leakage of the dye into the extracellular space reducing the intensity of the cell. The following observations have been made in sMN cultures treated sequentially with calcein AM over 48 hours as described above:

- There is a marked increase in the background intensity. Background fluorescence is present at very low levels initially and this is said to result from the activity of a low level of extracellular esterases, it is hypothesised that this increase in intensity reflects (activated) calcein AM released from compromised cells (**Figure 5**).

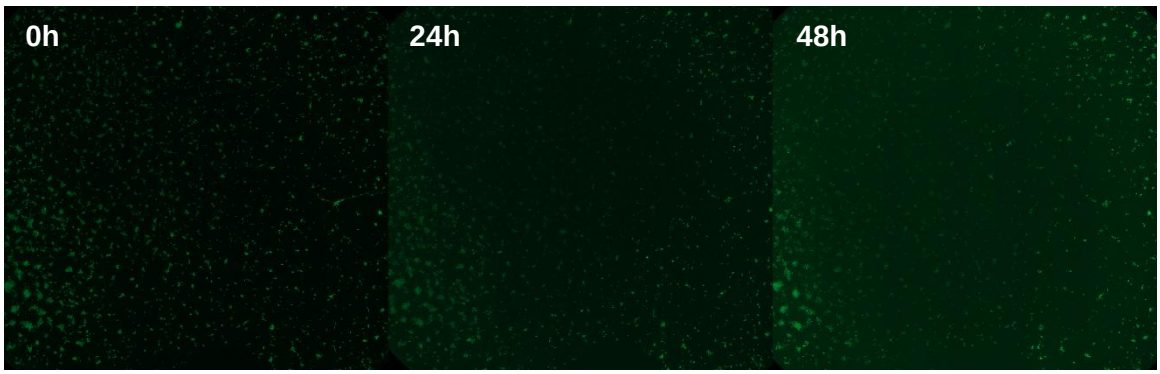


Figure 5: Increasing background over sequential imaging caused by leakage of calcein AM into the extracellular space or activation of added calcein by extracellular esterases. The time of imaging is shown in the top left-hand corner. Heterogeneity in fluorescence is also observed with a brighter signal towards the edges of the well.

- A subpopulation of cells increases in intensity over the 48h window. This population of cells is the only one easily visualised by eye (see **Figures 6** and **7**).
- The signal of a cell is rarely completely abolished. If cells do not increase in intensity as described above, they may remain at a stable intensity or decrease.
- Heterogeneity in fluorescence is observed across plates with some regions of the well (which correspond to those in which pipette tips have been introduced) less likely to house brightly fluorescing cells (see **Figure 5**).

These phenomena and others have important implications for the use of calcein in survival analysis. Two additional features that complicate survival analysis in our cultures are clustering and cell movement. The problem of clustering in MN cultures is well described and approaches to resolve the problem including dissociation of mature MNs remain lacking.

Advantages can be expected with the invention of new plate technologies but at present, we have reduced the levels of clustering with very low-density platedowns. This technique was developed for neurite outgrowth analysis and is suboptimal for survival analysis because survival in these low-density cultures does appear to be reduced. At a platedown density of 20,000 cells per well of a 96 well plate, we observe both clusters and single cells. In the cultures, cells that cannot be identified as discrete units using the calcein AM signal alone are ignored for manual cell tracking analysis. The use of live nuclear dyes has been considered and tested, however, these cultures are laden with a large amount of DAPI-positive debris. These bodies are small, circular and extremely brightly staining. They complicate identification of cells with nuclear markers for which they are often positive and this, in concert with cellular clustering, makes identification of single cells extremely challenging. I have made attempts to remove this debris using DNAase and washes but have not been successful. Additionally, designing computational algorithms that identify cell nuclei becomes extremely difficult. In addition to these issues, MNs move a small degree during culture and we have observed isolated cells moving to join clusters with longitudinal imaging. These factors make designing algorithms to track cells over the 48 hours very complex so manual cell tracking was used at the expense of the sample size.

Images were divided into forty-nine sampling regions using a seven by seven grid with a perimeter. The perimeter functioned to prevent inclusion of cells which might migrate beyond the image capture zone. One cell was sampled from each region containing cells. In the majority of wells, all regions contained cells and thus 49 cells were sampled per experiment. The number of experimental replicates completed varies among cells lines as some gene-edited controls were developed over the course of the MSc. The sample size for each line is shown in **Table 10**. Within the sampling region, discrete neuron-like cells were selected (astrocytic-type cells were very rarely observed but have a distinct morphology). Cells close to clusters were avoided to reduce the change of them being lost to follow up. Cells that were thought to join a cluster were treated as dead in keeping with the intention to treat analysis protocol. Cells were judged alive if the level of fluorescence was maintained or increased relative to the previous imaging session.

Line	Number of images	Number of cells
Con 1	17	817
Con 2	17	833
C9-1	17	833
C9-Δ1	17	833
C9-2	17	832
C9-Δ2	15	699
C9-3	12	588
C9-Δ3	5	385

Table 10: Table showing the number of images and cells for each line subject to manual survival analysis.

Command line and ImageJ scripts were written to organise images into stacks, blind images by removing slice labels and applying pseudo-random titles[F] and apply grids. Variable cell size precluded effective use of a rolling-ball background subtraction algorithm. Images in which there was thought to be a quality control issue were excluded before blinding was removed. Quality control issues were infrequent but included:

- Exposure time too low
- Pipette damage
- Lifting of cells from LMF

Mean Cellular Integrated Density

The fluorescence of cellular units (discrete and clumped neurons) over the course of the experiment provides an indication as to the integrity of their membrane and thus their viability. The addition of calcein before each imaging cycle means that cells should be able to take up new calcein and convert this to calcein AM increasing, or at least maintaining their fluorescence. Membrane compromise would allow calcein AM to escape. For this reason, the intensity of a cell provides an indication of viability. To overcome the limitation of counting calcein AM in the extracellular space, algorithms to measure cellular units were employed. To ensure reproducibility and throughput the 17 auto-thresholding algorithms on ImageJ were trialled on a minimum of fifteen images (0h, 24h and 48h) using a custom-made macro. Although all auto-thresholding algorithms were applied, some are clearly inappropriate because of their mathematical foundations. Algorithms were assessed by judging the cellular masks formed (Cellular overlays in **Figures 6 and 7**) at 0, 24 and 48 hours (below). The best

performing algorithm, Moments, is shown below applied to representative images of high and low survival. Below each image are tables detailing the mean cellular grey value, the mean cellular integrated density and survival as determined by manual cell tracking.

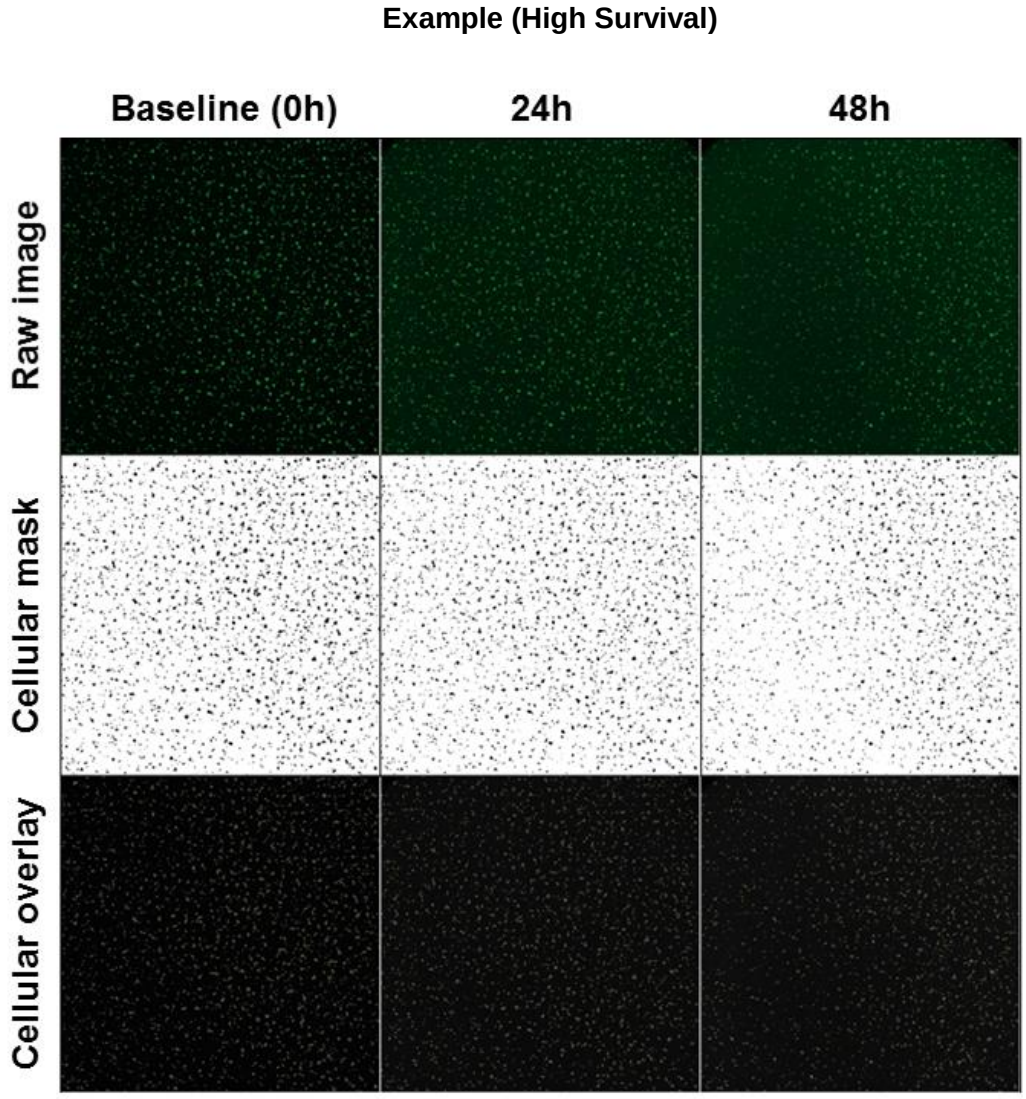


Figure 6: Representative images of a well with high survival over the 48h imaging period. The first row shows raw images. The second row shows region of interest (ROI) masks generated by the Moments algorithm. The last row shows the outline of the ROIs created by Moments applied to the raw image and requires visualisation at high magnification.

	Baseline (0h)	**24h**	**48h**
Mean Grey Value (Cellular)	100.00%	187.64%	208.66%
Mean Normalised Integrated Density (Cellular)	100.00%	75.14%	45.83%
Manual Cell Tracking Percentage Survival	100%	71.43%	28.57%

Table 11: Measures of viability over the 48h imaging period corresponding to Figure 6.

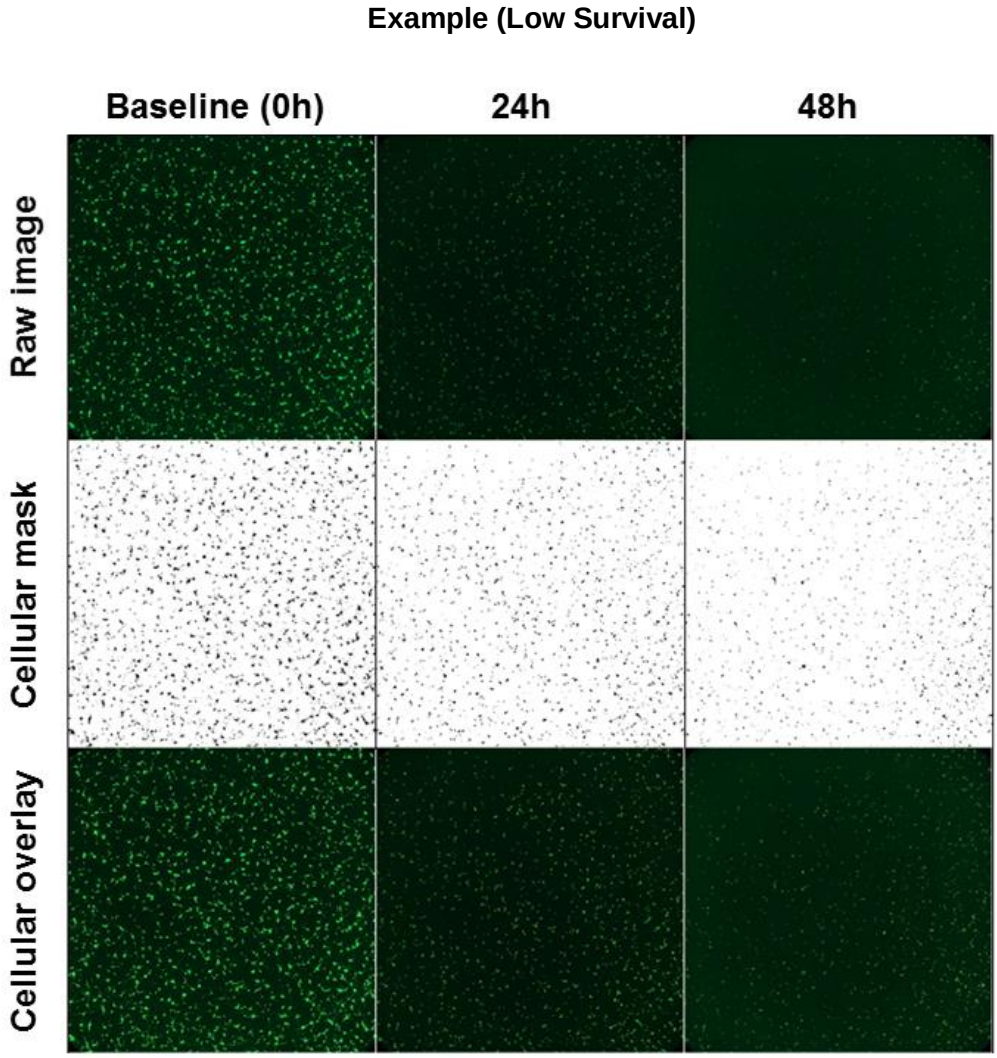


Figure 7: Representative images of a well with low survival over the 48h imaging period. The first row shows raw images. The second row shows region of interest (ROI) masks generated by the Moments algorithm. The last row shows the outline of the ROIs created by Moments applied to the raw image and requires visualisation at high magnification.

	Baseline (0h)	**24h**	**48h**
Mean Grey Value (Cellular)	100.00%	62.98%	81.75%
Mean Normalised Cellular Integrated Density	100.00%	63.47%	15.07%
Manual Cell Tracking Percentage Survival	100.00%	93.88%	6.12%

Table 12: Measures of viability over the 48h imaging period corresponding to Figure 7.

Integrated density is calculated by multiplying the mean grey value of a region of interest by its area. It can be seen from the tables above that mean grey value alone performs very poorly because it takes no account for the decreasing area of cells that follows death. Mean cellular integrated density proves a better surrogate for cellular viability. Given the subjectivity and variability of manual cell tracking described above it is argued that analysis by both manual cell tracking and mean normalised cellular integrated density (from this point on referred to as mnID) prove complementary. Manual cell tracking has the advantage of providing biologically meaningful units but suffers from subjectivity and is slow. mnID can be applied to a much larger sample size and takes into account the survival of clusters of neurons, not just individual cells. It is, however, less interpretable particularly considering it can increase over the course of 48h given that more calcein is added prior to each imaging session. Following calculation of the normalised mnID for all samples, every mask was screened. 33 of 422 masks were excluded (6.85%) because of poor fit.

Statistical Approach to Survival Data

The statistical analysis of this dataset is complex. Details are provided with results, but the general strategy has been to produce mixed linear regression models which allow important factors and covariates to be explored before construction of models with and without a dependent variable of interest. By constructing models with and without the dependent variable under investigation the effect of that variable on the performance of the model can be assessed. The best method of comparison is debated among statisticians – here the likelihood ratios test has been applied. It is acknowledged that the use of mixed linear regression models does not take into account the right censoring of this dataset. It does, however, allow multiple variables to be incorporated including random effects which is advantageous given the nature of the dataset and avoids problems associated with pseudoreplication.

TARDBP Genotype Validation PCR

DNA was extracted using the Wizard® SV Genomic DNA Purification System using centrifugation. Primers generating a 452bp amplicon within exon 6 spanning the mutation site were used (**Table 13**).

Primer	Sequence
TARDBP Forward	TTGCGCAGTCTCTTTGTGGA
TARDBP Reverse	GGATGCTGATCCCCAACCAA

Table 13: Primers used for the amplification of a 452bp region of TARDBP exon 6.

The PCR reaction mix consisted of 0.5U Q5® High-Fidelity DNA Polymerase, Q5® Reaction Buffer, 0.2mM dNTP, 1.25µL 10µM TARDBP Forward Primer, 1.25µL 10µM TARDBP Reverse Primer and 100ng cDNA made to 25µL with nuclease-free water.

Successful amplification was determined by gel electrophoresis (1% agarose in TAE Buffer) to identify the 452bp amplicon. Samples with successful amplification were sent for Sanger Sequencing (Source Biosciences).

Chapter 2: Materials and Methods Notes

A – Intron-spanning primers were used to reduce the chance of contamination through genomic DNA amplification.
B – The choice of the housekeeper was determined by the annealing temperature of the primer allowing target and housekeeping genes to be run on the same plate. These primers have previously been optimised by the Chandran lab.
C – Tunicamycin is constituted in DMSO.
D – The exact mass of protein used was equal to the minimum yield or 30µg (whichever was lower).
E – The expression pattern of calcein AM in neurons has been reported previously with relatively high emission from the soma than processes (Hancock et al. 2015).
F – The ImageJ macro used for obfuscation was written by Tiago Ferreira.

Chapter 3: Results

Culture Characterisation

When originally characterised the Maury sMN differentiation protocol produced approximately 70% HB9^{+}/ISLET1^{+} cells. In our hands over 95% of cells are MAP2^{+} and less than 5% of cells stain positive for the astrocyte marker GFAP at three weeks. At the three-week timepoint, ChAT expression is developing and becomes more abundant with time in culture (**Figure 8**).

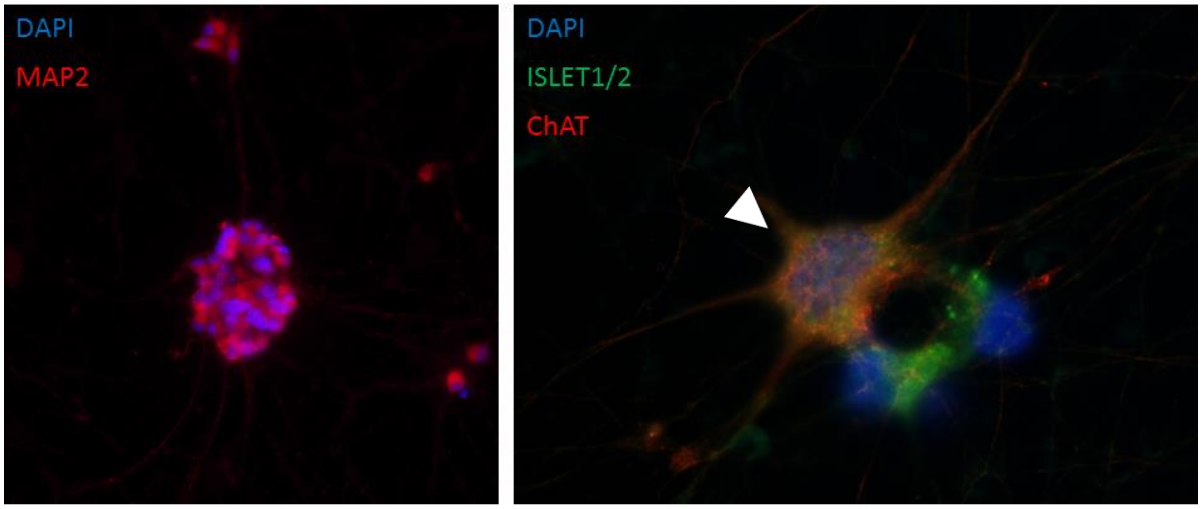


Figure 8: Representative images of cells positive for the pan-neuronal marker MAP2 and sMN markers ISLET1/2 and ChAT. Three-week-old sMN cultures were stained for the pan-neuronal marker MAP2 and sMN markers ISLET1/2 and ChAT along with the astrocyte marker GFAP. *Left:* > 95% of nucleic identified by DAPI (blue) colocalise with MAP2 (red), X10 magnification. *Right:* Positive staining for sMN markers ISLET1/2 (green) in all cells and ChAT (white arrow).

GO Analysis of RNAseq data

Lists of differentially expressed genes established by RNA sequencing (RNAseq) of three-week control, *C9ORF72* mutant and gene corrected sMN cultures were provided by Professor Hardingham. Gene ontology (GO) analysis was performed and of the 85 upregulated genes, four IDs were unmapped and correspond to non-protein coding genes (**Supplemental Table 1**). Of the 80 downregulated genes, 14 were non-protein coding (**Supplemental Table 1**). A complete list of RNAseq hits is provided by Dr B. T. Selvaraj *et al.*, (in review).

Data from PANTHER protein class GO analysis is shown below in **Figure 9**.

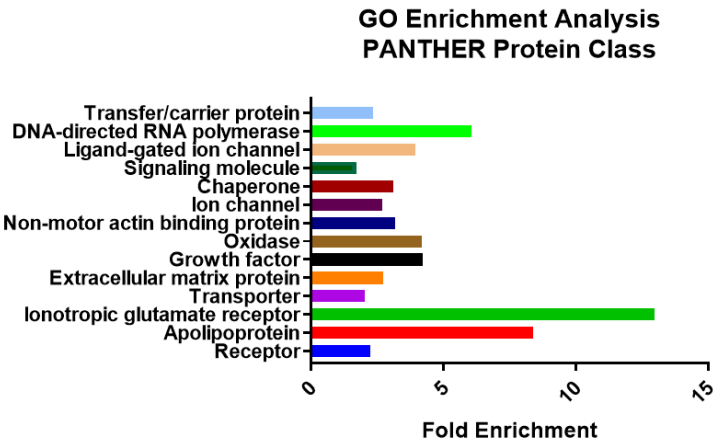


Figure 9: The fold enrichment of PANTHER protein class terms in *C9ORF72* mutant lines following GO analysis of RNAseq data. Note that data on up- and downregulated transcripts are presented together.

Transcripts coding for a wide range of protein classes were identified as being significantly dysregulated in *C9ORF72* mutant lines. As TDP-43 pathology is a common feature in most forms of MND and chaperone proteins were identified following GO analysis we decided to validate several candidate proteins involved in the unfolded protein response (UPR) by qPCR.

Validation of RNAseq Hits

Three transcripts were selected for validation based on the GO results – HSP90B1, HSPB8 and XPO1. RNA samples from three-week-old sMN cultures were subject to qPCR with appropriate housekeepers (see *Methods*).

