Stem cell models of *C9orf72*-linked Frontotemporal Dementia

Elisavet Preza

2020

A thesis submitted to the University College London for the degree of Doctor of Philosophy

Department of Neurodegenerative Disease

Queen Square Institute of Neurology

University College London

Declaration

I, Elisavet Preza, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: Elisavet Preza

Date: 02/12/2020

Statement of contribution

I have performed all procedures and experiments in this Thesis with the following contributions:

- Chapter 3: Dr Roberto Simone performed the RNA Fluorescence in situ hybridization (FISH) on iPSCs and cortical neurons for the *C9orf72* RNA expansion. Dr Thomas Moens performed the southern blot for the *C9orf72* repeat sizing on the iPSC and cortical neuron samples. Javier Gilbert Jaramillo and Mica Clarke performed the Meso Scale Discovery (MSD) immunoassays for the cortical neurons.
- Chapter 4: Bisulfite sequencing (BSP) and $(G_4C_2)_n$-methylation assay were performed in collaboration with Dr Ming Zhang and Prof Ekaterina Rogaeva at University of Toronto.

Publications arising from this thesis

Lines, G., Casey, J. M., Preza, E. & Wray, S. 2020. Modelling Frontotemporal Dementia using patient-derived induced pluripotent stem cells. *Mol Cell Neurosci***,** 103553.

Simone, R., Balendra, R., Moens, T. G., Preza, E., Wilson, K. M.*, et al.* 2018. G-quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *EMBO Molecular Medicine,* 10**,** 22-31.

Preza, E., Hardy, J., Warner, T. & Wray, S. 2016. Review: Induced pluripotent stem cell models of frontotemporal dementia. *Neuropathol Appl Neurobiol,* 42**,** 497-520.

Abstract

The GGGGCC repeat expansion in *C9orf72* is the most common genetic cause of frontotemporal dementia and amyotrophic lateral sclerosis (ALS). Expanded repeat-associated toxicity from either RNA foci or dipeptide protein repeats (DPRs) as well as a loss of C9orf72 function due to haploinsufficiency, have all been described as potential pathogenic mechanisms in *C9orf72*-linked FTD/ALS (C9-FTD/ALS). Moreover, the C9orf72 protein has been suggested to play an important role in autophagy and lysosomal trafficking.

The work presented in this thesis includes the investigation of FTD-related phenotypes and *C9orf72* haploinsufficiency in *C9orf72* patient iPSC-derived cortical neurons (iPSC-CNs), and the generation and study of *C9orf72* knockout iPSC-CNs. *C9orf72* patient iPSC-CNs exhibited: (i) sense RNA foci, (ii) poly-GP and poly-GR, (iii) *C9orf72* promoter hypermethylation, (iv) reduction in *C9orf72* total mRNA levels and (v) increased expression of intron1-retaining transcript, compared to control iPSC-CNs, with considerable variability between *C9orf72* lines. No reduction in C9orf72 protein levels was detected in *C9orf72* iPSC-CNs at 150 and 260 days *in vitro*. CRISPR/Cas9-mediated knockout of C9orf72 resulted in significant downregulation of autophagy-related genes *ULK1*, *LC3B* and *LAMP1* in *C9orf72*$^{-/-}$ iPSC-CNs compared to isogenic *C9orf72*$^{+/-}$ and/or WT iPSC-CNs, indicative of altered basal autophagy. Moreover, preliminary findings suggest that induction of mitophagy is impaired in *C9orf72*$^{-/-}$ iPSC-CNs. Future work will be required to replicate this finding and further examine the potential role of C9orf72 in autophagy and mitophagy in iPSC-CNs and how its loss may contribute to C9-FTD/ALS.

Impact Statement

The work presented in this thesis has contributed to a growing number of iPSC-modelling studies on *C9orf72*-linked FTD/ALS (C9-FTD/ALS). Importantly, the present work has focused on providing insight into the C9-FTD/ALS-related pathologies in iPSC-derived cortical neurons (iPSC-CNs) which are under-represented in published studies, the majority of which use iPSC-derived motor neuron (iPSC-MN) models. Part of this work has contributed to a scientific publication (Simone *et al.*, 2018). Moreover, apart from the investigation of the major pathological hallmarks in iPSC-CNs, other less commonly explored disease mechanisms are explored, with a particular focus in *C9orf72* haploinsufficiency and *C9orf72* promoter methylation. Importantly, many of the experiments were carried out on mature iPSC-CNs at extended timepoints that are not conventionally assessed in published studies, and also in a temporal manner, therefore providing a useful resource for future studies. The comprehensive analysis of *C9orf72* repeat expanision size, RNA foci, *C9orf72* mRNA expression and promoter methylation, collectively, has revealed strong correlations between these variables that recapitulate the associations from patient studies. Finally, the study of C9orf72 knockout iPSC-CNs has confirmed a potential autophagy phenotype and also pointed to impaired PINK1/Parkin-dependent mitophagy as a novel pathway potentially affected by C9orf72 loss of function in iPSC-CNs. This finding will be useful in the academic research community and will provide a basis for more studies to investigate and replicate this novel phenotype.

Table of Contents

Stem cell models of *C9orf72*-linked Frontotemporal Dementia 1
Declaration 2
Statement of contribution 3
Publications arising from this thesis 4
Abstract 5
Impact Statement 6
Table of Contents 7
List of Tables 12
List of Figures 13
List of abbreviations 15
Acknowledgements 16
Chapter 1 General introduction 17
1.1. Frontotemporal dementia 18
1.1.1. The FTD syndrome 18
1.1.2. Genetics of FTD 21
1.1.3. Neuropathology of FTLD 22
1.2. The *C9orf72* $(GGGGCC)_n$ repeat expansion in FTD and ALS 25
1.2.1. Identification of the *C9orf72* hexanucleotide repeat expansion 25
1.2.2. Repeat expansion length and somatic mosaicism 26
1.2.3. *C9orf72* gene and HRE molecular properties 27
1.2.4. Neuropathology of C9-FTD/ALS 30
1.2.5. Epidemiology and clinical presentation 32
1.2.6. Disease modifiers in C9-FTD/ALS 35
1.2.7. *C9orf72* promoter hypermethylation 37
1.2.8. *C9orf72* HRE disease mechanisms 40
1.2.9. C9orf72 function 49
1.3. iPSC modelling of C9-FTD/ALS 52
1.3.1. Patient-derived iPSCs in disease modelling 52
1.3.2. Genome editing in iPSCs 54
1.3.3. Cortical neuron development 59
1.3.4. Insights from *C9orf72* iPSC models 64
1.4. Mitophagy 79