Of the 313 Ct values obtained 25 outliers were removed (7.99%) as described in *Methods*. The expression fold changes for each investigated transcript are shown below (**Figures 10, 11** and **12**) and ΔΔCt values can be found in **Supplemental Table 2**. RNAseq hits were considered validated if the ΔΔCt varied significantly from the control ($\alpha = 0.05$) in *C9ORF72* mutant lines but not in gene-corrected control. Analysis has been undertaken on data both with and without outlier exclusion. Analysis of data without outlier exclusion is provided in **Supplemental Figures 1, 2** and **3** and **Supplemental Table 3**.

HSP90B1

HSP90B1, a chaperone involved in ER-associated degradation (ERAD) (Di et al. 2016), was upregulated on RNAseq. The RNAseq data was not validated by the qPCR (**Figure 10**) but there was a significant increase in HSP90B1 transcript expression in C9-3 (p = 0.0369) (**Table 14**).

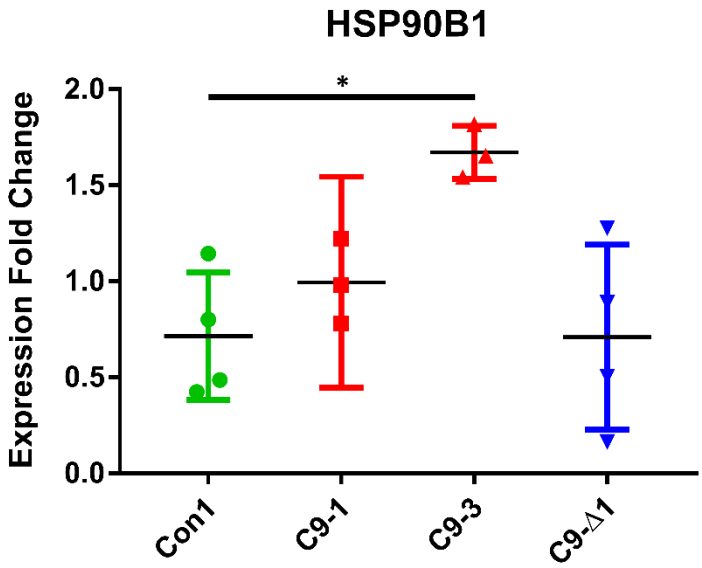


Figure 10: *C9ORF72* mutants show heterogeneity in HSP90B1 transcript expression. Expression fold change of HSP90B1 transcript in control (green), *C9ORF72* mutant (red) and gene corrected (blue). Outliers have been removed. Each point represents data from an RNA sample from an independent sMN conversion. Black bars mark the mean. Error bars show the 95% CI. Kruskal-Wallis refuted the null hypothesis with (p = 0.0406). Dunn's multiple comparisons test was performed and results are shown in **Table 14**. A significant increase in HSP90B1 expression was observed in C9-3 but not C9-1.

Comparison	P value
Con1 vs. C9-1	> 0.9999
Con1 vs. C9-3	0.0369
Con1 vs. C9-Δ1	> 0.9999

Table 14: Results of Dunn's multiple comparisons test for HSP90B1.

HSPB8

HSPB8, a chaperone implicated in the clearance of TDP-43 aggregations (Crippa et al. 2016), was also identified as upregulated on RNAseq. This result was confirmed by qPCR (**Figure 11**, **Table 15**). In addition to significant upregulation in C9-1 and C9-3, there was also rescue in the gene-corrected line C9-Δ1.

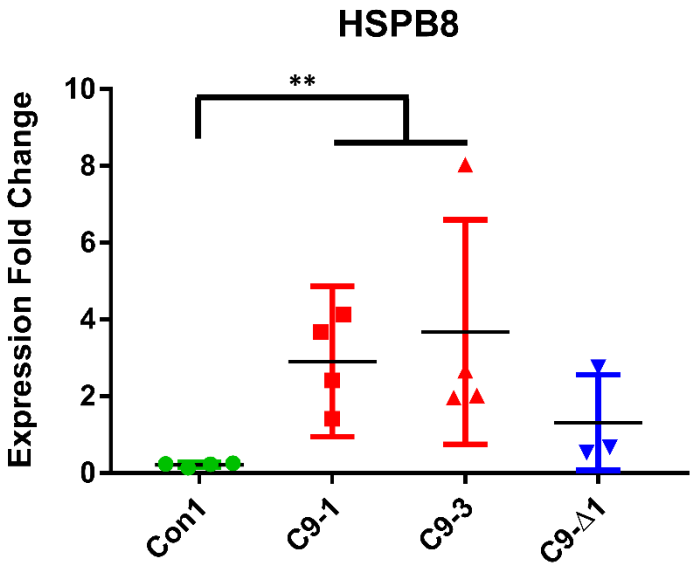


Figure 11: HSPB8 is expressed at significantly higher levels in *C9ORF72* mutant lines but at lower levels in gene-corrected lines. Plots showing expression fold change of HSPB8 transcript in control (green), *C9ORF72* mutant (red) and gene corrected (blue) sMN RNA. Each point represents the mean of three technical replicates from RNA samples from independent sMN conversions. Black bars mark the mean. Error bars show the 95% CI. Kruskal-Wallis refuted the null hypothesis with p = 0.0072. Dunn's multiple comparisons test was performed and results are shown in **Table 15**.

Comparison	P value
Con1 vs. C9-1	0.0216
Con1 vs. C9-3	0.0273
Con1 vs. C9-Δ1	0.3911

Table 15: Results of Dunn's multiple comparisons test for HSPB8.

XPO1

XPO1 functions in nucleocytoplasmic shuttling (Fornerod et al. 1997). Although XPO1 narrowly failed adjustment for multiple comparisons in the RNAseq analysis it was investigated because of recent evidence in support of nucleocytoplasmic transport deficits in iPSC and post-mortem brain tissue from individuals with C9-MND (Zhang et al. 2015). No significant variation in XPO1 transcript expression was observed between control, *C9ORF72* mutant and gene corrected lines (p = 0.3340) (**Figure 12**).

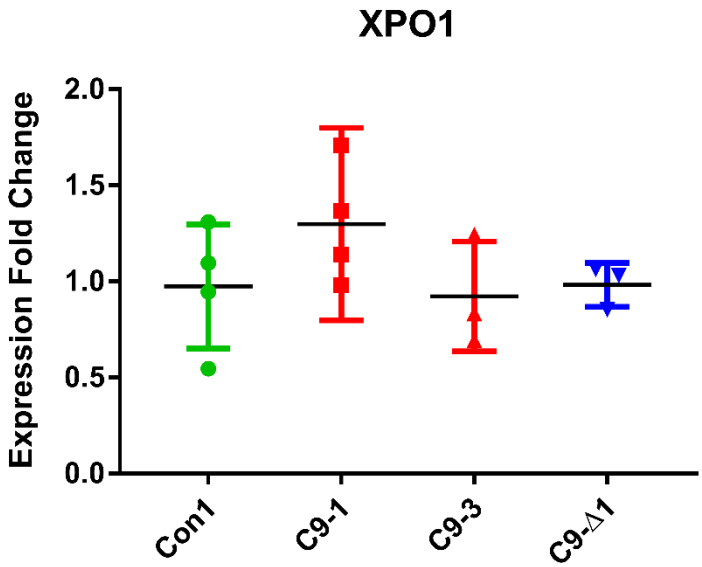


Figure 12: XPO1 is not differentially expressed in three-week sMN cultures with the *C9ORF72* HRE. Expression fold change of XPO1 transcript in control (green), *C9ORF72* mutant (red) and gene corrected (blue) sMN RNA. Each point represents the mean of three technical replicates from RNA samples from independent sMN conversions. Black bars mark the mean. Error bars show the 95% CI. Kruskal-Wallis failed to refute the null hypothesis with p = 0.3340.

Endoplasmic Reticulum Stress Sensors

Following validation of HSPB8 upregulation and heterogeneity among *C9ORF72* mutant lines in HSP90B1 expression further investigation into the possibility of ER stress in *C9ORF72* mutants was undertaken. The activation of three UPR stress sensors (IRE1, PERK and ATF6) was assessed under basal conditions and following induction of ER stress with tunicamycin.

XBP1

A 26 base pair long nucleotide sequence of the XBP1 transcript is removed by unconventional splicing by activated IRE1 (Yoshida et al. 2001). The spliced XBP1 transcript codes for a stable transcription factor which can bind to ER stress response elements (ERSE) with potentially pro-survival or apoptotic consequences. Inability to mount an XBP1 response appears to be neuroprotective in cellular and rodent models of MND (Hetz et al. 2009).

The ability to detect XBP1 splicing via the PCR technique described was assessed on two lines: HB9-GFP and G298S (**Figure 13**).

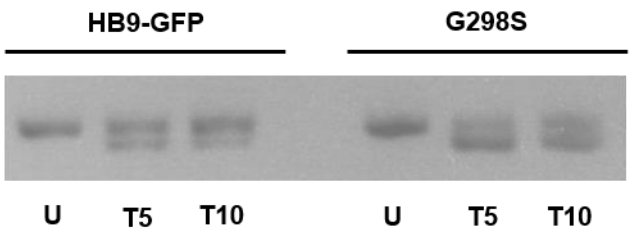


Figure 13: XBP1 splicing can be detected in sMN cultures treated with 5µg/mL or 10µg/mL tunicamycin for 24h. Data from two cell lines are shown: HB9-GFP (left) and G298S (right). U = untreated cells (full media change only), T5 = tunicamycin at 5µg/mL for 24 hours prior to RNA extraction, T10 = tunicamycin at 10µg/mL for 24 hours prior to RNA extraction. The upper band is at 283bp. The spliced 257bp product is observed below in cells treated with tunicamycin. In G298S there is no evidence of spliced XBP1 at rest which is interesting considering other markers the UPR are evident in these cells.

Following confirmation spliced XBP1 transcript could be detected, *C9ORF72* mutant and control sMN cultures under normal conditions (no treatment) were assessed for XBP1 splicing which would indicate ER stress (**Figure 14**).

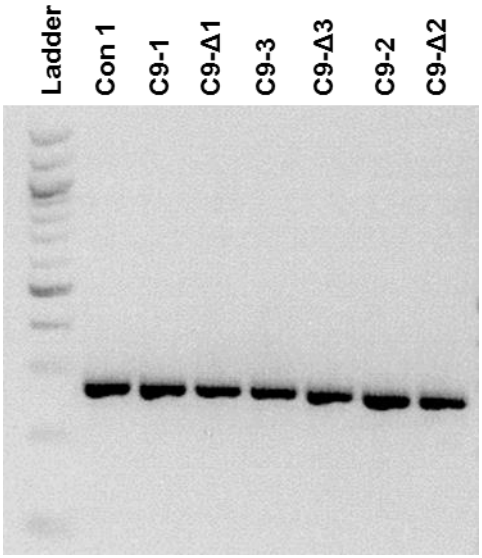


Figure 14: Unspliced XBP1 transcript is observed in control, *C9ORF72* mutant and gene-corrected controls but the active spliced fraction is not. Following the 100bp ladder cDNA samples from one control line and three *C9ORF72* mutants with gene-corrected controls have been subject to XBP1 PCR and run on a 2.5% agarose gel.

Figure 14 shows that in normal culture conditions *C9ORF72* mutant sMNs do not show evidence of XBP1 splicing suggesting *C9ORF72* mutants are not subject to a constitutive UPR.

The response of *C9ORF72* mutant sMNs to proteostatic insult via tunicamycin administration was assessed (**Figure 15**).

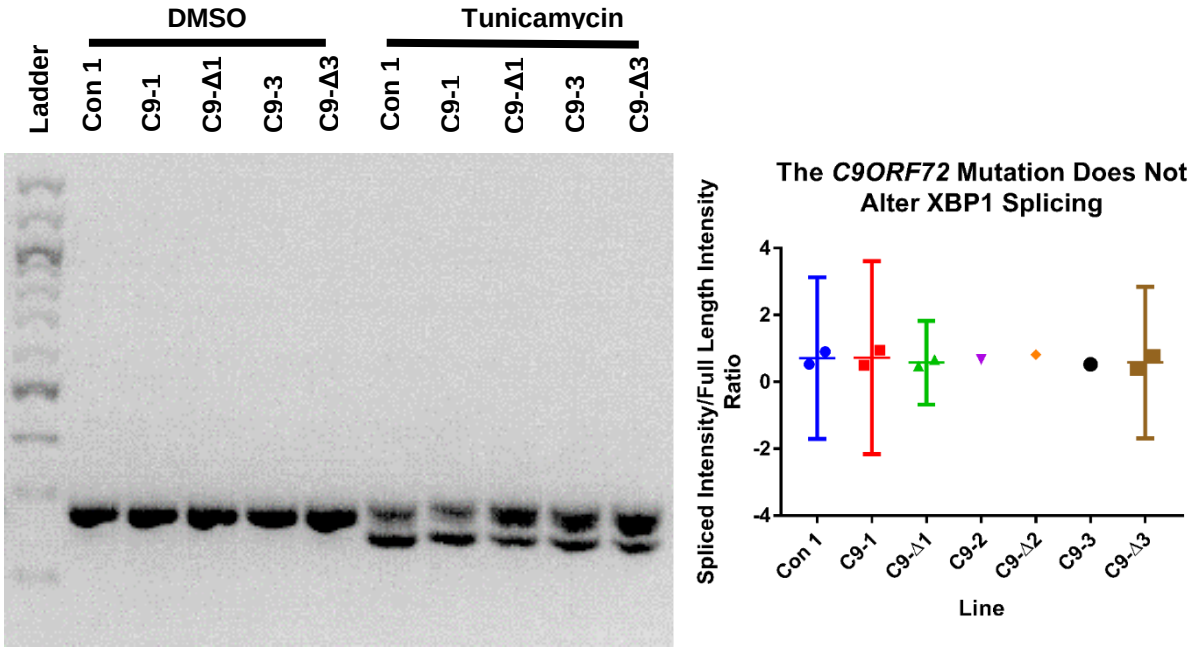


Figure 15: XBP1 splicing occurs in control, *C9ORF72* mutant and gene corrected sMN cultures at three weeks treated with 5µg/mL tunicamycin but not a DMSO by volume control. ***Left:*** Gel electrophoresis showing unspliced XBP1 transcript following DMSO treatment and both unspliced and spliced XBP1 transcript following treatment with tunicamycin in all lines. The far left lane houses a 100bp ladder. ***Right:*** Plot showing the ratio of spliced to unspliced XBP was established by densitometry and the ratio of the spliced band to the unspliced band calculated. This ratio allows normalisation for random variability in PCR product loading. Data from one control line, three *C9ORF72* mutants and their gene-corrected partners are shown. Each data point represents an independent sMN conversion. Error bars show the 95% CI. Further replication is required. Statistical analysis with such low n values would be inappropriate at this point, however, this data when expanded upon would likely be suitable for ANOVA or Kruskal-Wallis depending on whether or not data is normally distributed.

This data suggests that the XBP1 splicing response of *C9ORF72* mutant sMNs does not vary from that of controls.

PERK

Under conditions of ER stress, GRP78 dissociates from PERK leading to its oligomerization and autophosphorylation. Levels of total PERK have been demonstrated to be upregulated in the spinal cords of patients with MND at post-mortem (Atkin et al. 2008). Under basal conditions, PERK expression did not differ between control, *C9ORF72* mutant, *TARDBP*^{G298S} and gene corrected lines (**Figure 16**).

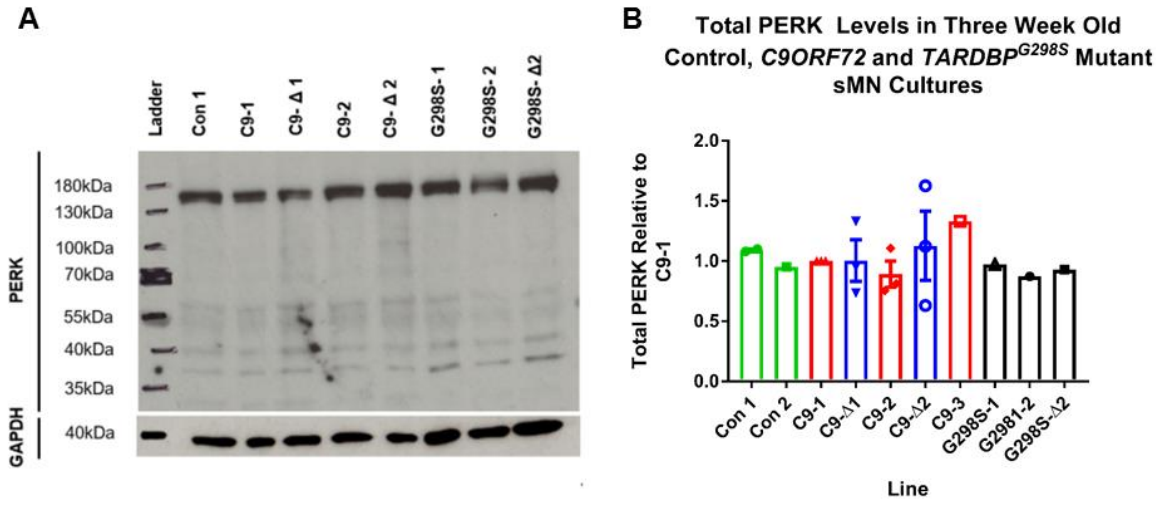


Figure 16: Total PERK levels do not vary between control, *C9ORF72* mutant and *TARDBP*^{G298S} mutant cell lines under normal conditions. A) Representative Western blot showing equal levels of total PERK (~140kDa) across control, *C9ORF72* and *TARDBP*^{G298S} lines. A GAPDH housekeeper is shown below. **B)** Plot showing band intensity from Western blotting determined by densitometry. Total PERK level was normalised to housekeepers and this ratio was then normalised to a single line (C9-1) and expressed as a fold change. Each data point represents an independent sMN conversion and subsequent independent experiments. N equals 1 for all lines except: Con 1 = 2, C9-1 = 3, C9-Δ1 = 3, C9-2 = 3 and C9-Δ3 = 3. Error bars show the 95% CI.

To confirm PERK upregulation upon ER stress could be detected, cells were treated with tunicamycin (5µg/mL) for 24h. **Figure 17** demonstrates a representative blot of PERK expression in control (DMSO) and treated (tunicamycin) cells across all lines.

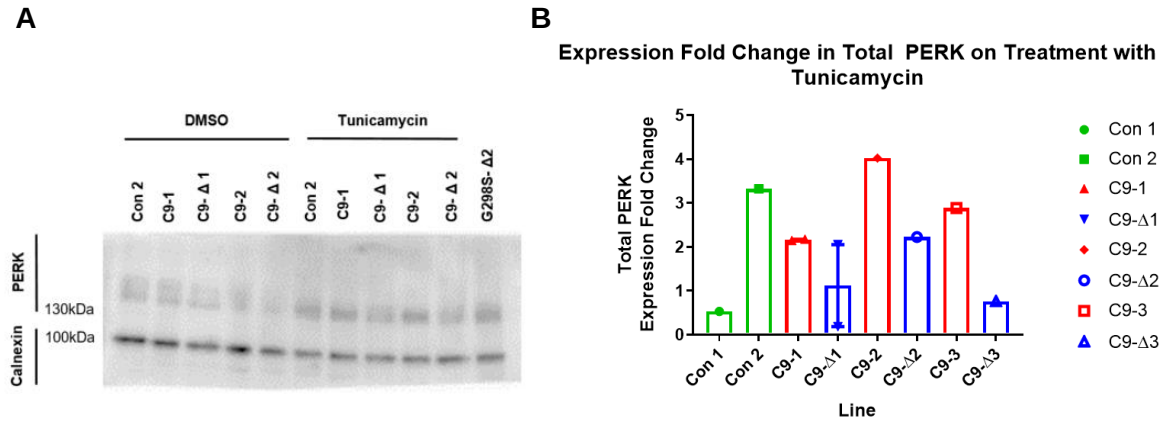


Figure 17: Upregulation of PERK on treatment with tunicamycin. A) Western blot showing upregulation of PERK with tunicamycin treatment. PERK is visible at ~ 140kDa. Calnexin was used as a loading control. **B)** Band intensity from Western blotting on samples treated with 5µg/mL tunicamycin or a DMSO control has been determined by densitometry. Total PERK level was normalised to housekeepers for both tunicamycin and DMSO cells. A ratio of normalised PERK levels in tunicamycin to DMSO treated groups was calculated – termed the Total PERK expression fold change upon tunicamycin treatment (y-axis). Each data point represents an independent sMN conversion with subsequent independent experiments. N equals 1 for all lines except: C9-1 = 2 and C9-Δ1 = 2. Error bars show the 95% CI.

Quantitative analysis of band intensity was carried out to determine the response of different cell lines to tunicamycin treatment. The interpretation of these results is limited due to the low sample size examined so far, however, there is some evidence of a differential PERK response in *C9ORF72* mutants relative to their gene corrected controls.

ATF6

Cleaved fractions of ATF6 in ER stress translocate to the nucleus where they bind to ER stress response elements (ERSEs) in gene promoters which are capable of inducing both pro-survival and pro-apoptotic changes (Yoshida et al. 1998).

Protein samples from three-week-old sMNs cultured in normal conditions were probed for ATF6 on Western blotting. Multiple positive bands were observed (**Figure 18**). The ~100kDa band corresponds to full-length ATF6. Spliced products are observed at ~80kDa, ~67kDa, ~57kDa, ~37kDa, ~30kDa, ~27kDa and ~15kDa. The intensity of these bands varied between samples. The analysis focused on the ratio of full-length ATF6 to the 37kDa cleaved fragment. The 37kDa band has been previously associated with cell stress using a different anti-ATF6 antibody (Mao et al. 2007).