1.4.1. PINK1/Parkin-dependent mitophagy 79
1.4.2. FTD/ALS-related genes in autophagy and mitophagy pathways 83
1.5. Research aims 86
Chapter 2 Materials and Methods 87
2.1. Ethics 88
2.2. Derivation of fibroblast lines from skin biopsies 88
2.3. Pluripotent stem cell lines used in this study 89
2.4. iPSC cultures 91
2.5. Cortical neuron differentiation 92
2.6. HEK293 cell cultures 93
2.7. Genome engineering 94
2.7.1. TALENs 94
2.7.2. CRISPR/Cas9n 97
2.8. Preparation of competent *E. Coli* 100
2.9. Transformation of competent *E. Coli* 101
2.10. Cloning 101
2.10.1. Small scale plasmid DNA purification 102
2.10.2. Large scale plasmid DNA purification 103
2.11. HEK293 transfections 104
2.12. T7 endonuclease I assay 104
2.13. Pluripotent media comparison for optimal single-cell survival post nucleofection 106
2.14. Nucleofection of iPSCs 107
2.15. Cell sorting and clonal expansion of iPSCs 107
2.16. Genomic DNA extraction 108
2.16.1. Genomic DNA extraction from CRISPR/Cas9n iPSC lines (96-well format) 108
2.16.2. Genomic DNA extraction using DNeasy Blood & Tissue Kit (QIAGEN) 109
2.17. PCR screening for CRISPR/Cas9n edited iPSC lines 109
2.18. Sanger sequencing 110
2.19. Agarose gel electrophoresis 111
2.20. Gel extraction 112
2.21. RT-qPCR genetic analysis for the detection of karyotypic abnormalities in iPSCs 113
2.22. Bisulfite sequencing 114
2.23. $(G_4C_2)_n$-methylation assay 116

2.24. RNA fluorescence *in situ* hybridisation (FISH) 121
2.25. Southern blot 122
2.26. Meso Scale Discovery (MSD) immunoassay 122
2.27. Immunocytochemistry 123
2.28. Lentiviral shRNA transduction 125
2.29. Western blot 125
2.30. Mitophagy induction with Oligomycin/Antimycin A treatment 127
2.31. RNA extraction 128
2.32. RT-PCR 129
2.33. Reverse transcription quantitative PCR (RT-qPCR) 129
2.33.1. SYBR Green gene expression analysis 129
2.33.2. Taqman assays for total *C9orf72* mRNA expression analysis 130
2.33.3. LNA probes and primers for *C9orf72* transcript analysis 131
2.34. Semi-quantitative RT-PCR 131
2.35. Allelic expression imbalance 132
2.36. Statistical analysis 133
Chapter 3 Generation and characterisation of cortical projection neurons from *C9orf72* HRE carriers and controls 134
3.1. Introduction 135
3.2. Aims 137
3.3. Results 138
3.3.1. Fibroblast derivation and characterisation 138
3.3.2. Characterisation of control and patient-derived iPSC lines 142
3.3.3. Differentiation and characterisation of control and *C9orf72* cortical neurons 148
3.3.4. Characterisation of CN cultures at 150 DIV 154
3.3.5. *C9orf72* iPSCs carry a stable repeat expansion during CN differentiation 157
3.3.6. *C9orf72* iPSC-CNs contain sense RNA foci 159
3.3.7. *C9orf72* iPSC-CNs exhibit elevated poly-GP levels 161
3.4. Discussion 163
Chapter 4 Analysis of *C9orf72* temporal expression and methylation 169
4.1. Introduction 170
4.1. Aims 172
4.2. Results 173
4.2.1. Methylation of the *C9orf72* HRE in fibroblasts 173

4.2.2. Methylation of the 5' and 3' CpG islands flanking the *C9orf72* HRE in fibroblasts 175
4.2.3. Temporal methylation of the *C9orf72* HRE during CN differentiation 177
4.2.4. Temporal methylation of the 5' and 3' CpG islands flanking the *C9orf72* HRE during cortical differentiation 183
4.2.5. Assessment of *C9orf72* RNA expression 189
4.2.6. Allele expression imbalance analysis of *C9orf72* in hypermethylated *C9orf72* iPSC-CNs 198
4.2.7. Characterisation of commercial antibodies for C9orf72 detection by western blot 200
4.2.8. Investigation of temporal C9orf72 protein levels in control iPSC-CNs at extended timepoints 202
4.2.9. C9orf72 protein levels in control and *C9orf72* iPSC-CNs 204
4.2.10. Relationships between *C9orf72* promoter methylation, RNA foci, total *C9orf72* mRNA and HRE length 207
4.3. Discussion 209
Chapter 5 Generation and characterisation of *C9orf72* knockout iPSCs 216
5.1. Introduction 217
5.2. Aims 219
5.3. Results 220
5.3.1. Generation and characterisation of TALEN and CRISPR-Cas9 nucleases targeting *C9orf72* exon 2 220
5.3.2. Optimisation of the protocol for the generation of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines 224
5.3.3. Characterisation of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines 228
5.3.4. Characterisation of differentiated *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs 239
5.3.5. Analysis of expression of autophagy-associated genes in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs 243
5.3.6. Investigation of mitophagy in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs 245
5.4. Discussion 248
Chapter 6 General discussion 252
6.1. Main conclusions 253
6.2. iPSC models in C9-FTD/ALS; challenges and limitations 255
6.3. Future directions 259
Appendix A 261
Appendix B 264

References 266

List of Tables

Table 1.1 Published protocols for directed differentiation of cortical neurons (Modified from Zirra *et al.* 2016). 63
Table 2.1 Summary of pluripotent stem cell lines used in this study. 90
Table 2.2 Oligo sequences used for the generation of CRISPR/Cas9n constructs for the pair 0 and pair 17. 98
Table 2.3 Bisulfite sequencing primers. 114
Table 2.4 $(G_4C_2)_n$-methylation assay primers. 120
Table 2.5 List of primary and secondary antibodies used for immunocytochemistry. 124
Table 2.6 List of primary and secondary antibodies used in western blots. 127
Table 2.7 Information on the TaqMan assays used in this thesis. 130
Table 2.8 Sequences of primers and FAM-labelled LNA probes for the detection of the three *C9orf72* transcripts 131
Table 5.1. Top 10 potential genomic off-targets for CRISPR pair 0 sgRNAs. 234