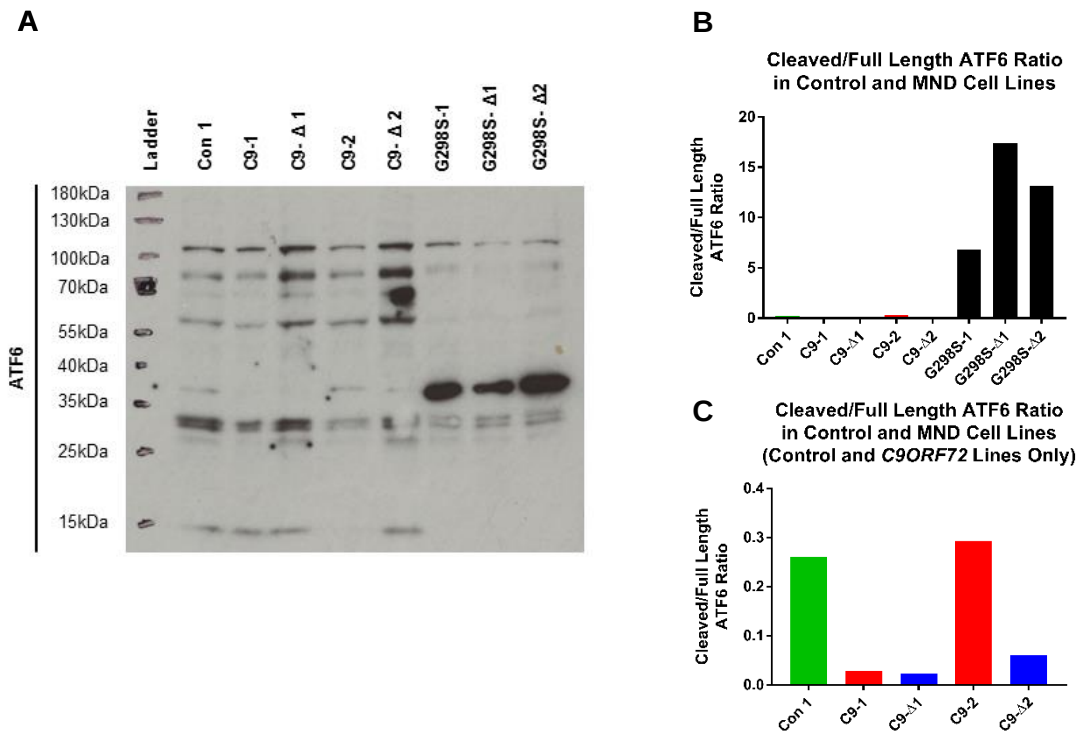


Figure 18: The 37kDa cleaved ATF6 fragment is present at higher levels in *TARDBP*^{G298S} mutants with heterogeneous ATF6 cleavage among *C9ORF72* mutants. **A)** Western blot for ATF6. Full-length ATF6 runs at approximately 100kDa. 60kDa, 50kDa and 37kDa cleaved fractions have also been reported. Additional fragments were observed around 80kDa, 67kDa, two further fractions at approximately 30kDa and a 15kDa fraction. Anti-ATF6 antibody was applied and the membrane probed prior to treatment anti-GAPDH and sodium azide excluding the possibility that the 36kDa fragment observed in the Anti-ATF6 Western blot is a GAPDH signal. **B)** Plot showing the densitometric ratio of 37kDa cleaved/full-length ATF6 in all lines. Controls are in green, *C9ORF72* mutants in red, *C9ORF72* gene corrections in blue and the *TARDBP* mutant and gene corrections in black. Levels of the spliced product are seen to be considerably higher in *TARDBP*^{G298S} mutants and also in their gene corrected partners suggesting the effect may be independent of the mutation – replication is required. **C)** A repeat of plot B with the *TARDBP* mutant and gene corrected lines removed to allow visualisation of this parameter in control and *C9ORF72* lines. While the ratio of ATF6 products is very similar in C9-1 it is interesting that C9-2 displays greater levels of the spliced product that its gene edited partner. The C9-2 levels are roughly comparable to that of Con 1.

Although the small sample size prohibits statistical analysis, there is a clear trend to an increased abundance of the 37kDa cleaved ATF6 protein in *TARDBP*^{G298S} mutants and gene corrected partners suggesting constitutive activation of this UPR stress sensor independent of the G298S mutation. In *C9ORF72* mutants heterogeneity is again observed with little AFT6 cleavage in C9-1 but greater quantities in C9-2, some of which appears to be dependent on the *C9ORF72* HRE as evidenced by the reduction in cleaved/full-length ATF6 ratio in the C9-Δ2.

Following examination of the constitutive state of ATF6 cleavage, this pathway was investigated in tunicamycin-treated cells (**Figure 19**).

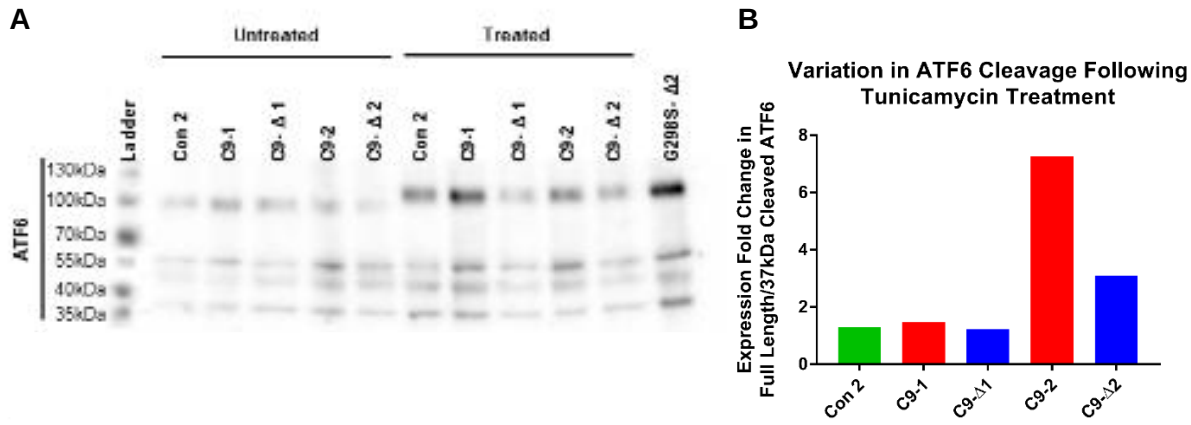


Figure 19: Heterogeneous ATF6 cleavage among *C9ORF72* mutant lines in response to tunicamycin may provide evidence for HRE toxicity modifiers. A) Western blot for ATF6 showing upregulation of full-length ATF6 on treatment with tunicamycin which appears to be more prominent in *C9ORF72* and *TARDBP*^{G298S} mutants. The first five samples are untreated. The next five samples were treated with 5µg/mL tunicamycin for 24 hours prior to protein extraction. G298S-Δ2 was included as a positive control. **B)** Graph showing the expression fold change in the ratio of the density of the 110kDa ATF6 product to that of the 37kDa product. The y-value is calculated in two steps. First, the ratio of full-length ATF6 to that of the 37kDa cleaved product is established in cells treated for 24 hours with 5µg/µL (5µL) tunicamycin (T) or a 5µL DMSO (D) control at three weeks of age. Each line was thus described with two values – indicating the abundance of the 37kDa spliced product relative to the full-length protein after DMSO or tunicamycin treatment. The tunicamycin treatment value was then divided by that of the DMSO value to determine the expression fold change of ATF6 cleavage following tunicamycin treatment. Only one experiment has been performed so variance cannot be assessed. Controls are in green, *C9ORF72* mutant lines in red and gene corrected partners in blue. Mirroring the results obtained on examination of constitutive ATF6 cleavage it can be seen that in C9-2 the *C9ORF72* HRE seems to confer an exaggerated response to tunicamycin which is significantly greater this time than that of the control line. This may suggest the presence of an HRE toxicity modifier active within C9-2.

On examination of **Figure 19**, upregulation of full-length ATF6 was apparent in tunicamycin-treated lines and in G298S-Δ2 that was used as a positive control. To allow comparison of samples from independent conversions, samples were first normalised to a housekeeper and then normalised to C9-Δ1^{A}. C9-Δ1 was chosen as it was present in all experiments.

The ATF6 cleavage response of the C9ORF72 lines to tunicamycin to some extent mirrors their basal ATF6 cleavage status. It can be seen that once again C9-1 and C9-2 respond differently.

Cleaved ATF6 in C9-1 is comparable to that of the control line and appears independent of *C9ORF72* repeat expansion mutation. In C9-2, the HRE appears to confer upon sMN cultures an increased response to tunicamycin in the form of greater production of the ATF6 37kDa cleaved fragment. This provides some evidence again for heterogeneity of *C9ORF72* pathophysiological processes though replication is required.

sMN culture survival analysis following tunicamycin-induced ER stress

With validation of HSPB8 upregulation in *C9ORF72* mutants and evidence that sMNs from patients carrying the *C9ORF72* HRE might respond differently to ER stress induced by tunicamycin, the effect of tunicamycin on the survival of these cultures was investigated. Evidence of ER stress in C9-MND iPSC-derived MNs is already reported in the literature (Haeusler et al. 2014) and we sought to replicate these findings. Tunicamycin at doses ranging from 1μg/mL to 10μg/mL was administered to three-week sMN cultures with viability assessed at 24 and 48 hours. Viability was assessed by two methods, which are presented as complementary – manual tracking of a sample population of cells and analysis of the mean normalised integrated density of wells (see *Methods*).

The primary goal of the analysis was to test the hypothesis that the *C9ORF72* mutation resulted in a progressive decrease in sMN viability after induction of ER stress. The data were analysed using mixed linear models. Construction of these models is iterative so before the hypothesis was tested, data on platedown variability and the relationship between tunicamycin dose and viability were first analysed to allow construction of an appropriate model. The covariates/factors considered in the analysis are detailed in **Table 16**. Independent variables are detailed in **Table 17**.

Covariates/Factor Name	Description
Genotype	Control, *C9ORF72* mutant or gene-corrected
Line	Cell line used: Con 1, Con 2, C9-1, C9-Δ1, C9-2, C9-Δ2, C9-3 or C9-Δ3
Dose	Tunicamycin dose in μg/mL: 0, 1, 3, 5 or 10
Place down	A letter assigned to each different sMN conversion and subsequent independent experiments: A – H

Table 16: Table showing covariate and factor names used in the tunicamycin survival analysis.

Independent Variable	Description
24h survival	The percentage survival at 24h as determined by manual cell tracking
48h survival	The percentage survival at 48h as determined by manual cell tracking
24h mean normalised integrated density(mnID)	The mnID at 24h determined using the Moments algorithm. A surrogate marker for survival.
48h mean normalised integrated density(mnID)	The mnID at 48h determined using the Moments algorithm. A surrogate marker for survival.

Table 17: Table showing the independent variables measured in the tunicamycin analysis. In brief, normalised integrated density (nID) is equivalent to the total quantity of activated calcein (calcein AM) within a cell and is calculated by multiplying the intensity of calcein AM fluorescence by the cell area and normalising this result to that of the baseline 0h image. The nID is reported calculated for each individual cell or cluster of cells. To avoid pseudoreplication these values are averaged and presented as the mean nID (mnID) per well (see *Methods*).

Platedown variability

The first factor considered was variability in platedown. It was hypothesised that data would be clustered by platedown which could have important implications for subsequent statistical analysis. The relationship between platedown and viability at 48h is displayed graphically below (**Figure 20**).

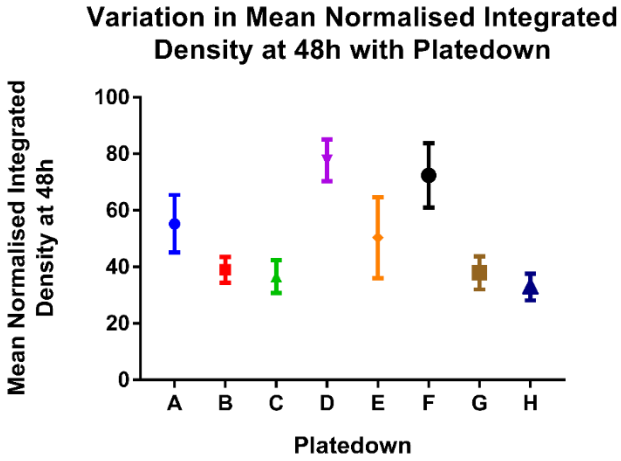


Figure 20: There is considerable variability in the mean normalised integrated density (mnID) between platedowns at 48h. Points show the mean of the mean normalised integrated density. Error bars show the 95% CI. The number of wells from which mean normalised integrated density for each platedown was calculated is: A = 61; B = 53, C = 60, D = 54, E = 21, F = 18.

Such variance between platedowns is unsurprising. Some of this variance likely results from other measured factors including variability between cell lines and tunicamycin doses used across experiments. Other confounding variables will include factors such as variance in the sMN differentiation protocol, calcein AM batch, plate batch, optical parameters and exact imaging time.

To investigate whether controlling for cell line, sMN age and dose could account for this variance a subset of data was sampled. With the above factors matched there was still a large amount of variability in mnID at 48h (data not shown[B]). It is therefore concluded that platedown should be treated as a random effect with respect to the statistical analysis of this dataset.

Tunicamycin dose

Having established the importance of the effect of platedown in explaining variability between experiments, the relationship between genotype, dose and the two measures of viability (manually determined survival and mnID) was investigated (**Figure 21**).

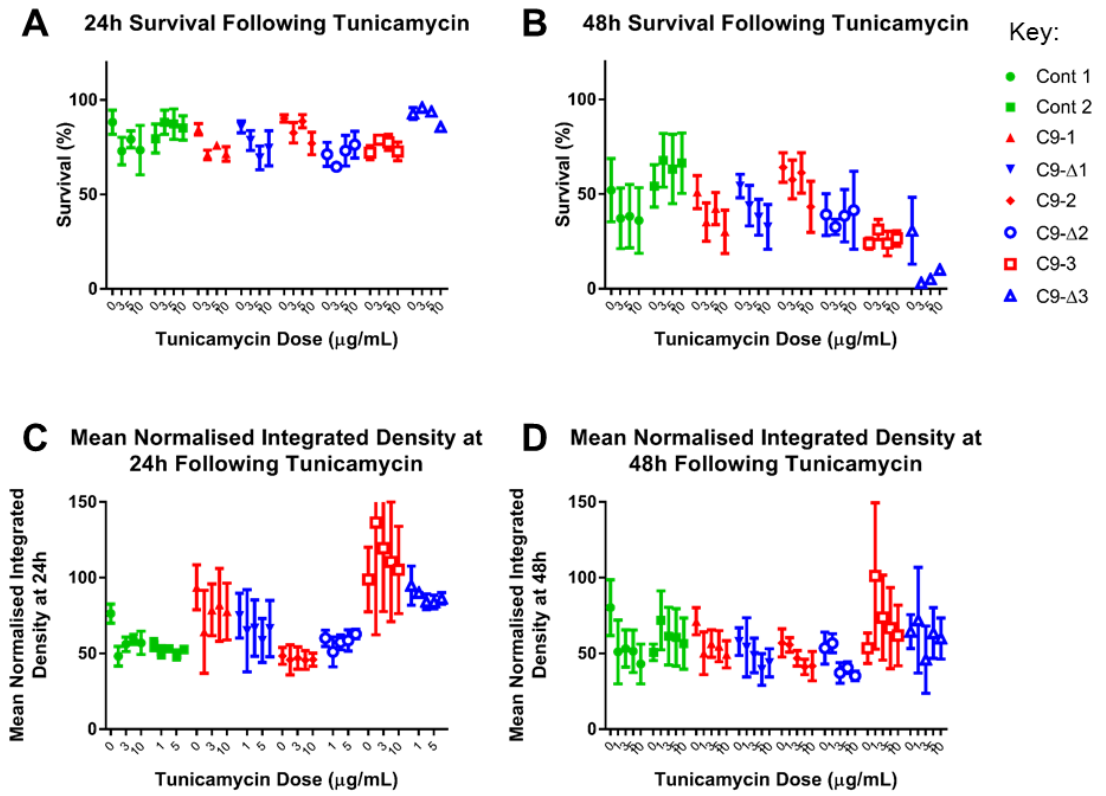


Figure 21: Plots showing two different assessments of the viability of three-week-old sMN cultures in response to a range of tunicamycin doses. Controls are shown in green, *C9ORF72* mutants in red and gene corrected controls in blue. A key at the top right allows assessment of the response of individual cell lines denoted by signals. Points denote the mean and error bars the SEM. The top row of plots shows the percentage of neurons that survived to 24 **(A)** and 48 **(B)** hours as determined by manual cell tracking. Survival in all lines at 24 hours is above 50% dropping to below 50% for most lines at 48 hours. **A)** At 24 hours there does not appear to be a consistent deficit in survival following administration of increasing doses of tunicamycin. **B)** By 48 hours, a trend towards decreased percentage survival with increasing doses of tunicamycin is observed. The bottom row of plots shows the relationship mean normalised integrated density (mnID), a survival proxy, and tunicamycin dose. In contrast to percentage survival, the mnID does not decrease significantly between 24 **(C)** and 48 hours **(D)**. The addition of calcein AM prior to imaging means the mnID of viable cells will remain constant or increase. **C)** Again, no clear relationship between mnID and increasing tunicamycin dose is evident at 24 hours. Note that SEM error bars in two cases have been excluded at the y-axis limit of 150[c]. An additional copy of **C** with a complete y-axis is shown in **Supplemental Figure 4 D)** By 48 hours there is a trend towards decreased mnID with increasing doses of tunicamycin.

Figure 21 shows a trend towards decreased survival in sMN cultures with increasing doses of tunicamycin which is apparent at 48 hours but not 24 hours, regardless of the analysis method used. The variability of the results denoted by the error bars can to some extent be addressed by including platedown in the statistical modelling of this data as described previously.

Statistical analysis (**Table 18**) demonstrates an increasing dose of tunicamycin decreases sMN survival. This effect is statistically significant 48 hours after tunicamycin administration but not at 24 hours ($\alpha = 0.05$).

Independent Variable	Coefficient	Standard Error	T value	Likelihood ratio test
24h survival	-0.4824	0.2630	-1.834	P = 0.06863
24h mnID	-0.235	0.646	-0.433	P = 0.6648
48h survival	-1.0575	0.4063	-2.603	P = 0.01043
48h mnID	-1.2850	0.4805	-2.674	P = 0.008024

Table 18: Output from four mixed linear regression models for the independent variables survival and mnID at 24 and 48h hours with dose and line as fixed effects and platedown as a random effect. Four additional models in which dose was removed as a covariate allowed assessment of hypothesis that survival would decrease with increasing tunicamycin dose. The likelihood ratio p values from this test are shown in the final column.

The prior data analysis was important to ensure that any association between genotype and cell viability was not confounded by the aforementioned variables.

The presence of the *C9ORF72* HRE does not increase decrease viability in sMN cultures treated with tunicamycin

The hypothesis that the *C9ORF72* mutation would decrease sMN survival at 48 hours after administration of tunicamycin, which induces ER stress and a UPR, was tested (**Figure 22**).

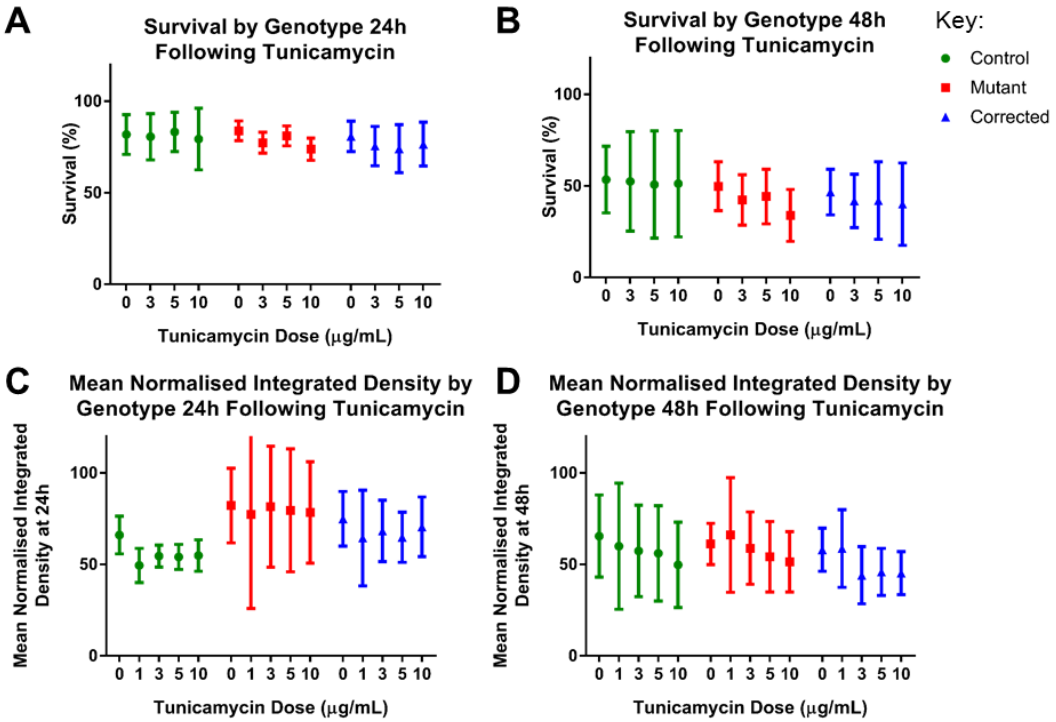


Figure 22: Plots demonstrating that the decreased viability of sMNs with increasing doses of tunicamycin does not vary between control, *C9ORF72* mutant and gene corrected lines. Controls are shown in green, *C9ORF72* mutants in red and gene corrected controls in blue. Points denote the mean and error bars the 95% CI. The top row of plots shows the percentage of neurons that survived to 24 **(A)** and 48 **(B)** hours as determined by manual cell tracking. As in related **Figure 21** survival can be seen to decrease between the 24 **(A)** and 48 hour **(B)** assessments. The bottom row of plots shows the relationship between mean normalised integrated density (mnID), a survival proxy, and tunicamycin dose. In contrast to percentage survival, the mnID does not decrease significantly between 24 **(C)** and 48 hours **(D)**. The addition of calcein AM prior to imaging means the mnID of viable cells will remain constant or increase. With both analysis protocols the trend towards decreased survival with increasing doses of tunicamycin can again be seen at 48 hours **(B and D)** mirroring the result observed in **Figure 21** where the responses of individual lines were assessed. There does not appear to be a clear difference in the response of any one genotype though the data cannot be displayed in its full dimensionality. Statistical modelling of the effect of genotype is shown in **Table 19**.

Independent Variable	Genotype	Coefficient	Standard Error	T value	Likelihood ratio test
24h survival	Mutant	1.1734	2.3924	0.490	0.4524
	Corrected	-1.6582	2.4727	-0.671	
24h mnID	Mutant	21.795	6.343	3.436	0.002362
	Corrected	10.332	6.378	1.620	
48h survival	Mutant	1.4287	3.7039	0.386	0.8579
	Corrected	-0.3842	3.8239	-0.100	
48h mnID	Mutant	4.5033	4.7186	0.954	0.2314
	Corrected	-2.1858	4.7444	-0.461	

Table 19: Output from four mixed linear regression models for the independent variables survival and mnID at 24 and 48h hours with dose as a fixed effect and platedown as a random effect. Four additional models in which genotype was removed as a factor allowed assessment of hypothesis that survival would decrease with increasing tunicamycin dose. The likelihood ratio p values from this test are shown in the final column.

The visual data, the absence of negative correlations in *C9ORF72* mutant lines in mixed linear regression modelling and insignificant likelihood ratios test suggest that the *C9ORF72* mutation does not result in decreased viability in response to tunicamycin relative to controls.