List of Figures

Figure 1.1. Overview of the early symptoms distinguishing the four clinical variants of FTD. 20
Figure 1.2 Molecular classification of frontotemporal lobar degeneration. 24
Figure 1.3 Schematic representation of *C9orf72* gene and HRE-linked disease mechanisms. 29
Figure 1.4 DPR and RNA foci pathology in *C9orf72* HRE carriers. 31
Figure 1.5 The *C9orf72* HRE is flanked by two CpG islands. 38
Figure 1.6 Overview of genome editing technologies. 58
Figure 1.7 Summary of Parkin/PINK1-dependent mitophagy. 81
Figure 2.1 Schematic overview of TALEN hierarchical modular assembly protocol (Sanjana *et al.*, 2012). 96
Figure 2.2 Vector maps for the CRISPR/Cas9 D10A nickase plasmids. 99
Figure 2.3 Overview of the T7 endonuclease I mutation detection assay for the functional testing of CRISPR/Cas9n and TALEN nucleases. 105
Figure 2.4 Schematic illustration of the 5' CpG island upstream of the *C9orf72* HRE. 115
Figure 2.5 Schematic overview of the $(G_4C_2)_n$-methylation assay. 117
Figure 3.1 Phase contrast images of control and *C9orf72* HRE patient fibroblast lines. 139
Figure 3.2 *C9orf72* HRE patient fibroblasts express fibroblast-specific marker S100 calcium-binding protein A4 (S100A4). 140
Figure 3.3 *C9orf72* fibroblast lines express fibroblast specific markers *S100A4* and *VIM*. 141
Figure 3.4 Phase contrast images of iPSC lines. 143
Figure 3.5 Expression of the pluripotency markers OCT4 and SSEA4 in iPSCs. . 144
Figure 3.6 Expression of the pluripotency markers TRA-1-81 and NANOG in iPSCs. 145
Figure 3.7 iPSC express pluripotency markers but not lineage-specific genes..... 146
Figure 3.8 iPSC lines have normal karyotypes. 147
Figure 3.9 Overview of the Shi *et al*. 2012 CN differentiation protocol. 149
Figure 3.10 Control and *C9orf72* repeat expansion cortical rosettes at 15 DIV. ... 150
Figure 3.11 Cortical rosettes at day 15 express early forebrain marker OTX2. 151
Figure 3.12 Early-born TUJ1 positive neurons emerge at the periphery of the cortical rosettes at 15-20 DIV. 152
Figure 3.13 iPSC-CNs at 90 DIV express BRN2 marker of late cortical progenitors and upper layer (II-Vb) neurons. 153
Figure 3.14 No difference in the levels of GLAST+ astrocytes and GLAST:TUJ1 ratios in 150 DIV iPSC-CNs of *C9orf72* patients and controls. 155
Figure 3.15 C9-1 150 DIV iPSC-CN exhibit significantly upregulated *GFAP* levels. 156
Figure 3.16 *C9orf72* HRE iPSCs and differentiated iPSC-CNs carry a stable HRE. 158
Figure 3.17 Sense RNA foci are detected in *C9orf72* iPSC-CNs. 160
Figure 3.18 *C9orf72* iPSC-CNs exhibit elevated levels of poly-GP. 162
Figure 4.1 The *C9orf72* HRE is unmethylated in patient fibroblast lines. 174

Figure 4.2 The 5' and 3' CpG islands flanking the *C9orf72* HRE are not hypermethylated in patient fibroblasts. 176
Figure 4.3 The *C9orf72* HRE is methylated in patient iPSC-CNs. 179
Figure 4.4 The 5' CpG island upstream of *C9orf72* is hypermethylated in *C9orf72* HRE patient iPSC-CNs. 184
Figure 4.5 The 3' CpG island downstream of *C9orf72* is not hypermethylated in *C9orf72* HRE patient iPSC-CNs. 187
Figure 4.6 *C9orf72* total mRNA expression is reduced in C9-2 and C9-3 iPSC-CNs. 191
Figure 4.7 Schematic illustration of the semi-quantitative RT-PCR primer binding sites for the detection of the three *C9orf72* transcript variants. 193
Figure 4.8 *C9orf72* transcripts V1 and V3 are significantly upregulated in C9-1 patient 150 DIV iPSC-CNs compared to control and other patient CNs. 194
Figure 4.9 *C9orf72* V3 is significantly upgregulated in C9-1 150 DIV iPSC-CNs. . 197
Figure 4.10 CTRL-3 and C9-3 lines are heterozygous for rs10122902 and C9-3 line is heterozygous for rs17769294. 199
Figure 4.11 No significant *C9orf72* allelic expression imbalance is observed for hypermethylated C9-3 line 150 DIV iPSC-CNs. 200
Figure 4.12. C9-L isoform is detected as a single band at 50 kDa via western blot. 201
Figure 4.13. C9-L is stably expressed throughout *in vitro* cortical development. .. 203
Figure 4.14 No evidence of haploinsufficiency in *C9orf72* 150 DIV iPSC-CNs compared to controls. 205
Figure 4.15. No evidence of haploinsufficiency in *C9orf72* 260 DIV iPSC-CNs compared to controls. 206
Figure 4.16 5' CpG methylation is inversely correlated with percentage of RNA foci and *C9orf72* total mRNA levels and positively correlated with the *C9orf72* HRE size. 208
Figure 5.1 CRISPR/Cas9n and TALEN nucleases DNA binding sequences on *C9orf72* exon 2. 221
Figure 5.2 All generated nucleases successfully generate double strand breaks in *C9orf72* exon 2. 223
Figure 5.3 Preliminary data suggest mTeSR1 maintenance media offers better survival of single iPSCs post-sorting. 226
Figure 5.4 CRISPR/Cas9n nucleofection efficiencies. 227
Figure 5.5 High resolution electrophoresis screening of edited iPSCs for NHEJ-induced indels. 230
Figure 5.6 Monoallelic and biallelic frameshift indels in iPSCs. 231
Figure 5.7 *C9orf72* knockout iPSC lines display typical ESC morphology. 235
Figure 5.8 *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines maintain pluripotency. 236
Figure 5.9 Karyotype analysis of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines. 237
Figure 5.10 Confirmation of C9-L loss in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines. 238
Figure 5.11 Cortical rosettes at 25 DIV express early forebrain marker FOXG1... 240
Figure 5.12 Cortical neurons at 40 DIV express deep-layer CN marker CTIP2. ... 241
Figure 5.13 Confirmation of C9-L loss in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs. 242
Figure 5.14 Altered expression of autophagy-related genes in *C9orf72*$^{-/-}$ iPSC-CNs. 244
Figure 5.15 Impaired mitophagy in *C9orf72*$^{-/-}$ iPSC-CNs. 246

List of abbreviations

AD	Alzheimer's disease	**HEK**	Human embryonic kidney cells
ALS	Amyotrophic Lateral Sclerosis	**hnRNP**	Heterogeneous nuclear ribonucleoprotein
ASO	Antisense oligonucleotide	**HR**	Homologous recombination
ATP	Adenosine triphosphate	**HRE**	Hexanucleotide repeat expansion
bp	Base pair	**iPSCs**	Induced pluripotent stem cells
BSP	Bisulfite Sequencing PCR	**iNs**	Induced neurons
bvFTD	Behavioural variant frontotemporal dementia	**kDa**	Kilodaltons
C9-L	C9orf72 long protein isoform	**MN**	Motor neuron
C9orf72	Chromosome 9 open reading frame 72	**MSD**	Meso scale discovery
C9-S	C9orf72 short protein isoform	**MSP**	Methylation-specific PCR
Cas9n	Cas9 nickase	**NHEJ**	Non homologous end-joining
CBS	Corticobasal syndrome	**PBS**	Phosphate-buffered saline
CN	Cortical neuron	**PCR**	Polymerase chain reaction
CNS	Central nervous system	**PNK**	Polynucleotide kinase
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	**PPA**	Primary progressive aphasia
DAPI	4′,6-diamidino-2-phenylindol	**PSP**	Progressive supranuclear palsy
DIV	Days *in vitro*	**qPCR**	Quantitative PCR
DPRs	Dipeptide repeat proteins	**RAN**	Repeat-associated non-ATG
DSB	Double strand break	**rpPCR**	Repeat-primed PCR
ESCs	Embryonic stem cells	**RT**	Room temperature
EDTA	Ethylenediaminetetraacetic acid	**RT-PCR**	Reverse transcriptase PCR
FACS	Fluorescence-activated cell sorting	**sgRNA**	Single guide RNA
FBS	Fetal bovine serum	**shRNA**	Short hairpin RNA
FTD	Frontotemporal dementia	**SNP**	Single nucleotide polymorphism
FTLD	Frontotemporal lobar degeneration	**SSC**	Saline-sodium citrate
GEF	GDP/GTP exchange factor	**TALENs**	Transcription Activator-Like Effector Nucleases
GFP	Green fluorescent protein	**TBE**	Tris-borate EDTA
GTP	Guanosine-5'-triphosphate	**Wt**	Wild-type
HDR	Homology directed repair	**ZFNs**	Zinc finger nucleases