Variability between individual cell lines is important to consider given the heterogeneity among lines demonstrated already in this work. Further investigation showed the increased survival in *C9ORF72* mutants at 24 hours following tunicamycin determined by mnID to be heterogeneous. Of the three *C9ORF72* mutant lines, two showed increased survival with coefficients of 41.9863 and 14.4327. The third line showed decreased survival with a coefficient of –22.3029. Further methodological and statistical parameters will have to be considered.

It is concluded that the *C9ORF72* HRE does not make three week sMNs more vulnerable to death following tunicamycin up to 48 hours following administration.

Insoluble TDP-43 is present in *TARDBP* mutant and corrected lines but not in *C9ORF72* HRE carriers or controls

TDP-43 pathology is evident in most cases of MND and is of relevance in C9-MND because it correlates well with neurodegeneration unlike DPR proteins (MacKenzie et al. 2013). Tunicamycin has also been demonstrated to lead to cytoplasmic TDP-43 inclusions in rodent slice cultures (Leggett et al. 2012).

As C9-MND is also associated with TDP-43 pathology we sought to assess the TDP-43 status of these lines. Initially, experiments were performed on *TARDBP* mutant lines as a positive control to optimise the technique.

Soluble and insoluble protein fractions were probed for TDP-43 from three-week-old sMN cultures (**Figure 23**).

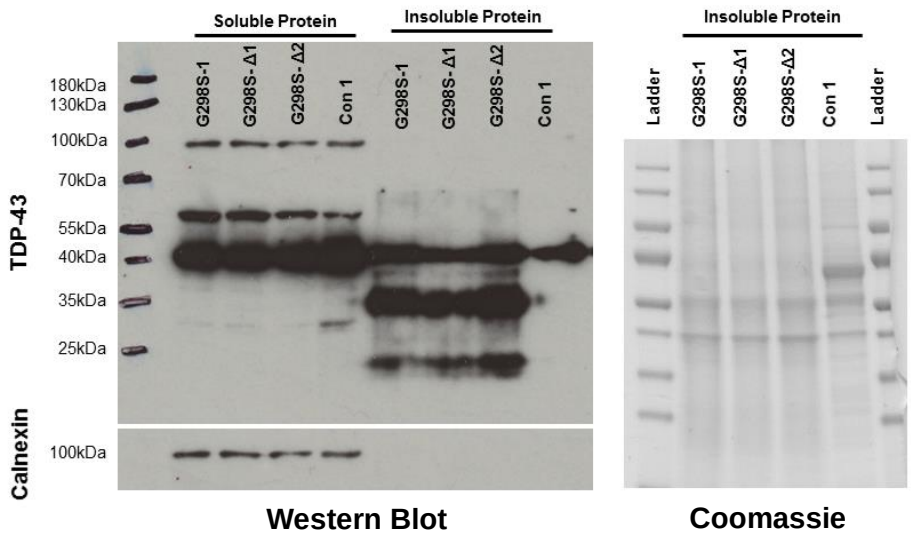


Figure 23: G298S-1 and gene corrected partners show increased expression of ~58kDa, ~35kDa and ~25kDa TDP-43 products thought to be pathological from previous studies. ***Left:*** Western blot for TDP-43 in a *TARDBP*^{G298S} mutant, gene corrected and control lines. Upper black horizontal bars indicate whether protein samples were derived from the soluble or insoluble (urea) fractions. A protein ladder has been marked on in pen (far left column). Calnexin has been used as a housekeeper. Note that an overexposed image has been presented to demonstrate all products identified. Where semi-quantitative densitometry has been performed Western blots analysed were insured to have grey values < 255. ***Right*****:** Coomassie blue gel (greyscale) showing comparable protein loading for insoluble fraction samples which

cannot be normalised using a housekeeper. The samples are flanked on either side by protein ladders which have also taken up the Coomassie dye.

Physiological TDP-43 at 43kDa was observed in control, *TARDBP*^{G298S} and gene corrected lines in both soluble (n = 1) and insoluble protein fractions (n = 2). When levels of the 43kDa TDP-43 in the soluble fraction have been normalised to a calnexin housekeeper to normalised for loading (**Figure 24**).

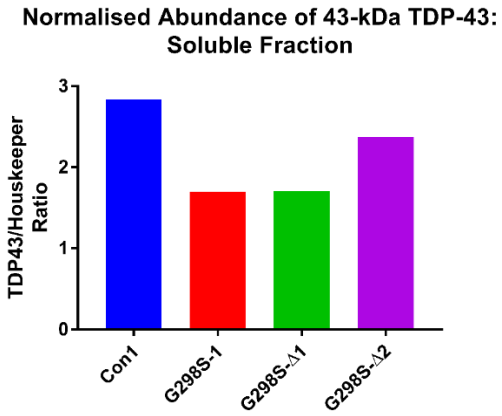


Figure 24: Normalised abundance of TDP-43 to a calnexin housekeeper in the soluble fraction of three-week-old sMN cultures. n = 1. Replication is required before interpretation.

A ~58kDa protein was observed in the soluble fraction only. The abundance of the ~58kDa product is lower in the control line when intensity is normalised to the calnexin housekeeper (**Figure 25**), however, a higher sample size is required to validate these findings.

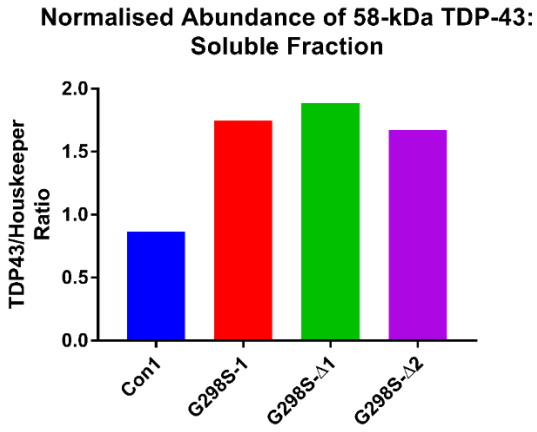


Figure 25: Normalised abundance of the 58kDa TDP-43 product to a calnexin housekeeper in the soluble fraction of three-week-old sMN cultures. n = 1.

While repetition and additional control lines are required this 58kDa product has been previously reported to represent phosphorylated TDP-43. Hyperphosphorylation of TDP-43 is a pathological hallmark of MND and the 58kDa product is more abundant in G298S-1 and its gene corrected partners which mirrors pathology with respect to truncated TDP-43.

G298S-1 and gene corrected partners showed additional products in the insoluble fraction at ~35kDa and ~25kDa (**Figure 23**) and this was validated upon replication (n = 2). To a much lesser degree, some amount of the ~35kDa product is also present in the soluble fraction of a control line. Replication validated the presence of these products in the insoluble fraction of G298S-1 and G298S-Δ1 but provided no evidence of these products in controls. All the fragments described have also be observed in a *TARDBP*^{M337V} mutant (data not shown). Gene correction does not abolish these pathological truncated TDP-43 products.

Having demonstrated the ability to detect TDP-43 pathology *TARDBP* lines, the same experiments were repeated on lines harbouring the *C9ORF72* HRE. The TDP-43 levels in the soluble fraction of control, *C9ORF72* mutants and corrected lines were assessed normalised to the housekeeper protein - calnexin (**Figure 26**). G298S-1 was used as a positive control. Data from a single experiment did not show any relationship between soluble TDP-43 levels and genotype (n = 1).

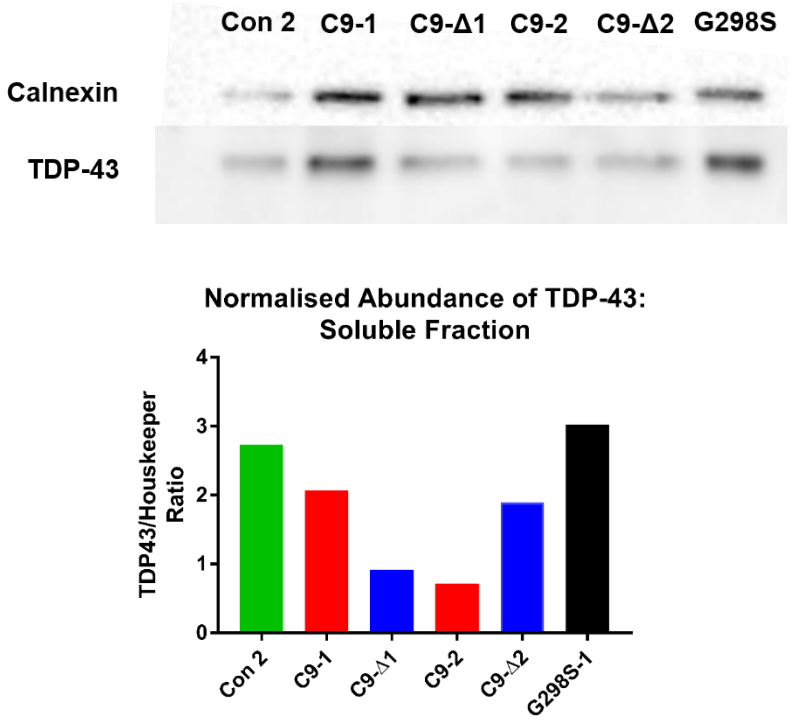


Figure 26 Preliminary data suggests TDP-43 abundance in the soluble fraction varies between lines but is not associated with genotype. ***Top*****:** Western blot for TDP-43 and calnexin in the soluble protein fraction from three-week-old sMN cultures. n = 1 for all samples. ***Bottom*****:** The density of the 43kDa band was normalised to that of the calnexin housekeeper and is expressed as a ratio on the y-axis. Further replication is required before conclusions can be drawn however the heterogeneity among *C9ORF72* mutant lines would seem to indicate the abundance of soluble TDP-43 does not alter in lines carrying the *C9ORF72* HRE.

In the insoluble fraction of the *C9ORF72* lines a small quantity of ~43kDa product was observed (n = 2). The G298S-1 positive control showed the expected products at ~35kDa and ~25kDa as expected which were not evident in *C9ORF72* mutants or gene corrected controls (**Figure 27**). Coomassie dye was used to show equivalent loading within wells.

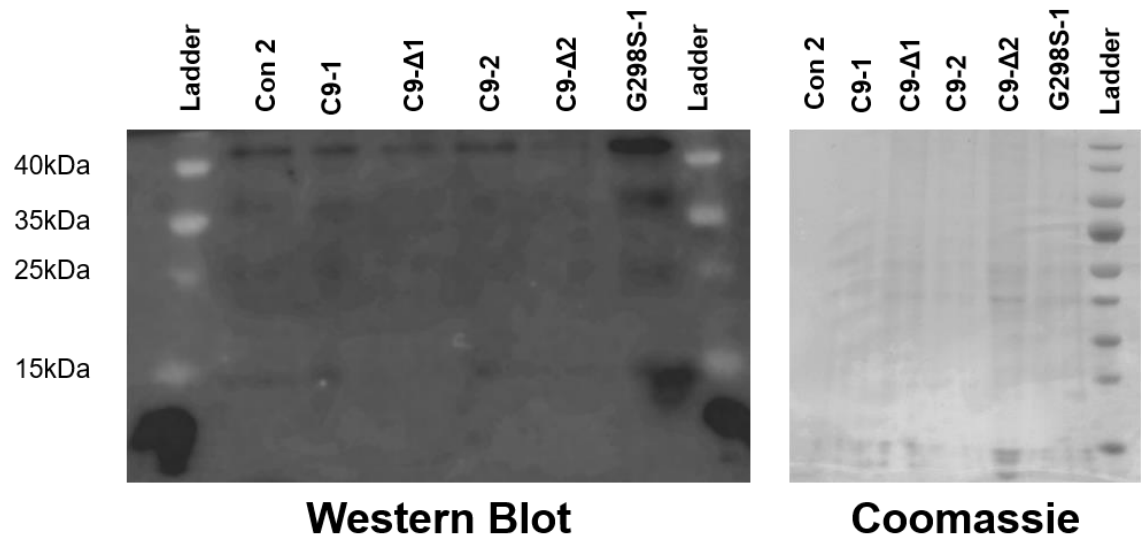


Figure 27: *C9ORF72* mutant sMNs cultures at three weeks do show truncated TDP-43 products in the insoluble protein fraction. ***Left*****:** Western blot showing control, *C9ORF72* mutant and gene corrected insoluble protein samples with a G298S-1 positive control. Images have been subject to processing in an attempt to reveal low-intensity fragments. In all lines TDP-43 can be seen at its expected mass. Addition products are visible at ~35kDa and ~25kDa in the G298S-1 line. ***Right*****:** Coomassie stained membrane showing protein loading.

Nuclear evacuation of TDP-43 has been described in human tissue from patients with MND and FTD (Arai et al. 2006; Neumann et al. 2006). Work from rodent primary neuronal cultures has suggested nuclear evacuation of TDP-43 is a better indicator of impending cell death than cytoplasmic aggregations (Barmada et al. 2010).

After quantifying TDP-43 expression the next step was to use immunocytochemistry to investigate TDP-43 mislocalisation – another hallmark of MND pathology. On examination of control (n = 2), *C9ORF72* mutant (n = 6) and gene corrected (n= 5) lines, there was no evidence of aggregation or nuclear evacuation. **Figure 28** shows a representative image.

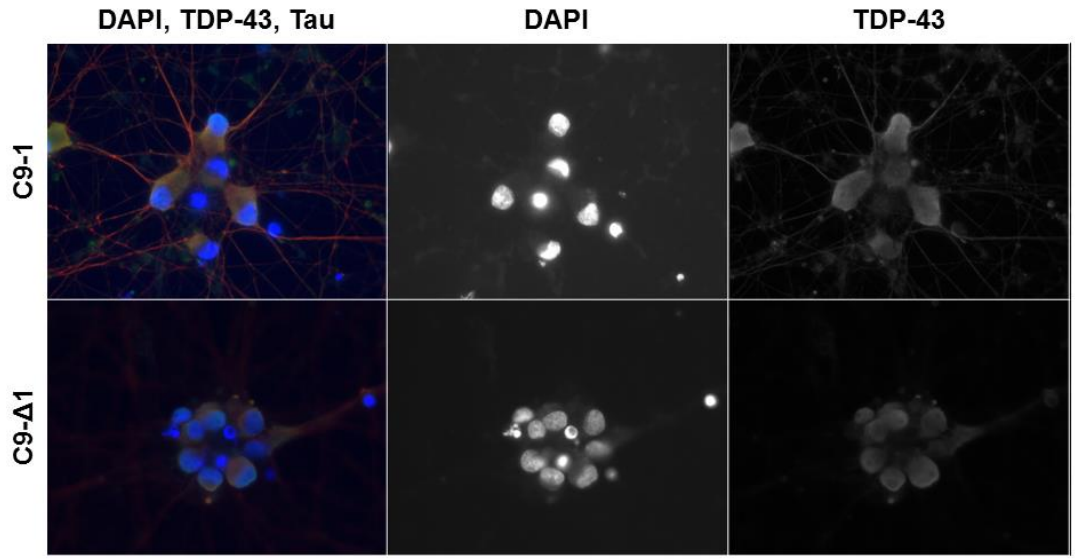


Figure 28: Representative image of ICC for TDP-43 and tau with the nuclear dye DAPI showing homogenous nucleocytoplasmic distribution of TDP-43. Representative images from three week old sMN cultures (coverslip) imaged at x63. ***Left column*:** Image showing all channels: tau (red), TDP-43 (green) and DAPI (blue). ***Middle column:*** DAPI signal alone. ***Right column*:** TDP-43 signal alone.

TARDBP Genotyping

The identification of insoluble TDP-43 in the gene corrected lines was unexpected. To confirm the identity of the lines DNA was purified and PCR performed generating a 452bp amplicon of exon 6 spanning the mutation site. Confirmation of successful amplification was obtained by gel electrophoresis before the samples were sent for Sanger Sequencing (**Figure 29**).

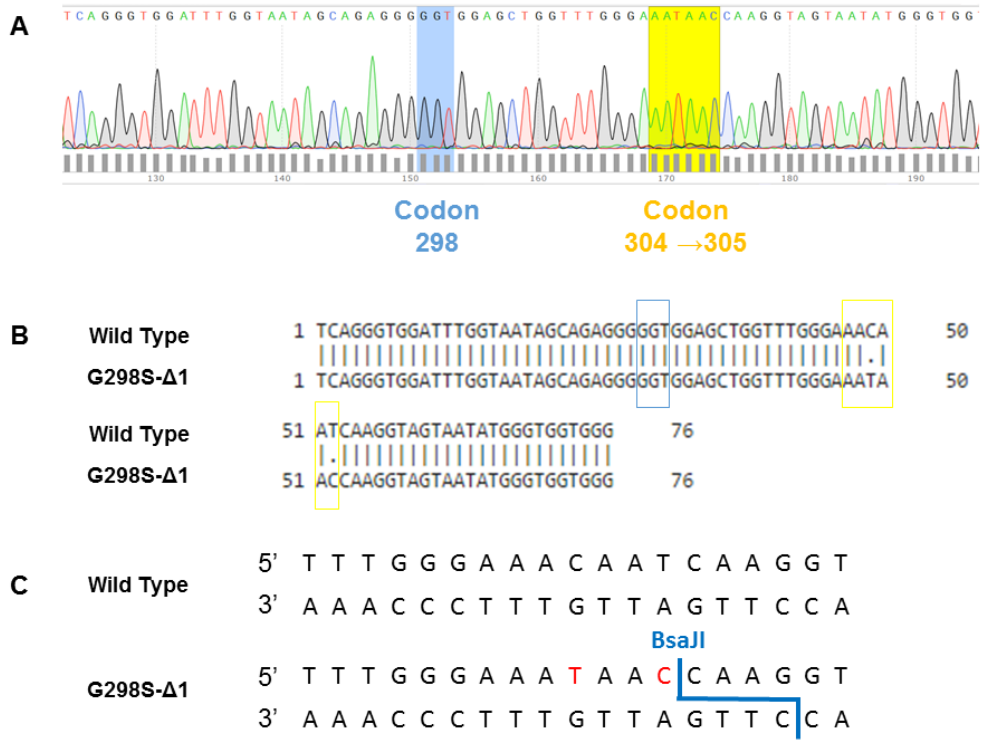


Figure 29: Genotyping of G298S-Δ1. **A)** Sanger Sequence of G298S-Δ1**. B)** Sequence alignment of the wild-type (upper) and G298S-Δ1 sample were assessed using the EMBOSS local alignment tool (Water). Codon 298 does not vary refuting the hypothesis that the insoluble TDP-43 observed in this line was the result of the line not having a control identity. Silent mutations inserted at the time of gene-correction in codons 304 and 305 are also detected confirming the line has been gene corrected and is not a wild-type control. **C)** The silent mutation introduces a BsaJI cut site.

The above evidence suggests that both G298S-1 and its gene corrected controls show evidence of TDP-43 pathology. This throws into question the relationship between *TARDBP* mutations and TDP-43 pathology.

Chapter 3 – Results Notes

A – On reflection, two control lines should have been included in each experiment. One would have been able to serve as a constant for normalisation while the addition of the second would allow the variability of protein expression in control lines to be expressed. Cell culture logistics though and a lack of experience and forward thinking with respect to data analysis on my behalf precluded this.
B – This data set is large and noisy but the variation between platedown is prominent in a number of matched subsamples.
C – Failure to limit the y-axis would have resulted in further compression of the majority of datasets making the visualisation of results more challenging.

Chapter 4: Discussion

A cellular model of MND has been used to investigate the role of the *C9ORF72* HRE on sMNs. The inclusion of both healthy controls and isogenic controls for each of the three C9 lines studied allows several different types of conclusion to be made. A hypothetical pathological phenotype is considered in **Table 20**. Possible distributions of this phenotype among control, mutant and corrected lines are considered as well as the inferences which can be made from such distributions.

Control	Mutant	Corrected	Inference
Absent	Present	Absent	The phenotype is a function of the *C9ORF72* HRE specifically
Absent	Present	Present	The phenotype is a function of diseased lines independent of the *C9ORF72* mutation
Absent	Absent	Absent	The phenotype is not associated with C9-MND in this study

Table 20: Table showing the distribution of a hypothetical pathological phenotype among control, mutant and gene corrected lines with an inference about the phenotype based on this distribution. The phenotype in each case is described as present or absent.

With these distinctions possible, it was sought to investigate the role of the UPR in C9-MND and specifically the role of the *C9ORF72* HRE.

Transcriptomic data identified the *C9ORF72* HRE perturbs expression of a number of transcripts many of are associated with pathways already associated with MND

RNAseq data was used to identify transcripts in which gene expression was significantly altered by the *C9ORF72* mutation. Whole transcriptome profiling was performed on three-week-old sMNs derived from for 2 independent healthy controls, 2 C9-MND cases and 1 isogenic control line. Differentially expressed genes were identified by following criteria:

1. Genes must have significant differential expression in both independent mutant *C9ORF72* lines ($p < 0.05$) when compared to independent control lines
2. There must be a significant difference between C9-1 and C9-Δ1 ($p < 0.05$)
3. There must be no significant difference observed between C9-Δ1 and controls ($p > 0.05$).

Assessment of protein class using PANTHER showed enrichment for many pathways which have already been implicated in C9-MND and more broadly MND as a whole. These proteins map to pathways involved in RNA metabolism, excitotoxicity, nucleocytoplasmic shuttling, axonal transport, metabolism, oxidative stress and the UPR. This is a valuable dataset and though it shall not be considered further here it will provide a surplus of candidates for further investigation. With a decrease in price, transcriptomic datasets are becoming more common and a number of such datasets of relevance to the study of MND have been made available online (Ferraiuolo et al. 2007; de Oliveira et al. 2013; Donnelly et al. 2013; Groen et al. 2013; Heath et al. 2013; Maximino et al. 2014; Aronica et al. 2015; Noristani et al. 2015; Patin et al. 2016; Riva et al. 2016; Wang et al. 2017).

In the future, it would be interesting to compare these datasets. Though often protocols, species and timepoints vary substantially if experiments have been independently well performed it may prove an excellent resource for generating hypotheses in an economically efficient manner.

Of the three candidate genes examined, only HSPB8 mRNA upregulation was validated

Three candidate genes from the RNAseq findings involved in the UPR were selected for validation by qPCR.

HSP90B1 has been implicated in many cellular functions but of particular relevance here is the role it plays as an ER chaperone. Its role seems to have been most heavily studied in immunological protein homeostasis but recently it has been shown to play a role in ERAD (Di et al. 2016). HSP90B1 status does not seem to have been described in MND but has been implicated in another of the motor neuron diseases – SMA. SMA is caused by mutations in the *SMN1* gene which codes for survival motor neuron (SMN) protein (Lorson et al. 1999). *HSP90B1* is downregulated in cellular models of SMN knockout (KO) (Wu et al. 2011). *HSP90B1* mutations have not been associated with any human disease. Failure to validate this RNAseq hit is disappointing given it had one of the greatest effect sizes. The transcript does appear to be significantly upregulated in C9-3 but not in C9-1 (**Figure 10**). This provides the

first evidence presented of heterogeneity among *C9ORF72* mutants which will be discussed later.

HSPB8 has been previously implicated in MN pathology. Two diseases, both of which affect motor neurons, have been attributed to *HSPB8* mutations: hereditary distal motor neuropathy type IIA (Irobi et al. 2010) and hereditary motor and sensory neuropathy type 2L (Tang et al. 2005). HSPB8 known to have chaperone activity and has recently been reported to play a role in combating the accumulation of truncated TDP-43 products (Crippa et al. 2016). Work in primary neuronal cultures has shown that neurons subject to lentiviral transfection with mutant HSPB8 show evidence of neurite degeneration (Irobi et al. 2010); specifically, a decreased number of neurites per cell and decreased neurite length. This was not seen in cells transfected with wild-type HSPB8 and the phenotype was most marked in motor neurons, with minor involvement of sensory neurons and sparing of cortical neurons. In the current study, upregulation of *HSPB8* expression is demonstrated in *C9ORF72* mutants (**Figure 11**) with restoration to within normal limits in gene corrected controls suggesting the upregulation is a direct result of the *C9ORF72 HRE*. This could result from several mechanisms including transcriptional changes caused by pathological RNA foci or as a response to DPR protein aggregates within these cells. Considering the absence of insoluble TDP-43 pathology in these lines, the presence of DPR proteins and RNA foci and restoration of the phenotype in gene corrected controls, it is speculated that *HSPB8* upregulation may be independent of TDP-43. It could be argued though that TDP-43 pathology is driving the upregulation of *HSPB8* in a response which will eventually become decompensated leading to aggregation and neuronal death. This possibility could be investigated by examining the C9ORF72 lines at later time points. If TDP-43 pathology can be identified at later developmental stages it could be hypothesised that *HSPB8* KO would lead to earlier onset TDP-43 pathology. It should be remembered that there is increasing evidence in neurodegenerative diseases associated with protein aggregates that soluble forms of the proteins that cause toxicity (Haass and Selkoe 2007).

XPO1 was of interest because of its role in nucleocytoplasmic shuttling (Fornerod et al. 1997). It functions in the export of over 200 proteins from the nucleus (Fu et al. 2013) and evidence

has implicated nucleocytoplasmic transport as a potential driver of pathophysiology in *C9ORF72* HRE disease recently reviewed by Zhang et al. (Zhang et al. 2015). No change in XPO1 levels was demonstrated in C9-MND lines (**Figure 12**). While, to our knowledge, no previous reports of alterations in XPO1 status have been associated with C9-MND, XPO1 levels have been shown to be associated with those of TDP-43. In a cellular TDP-43 KO model increased levels of XPO1 have been demonstrated (Štalekar et al. 2015). It may be that it is too early to see changes associated with XPO1 which may come with more advanced TDP-43 pathology.

C9ORF72 mutant sMNs did not show evidence of a constitutively active unfolded protein response

Given the link between HSPB8, TDP-43 and *C9ORF72*, the status of the UPR in *C9ORF72* mutant sMNs was investigated. Previous work by Dafinca et al. has demonstrated the presence of ER stress granules in iPSC-derived sMNs between two and four weeks post platedown (Dafinca et al. 2016) so it was hypothesised that increased ER stress sensor activation would be observed in these lines. The UPR was investigated in three-week-old *C9ORF72* mutant sMNs by studying the three ER stress sensors – IRE1, PERK and ATF6.

IRE1 activation was assessed by looking downstream at XBP1 splicing. The ratio between spliced and unspliced *XBP1* transcripts was used as an indicator of ER stress and IRE1 activation. The spliced *XBP1* transcript was not observed in any cell line examined in basal conditions (control, *C9ORF72* mutant, *C9ORF72* gene corrected and $TARDBP^{G298S}$ gene corrected) (**Figure 14**). To determine if a dynamic response to tunicamycin might be observed, cells treated with tunicamycin (5ug/mL) or DMSO as a control were subject to the same analysis. *XBP1* transcript splicing was observed in tunicamycin-treated cells and not in the DMSO control group but no differences in response to tunicamycin were observed in any cell line tested (**Figure 15**). In MND disease and Huntington's disease, XBP1 deficiency appears to be neuroprotective (Hetz et al. 2009; Vidal et al. 2012). In a transgenic rat model (Camk2α-tTA) progressive depletion of both full length and spliced XBP1 protein have been demonstrated with increasing TDP-43 pathology in neurons. The XBP1 response to disease was cell type specific with upregulation in reactive microglia and not in reactive astrocytes (Tong et al.

2012). This XBP1 depletion was found before the formation of ubiquitin inclusion bodies or Golgi fragmentation. If TDP-43 pathology is identified in *C9ORF72* mutant sMNs at later time points in the future it would be interesting to then re-examine the status of XBP1 splicing in these lines.

PERK upregulation has previously been demonstrated in post-mortem tissue from patients with MND (Atkin et al. 2008) and PERK overexpression has been shown to reduce tau pathology associated with progressive supranuclear palsy (PSP) (Bruch et al. 2017). In the $SOD1^{G85R}$ mouse model of MND heterozygotes for *PERK* KO showed earlier disease onset and a decreased lifespan (Wang et al. 2011). Together, these results can be used to hypothesise that the upregulation of total PERK seen in end-stage MND may represent a decompensated neuroprotective response. Here, there is no evidence of PERK upregulation under basal conditions in *C9ORF72* mutants in three-week-old sMNs (**Figure 16**). There is, however, when isogeneic pairs are considered, a trend towards increased PERK expression in cells which have been treated with tunicamycin associated with the *C9ORF72* HRE (**Figure 17**). Variation among this response in the two control lines examined though was high so this may represent natural variation and repetition is required before conclusions can be made. Nevertheless, this trend towards increased total PERK as a result of the *C9ORF72* HRE was observed in both *C9ORF72* mutant lines examined (C9-1 and C9-2).

AFT6, following ER stress, is subject to regulated intramembrane proteolysis with cleaved fragments translocating to the nucleus to modify gene expression (Haze et al. 1999). The greatest alteration in ATF6 observed in these iPSC-derived sMN lines was markedly increased production of a 37kDa spliced fraction of ATF6 in $TARDBP^{G298S}$ mutants (**Figure 18**). This protein fragment was observed in both mutant and gene corrected lines. This suggests that the individual from whom these lines were derived was prone to ER stress at a cellular level in a manner independent of the disease mutation. This is in keeping with the finding of $TARDBP^{G2983}$-independent insoluble pathogenic TDP-43 in the gene-corrected lines of this individual (**Figure 23**).

In *C9ORF72* mutants the C9-1 isogenic pair does not show any alteration in the ratio of full length to cleaved ATF6, either constitutively, or following treatment with tunicamycin (**Figures 18** and **19**). In comparison, C9-2 showed an HRE-dependent increase in ATF6 cleavage basally and an exaggerated response after tunicamycin treatment which is only partly explained by the *C9ORF72* HRE. This requires reconciliation with the evidence that even in C9-2 under basal conditions the ratio of full length to cleaved ATF6 is comparable to that of the control line. Further repetition of this experiment with the addition of another control line will aid interpretation. If divergence of values is demonstrated in C9-1 and it's isogenic control it could be postulated that an environmental variable is driving this divergence and the increased ATF6 cleavage in C9-2 cannot be attributed to the *C9ORF72* HRE in its entirety. If, however, the values remain consistent for C9-1 we might argue that gene correction is proving effective in reducing interline variability and in C9-2 the HRE is driving increased ATF6 cleavage. This suggests an additional factor in C9-2 modifies the response to the HRE. This could be a 'juxtamutation' factor (or factors) in concert with the *C9ORF72* HRE in the C9-2 genome have led to a phenotype of increased ATF6 cleavage or alternatively, a difference in the HRE itself, such as length, leads to different phenotypes. The former hypothesis seems more likely, however, it should be noted that the repeat length of C9-2 is known to be longer than that of C9-1. In the case of basal ATF6 cleavage, it would seem that this observation could not account alone for the individuals' MND because the healthy control also shows a similar relative abundance of 37kDa ATF6. The situation following treatment with tunicamycin is different again. This time, C9-2 shows a markedly increased upregulation of 37kDa ATF6 in response to tunicamycin in excess of that of the control line. An important limitation of the dataset at present is that the control line is not the same as that used to investigate the constitutional activity of ATF6; this will be corrected in future experiments.

Convergence of pathways implicated in MN degeneration in MND is evidenced by considering the interaction between excitotoxicity and the unfolded protein response. Evidence has been acquired in cellular models of *SOD1* mutations showing XBP1 splicing increases with MN electrical activity (Wang et al. 2017). Interestingly, this relationship is bidirectional and induction of XBP1 splicing with dichlorodiphenyltrichloroethane (DDT) increased the firing rate of MNs.

C9ORF72 HRE sMNs are not more vulnerable to tunicamycin than controls

Ultimately, the end result of pathology in MND is the death of MNs. Thus, after the above mechanistic investigation, the viability of *C9ORF72* sMNs in response to the ER stressor tunicamycin was undertaken.

Data from the two analysis approaches applied to survival data both suggest that the *C9ORF72* HRE in sMNs does not confer susceptibility to death following tunicamycin-induced ER stress. This finding contrasts with a previous report in the literature (Haeusler et al. 2014). This study shows the increased vulnerability of iPSC-derived motor neurons to tunicamycin determined by increased uptake of the dye propidium iodide (PI) which they present in their Extended Data Figure 9. In an attempt to understand this discrepancy the methods used by Haeusler et al. are compared to the manual cell tracking survival analysis presented here. The manual cell tracking survival analysis is compared because it is more technically similar to the method used by Haeusler.

PI cannot cross intact cell membranes so is taken up by compromised cells providing a means to assess viability. Calcein AM also reports primarily on membrane integrity as active AM metabolite cannot escape from the cell across intact membranes. Haeusler et al. show data from one *C9ORF72* control line and one *C9ORF72* HRE line. Both these lines were independently differentiated twice compared to the four-plus number of replicates here. MNs were differentiated using a different protocol and grown using different surfaces and coatings (poly-L-lysine and laminin on glass verses LMF on plastic). Anecdotally, the Chandran lab has observed diverge in phenotype based on the type of coating used (LMF verses laminin) with resistance to a 'pathological' phenotype on LMF as we used here. This could provide some explanation for the divergence. Though not explicitly stated, based on the timing of other experiments in the paper and the previous work of the lab it seems feasible experiments were conducted on MNs around 50 days old (slightly older than the 37-day old neurons used here). The doses of tunicamycin are reported as varying from 0.1μM to 3μM. Tunicamycin is a mixture of homologous compounds with a variable atomic mass. In the absence of more information, we shall assume the four major homologs are present in equal quantities giving an average atomic mass of 840.25. This equates to 84.025μg to 2.52075mg per litre (**Table 21**).

Work	Minimum Tunicamycin Dose (μg/mL)	Maximum Tunicamycin Dose (μg/mL)
This thesis	1	10
Haeusler et al. 2014	0.084025	2.52075

Table 21: Comparison of tunicamycin doses used by Gane and Haeusler.

The mean number of cells sampled per condition was 59.8 by Haeusler in contrast to 48.5 in the manual cell tracking experiments described here. The effect size of the survival deficit is large.

Given the greater number of *C9ORF72* mutant lines and the increased number of technical replicates presented here the work of Haeusler at al. does not decrease my confidence in the concluding that iPSC-derived MNs from C9-MND patients do not have increased vulnerability to ER stress. I would speculate the differences observed most likely be due to line variation independent of *C9ORF72* mutation status which could be missed using only one iPSC line or, alternatively, an effect of ageing. MND is an adult-onset disease and the older MN cultures of Haeusler at al. have an advantage in that we might hope, for the purposes of accurate disease modelling, not to observe these pathophysiological phenomena in 'young' neurons. Having said this, the difference in age is not large (though important developmental events do happen over a short space of time) and with the higher doses used here we might hope to observe a phenotype in younger cells because of the greater degree of stress applied.

Fitting abnormal protein homeostasis into the bigger picture of MND

A summary of key findings is shown in **Table 22**. With the analysis of UPR stress sensors, tunicamycin survival assay, TDP-43 and DPR status of control, mutant and gene corrected lines we can consider how protein homeostasis may fit into the bigger picture of MND. The picture is complex and I do not believe given the available evidence a strong conclusion can be made regarding the relationship between the *C9ORF72* HRE and UPR – repetition of all the biochemical experiments is required with the addition of lines that were missing so that data on all eight lines are available. This is essential because there is heterogeneity among the *C9ORF72* mutants.

C9-1 shows alterations in of HSPB8 transcript and PERK function, but not HSP90B1 or ATF6. C9-2 shows evidence not only of ER-stress as determined by PERK and ATF6 cleavage. Data regarding the HSP90B1 status of C9-2 has not been acquired – this shall be sought. C9-3 shows upregulation of HSP90B1 and PERK. Of the three *C9ORF72* mutants, C9-2 was the only one not to show a pro-survival phenotype at 24h in the mnID survival analysis. C9-3 showed evidence of HSPB8 and HSP90B1 upregulation with PERK changes but data regarding the ATF6 data of this line is missing. This heterogeneity is in keeping with the hypothesis that MND is a multistep disease. Major genetic risk factors alone are insufficient to cause the disease – as evidenced by the fact that a *C9ORF72* HRE is present around 0.1% of the healthy population (Murphy et al. 2017).

Line	HSP90B1	HSPB8	XPO1	PERK	ATF6	XBP1	Survival	Insoluble TDP-43	DPR Protein *Aggregates*
C9-1	Normal	Upregulated	Normal	Exaggerated response to tunicamycin	Normal	Normal	May delay death	None detected	Increased
C9-Δ1	Normal	Normal	Normal	Rescue	Normal	Normal	No deficit	None detected	Rescued
C9-2	-	-	-	Exaggerated response to tunicamycin	Basal stress associated with HRE. Exaggerated response to tunicamycin	Normal	No deficit	None detected	-
C9-Δ2	-	-	-	Rescue	Partial rescue	Normal	No deficit	None detected	-
C9-3	Upregulated	Upregulated	Normal	Exaggerated response to tunicamycin	-	Normal	May delay death	None detected	Increased
C9-Δ3	-	-	-	Rescue	-	Normal	No deficit	None detected	-
G298S	-	-	-	No basal stress	Basal stress	Normal under basal conditions	-	Present	-
G298S-Δ1	-	-	-	No basal stress	Basal stress	-	-	Present	-
G298S-Δ1	-	-	-	No basal stress	Basal stress	-	-	Present	-
M337V	-	-	-	-	-	-	-	Present	-

Table 22: Summary of experimental findings. Factors not yet assessed are denoted by a hyphen. Data on DPR protein aggregates were obtained by Dr B. T. Selvaraj and is shown in **Supplemental Figure 6**.

Heterogeneity in C9-MND iPSC-derived sMN is also evident in the literature. Dafinca et al. have recently reported abnormalities in ER calcium handling and stress granule formation (Dafinca et al. 2016). Using six lines derived from three patients with C9-MND, it was demonstrated that on treatment with thapsigargin C9-MND lines retain Ca^{2+}. Elevated Ca^{2+} within the ER has been associated with ER stress sensor activation and the authors go on to investigate the level of GRP78. Of the three lines examined, only one shows significantly upregulated GRP78 supporting the findings here of heterogeneity amount C9-MND lines. Interestingly, the pair of lines associated with the greatest retention of ER Ca^{2+} was not the pair in which GRP78 was significantly upregulated. It is also reassuring to note that in this data values of GRP78 normalised to β-actin vary little between lines derived from the same patient, suggesting this heterogeneity is a genuine function the patient genomes and not an artefact of reprogramming or cell culture.

These results lend further support for a multistep hypothesis of MND. Such a model may explain clinical phenomena such as the clinical variability of MND within families (Siddique et al. 1991) and marked alterations in genetic contributions in different populations (Veldink 2017), as well as differences in cellular pathophysiology between patients as is described here. Studying these diseases in models with a limited number of control or disease lines may result in the generation of false dichotomies when in reality different individuals have MND because of different 'hits' which converge to cause MND. It seems likely because of the features shared among motor neuron diseases that these hits converge in a pathway or combination of pathways uniquely important to MNs. A simplified model is shown in **Figure 30**.

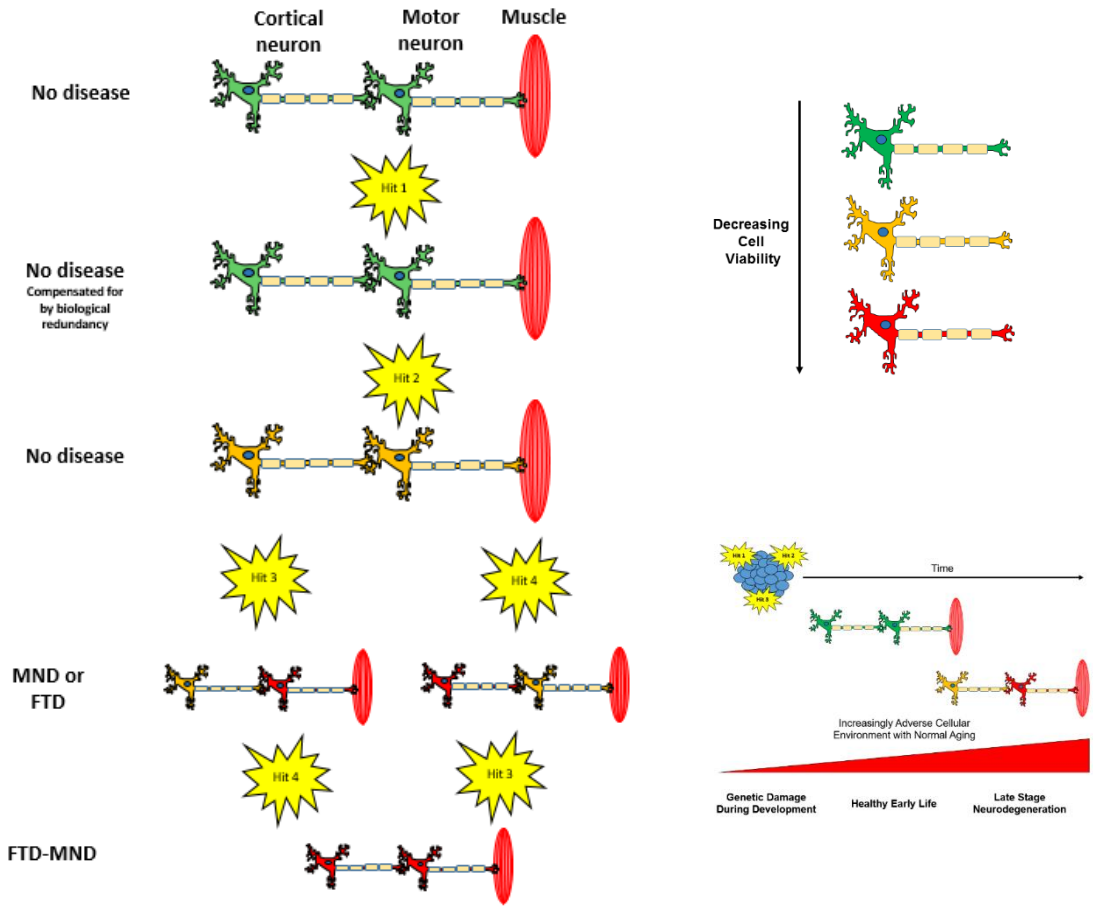


Figure 30: Model hypothesis by which multiple hits result in the development of MND, FTD or both in an individual. From left to right graphics show cortical neurons, motor neurons and muscle. Degeneration of motor neurons results in MND. Degeneration of cortical neurons results in FTD. The neuronal viability is shown in colour with degeneration from green (healthy) through orange to red (non-viable). Hits 1 to 4 are added in series but in this simplified scenario could occur in any order. The order may be important *in vivo*, for example, KO of a DNA repair gene A followed by subsequent mutagenesis of gene B may be more detrimental than if gene B was damaged and subject to DNA repair mechanisms before KO of DNA repair gene A. The timing of these hits is also important. For hits acquired in a large number of somatic cells, it is most conceivable that DNA damage occurred early in development in a pluripotent cell affecting all subsequent offspring (*bottom right*). Age-related factors or subsequent environmental hits transform this developmental susceptibility to adult-onset disease. An alternative scenario is that hits are acquired throughout life but this requires a mechanism by which a neuronal subpopulation can become selectively vulnerable. A major issue when seeking to identify necessary and sufficient hits is biological redundancy – demonstrated here by the lack of discernible effect on the system following Hit 1. Assuming Hit 1 is genetic - whole genome sequencing (WGS) of the system at this time though may lead to the identification of evidence a pathological hit has been obtained in the absence of a phenotype.

The data obtained from the *TARDBP*^{G298S} line similarly suggests a complex disease process with at least some 'classical' features of MND including ER stress (as determined by increased ATF6 cleavage) and the presence of insoluble TDP-43 are independent of the G298S mutation. The ultimate test of the sufficiency of the G298S mutation would be to insert it into a healthy control line and investigate for signs of pathology. This could be achieved using a strategy very similar to the *TARDBP*^{M337V} gene correction described in *Chapter 5*. If the pathological aspects of MND are again not evident it will suggest that something beyond the G298S mutation is producing a pathological signature.

Deficits in how cells handle ER stress may constitute a possible 'hit' in the multistep process that is MND and these hits may fall within different components of the UPR leading to heterogeneity in patient lines. With this data, we are not in a position to say whether a hit to the UPR is necessary for MND to develop. This would be an interesting question to address in the future.

The power of studying a multistep disease with gene editing

Genetic engineering of mammalian cell lines offers immense potential in the dissection of diseases with complex genetics. The introduction of the CRISPR/Cas system for genetic engineering has already lead to publications in the field of MND using gene corrected control lines (Kiskinis et al. 2014; Wang et al. 2017). It is interesting that RNAseq experiments performed on these datasets so far have compared mutant cells lines to gene corrected controls only. As described above, while this lends power to the study of the specific effect of the mutation information is also lost – any other alterations in the 'juxtamutation' genome which predisposed the individuals to develop this complex genetic disease. These concerns are arguably greater for mutations like the *C9ORF72*, which has been associated with a number of different phenotypes.

Limitations

The limitations of this research can be considered as specific limitations of this work and broader limitations of iPSC disease models.

Specific Limitations

The major limitation of this work is the limited number of times certain experiments have been independently repeated. Until this is achieved, confident conclusions as to the UPR phenotype of these lines is impossible.

It will also be necessary to confirm PERK activation by examination of phosphorylation or investigation of downstream targets. The status of phospho-PERK in MND patients has not yet been reported in the literature. The fact we observe an increase in total PERK on administration though is a positive sign that a functional readout is being obtained. It should also be remembered that ER stress which has been simulated with tunicamycin here whereas ER stress in human C9-MND is likely to result from a number of different factors.