Acknowledgements

I am grateful to my supervisors Professors Selina Wray, John Hardy and Adrian Isaacs for their continuing guidance, patience and support. It has been a great priviledge knowing them, working with them and learning from them. Special thanks go to Selina Wray for being a caring supervisor, supportive boss and amazing mentor throughout the years I have been working in her lab. She has offered her support in challenging times for me and I will always be grateful.

I have had the priviledge to work with some amazing people in the Wray lab and I would like to thank all of them for their help and support as well as for all the fun moments we have shared. Special thanks to Dr Charlie Arber, Argyro Alatza, Jackie Casey, Georgie Lines and Chris Lovejoy for all the help they have generously offered throughout the years.

I would also like to thank our collaborators Dr Ming Zhang and Prof Ekaterina Rogaeva at the university of Toronto, Prof Helene Plun-Favreau, and lab members and alumni of the Isaacs lab, especially Dr Rubika Balendra, Dr Katie Wilson, Dr Thomas Moens, Mica Clarke, Javier Gilbert Jaramillo and Dr Roberto Simone, for offering their expertise, sharing their knowledge and contributing to this work.

This work would have not been possible without the participation of patients who kindly donated the skin fibroblasts. I would like to thank them as well as the National Institute for Health and Research (NIHR) for funding my work.

Finally, I would like to thank my parents Thanasis and Lina, brother Tasos, as well as my husband Christos and my beloved son Anastasis for their love and support.

Chapter 1 General introduction

1.1. Frontotemporal dementia

1.1.1. The FTD syndrome

Frontotemporal dementia is a clinically, genetically and neuropathologically heterogeneous clinical syndrome presenting as a diverse group of progressive neurodegenerative diseases affecting behaviour, cognition and language as well as certain types of movement (Neary *et al.*, 1998, McKhann *et al.*, 2001, Bird *et al.*, 2003). The common feature of these diseases is the selective degeneration of the frontal and anterior temporal lobes of the brain (Bang *et al.*, 2015). The first case of FTD was described in 1892 by Arnold Pick, in a patient who had aphasia, lobar atrophy and presenile dementia. The term “Pick’s disease” was introduced later to describe the clinical syndrome of FTD and has remained a synonym term. The term frontotemporal lobar degeneration is used to describe the pathology underlying the heterogeneous group of FTD neurodegenerative diseases.

FTD is the most common form of dementia in the presenile age group (under 65 years) and the third most common form of dementia across all age groups (Vieira *et al.*, 2013). Its prevalence in the population is approximately 15-22 per 100,000 (Onyike and Diehl-Schmid, 2013). According to epidemiological data from the WHO, dementia incidence is expected to double every 20 years and is predicted to reach 115.4 million by 2050 (Prince *et al.*, 2015).

Clinically, FTD is divided into four subtypes (Figure 1.1); behavioural variant FTD (bvFTD) which is characterised by behavioural changes, disinhibition and executive dysfunction, and three variants of the primary progressive aphasia (PPA), the semantic dementia (SD) which affects naming and word comprehension, the progressive nonfluent aphasia (PNFA) which affects speech and grammar output and the logopenic aphasia (LPA) which affects speech, word retrieval and sentence repetition (Neary *et al.*, 1998, Gorno-Tempini *et al.*, 2004, Rascovsky *et al.*, 2011).

FTD can also overlap with other movement disorders such as progressive supranuclear palsy (PSP), corticobasal syndrome (CBS) and amyotrophic lateral sclerosis (ALS) (Lomen-Hoerth *et al.*, 2002, Rohrer *et al.*, 2009). As FTD progresses and gradually spreads through the frontal and temporal lobes, the clinical symptoms converge, therefore examination of the early clinical features is crucial for the accurate diagnosis of the FTD subtypes. Currently there is no available treatment for FTD, instead the regime is focused on the management of the behavioural symptoms.