Statistical limitations regarding the assumptions for mixed linear regression modelling in the survival analysis also require consideration. Of these, right-censoring of the data is the most marked but heteroscedasticity at high mnID48 was also observed.

Model Limitations

The isolated study of sMNs has its limitations. Non-cell autonomous effects are increasingly being reported with the contributions of glial cells to the pathophysiology of MND increasingly recognised (Ilieva et al. 2009). These cultures lack cellular input from cortical neurons and cannot provide output to muscle. A lack of cortical input limits our ability to study the dying-forward hypothesis of MND which has recently been reviewed (Geevasinga et al. 2016). Briefly, the dying forward

hypothesis postulates that sMN death in MND results from dysfunction of upstream cortical neurons (Eisen et al. 1992). The key mechanism proposed currently for a dying-forward mechanism is excitotoxicity and since the publication of this review new evidence of upstream pathology has arisen. Unpublished observations by Dr B. T. Selvaraj et al. (in review) demonstrate AMPA-mediated excitotoxicity in C9-MND. This has not yet been demonstrated in a cortical-sMN co-culture model *in vitro* though without the addition of depolarising agents. The role of muscle in MND is also under investigation and advances have been made recently which allow investigation of NMJ function in health and disease in all human systems (Puttonen et al. 2015).

Conclusion

In conclusion, the status of the UPR of sMNs has been found to vary across three different C9-MND patient lines and it has been demonstrated that *C9ORF72* mutant neurons are not more vulnerable to tunicamycin-induced death in this system. Additionally, it is evident that specific arms of the UPR can be activated differentially. A common feature of all three lines was upregulation of HSPB8 transcript, known to play a role in the disassembly of protein aggregates, including TDP-43. TDP-43 pathology, however, is not present in *C9ORF72* mutant sMN cultures at this time though DPR protein inclusions are. Further replication is required before conclusions can be made regarding the status of individual ER stress sensor pathways.

In contrast, there is evidence that sMNs derived from a patient with *TARDBP*^{G298S} mutation show evidence of constitutive ER stress determined by ATF6 cleavage and insoluble TDP-43 pathology independent of the G298S mutation.

Given the association between the *C9ORF72* HRE and disruption to cell proteostasis, further dissection of these pathways will be of relevance to MND and other neurodegenerative conditions more broadly. These findings are in keeping with a multistep hypothesis of C9-MND.

Chapter 5: Correction of the *TARDBP*^{M337V} mutation in iPSCs

Background

Adult-onset ALS has been reported with the *TARDBP*^{M337V} mutation (Tamaoka et al. 2010). An adenine (A) to guanine (G) substitution at position 1009 in exon 6 of *TARDBP* results in codon 337 coding for valine instead of methionine. The Chandran lab has an iPSC line from an individual with ALS who carries the *TARDBP*^{M337V} mutation. An isogenic control for this line has not yet been generated. To demonstrate an understanding of the gene editing process I outline a strategy here for the generation of a corrected cell line using the CRISPR-Cas9 system. During my time in the lab, I assisted in gene editing projects and some of these experiences are included as notes where relevant.

In brief, the CRISPR-Cas9 system constitutes part of the bacterial acquired immune system (Barrangous et al. 2007). Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) interspaced with hypervariable sequences (protospacers) were identified following sequencing of the bacterial genome and were identified as structurally peculiar. It was not till later on that it was demonstrated that these hypervariable spacer sequences are derived from external nucleic acids such as those of bacteriophages (Barrangous et al. 2007). Three different types of CRISPR system have been described with the type II system now commonly used for gene editing in many species. The different CRISPR systems have different properties but the most commonly used system is derived from *Streptococcus pyogenes* in which the target sequence must precede 5′-NGG protospacer adjacent motif (PAM) (Ran et al. 2013). The PAM sequence is specific to different CRISPR systems.

In its essence, the CRISPR-Cas9 system in *Streptococcus pyogenes* consists of three components:

1. Cas9 – an RNA-guided nuclease which generates double-stranded breaks (DSBs)
2. crRNA – a 20nt guide sequence preceding the PAM which is complementary to the target sequence
3. tracrRNA – a nucleic acid through which the crRNA interacts with Cas9

Different genomic locations can be targeted by changing the sequence of the crRNA without modifying the other components. This feature gives the CRISPR-Cas9 system an advantage over its predecessors, zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), in which a nuclease domain is attached to a DNA-binding protein as it is much easier to design and produce the targeting crRNA sequence than it is to create customised DNA-binding proteins (Ran et al. 2013). Furthermore, crRNA and tracrRNA sequences can be subcloned as a single construct termed a single guide RNA (sgRNA) (Jinek et al. 2012) or synthesised directly following annealing of oligonucleotides. CRISPR-Cas9 genome targeting is, however, constrained in that sgRNAs must target 5′ to the specific PAM sequence – NGG for *Streptococcus pyogenes*. Practically, this poses little constraint because of the regular occurrence of these three nucleotide sequences throughout the genome. Cas9 cuts ~3bp upstream of the PAM.

The most simple form of gene editing using CRISPR-Cas9 requires the introduction of two constructs within the cell – the Cas9 and the sgRNA. Cas9 can be introduced either as a protein or as plasmid expressing Cas9 cDNA under a mammalian promoter. Although early protocols used Cas9-expressing plasmids to synthesis Cas9 *de novo* more recent studies show more favourable kinetics with protein transfection. The reduced half-life of Cas9 following protein transfection compared to plasmid transfection has been reported to reduce off-target effects (Liang et al. 2015) with

comparable HDR frequency (Sentmanat and Miller)[A]. Introduction of Cas9 and sgRNA into a cell will result in the introduction of double-stranded DNA (dsDNA) breaks at target sites. Unmodified Cas9 cleaves DNA with blunt ends and DNA repair mechanisms within the cell then reconstitute cleaved DNA.

When DSBs are introduced in DNA two principle repair pathways exist – non-homologous end joining (NHEJ) or homology-directed repair (HDR) (Kanaar et al. 2008). Of these, HDR is a more advanced process in which a homology repair template (HRT) informs the product of the DNA repair machinery. The editing of *TARDBP*^{M337V} will require a strategy compatible with HDR. In cases where HDR is required the repair template can be introduced either in a plasmid or as single-stranded oligodeoxynucleotides (ssODNs). ssODNs can be easily synthesised using the phosphoramidite method – the major constraint being the maximum length of an oligonucleotide that can be sequenced. In this project, a relatively short (<100bp) HRT is required and thus a ssODN repair template has been used.

Briefly, NHEJ has found application for several functions. Some utility comes from the error-prone nature of NHEJ which can lead to random mutations including insertions, deletions (in-dels) at the site of DSB. This strategy has been used to create knockout (KO) models in which in-dels lead to the formation of truncated products through direct synthesis of a stop codon or more commonly shifting the reading frame. Pairs of sgRNAs have also been used to excise portions of genomes and this strategy was adopted for the generation of the *C9ORF72* gene corrected controls (see *Methods*). The factors that influence whether a cell undergoes NHEJ or HDR are of importance as in both cases the efficiency of genome editing in mammalian cells is low. Factors such as the type of nuclease, guide orientation, target locus and cell type all affect the efficiency of the various different gene editing pathways (Miyaoka et al. 2016). Interestingly, the addition of non-homologous DNA during transfection increases in-del rate so even in genome editing strategies in which HDR is not

required it may be advisable to include additional oligonucleotides in the transfection mix (Richardson et al. 2016).

An overview of a strategy by which the *TARDBP*^{M337V} correction could be performed is now provided. Following this, the specific constructs required are detailed.

Strategy Overview

Homology-directed repair (HDR) is required to correct the A → G missense mutation at position c1009 of *TARDBP* resulting in a substitution of methionine to valine at codon 337 (pM337V). The mutation lies within exon 6 which can be amplified via PCR as described in *Methods*. This amplified region can then be subject to Sanger sequencing (Source Biosciences). Sequencing the target genome served both to confirm the nature of the A to G missense mutation at codon 337 as well as identifying variants in the surrounding sequences. Failure to identify such variants could reduce the efficiency of targeting constructs and was thus preferable to relying on reference sequences.

An online tool by the Zhang lab allows identification of possible crRNA sequences (http://crispr.mit.edu/). The advantage of this approach over manual identification of PAM sites (5′-NGG-3′ in the case of *Streptococcus pyogenes*) near the mutation and gRNAs is that it allows rapid comparison of many guides with information on their likely off-target effects. If specificity cannot be achieved with a crRNA a nickase system can be considered. If a silent mutation can be introduced into the PAM this is ideal as these changes can be incorporated into the HRT to prevent targeting of the repair template by Cas9. These mutations must be silent as they will also be introduced to the target genome. If this is not the case silent mutations will need to be inserted into the region of the HRT corresponding to the seed region of the sgRNA – the 12 nucleotides 5′ of the PAM. With this in mind, the HRT can then be designed.

The HRT must not be targeted by Cas9 so mutations must either be introduced to the PAM or seed as mentioned above, these mutations will be transferred to the target genome though so should be silent. In practice, this is easy to achieve because of redundancy in the amino acid code. More challenging is configuring these silent mutations so that a restriction enzyme cut site is created or destroyed allowing screening via restriction fragment length polymorphism (RFLP) analysis. Ideally, the banding pattern expected should be easy to resolve and interpret when used for RFLP screening. Synthesis of the HRT for this specific edit is reasonably straightforward as the short length of the HRT permits direct industrial synthesis. Consideration must then be given to the delivery system. Cas9 is delivered as a protein here because of the associated benefit of reduction of off-target cutting and increased efficiency. The sgRNA here is synthesised by annealing together two oligonucleotides – the custom designed crRNA with an additional 3' direct repeat sequence and the standard tracrRNA scaffold. The transfection will be carried out by nucleofection. Initially, the efficiency of sgRNA targeting should be assessed via the T7 endonuclease assay. If the cutting efficiency is low an alternative crRNA should be tried. If sufficient, the targeting experiment can be performed with Cas9, gRNA and HRT. Following transfection and recovery of iPSCs, cells are re-plated down at a low density in a 10cm dish to allow clonal selection. Following clonal selection colonies should be expanded and passaged with one sample used for screening, in this case using RFLP, and the other maintained so that in clones in which screening is promising cultures can be expanded and sequenced for confirmation.

Construct Design

A 452bp amplicon of *TARDBP* exon 6 of a *TARDBP*^{M337V} line was sequenced as described in *Methods.* This served to both confirm the nature of the A to G missense mutation at codon 337 as well as identifying variants in the surrounding sequences. Failure to identify such variants could reduce the efficiency of targeting constructs and was thus preferable to relying on reference sequences.

PAM sites surrounding codon 337 were identified and are shown in yellow in **Figure 30**.

Figure 30: Figure showing PAM sites surrounding *TARDBP* codon 337. Codon 337 is marked with a grey box. PAM (NGG) sites are shown in yellow.

The closest possible PAM site at which a silent mutation could be introduced was over 30bp from codon 337. At this distance, HDR is likely to be very inefficient. As the HRT would require mutations within the seed region there was no known disadvantage to choosing the closest possible PAM to the target site – a TGG sequence starting within codon 337 at position c1010. Following identification of this PAM, a 20bp sequence upstream (5'-ACAGAGCAGTTGGGGTATGA-3'**)** was screened (crRNA-1) to identify potential off-target effects using the Zhang lab online CRISPR design tool (http://crispr.mit.edu/). The ten highest scoring off-target sites are shown in **Table 23**.

Score	Mismatches	UCSC gene	Locus
31.5	1MMs [19]		chr8:-64049192
2.3	3MMs [2:3:5]		chr12:-131021031
1.5	3MMs [3:4:9]		chr1:-177306588
1.2	3MMs [1:10:17]		chr18:-74392586
1.1	3MMs [1:5:19]		chr12:-125358101
0.6	4MMs [2:3:10:20]		chr3:+124103400
0.6	3MMs [1:10:18]		chr8:-93078163
0.6	4MMs [1:7:8:9]		chr5:+31676555
0.6	4MMs [1:5:10:17	NM_001406	chr17:-7612097
0.5	3MMs [3:17:20]		chr4:+176460365

Table 23: Ten highest off-target site scores for sgRNA-1.

Given the lack of affinity for crRNA-1 for most other genomic targets with no protein-coding genes affected with less than four mismatches, the sequence appears a good candidate. Had off-target scores been higher, a dual nickase strategy could have been considered. Shifting the sgRNA upstream just one basepair resulted in a markedly less favourable off target screen (data not shown).

With a theoretically acceptable crRNA, consideration was then given to the design of the HRT. An HRT with 40bp homology arms flanking the target nucleotide at position 1009 was designed (**Table 24**).

M337V	5′- GGCTGCCGCCCAGGCAGCACTACAGAGCAGTTGGGGTATGGTG GGCATGTTAGCCAGCCAGCAGAACCAGTCAGGCCCATCGG-3′
HRT	5′- GGCTGCCGCCCAGGCAGCACTACAGAGCAGcTGG|GGAATCAT GGGCATGTTAGCCAGCCAGCAGAACCAGTCAGGCCCATCGG′-3′

Table 24: Table showing the endogenous M337V and HRT sequences. Codon 337 is denoted by bold black. Position 1009 is underlined. The PAM is highlighted yellow. The seed is highlighted in blue. Silent mutations are shown in bold red. The silent mutation responsible for the generation of a PspFI RE digest site is shown in lower case red with the cut site itself denoted by '|'.

Three silent mutations shown in red in **Figure 24** have been introduced with the HRT in the gRNA binding sequence in addition to the A → G correction in position c1009. Of the silent mutations, T → C at position c998 creates a PspFI cut site. The recognition sequence of these restriction enzymes is shown in **Figure 31**.

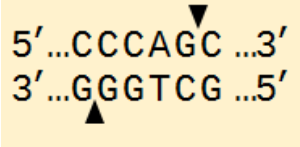


Figure 31: Diagram showing the recognition sequence of the PspFI restriction enzymes. Arrows denote the cut site which produces a sticky end with a 3bp overhang.

Introduction of PspFI restriction site is ideal since within the exon 6 PCR amplicon this enzyme will only cut once within lines in which HDR has successfully occurred and produces fragments with over a 100bp difference in size which will be easy to detect via gel electrophoresis during the screening stage.

In this case, the repair template can be obtained relatively cheaply from industry suppliers (Sigma Aldrich or IDT as just two examples). This is because constructs are within lengths obtainable with reasonable yield via the phosphoramidite method of oligonucleotide synthesis.

Using the same resource from Sigma-Aldrich used to design primers earlier in this thesis the HDR template can be synthesised. The following information is required:

Parameter	Options
Synthesis scale	0.05µM 0.2µM 1µM 10µM 15µM Larger
Purification	Desalt Cartridge HPLC PAGE
Format	Dry Solution (in water) Solution (in TE) Solution (in TRIS)

Table 25: Table showing decisions required for the ordering of the HRT from Sigma-Aldrich.

The Liu protocol recommends 15µg of Cas9 protein be incubated with 7.5-10ug of sgRNA at room temperature for 10 minutes to allow the formation of ribonucleoproteins (RNPs). This represents a very similar volume of Cas9 protein (16µg) recommended in a white paper by Lonza who use 200pmol sgRNA (Sentmanat and Miller)[A]. Transfection efficiency is reported to decrease if the volume of RNP solution exceeds 10% of the transfection volume.

Consideration is now given to the manner in which these constructs will be delivered. Cas9 protein will be used instead of a Cas9 encoding plasmid for the advantages described in methods. The ssODN repair template can be transfected directly as oligonucleotides. The crRNA and tracrRNA can be ordered directly relatively inexpensively for industrial suppliers such as Integrated DNA Technologies. The necessary scaffold is automatically added to the crRNA sequence with this approach

and this strategy is cost efficient for labs in which editing is infrequent. These companies also provide protocols describing the assembly of ribonucleoprotein (RNP) complexes before transfection.

Transfection and Determination of Cutting Efficiency

Constructs can be introduced into target cells in a number of ways. Electroporation and lipofection are commonly used[B]. I would attempt electroporation first for several reasons, all anecdotal. Firstly, I have some experience of electroporation following nucleofection of iPSCs in a BAC-transposon system. Secondly, previous success has been had by the Chandran lab on iPSCs using the CA137 protocol of the Lonza 4D-Nucleofector and discussion at the Inaugural MRC Stem Cell Partnership Workshop in Edinburgh (06/06/2017-07/04/2017) confirmed that other groups were not only using the Lonza 4D-Nucleofector but also the very same protocol. Thirdly, a representative from Merck themselves at the Edinburgh University CRISPR/Cas9 Workshop (16/05/2017) confirmed that they use Lonza's (a competitor company) transfection systems for gene editing experiments. While there is a paucity of quality evidence for the advantages of electroporation over lipofection it is important to acknowledge that wisdom from those in the field is valuable and provides the potential to save time and money.

The nucleofection of human embryonic stem cells (hESCs) has been recently described by Liu et al. (Liu et al. 2015). Following dissociation into single cells using a procedure identical to that that used at the start of the sMN generation protocol described in *Methods*. It is recommended by Liu that mTeSR1 medium is added to cells to stop the accutase digestion but previous success in our lab has not found this necessary. It is noted though that in the case of a cell line in which a large degree of cell death during transfection occurs this presents a modifiable step which may improve survival. Addition of the ssODN repair template to this mixture is also

required and concentrations used in the published literature vary at least from 0.5 – 10μM for each microliter of nucleofection reaction mix volume.

Following transfection, iPSCs should be plated at a density of 200,000 cells/well in a MG coated six-well plate in E8 Media[C]. The E8 media should be supplemented with 10μM Rock Inhibitor.

Incubation at 37°C for several days allows recovery and upon reaching approximately 80% confluence cells are dissociated with accutase and plated at a low density (~5000 cells/10cm dish) to facilitate the picking of clones[D,E]. Before clonal colonies are picked it is advisory to make some assessment as to the CRISPR-Cas9 activity and targeting efficiency of the RNP complex. CRISPR-Cas9 activity can be assessed by T7 endonuclease assay on leftover transfected cells following the platedown to 10cm dishes. In brief, the T7 endonuclease assay is based on the principle that T7 endonuclease specifically cleaves the heteroduplex DNA structures. DNA should be extracted from transfected cells and exon 6 amplified using previously described protocols (see *Methods*). The PCR amplicon is denatured and re-annealed resulting in the formation of heteroduplex structures if mutations are introduced due to the error-prone NHEJ after Cas9 nuclease activity. Sites of mutagenesis are cleaved by T7 endonuclease producing a banding pattern that can be resolved by agarose gel electrophoresis. In this case, in addition to the 452bp PCR amplicon, two large bands of ~290 and ~165bp would be expected with a number of much smaller cleavage products. If this banding pattern is not observed, it is likely the RNP complex has poor cutting efficiency and continuation with clonal screening is likely to fail. A new RNP should be designed or transfection parameters optimised.

Even though the efficiency is observed to be < 15 % it is encouraged to check if the HDR *per-se* is successful by performing PspF1 RFLP analysis on the transfected cells prior to colonies are picked. An additional band at ~288bp, if observed, indicates successful gene targeting.

Clonal Selection and Screening

Assuming a positive T7 endonuclease activity and PspF1 RFLP results, the clonal colonies are picked using pipette tips and a microscope in sterile conditions. Previous work in the lab suggests around 500 clones should be picked but it is acknowledged that this efficiency varies from experiment to experiment. Clonal colonies are plated in individual wells of a MG-coated 96-well plate and grown in E8 (Essential 8™ Medium, ThermoFisher Scientific, A1517001) with RI. Following further culture for several days in these plates cells should be passaged with 0.5mM EDTA to generate at least two replicates of each plate. One of these will be subject to RFLP screening and the other maintained in culture so that candidate clones with promising RFLP results can be selected.

Remaining cells should be subject to RFLP analysis with PspFI to determine if transfection has been successful. It should be noted that the primers used to generate the amplicon subject to RFLP must lie out with the HRT as is the case here. Failure to recognise this can result in amplification of the repair template itself and/or random targeting events leading to false-positive screening. Successful HDR targeting can be either be mono-allelic or bi-allelic, RFLP analysis with PspF1 will be able to determine both the scenarios. Clones with mono-allelic targeting will have ~450bp, 288bp and 165bp bands whereas clones with bi-allelic targeting event will only have 288bp and 165bp band in on an agarose gel. If RFLP analysis with PspFI does not demonstrate the expected dual banding pattern (~288bp and ~165bp) the transfection is highly unlikely to have been successful and transfection should be reattempted with altered parameters.

In clones with promising RFLP results sequencing of exon 6 can then be undertaken to confirm whether gene correction (pM337M) has been successful. In case of clones with a mono-allelic targeting event, it is important to sequence both the alleles using

TA cloning to confirm the untargeted allele has not acquired other mutations due to the Cas9 nuclease activity.

Searching for Off-Target Effects

Attempts should then be made to identify off-target effects. One approach is via whole genome sequencing (WGS) though this is time-consuming, expensive and coverage may be a concern. An alternative approach is selectively sequencing genomic regions in which off-target cuts might be expected using bioinformatics techniques such as that described above by the Zhang lab. WGS has the added advantaged of detecting not only CRISPR-Cas9 off-target effects but also other sporadic mutations which may have occurred throughout the gene editing process.

It may be advisable to also culture non-edited clones that have been subject to the entire CRISPR-Cas9 gene editing process. It has been argued that such cell lines may provide a more appropriate control than the original parent line. In reality, information is likely to be gained from a comparison between all lines available.

Chapter 5: Correction of the TARDBP^{M337V} Mutations Notes

A – Not peer reviewed.
B – Within the field there is anecdotal evidence that lipofection has poor efficiency for motor neuron transfection (though here iPSCs are transfected).
C – MG is not the only option here and we have also assessed the survival of post-electroporation iPSCs on laminin 521. In my hand's survival is better on MG-coated plates though in difficult to transfect lines alternative approaches could be tried.
D – Accutase is used rather than trypsin or collagenase to yield single cells allowing truly clonal colony selection.
E – I have had some experience picking clones during the generation of C9-Δ3.

References

Al-Chalabi A, Calvo A, Chio A, Colville S, Ellis CM, Hardiman O, et al. Analysis of amyotrophic lateral sclerosis as a multistep process: A population-based modelling study. Lancet Neurol [Internet]. Elsevier Ltd; 2014;13(11):1108–13. Available from: http://dx.doi.org/10.1016/S1474-4422(14)70219-4

Arai T, Hasegawa M, Akiyama H, Ikeda K, Nonaka T, Mori H, et al. TDP-43 is a component of ubiquitin-positive tau-negative inclusions in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Biochem Biophys Res Commun. 2006;351(3):602–11.