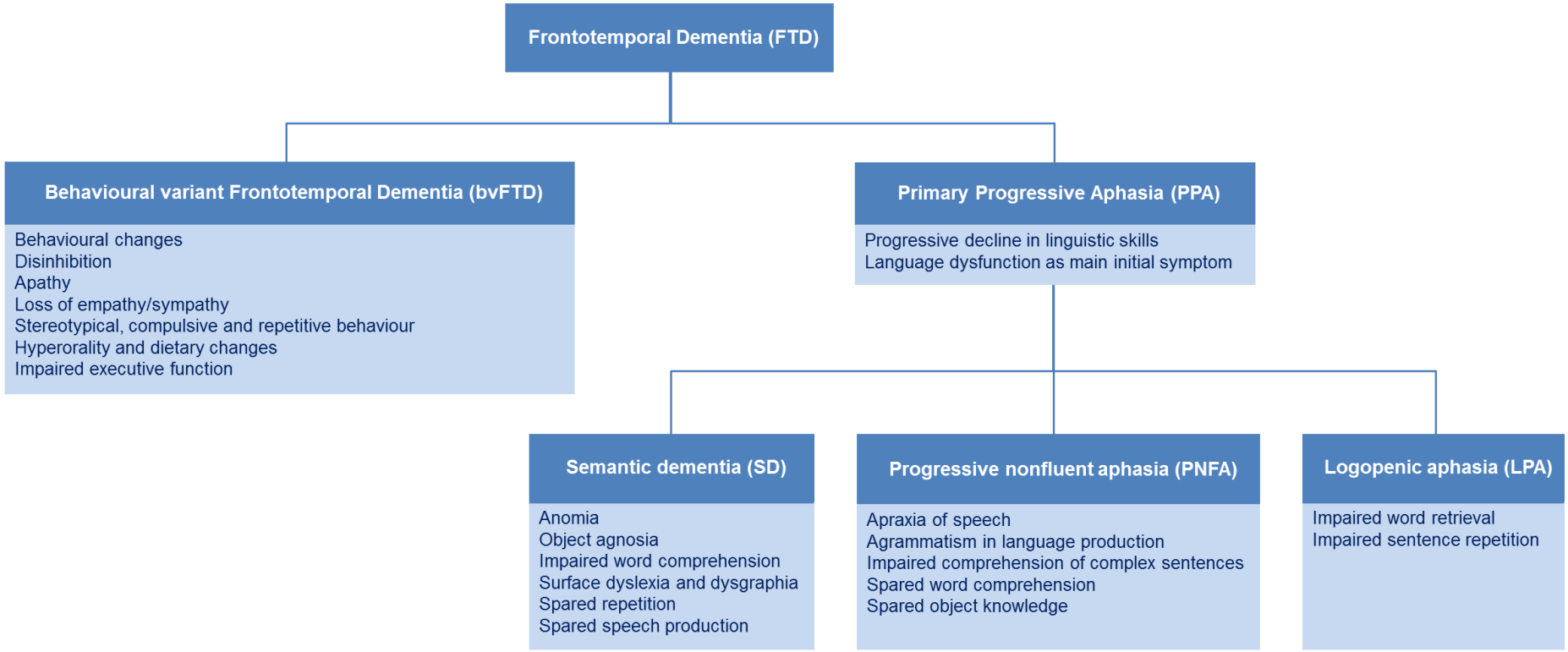


Figure 1.1. Overview of the early symptoms distinguishing the four clinical variants of FTD.

FTD is divided into four clinical subtypes; the behavioural variant FTD (bvFTD) which is characterised by behavioural changes, disinhibition and executive dysfunction, and three variants of the primary progressive aphasia (PPA), the semantic dementia (SD) which affects naming and word comprehension, the progressive nonfluent aphasia (PNFA) which affects speech and grammar output and the logopenic aphasia (LPA) which affects speech, word retrieval and sentence repetition.

1.1.2. Genetics of FTD

FTD is a highly heritable neurodegenerative disease with a family history in approximately 41% of the patients (Chow *et al.*, 1999, Rohrer *et al.*, 2009). However, only 10% of the cases show a clear autosomal dominant inheritance pattern (Rohrer *et al.*, 2009). Heritability varies among the different FTD clinical subtypes, with bvFTD showing the highest degree of heritability and SD and LPA the lowest (Seelaar *et al.*, 2008, Rohrer *et al.*, 2009). The most common cause of familial FTD is a hexanucleotide repeat expansion in *chromosome 9 open reading frame 72* (*C9orf72*) gene located on chromosome 9p which accounts for 25% of familial FTD cases (DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011). Other frequent causes of familial FTD are mutations in *microtubule-associated protein tau* (*MAPT*) (Hutton *et al.*, 1998, Spillantini *et al.*, 1998) and *progranulin* (*GRN*) (Baker *et al.*, 2006, Cruts *et al.*, 2006) genes, both located on chromosome 17 and each accounting for 5–20% of familial FTD cases. Collectively, mutations in these three genes account for approximately 60% of familial FTD cases (Le Ber, 2013). Other less frequent mutations have been reported in *charged multivesicular body protein 2B* (*CHMP2B*) (Skibinski *et al.*, 2005), *valosin containing protein* (*VCP*) (Watts *et al.*, 2004), *optineurin* (*OPTN*) (Maruyama *et al.*, 2010, Pottier *et al.*, 2015), *TANK-binding kinase 1* (*TBK1*) (Freischmidt *et al.*, 2015, Pottier *et al.*, 2015) and *sequestosome 1* (*SQSTM1*) (Fecto *et al.*, 2011, Rubino *et al.*, 2012) genes that cause FTD and ALS. Finally, mutations in TAR DNA binding protein (*TARDBP*) (Sreedharan *et al.*, 2008) on chromosome 1 and fused in sarcoma (*FUS*) (Kwiatkowski *et al.*, 2009, Vance *et al.*, 2009) on chromosome 16 are rare causes of ALS (Lattante *et al.*, 2013). However these proteins are found mislocalised and aggregated into characteristic neuropathological inclusions in FTD (see 1.1.3) (Neumann *et al.*, 2009, Mackenzie *et al.*, 2010), strengthening the notion that ALS and FTD are the opposites of the same spectrum with overlapping molecular aetiology.

1.1.3. Neuropathology of FTLD

Frontotemporal lobar degeneration describes the neuropathology of the heterogeneous FTD clinical syndrome characterised by selective degeneration of the frontal and temporal lobes (Neary *et al.*, 1998). Despite the relatively consistent brain degeneration pattern, the neuropathology of FTLD is also heterogeneous (Cairns *et al.*, 2007). The classification of FTLD is based on characteristic patterns of abnormal deposition of proteins in neurons and/or glia of FTD patients. FTLD is classified into four main subgroups (Figure 1.2); FTLD-tau (further divided into 3R, 4R and 3R&4R tauopathies), FTLD-TDP (further divided into types A, B, C and D), FTLD-FUS (further divided into atypical FTLD-U (aFTLD-U), neuronal intermediate filament inclusion disease (NIFID) and basophilic inclusion body disease (BIBD)) and FTLD-UPS (Lashley *et al.*, 2015, Mackenzie and Neumann, 2016).

FTLD-TDP is the most common FTLD subtype, representing approximately 50% of cases. FTLD-TDP pathology is present in FTD patients carrying mutations in *C9orf72* (Hsiung *et al.*, 2012, Mackenzie *et al.*, 2013b), *GRN* (Cairns *et al.*, 2007, Mackenzie, 2007), *VCP* (Cairns *et al.*, 2007, Neumann *et al.*, 2007), *TBK1* (Pottier *et al.*, 2015), *OPTN* (Pottier *et al.*, 2015), *SQSTM1* (Kovacs *et al.*, 2015), and *TARDBP* (Kovacs *et al.*, 2009). It is characterised by intracellular pathological inclusions that are positive for TDP-43 and ubiquitin. FTLD-tau is the second most common proteinopathy and it accounts for 40% of all FTLD cases (Mackenzie and Neumann, 2016). It is characterised pathologically by the presence of intracellular inclusions of hyperphosphorylated protein tau which is the neuropathological hallmark of FTD patients with mutations in the *MAPT* gene (Lee *et al.*, 2001). FTLD-FUS (also referred to as FTLD-FET) represents 10% of all FTLD and comprises of pathological inclusions that stain positive for fused in sarcoma (FUS), Ewing's sarcoma (EWS), and TATA-binding protein-associated factor 15 (TAF15) (collectively FET protein family) proteins and ubiquitin (Neumann *et al.*, 2009, Neumann *et al.*, 2011,

Mackenzie and Neumann, 2016). FTLD-UPS occurs in a few FTLD cases with mutations in *CHMP2B* that are characterised by inclusions that are only positive for markers of the ubiquitin proteasome system (UPS), but are negative for tau, TDP-43, and FUS (Holm *et al.*, 2007, Holm *et al.*, 2009, Mackenzie and Neumann, 2016).

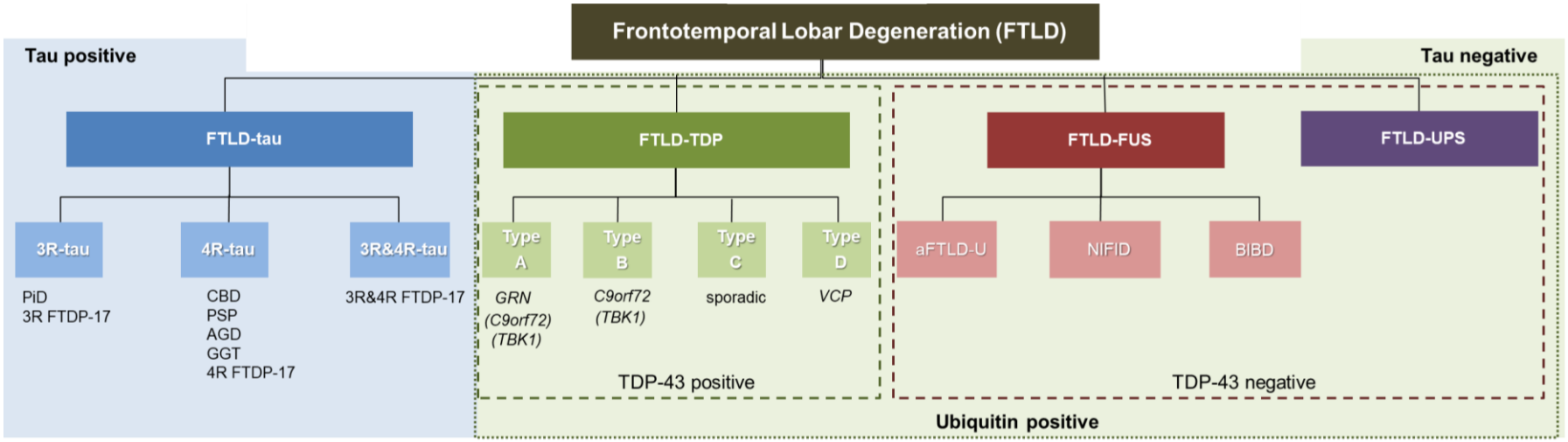


Figure 1.2 Molecular classification of frontotemporal lobar degeneration.

FTLD is divided into four molecular subtypes; (i) FTLD-tau, characterised by tau-positive inclusions, (ii) FTLD-TDP, characterised by tau-negative, TDP-43- and ubiquitin-positive inclusions, (iii) FTLD-FUS, characterised by tau- and TDP-43 negative but FUS-, EWS-, TAF15-and ubiquitin-positive inclusions, and (iv) FTLD-UPS, characterised by tau- and TDP-43 negative but ubiquitin-positive inclusions. FTLD-tau is further subdivided into three groups; 3R (including PiD and 3R FTDP-17), 4R (including CBD, PSP, AGD, GGT and 4R FTDP-17) and 3R&4R tauopathies (including 3R&4R FTDP-17). FTLD-TDP is further subdivided into four types: A (mainly associated with mutations in *GRN,* and less often in *C9orf72 and TBK1*), B (mainly associated with mutations in *C9orf72* and less often in *TBK1*), type C (not associated with familial mutations but typically with sporadic FTD) and D (associated with mutations in *VCP*). FTLD-FUS is further subdivided into three groups; aFTLD-U, NIFID and BiBD. Abbreviations: FUS, fused in sarcoma; EWS, Ewing's sarcoma, TAF15, TATA-binding protein-associated factor 15, PiD, Pick's disease; CBD, corticobasal degeneration; PSP, progressive supranuclear palsy; AGD, argyrophilic grain disease; GGT, Globular glial tauopathy; aFTLD-U, atypical FTLD-U; NIFID, neuronal intermediate filament inclusion disease; BiBD, basophilic inclusion body disease.

1.2. The *C9orf72* $(GGGGCC)_n$ repeat expansion in FTD and ALS

1.2.1. Identification of the *C9orf72* hexanucleotide repeat expansion

Several linkage analysis studies in families with autosomal dominant forms of FTD, ALS or both FTD and ALS suggested the existence of a major locus for these diseases on chromosome 9p (Morita *et al.*, 2006, Vance *et al.*, 2006, Boxer *et al.*, 2011, Pearson *et al.*, 2011), however the causative gene on this locus remained elusive for a number of years. Independent genome-wide association studies (GWAS) in FTD and ALS have also identified chromosome 9p21 and highlighted the importance of this locus in sporadic forms of these diseases (van Es *et al.*, 2009, Laaksovirta *et al.*, 2010, Shatunov *et al.*, 2010, Van Deerlin *et al.*, 2010). A GWAS study in the Finnish population, a genetically homogeneous population isolate, narrowed the association signal to a 232 kb linkage disequilibrium (LD) block that contained three genes: *MOBKL2B*, *IFNK*, and *C9orf72* and identified a 42 SNP founder haplotype that increased risk of disease by over 20-fold (Laaksovirta *et al.*, 2010). The association signal was further narrowed to a 140 kb region covering the same three genes by a large meta-analysis which also identified a 20 SNP founder haplotype common in all European-ancestry populations (Mok *et al.*, 2012).

In 2011, the $(GGGCC)_n$ hexanucleotide repeat expansion (HRE) in the first intron of *C9orf72* gene was discovered as the most common genetic cause of FTD and ALS (DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011). The highly repetitive and GC-rich sequence of the HRE and its localisation in a non-coding region of the *C9orf72* gene have been the main obstacles in the efforts to identify the mutation. Conventional sequencing of coding regions for potentially deleterious mutations missed the pathogenic expansion and it was the use of next generation sequencing together with repeat-primed long read PCR targeting the gene region that led to its identification.

1.2.2. Repeat expansion length and somatic mosaicism

FTD/ALS patients usually carry *C9orf72* alleles of hundreds or thousands of expanded repeats, whereas in most unaffected individuals the length is 2-24 repeats (Rutherford *et al.*, 2012, van der Zee *et al.*, 2013, Van Mossevelde *et al.*, 2017a). Currently, there is not yet a well-established minimal pathogenic HRE length and an arbitrary cut off of 30 repeats is used as the pathogenic repeat size threshold (DeJesus-Hernandez *et al.*, 2011). Intermediate expansions of 20-22 repeats have been described in FTD patients, and similarly, expansions as short as 24-28 repeats have been identified in ALS patients (Byrne *et al.*, 2012, Millecamps *et al.*, 2012, Ratti *et al.*, 2012, Gomez-Tortosa *et al.*, 2013). Meta-analysis studies have shown an association between ALS and intermediate expansions of 24 to 30 repeats (Chen *et al.*, 2016, Iacoangeli *et al.*, 2019). On the other hand, 32 and 35 hexanucleotide repeats, and even larger expansions of 400 repeats, have been reported in unaffected individuals (Ratti *et al.*, 2012, Simon-Sanchez *et al.*, 2012, Beck *et al.*, 2013). Post-mortem analysis in a cohort of control brains identified an unaffected carrier of 30 repeats with RNA foci and DPR lesions but not TDP-43 inclusions or clinical symptoms until death at 84 years (Gami *et al.*, 2015). Despite the advances, efforts to establish a minimal pathogenic repeat length are ongoing.