Arnold ES, Ling S-C, Huelga SC, Lagier-Tourenne C, Polymenidou M, Ditsworth D, et al. ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. Proc Natl Acad Sci [Internet]. 2013;110(8):E736-45. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=23382207&retmode=ref&cmd=prlinks%5Cnpapers2://publication/doi/10.1073/pnas.1222809110

Aronica E, Baas F, Iyer A, ten Asbroek ALMA, Morello G, Cavallaro S. Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. Neurobiol Dis [Internet]. Elsevier B.V.; 2015;74:359–76. Available from: http://dx.doi.org/10.1016/j.nbd.2014.12.002

Atkin JD, Farg MA, Walker AK, McLean C, Tomas D, Horne MK. Endoplasmic reticulum stress and induction of the unfolded protein response in human sporadic amyotrophic lateral sclerosis. Neurobiol Dis [Internet]. Elsevier Inc.; 2008;30(3):400–7. Available from: http://dx.doi.org/10.1016/j.nbd.2008.02.009

Barmada SJ, Skibinski G, Korb E, Rao EJ, Wu JY, Finkbeiner S. Cytoplasmic mislocalization of TDP-43 is toxic to neurons and enhanced by a mutation associated with familial ALS. J Neurosci. 2010;30(2):639–65.

Barrangous R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR Provides Acquired Resistance against Viruses in Prokaryotes. Science (80-). 2007;315(5819):1709–12.

Ben-Shushan E, Feldman E, Reubinoff BE. Notch signalling regulates motor neuron differentiation of human embryonic stem cells. Stem Cells [Internet]. 2015;33(2):403–15. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25335858

Bertolotti A, Zhang Y, Hendershot LM, Harding HP, Ron D. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nat Cell Biol. 2000;2(6):326–32.

Bruch J, Xu H, Rösler TW, De Andrade A, Kuhn P, Lichtenthaler SF, et al. PERK activation mitigates tau pathology *in vitro* and *in vivo*. EMBO Mol Med [Internet]. 2017;e201606664. Available from: http://embomolmed.embopress.org/lookup/doi/10.15252/emmm.201606664

Celli J, Tsolis RM. Bacteria, the endoplasmic reticulum and the unfolded protein response: friends or foes? Nat Rev Microbiol [Internet]. Nature Publishing Group;

2014;13(2):71–82. Available from: http://www.nature.com/doifinder/10.1038/nrmicro3393

Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, et al. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. Ann Neurol. 2013;74(2):180–7.

Crippa V, Cicardi ME, Ramesh N, Seguin SJ, Ganassi M, Bigi I, et al. The chaperone HSPB8 reduces the accumulation of truncated TDP-43 species in cells and protects against TDP-43-mediated toxicity. Hum Mol Genet. 2016;25(18):3908–24.

Dafinca R, Scaber J, Ababneh N, Lalic T, Weir G, Christian H, et al. C9orf72 Hexanucleotide Expansions are Associated with Altered Endoplasmic Reticulum Calcium Homeostasis and Stress Granule Formation in Induced Pluripotent Stem Cell-Derived Neurons from Patients with Amyotrophic Lateral Sclerosi. Stem Cells. 2016;34:2063–78.

Davidson YS, Barker H, Robinson AC, Thompson JC, Harris J, Troakes C, et al. Brain distribution of dipeptide repeat proteins in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. Acta Neuropathol Commun [Internet]. England: Mann,David M. Clinical and Cognitive Sciences Research Group, Institute of Brain, Behaviour and Mental Health, Faculty of Medical and Human Sciences, University of Manchester, Salford Royal Hospital, Salford M6 8HD, UK. david.mann@manchester.ac.uk.; 2014;2:70. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24950788

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron [Internet]. United States: DeJesus-Hernandez, Mariely. Department of Neuroscience, Mayo Clinic Florida, Jacksonville, FL 32224, USA.; 2011;72(2):245–56. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=21944778

Dejesus-Hernandez M, Finch NA, Wang X, Gendron TF, Kevin ·, Bieniek F, et al. In-depth clinico-pathological examination of RNA foci in a large cohort of C9ORF72 expansion carriers. Acta Neuropathol. Springer Berlin Heidelberg; 2017;

Di XJ, Wang YJ, Han DY, Fu YL, Duerfeldt AS, Blagg BSJ, et al. Grp94 protein delivers γ-aminobutyric acid type A (GABAA) receptors to Hrd1 protein-mediated endoplasmic reticulumassociated degradation. J Biol Chem. 2016;291(18):9526–39.

van Dis V, Kuijpers M, Haasdijk ED, Teuling E, Oakes SA, Hoogenraad CC. Golgi fragmentation precedes neuromuscular denervation and is associated with endosome abnormalities in SOD1-ALS mouse motor neurons. Acta Neuropathol Commun [Internet]. England: Jaarsma, Dick. Department of Neuroscience, Erasmus Medical Center, P,O,Box 2040, 3000 CA Rotterdam, The Netherlands. d.jaarsma@erasmusmc.nl.; 2014;2:38. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24708899

Dobson CM. Protein folding and misfolding. Nature. 2003;426(6968):884–90.

Dobson CM, Ellis RJ. Protein folding and misfolding inside and outside the cell. EMBO J. 1998;17(18):5251–4.

Donnelly CJ, Zhang PW, Pham JT, Heusler AR, Mistry NA, Vidensky S, et al. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. Neuron [Internet]. Elsevier; 2013;80(2):415–28. Available from: http://dx.doi.org/10.1016/j.neuron.2013.10.015

Echeverria G V., Cooper TA. RNA-binding proteins in microsatellite expansion disorders: Mediators of RNA toxicity. Brain Res [Internet]. Elsevier B.V.; 2012;1462:100–11. Available from: http://dx.doi.org/10.1016/j.brainres.2012.02.030

Eisen A, Kim S, Pant B. Amyotrophic lateral sclerosis (ALS): A phylogenetic disease of the corticomotoneuron? Muscle Nerve. 1992;15(2):219–24.

Esanov R, Belle KC, van Blitterswijk M, Belzil V V., Rademakers R, Dickson DW, et al. C9orf72 promoter hypermethylation is reduced while hydroxymethylation is acquired during reprogramming of ALS patient cells. Exp Neurol [Internet]. Elsevier Inc.; 2016;277:171–7. Available from: http://dx.doi.org/10.1016/j.expneurol.2015.12.022

Fauvet B, Mbefo MK, Fares MB, Desobry C, Michael S, Ardah MT, et al. α-Synuclein in central nervous system and from erythrocytes, mammalian cells, and Escherichia coli exists predominantly as disordered monomer. J Biol Chem. 2012;287(19):15345–64.

Ferraiuolo L, Heath PR, Holden H, Kasher P, Kirby J, Shaw PJ. Microarray Analysis of the Cellular Pathways Involved in the Adaptation to and Progression of Motor Neuron Injury in the SOD1 G93A Mouse Model of Familial ALS. J Neurosci [Internet]. 2007;27(34):9201–19. Available from: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1470-07.2007

Fornerod M, Ohno M, Yoshida M, Mattaj IW. CRM1 is an export receptor for leucine-rich nuclear export signals. Cell. 1997;90(6):1051–60.

Fratta F, Poulter M, Lashley T, Rohrer JD, Polke JM, Beck J, et al. Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. Acta Neuropathol [Internet]. Germany: Fratta,Pietro. Department of Neurodegenerative Disease, UCL Institute of Neurology, Queen Square, London, WC1N 3BG, UK. p.fratta@prion.ucl.ac.uk; 2013;126(3):401–9. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23818065

Fu SC, Huang HC, Horton P, Juan HF. ValidNESs: A database of validated leucine-rich nuclear export signals. Nucleic Acids Res. 2013;41(D1):338–43.

Geevasinga N, Menon P, Özdinler PH, Kiernan MC, Vucic S. Pathophysiological and diagnostic implications of cortical dysfunction in ALS. Nat Rev Neurol [Internet]. Nature Publishing Group; 2016;12(11):651–61. Available from: http://www.nature.com/doifinder/10.1038/nrneurol.2016.140

Gitler AD, Tsuiji H. There has been an awakening: Emerging mechanisms of C9orf72 mutations in FTD/ALS. Brain Res [Internet]. Elsevier; 2016;1647:19–29. Available from: http://dx.doi.org/10.1016/j.brainres.2016.04.004

Gomez-Deza J, Lee Y-B, Troakes C, Nolan M, Al-Sarraj S, Gallo J-M, et al. Dipeptide repeat

protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. Acta Neuropathol Commun [Internet]. Acta Neuropathologica Communications; 2015;3(1):38. Available from: http://dx.doi.org/10.1186/s40478-015-0218-y%5Cnpapers2://publication/doi/10.1186/s40478-015-0218-y%5Cnhttp://actaneurocomms.biomedcentral.com/articles/10.1186/s40478-015-0218-y

Gonatas NK, Stieber a, Mourelatos Z, Chen Y, Gonatas JO, Appel SH, et al. Fragmentation of the Golgi apparatus of motor neurons in amyotrophic lateral sclerosis. Am J Pathol [Internet]. 1992;140(3):731–7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1886164&tool=pmcentrez&rendertype=abstract

Groen EJN, van Vught PWJ, van Es MA, Blauw HM, Veldink JH, van den Berg LH. Gene expression profile of SOD1-G93A mouse spinal cord, blood and muscle. Amyotroph Lateral Scler Frontotemporal Degener [Internet]. England: Saris,Christiaan G J. Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Center Utrecht, Utrecht, The Netherlands.; 2013;14(3):190–8. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23298163

Haass C, Selkoe DJ. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid beta-peptide. Nat Rev Mol Cell Biol [Internet]. 2007;8(2):101–12. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17245412

Haeusler AR, Donnelly CJ, Periz G, Simko E a J, Shaw PG, Kim M-S, et al. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. Nature [Internet]. England: Haeusler,Aaron R. 1] Department of Biochemistry and Molecular Biology, Johns Hopkins University Baltimore, Maryland 21205, USA [2] Department of Neuroscience, Johns Hopkins University Baltimore, Maryland 21205, USA.; 2014;507(7491):195–200. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24598541

Haeusler AR, Donnelly CJ, Rothstein JD. The expanding biology of the C9orf72 nucleotide repeat expansion in neurodegenerative disease. Nat Rev Neurosci [Internet]. Nature Publishing Group; 2016;17(6):383–95. Available from: http://www.nature.com/doifinder/10.1038/nrn.2016.38

Hancock MK, Kopp L, Kaur N, Hanson BJ. A Facile Method for Simultaneously Measuring Neuronal Cell Viability and Neurite Outgrowth. Curr Chem Genomics Transl Med [Internet]. 2015;9(1):6–16. Available from: http://benthamopen.com/ABSTRACT/CCGTM-9-6

Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, et al. Regulated Translation Initiation Controls Stress-Induced Gene Expression in Mammalian Cells. Mol Cell [Internet]. 2000;6(5):1099–108. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1097276500001088

Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature [Internet]. 1999;397(6716):271–4. Available from: http://www.nature.com/doifinder/10.1038/16729

Haze K, Yoshida H, Yanagi H, Yura T, Mori K. Mammalian Transcription Factor ATF6 Is

Synthesized as a Transmembrane Protein and Activated by Proteolysis in Response to Endoplasmic Reticulum Stress. Mol Biol Cell [Internet]. 1999;10(11):3787–99. Available from: http://www.molbiolcell.org/cgi/doi/10.1091/mbc.10.11.3787

Heath PR, Kirby J, Shaw PJ. Investigating cell death mechanisms in amyotrophic lateral sclerosis using transcriptomics. Front Cell Neurosci [Internet]. 2013;7(December):1–10. Available from: http://journal.frontiersin.org/article/10.3389/fncel.2013.00259/abstract

Hensman DJ, Poulter M, Beck J, Hehir J, Polke JM, Campbell T, et al. C9orf72 expansions are the most common genetic cause of Huntington disease phenocopies. Neurology. 2014;82(4):292–9.

Hetz C, Thielen P, Matus S, Nassif M, Court F, Kiffin R, et al. XBP-1 deficiency in the nervous system protects against amyotrophic lateral sclerosis by increasing autophagy. Genes Dev. 2009;23(19):2294–306.

Ilieva H, Polymenidou M, Cleveland DW. Non-cell autonomous toxicity in neurodegenerative disorders: ALS and beyond. J Cell Biol. 2009;187(6):761–72.

Ingre C, Roos PM, Piehl F, Kamel F, Fang F. Risk factors for amyotrophic lateral sclerosis. Clin Epidemiol. 2015;7:181–93.

Irobi J, Almeida-Souza L, Asselbergh B, De Winter V, Goethals S, Dierick I, et al. Mutant HSPB8 causes motor neuron-specific neurite degeneration. Hum Mol Genet [Internet]. 2010;19(16):3254–65. Available from: http://www.hmg.oxfordjournals.org/cgi/doi/10.1093/hmg/ddq234

Ji J, Ng SH, Sharma V, Neculai D, Hussein S, Sam M, et al. Elevated Coding Mutation Rate During the Reprogramming of Human Somatic Cells into Induced Pluripotent Stem Cells. Stem Cells [Internet]. 2012;30(3):435–40. Available from: http://doi.wiley.com/10.1002/stem.1011

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science (80-). 2012;337(August):816–22.

Kanaar R, Wyman C, Rothstein R. Quality control of DNA break metabolism: in the “end”, it’s a good thing. EMBO J [Internet]. 2008;27(4):581–8. Available from: http://emboj.embopress.org/cgi/doi/10.1038/emboj.2008.11

Kiskinis E, Sandoe J, Williams LA, Boulting GL, Moccia R, Wainger BJ, et al. Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. Cell Stem Cell. 2014;14(6):781–95.

Koppers M, Blokhuis AM, Westeneng HJ, Terpstra ML, Zundel CAC, Vieira De Sá R, et al. C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. Ann Neurol. 2015;78(3):426–38.

Kostic VS, Dobricic V, Stankovic I, Rali?? V, Stefanova E. C9orf72 expansion as a possible genetic cause of Huntington disease phenocopy syndrome. J Neurol. 2014;261(10):1917–21.

Kwon M-J, Choi WJ, Oh K-W, Koh S-H, Ki C-S, Kim SH. Analysis of the C9orf72

hexanucleotide repeat expansion in Korean patients with familial and sporadic amyotrophic lateral sclerosis. Neurobiol Aging [Internet]. United States: Jang,Ja-Hyun. Department of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea.; 2013;34(4):1311.e7-9. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=23088937

Lagier-Tourenne C, Baughn M, Rigo F, Sun S, Liu P, Li H-R, et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. Proc Natl Acad Sci [Internet]. 2013;110(47):E4530–9. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1318835110

Leblond CS, Kaneb HM, Dion P a, Rouleau G a. Dissection of genetic factors associated with amyotrophic lateral sclerosis. Exp Neurol [Internet]. United States: Leblond,Claire S. Department of Human Genetics, McGill University, Montreal, Qc, Canada; Montreal Neurological Institute and Hospital, McGill University, Montreal, Qc, Canada; Department of Neurology and Neurosurgery, McGill University, Montreal, Qc, Cana; 2014;262 Pt B:91–101. Available from: http://www.sciencedirect.com/science/article/pii/S0014488614001150

Leggett C, McGehee DS, Mastrianni J, Yang W, Brorson JR. Tunicamycin produces TDP-43 cytoplasmic inclusions in cultured brain organotypic slices. J Neurol Sci. 2012;317(1–2):66–73.

Li Z, Vuong JK, Zhang M, Stork C, Zheng S. Inhibition of nonsense-mediated RNA decay by ER stress. RNA [Internet]. 2016;rna.058040.116. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27940503

Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, et al. Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol [Internet]. Elsevier B.V.; 2015;208:44–53. Available from: http://dx.doi.org/10.1016/j.jbiotec.2015.04.024

Liu EY, Russ J, Wu K, Neal D, Suh E, McNally AG, et al. C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. Acta Neuropathol. 2014;128(4):525–41.

Liu J, Gaj T, Yang Y, Wang N, Shui S, Kim S, et al. Efficient delivery of nuclease proteins for genome editing in human stem cells and primary cells. Nat Protoc [Internet]. Nature Publishing Group; 2015;10(11):1842–59. Available from: http://www.nature.com/doifinder/10.1038/nprot.2015.117

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25:402–8.

Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. Proc Natl Acad Sci [Internet]. 1999;96(11):6307–11. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.96.11.6307

Mackenzie IRA, Bigio EH, Ince PG, Geser F, Neumann M, Cairns NJ, et al. Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral

sclerosis with SOD1 mutations. Ann Neurol. 2007;61(5):427–34.

MacKenzie IR, Arzberger T, Kremmer E, Troost D, Lorenzl S, Mori K, et al. Dipeptide repeat protein pathology in C9ORF72 mutation cases: Clinico-pathological correlations. Acta Neuropathol. 2013;126(6):859–79.

Majounie E, Abramzon Y, Renton A, Perry R, Bassett S, Pletnikova O, et al. Repeat Expansion in C9ORF72 in Alzheimer ' s Disease. N Engl J Med. 2012;366(3):283–4.

Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, Schymick JC, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron [Internet]. United States: Renton,Alan E. Neuromuscular Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD 20892, USA.; 2011;72(2):257–68. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=21944779

Mao W, Fukuoka S, Iwai C, Liu J, Sharma VK, Sheu S-S, et al. Cardiomyocyte apoptosis in autoimmune cardiomyopathy: mediated via endoplasmic reticulum stress and exaggerated by norepinephrine. Am J Physiol Heart Circ Physiol [Internet]. 2007;293(3):H1636-45. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17545481

Maury Y, Côme J, Piskorowski RA, Salah-Mohellibi N, Chevaleyre V, Peschanski M, et al. Combinatorial analysis of developmental cues efficiently converts human pluripotent stem cells into multiple neuronal subtypes. Nat Biotechnol [Internet]. 2014;33(1):89–96. Available from: http://dx.doi.org/10.1038/nbt.3049

Maximino JR, de Oliveira GP, Alves CJ, Chadi G. Deregulated expression of cytoskeleton related genes in the spinal cord and sciatic nerve of presymptomatic SOD1(G93A) Amyotrophic Lateral Sclerosis mouse model. Front Cell Neurosci [Internet]. 2014;8(May):148. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4033281&tool=pmcentrez&rendertype=abstract

McMillan CT, Russ J, Wood EM, Irwin DJ, Grossman M, Mccluskey L, et al. C9orf72 promoter hypermethylation is neuroprotective: Neuroimaging and neuropathologic evidence. Neurology. 2015;84:1622–30.

Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. Nat Protoc. 2013;8(8):1551–66.

Miyaoka Y, Berman JR, Cooper SB, Mayerl SJ, Chan AH, Zhang B, et al. Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. Sci Rep [Internet]. Nature Publishing Group; 2016;6(1):23549. Available from: http://www.nature.com/articles/srep23549

Mulder D, Kurland L, Otford K. Familial adult motor neuron disease: amyotrophic lateral sclerosis. Neurology. 1986;36:511–7.

Murphy NA, Arthur KC, Tienari PJ, Houlden H, Chiò A, Traynor BJ. Age-related penetrance of the C9orf72 repeat expansion. 2017;(February):1–7. Available from: http://download.springer.com/static/pdf/629/art%253A10.1038%252Fs41598-017-

02364-1.pdf?originUrl=http%3A%2F%2Fwww.nature.com%2Farticle%2Fs41598-017-02364-1&token2=exp=1495268212~acl=%2Fstatic%2Fpdf%2F629%2Fart%25253A10.1038%25252Fs41598-017-02364-1.pdf*~h

Neumann M, Sampathu DM, Kwong L, Truax AC, Micsenyi MC, Chou TT, et al. Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Science (80-). 2006;314(5796):130–3.

Nishitoh H. CHOP is a multifunctional transcription factor in the ER stress response. J Biochem. 2012;151(3):217–9.

Noristani HN, Sabourin JC, Gerber YN, Teigell M, Sommacal A, dM Vivanco M, et al. Brca1 is expressed in human microglia and is dysregulated in human and animal model of ALS. Mol Neurodegener [Internet]. Molecular Neurodegeneration; 2015;10(1):34. Available from: http://www.molecularneurodegeneration.com/content/10/1/34

Novoa I, Zeng H, Harding HP, Ron D. Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2alpha. J Cell Biol. 2001;153(5):1011–22.

O'Rourke JG, Meera P, Muhammad AKMG, Grant S, Simpkinson M, Bell S, et al. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. Sci Transl Med [Internet]. United States: Sareen,Dhruv. Regenerative Medicine Institute, Cedars-Sinai Medical Center, 8700 Beverly Boulevard, Los Angeles, CA 90048, USA.; 2013;5(208):208ra149. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=24154603

de Oliveira GP, Alves CJ, Chadi G. Early gene expression changes in spinal cord from SOD1G93A Amyotrophic Lateral Sclerosis animal model. Front Cell Neurosci [Internet]. 2013;7(November):1–17. Available from: http://journal.frontiersin.org/article/10.3389/fncel.2013.00216/abstract

Oyanagi K, Yamazaki M, Takahashi H, Watabe K, Wada M, Komori T, et al. Spinal anterior horn cells in sporadic amyotrophic lateral sclerosis show ribosomal detachment from, and cisternal distention of the rough endoplasmic reticulum. Neuropathol Appl Neurobiol. 2008;34(6):650–8.

Patin F, Baranek T, Vourc'h P, Nadal-Desbarats L, Goossens JF, Marouillat S, et al. Combined Metabolomics and Transcriptomics Approaches to Assess the IL-6 Blockade as a Therapeutic of ALS: Deleterious Alteration of Lipid Metabolism. Neurotherapeutics [Internet]. Neurotherapeutics; 2016;13(4):905–17. Available from: http://dx.doi.org/10.1007/s13311-016-0461-3

Peters OM, Cabrera GT, Tran H, Gendron TF, Jeanne E, Metterville J, et al. Expression of human C9ORF72 hexanucleotide repeat expansion reproduces RNA foci and dipeptide repeat proteins but not neurodegeneration in BAC transgenic mice. Neuron. 2015;88(5):902–9.

Puttonen KA, Ruponen M, Naumenko N, Hovatta OH, Tavi P, Koistinaho J. Generation of Functional Neuromuscular Junctions from Human Pluripotent Stem Cell Lines. Front Cell Neurosci [Internet]. 2015;9(December):1–9. Available from:

http://journal.frontiersin.org/Article/10.3389/fncel.2015.00473/abstract

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc [Internet]. 2013;8(11):2281–308. Available from: http://www.nature.com/doifinder/10.1038/nprot.2013.143

Richardson CD, Ray GJ, Bray NL, Corn JE. Non-homologous DNA increases gene disruption efficiency by altering DNA repair outcomes. Nat Commun [Internet]. Nature Publishing Group; 2016;7:12463. Available from: http://www.nature.com/doifinder/10.1038/ncomms12463

Riva N, Clarelli F, Domi T, Cerri F, Gallia F, Trimarco A, et al. Unraveling gene expression profiles in peripheral motor nerve from amyotrophic lateral sclerosis patients: insights into pathogenesis. Sci Rep [Internet]. 2016;6(December):39297. Available from: http://www.nature.com/articles/srep39297

Russ J, Liu EY, Wu K, Neal D, Suh ER, Irwin DJ, et al. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. Acta Neuropathol. 2015;129(1):39–52.