The main obstacles in accurately sizing the HRE length include the large repeat sizes, very GC rich composition, repetitive sequence and somatic instability of the HRE. The *C9orf72* HRE is characterised by a high degree of somatic instability resulting in mosaicism and significant variation in HRE sizes across different tissues in *C9orf72* HRE carriers (Beck *et al.*, 2013). Indeed, the HRE length in peripheral blood does not correlate with the HRE length in the central nervous system (CNS) (Beck *et al.*, 2013, Fratta *et al.*, 2015, Nordin *et al.*, 2015, Gijselinck *et al.*, 2016). Moreover, mosaicism in HRE length is also evident across different brain regions within the CNS (Beck *et al.*, 2013, van Blitterswijk *et al.*, 2013b). Somatic mosaicism seems to be age-

dependent, therefore HRE lengths may also be confounded by the age at collection of the DNA samples (van Blitterswijk *et al.*, 2013b).

1.2.3. *C9orf72* gene and HRE molecular properties

The *C9orf72* gene consists of 11 exons in total, including two non-coding exons 1a and 1b at the 5' UTR (Figure 1.3). Upon transcription, depending on the alternative splicing of exons 1a and 1b, there are three main pre-mRNA splice variants of the gene produced; V1 (NM_145005.6), V2 (NM_018325.4) and V3 (NM_001256054.2). *C9orf72* encodes two protein isoforms; a 222 amino acid short isoform C9-S, with a molecular weight of 24 kDa, produced upon translation of V1 and a 481 amino acid long isoform C9-L, with a molecular weight of 54 kDa produced upon translation of V2 and V3. The structure of *C9orf72* gene and its splicing are summarised in Figure 1.3.

The $(GGGGCC)_n$ HRE is located in the first intron of *C9orf72* and is flanked by the non-coding exons 1a and 1b (Figure 1.3). Upon transcription and depending on the alternative splicing of 1a and 1b, the HRE is located either in the intronic region of V1 and V3 pre-mRNA variants or in the promoter region of the V2 pre-mRNA variant. The levels of V2 are higher than V1 and V3 in the CNS (Tran *et al.*, 2015, Rizzu *et al.*, 2016). The presence of the HRE in different regions of the resulting pre-mRNA splice variants is associated with different downstream pathogenic mechanisms.

The presence of the HRE in the intronic region of V1 and V3 pre-mRNA variants is associated with toxic gain of function (GOF) pathogenic mechanisms via (i) the bi-directional transcription of the expanded repeat and the subsequent formation of intranuclear sense and antisense RNA foci in patients' neurons which sequester RNA-binding proteins (DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011, Lee *et al.*, 2013), and (ii) the production of aggregation-prone dipeptide repeat proteins

(DPRs) via non-canonical, repeat-associated, non-ATG (RAN) translation of both sense and anti-sense expanded RNA transcripts in all reading frames; poly-glycine-alanine (poly-GA), poly-glycine-proline (poly-GP) and poly-glycine-arginine (poly-GR) in the sense frames and poly-glycine-proline (poly-GP), poly-alanine-proline (poly-AP) and poly-proline-arginine (poly-PR) in the antisense frames (Ash *et al.*, 2013, Gendron *et al.*, 2013, Mori *et al.*, 2013c). Alternatively, the presence of the HRE in the promoter region of the V2 pre-mRNA variant results in reduced expression of the corresponding mRNA transcript (DeJesus-Hernandez *et al.*, 2011, Gijselinck *et al.*, 2012, van Blitterswijk *et al.*, 2015), consistent with haploinsufficiency and loss of *C9orf72* function.

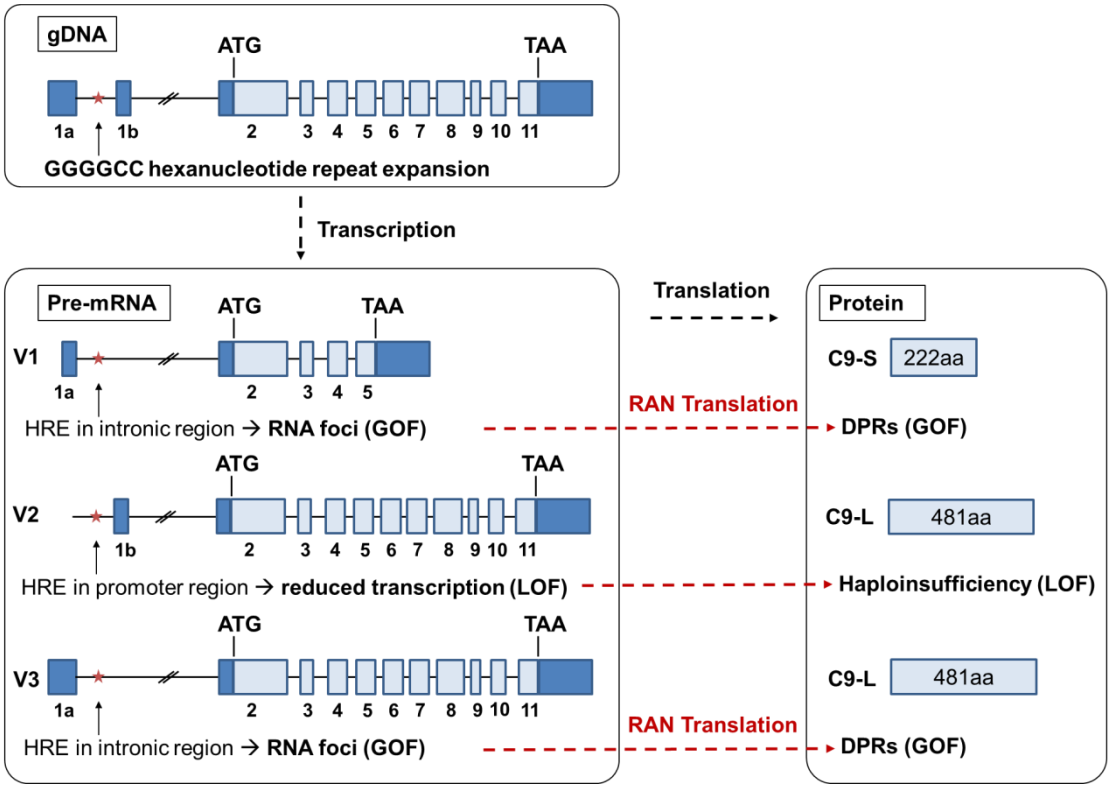


Figure 1.3 Schematic representation of *C9orf72* gene and HRE-linked disease mechanisms.