Samali a, Fitzgerald U, Deegan S, Gupta S. Methods for monitoring endoplasmic reticulum stress and the unfolded protein response. Int J Cell Biol [Internet]. 2010;2010:830307. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20169136

Sentmanat M, Miller S. Efficient CRISPR/Cas9 Delivery Using Nucleofector(TM) Technology: Comparison of Plasmid and RNP-based Editing, [Internet]. Available from: http://bio.lonza.com/uploads/tx_mwaxmarketingmaterial/Lonza_WhitePapers_Efficient_CRISPR_Cas9_Delivery_Using_Nucleofector_Technology_Comparison_of_Plasmid-_and_RNP-based_Editing.pdf

Shen J, Chen X, Hendershot L, Prywes R. ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of golgi localization signals. Dev Cell. 2002;3(1):99–111.

Siddique T, Figlewicz DA, Pericak-Vance MA, Haines JL, Rouleau G, Jeffers AJ, et al. Linkage of a gene causing familial amyotrophic lateral sclerosis to Chromosome 21 and evidence of genetic-locus heterogenity. N Engl J Med. 1991;324(20):1381–4.

Siddique T, Pericak-Vance M. Linkage analysis in familial amyotrophic lateral sclerosis. Neurology. 1989;39:919–26.

Soldner F, Jaenisch R. iPSC Disease Modeling. Science (80-) [Internet]. 2012;338(6111):1155–6. Available from: http://www.sciencemag.org/cgi/doi/10.1126/science.1227682

Sposito T, Preza E, Mahoney CJ, Set??-Salvia N, Ryan NS, Morris HR, et al. Developmental regulation of tau splicing is disrupted in stem cell-derived neurons from frontotemporal dementia patients with the 10 + 16 splice-site mutation in MAPT. Hum Mol Genet. 2015;24(18):5260–9.

Štalekar M, Yin X, Rebolj K, Darovic S, Troakes C, Mayr M, et al. Proteomic analyses reveal that loss of TDP-43 affects RNA processing and intracellular transport. Neuroscience [Internet]. IBRO; 2015;293:157–70. Available from: http://dx.doi.org/10.1016/j.neuroscience.2015.02.046

Tamaoka A, Arai M, Itokawa M, Arai T, Hasegawa M, Tsuchiya K, et al. TDP-43 M337V mutation in familial amyotrophic lateral sclerosis in Japan. Intern Med [Internet]. 2010;49(4):331–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/20154440

Tang BS, Zhao G hua, Luo W, Xia K, Cai F, Pan Q, et al. Small heat-shock protein 22 mutated in autosomal dominant Charcot-Marie-Tooth disease type 2L. Hum Genet. 2005;116(3):222–4.

Therrien M, Rouleau GA, Dion PA, Parker JA. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in C. elegans. PLoS One. 2013;8(12):1–10.

Tong J, Huang C, Bi F, Wu Q, Huang B, Zhou H. XBP1 depletion precedes ubiquitin aggregation and Golgi fragmentation in TDP-43 transgenic rats. J Neurochem. 2012;123(3):406–16.

Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. Science (80-) [Internet]. 2009;323(5918):1208–11. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4516382&tool=pmcentrez&rendertype=abstract

Vaskova E. ., Stekleneva AE, Medvedev SP, Zakian SM. "Epigenetic Memory" Phenomenon in Induced Pluripotent Stem Cells. Acta Naturae [Internet]. 2013;5(19):15–21. Available from: http://dx.doi.org/10.1038/nature09342

Veldink JH. ALS genetic epidemiology "How simplex is the genetic epidemiology of ALS?" J Neurol Neurosurg Psychiatry [Internet]. 2017;jnnp-2016-315469. Available from: http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2016-315469

Vidal RL, Figueroa A, Court FA, Thielen P, Molina C, Wirth C, et al. Targeting the UPR transcription factor XBP1 protects against Huntington's disease through the regulation of FoxO1 and autophagy. Hum Mol Genet. 2012;21(10):2245–62.

Wang L, Popko B, Roos RP. The unfolded protein response in familial amyotrophic lateral sclerosis. Hum Mol Genet [Internet]. England: Wang,Lijun. Department of Neurology/MC2030, The University of Chicago Pritzker School of Medicine, 5841 S. Maryland Avenue, Chicago, IL 60637, USA.; 2011;20(5):1008–15. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=medl&NEWS=N&AN=21159797

Wang L, Yi F, Fu L, Yang J, Wang S, Wang Z, et al. CRISPR/Cas9-mediated targeted gene correction in amyotrophic lateral sclerosis patient iPSCs. Protein Cell. Higher Education Press; 2017;8(5):1–14.

Wen X, Tan W, Westergard T, Krishnamurthy K, Shi Y, Lin S, et al. Antisense Proline-Arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. Neuron. 2014;84(6):1213–25.

Wilkinson B, Gilbert HF. Protein disulfide isomerase. Biochim Biophys Acta - Proteins Proteomics [Internet]. 2004;1699(1–2):35–44. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1570963904000639

Wu C-Y, Whye D, Glazewski L, Choe L, Kerr D, Lee KH, et al. Proteomic assessment of a cell model of spinal muscular atrophy. BMC Neurosci [Internet]. 2011;12(1):25. Available

from: http://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-12-25

Yoshida H, Haze K, Yanagi H, Yura T, Mori K. Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins. Involement of basic leucin zipper transcription factors. J Biol Cell [Internet]. 1998;273(50):33741–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9837962

Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell. 2001;107(7):881–91.

Yoshida H, Okada T, Haze K, Yanagi H, Yura T, Negishi M, et al. ATF6 Activated by Proteolysis Binds in the Presence of NF-Y (CBF) Directly to the cis-Acting Element Responsible for the Mammalian Unfolded Protein Response. Mol Cell Biol [Internet]. 2000;20(18):6755–67. Available from: http://mcb.asm.org/cgi/doi/10.1128/MCB.20.18.6755-6767.2000

Zhang K, Donnelly CJ, Haeusler AR, Grima JC, Machamer JB, Steinwald P, et al. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. Nature [Internet]. 2015;525(7567):56–61. Available from: http://www.nature.com/doifinder/10.1038/nature14973

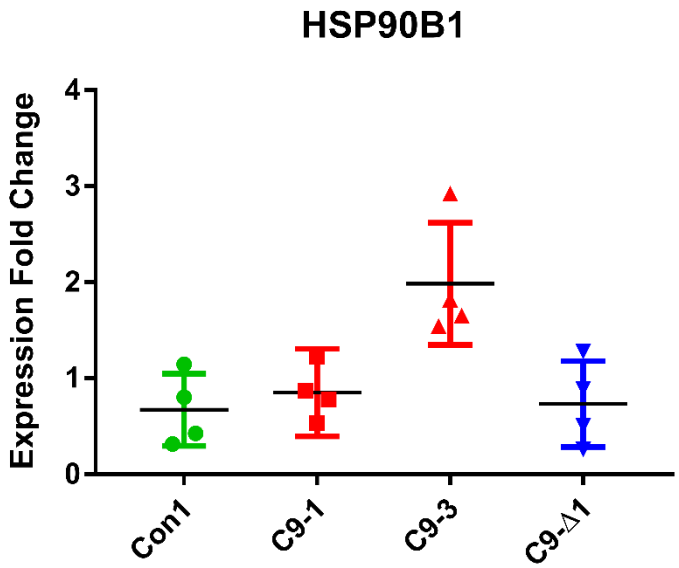
Zhang Y, Gendron TF, Grima JC, Sasaguri H, Xu Y, Katzman RB, et al. C9ORF72 poly(GA) aggregates sequester and impair H23 and nucleocytoplasmic transport proteins. Nat Neurosci. 2016;19(5):668–77.

Zhou H-X. Influences of Crowded Cellular Environments on Protein Folding, Binding, and Oligomerization: Biological Consequences and Potentials of Atomistic Modeling. FEBS Lett. 2013;8(578):1053–61.

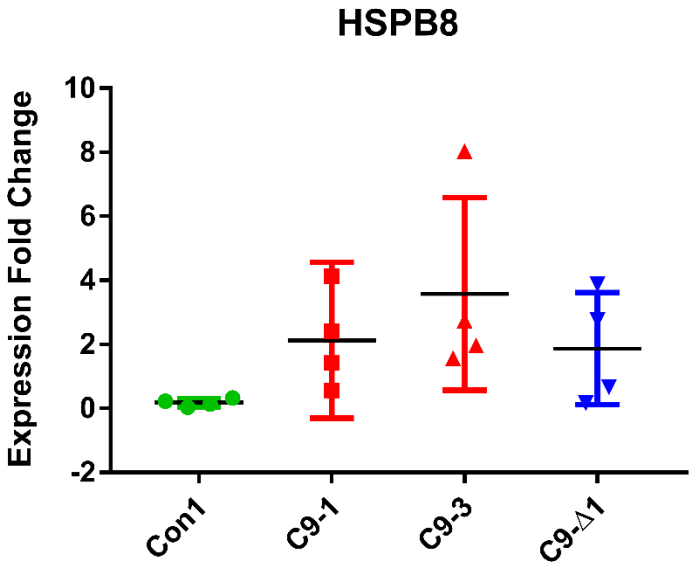
Zou Z-Y, Zhou Z-R, Che C-H, Liu C-Y, He R-L, Huang H-P. Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis. J Neurol Neurosurg Psychiatry [Internet]. 2017;jnnp-2016-315018. Available from: http://jnnp.bmj.com/lookup/doi/10.1136/jnnp-2016-315018

Supplemental Figures

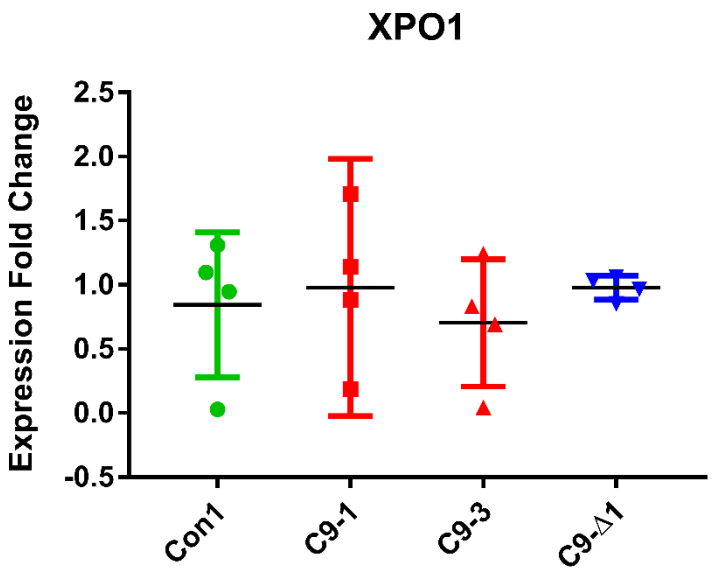
Supplemental Figure 1: HSP90B1 transcript levels in three-week-old sMNs are increased in C9-3 but consistent between other lines examined. Expression fold change of HSP90B1 transcript in control (green), *C9ORF72* mutant (red) and gene-edited (blue) MN RNA. Each point represents the mean of three technical replicates from RNA samples from independent sMN conversions. Black bars mark the mean. Error bars show the 95% CI. Kruskal-Wallis refuted the null hypothesis with (p=0.0151). Dunn's multiple comparisons test was performed and results are shown in **Supplementary Table 4**.



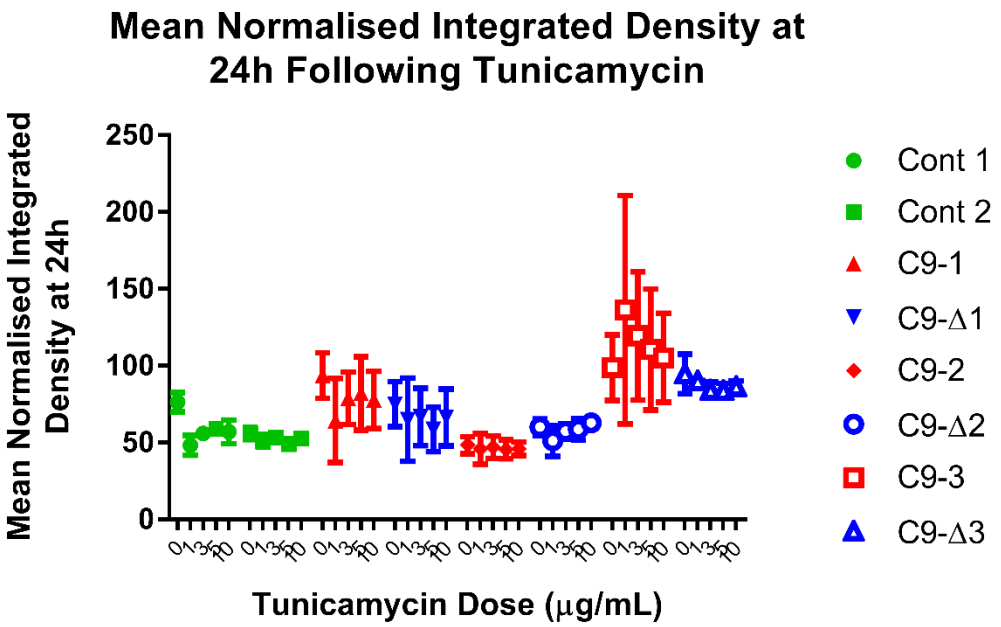
Supplemental Figure 2: Without outlier exclusion, significant upregulation of HSPB8 was observed only in C9-3. Expression fold change of HSPB8 transcript in control (green), *C9ORF72* mutant (red) and gene-edited (blue) MN RNA. Each point represents the mean of three technical replicates from RNA samples from independent sMN conversions. Black bars mark the mean. Error bars show the 95% CI. Kruskal-Wallis refuted the null hypothesis with (p=0.0364). Dunn's multiple comparisons test was performed and results are shown in **Supplemental Table 5**.



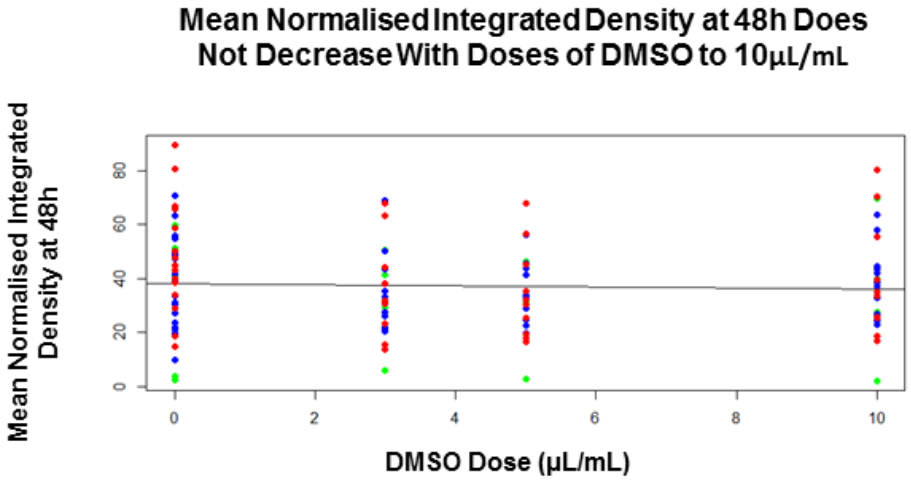
Supplemental Figure 3: XPO1 transcript is not differentially expressed among control, *C9ORF72* mutant and gene corrected lines. Expression fold change of XPO1 transcript in control (green), *C9ORF72* mutant (red) and gene-edited (blue) MN RNA. Each point represents the mean of three technical replicates from RNA samples from independent sMN conversions. Black bars mark the mean. Error bars show the 95% CI. Kruskal-Wallis failed to refute the null hypothesis with (p=0.7728).



Supplemental Figure 4: Full y axis plot of Figure 21C.



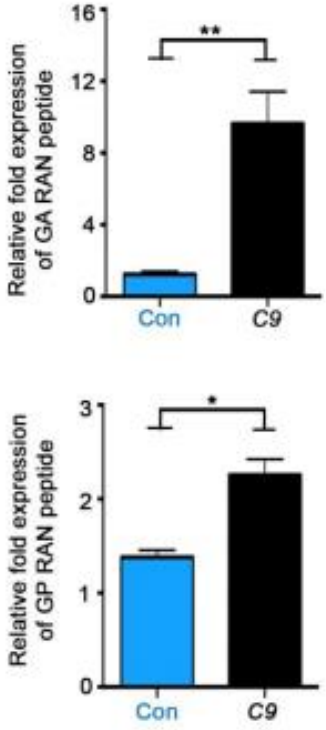
Supplemental Figure 5: Plot showing that mean normalised integrated density (mnID) does not decrease significantly with increasing doses of DMSO up to 10gL/mL. Each point denotes the mnID for one well. The colour of data points corresponds to genotype: green = control, red = mutant and blue = gene corrected. Data is clustered by platedown (dimension not displayed) which has been accounted for in the subsequent statistical analysis.



The effect of DMSO on cell survival up to the highest concentration used (0.1%) was investigated this time using mnID as a survival surrogate. The mean normalised integrated density (mnID) (a surrogate for viability) at 48h was investigated as the independent variable and its relationship with increasing volumes of DMSO is shown below. A volume of 10µL DMSO corresponds to 0.1% DMSO – the highest dose used in any experiment. The results of the likelihood ratio test are shown in **Supplemental Table 6**.

Supplemental Figure 5 suggests that increasing doses of DMSO are not associated with decreased sMN viability as determined by mnID and this is reflected in the statistical analysis (**Supplemental Table 6**).

Supplemental Figure 6: GA and GP DPR protein aggregates are present at higher levels in a *C9ORF72* mutant line relative to control lines in three-week-old sMN cultures. Data acquisition, analysis and display by Dr B. T. Selvaraj.



Supplemental Tables

Supplemental Table 1: Non-protein coding RNAseq hits not subject to GO analysis.

Non-Protein Coding RNAseq hits	
Upregulated genes	**Downregulated genes**
RP11-436D23	THRA1/BTR
RP11-436D23	RP11-395G23
RP11-247C2	PTGES3P1
FLJ12825	RP11-20D14
ZBTB11-AS1	DRAXINP1
	RP11-369C8
	RP11-219G17
	KCNJ2-AS1
	RP11-909N17
	RP13-977J11
	RP11-277P12.20
	ATP5F1P5
	CTD-2201-E18
	RP11-83N9.5

Supplemental Table 2: Delta-delta Ct values for all qPCR replicates with outliers removed.

Delta-delta Ct values for data in which outliers have been removed					
Transcript	**Cell line**	**Replicate 1**	**Replicate 2**	**Replicate 3**	**Replicate 4**
HSP90B1	Con2	0	0	0	0
	Con1	1.236306	0.319898	-0.195069072	1.04324
	C9-1	NA	0.35825	-0.2888794944351	0.027117
	C9-3	NA	-0.72484	-0.862288496	-0.62477
	C9-Δ	0.990026	0.169191	-0.35358019	2.60853
HSPB8	Con2	0	0	0	0
	Con1	2.056279	1.960721	2.856852395	2.135936
	C9-1	-2.04378	-1.87567	-1.271945704	-0.49843
	C9-3	-3.00577	-1.41333	-0.971130567	-1.01401
	C9-Δ1	-1.46253	NA	0.584017102	0.925905
XPO1	Con2	0	0	0	0
	Con1	-0.38903	0.875346	-0.131471104	0.08098
	C9-1	-0.44948	0.026703	-0.771250647	-0.18782
	C9-2	-0.31302	NA	0.264785287	0.531315
	C9-Δ1	-0.08703	NA	-0.044589753	0.231451
Number of outliers removed		8	12	0	5

Supplemental Table 3: Delta-delta Ct values for all qPCR replicates.

Delta-delta Ct values for all data					
Transcript	**Cell line**	**Replicate 1**	**Replicate 2**	**Replicate 3**	**Replicate 4**
HSP90B1	Con2	0	0	0	0
	Con1	1.236306	0.319898	-0.195069072	1.661096
	C9-1	0.200944	0.35825	-0.2888794944351	0.910776
	C9-3	-1.54676	-0.72484	-0.862288496	-0.62477
	C9-Δ1	0.990026	0.169191	-0.35358019	1.951242
HSPB8	Con2	0	0	0	0
	Con1	1.615841	4.800074	2.856852395	2.135936
	C9-1	-2.04378	0.829417	-1.271945704	-0.49843
	C9-3	-3.00577	-0.64785	-0.971130567	-1.45578
	C9-Δ1	-1.46253	-1.95477	0.584017102	2.603072
XPO1	Con2	0	0	0	0
	Con1	-0.38903	5.163168	-0.131471104	0.08098
	C9-1	0.179312	2.417937	-0.771250647	-0.18782
	C9-2	-0.31302	4.511819	0.264785287	0.531315
	C9-Δ1	-0.08703	0.049069	-0.044589753	0.231451

Supplemental Table 4: Results from Dunn's multiple comparisons testing post-Kruskal-Wallis test on HSP90B1 qPCR ΔΔCt values with outliers not removed. Interpretation of HSP90B1 transcript expression does therefore not vary with outlier removal. In both cases a statistically significant ($\alpha = 0.05$) upregulation was observed in C9-2 only.

Comparison	**P value**
Con1 vs. C9-1	>0.9999
Con1 vs. C9-3	0.0225
Con1 vs. C9-Δ1	>0.9999

Supplemental Table 5: Results from Dunn's multiple comparisons testing post-Kruskal-Wallis test on HSPB8 qPCR ΔΔCt values with outliers not removed. Interpretation of HSPB8 transcript expression varies outlier removal. Without outlier removal, a significant upregulation of HSPB8 is observed in C9-2 only.

Comparison	**P value**
Con1 vs. C9-1	0.1128
Con1 vs. C9-3	0.0280
Con1 vs. C9-Δ1	0.1901

Supplemental Table 6: Output from a mixed linear regression model and likelihood ratio test investigating the relationship between 48 mnID with dose and genotype as fixed effects and platedown modelled as a random effect. Construction of a model testing the null hypothesis allowed calculation of the likelihood ratio test p value which rejects the null hypothesis that DMSO does influences viability as determined by 48h mnID.

Independent Variable	Coefficient	Standard Error	T value	Likelihood ratio test
48h mnID	-0.297	4.81840	0.092	P = 0.9266