C9orf72 gene consists of 11 exons, including two non-coding exons 1a and 1b. During alternative splicing, exons 1a and 1b are removed, and mature mRNA transcripts start from exon 2. The HRE in *C9orf72* is located between the non-coding exons 1a and 1b. Upon transcription, three pre-mRNA splice variants are generated; V1 (NM_145005.6), V2 (NM_018325.4) and V3 (NM_001256054.2). The presence of the HRE in the promoter of V2 is associated with reduction in the expression of this transcript, whereas the localisation of the HRE in the first intron of transcripts V1 and V3 is associated with the presence of toxic RNA foci. The three main *C9orf72* transcripts are translated to generate two protein isoforms, a 481aa long isoform (C9-L) and a 222aa short isoform (C9-S). When the HRE is present, V2 reduced expression leads to reduction in C9-L, whereas potentially toxic DPRs may be produced via Repeat associated non-ATG (RAN) translation of the HRE from V1 and V3.

1.2.4. Neuropathology of C9-FTD/ALS

The pathology associated with C9-FTD/ALS is consistent with the presence of neuronal and glial TDP-43 cytoplasmic inclusions spreading across different brain regions including the frontal, temporal and primary motor cortices, hippocampus, amygdala, basal ganglia, thalamus and midbrain (Murray *et al.*, 2011, Mahoney *et al.*, 2012). Moreover, a unique and characteristic pathological hallmark of the C9-FTD/ALS is the presence of RNA foci and DPR inclusions (Figure 1.4). Sense and/or antisense RNA foci are mainly detected in the nuclei (and less frequently in the cytoplasm) of neurons in the frontal and motor cortices, hippocampus, cerebellum and spinal cord (Mizielinska *et al.*, 2013, DeJesus-Hernandez *et al.*, 2017). The DPR inclusions are ubiquitin and p62 positive but TDP-43 negative (Mann *et al.*, 2013, Mori *et al.*, 2013a) and are most frequently observed in the cytoplasm of neurons but may also be found as neuritic or intranuclear inclusions (Ash *et al.*, 2013, Gendron *et al.*, 2013, Mori *et al.*, 2013c). They are detected in the cerebellum, hippocampus and neocortex, less frequently in subcortical regions and are rare in the brainstem and spinal cord (Gomez-Deza *et al.*, 2015). All DPRs have been detected at *post mortem*, and according to neuropathological studies, poly-GA is the most frequently detected DPR (Mori *et al.*, 2013a, Mori *et al.*, 2013c, Gomez-Deza *et al.*, 2015, Mackenzie *et al.*, 2015, Davidson *et al.*, 2016), followed by poly-GP and poly-GR, whereas poly-PA and poly-PR are the least frequently detected DPRs (Gendron *et al.*, 2013, Mori *et al.*, 2013a, Gomez-Deza *et al.*, 2015, Mackenzie *et al.*, 2015, Davidson *et al.*, 2016). As revealed by a *post-mortem* neuropathological examination in three *C9orf72* HRE carriers who died prematurely of unrelated causes to FTD, DPR pathology may precede that of TDP-43 (Baborie *et al.*, 2014). The neuroanatomical distribution of TDP-43 pathology has been shown to correlate with neurodegeneration in C9-FTD/ALS cases (Mackenzie *et al.*, 2013a, Mann *et al.*, 2013). On the contrary, DPR pathology is consistent among carriers, regardless of the clinical phenotype, and does

not correlate with TDP-43 pathology and neurodegeneration (Mackenzie *et al.*, 2013a, Mann *et al.*, 2013). However, recent evidence suggests that the distribution of poly-GR correlates with neurodegeneration (Saberi *et al.*, 2017, Sakae *et al.*, 2018), and that out of all DPRs, poly-GR only was found to colocalise with TDP-43 pathology in a subset of *C9orf72* HRE brains (Saberi *et al.*, 2017).

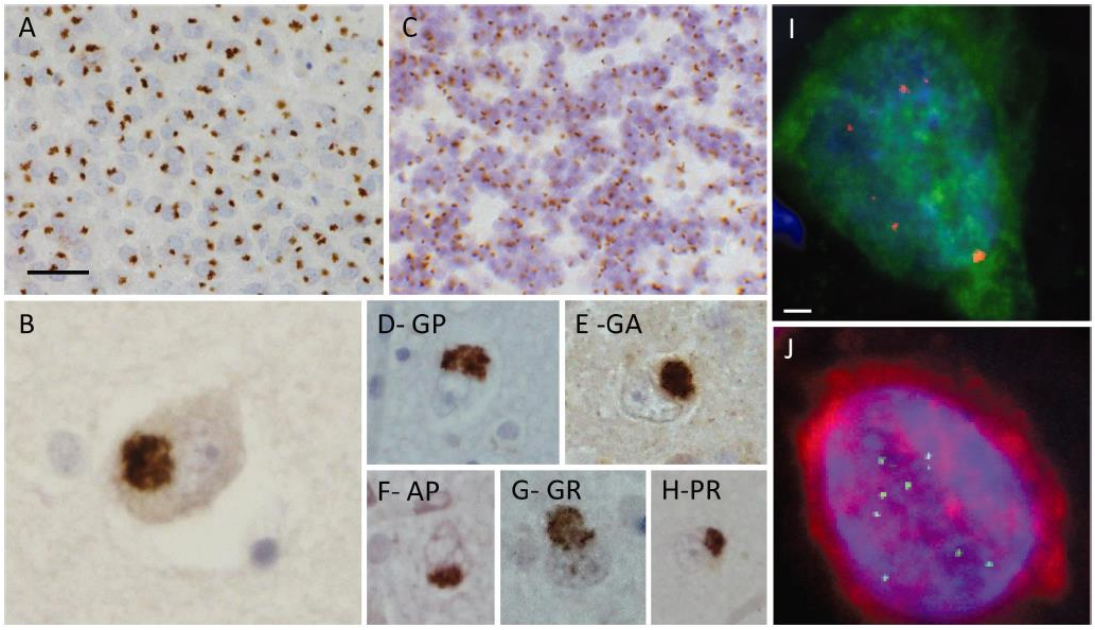


Figure 1.4 DPR and RNA foci pathology in *C9orf72* HRE carriers.
C9orf72 HRE carriers exhibit p62-positive and TDP-43-negative neuronal cytoplasmic inclusions in the granule cell layer of the dentate gyrus (A), the CA4 hippocampal subregion (B) and the cerebellar granule cells (C). All five DPRs, poly-GP (D), poly-GA (E), poly-AP (F), poly-GR (G) and poly-PR (H), can be detected in the p62-positive inclusions. Moreover, *C9orf72* HRE carriers exhibit sense (I) and anti-sense (J) intranuclear RNA foci. Scale bar in A; 40 μm on A and C; 20 μm on D–H;10 μm on B. Scale bar in I; 2 μm on I and J. Adapted from (Lashley *et al.*, 2015).

1.2.5. Epidemiology and clinical presentation

The hexanucleotide repeat expansion in *C9orf72* gene is the most common genetic cause of both familial ALS and FTD in Caucasian populations (Majounie *et al.*, 2012). The transmission in families is consistent with an autosomal dominant pattern with incomplete penetrance (DeJesus-Hernandez *et al.*, 2011, Majounie *et al.*, 2012, Murphy *et al.*, 2017). The incomplete penetrance of the HRE is supported by the remarkable variability in the age of disease onset among the carriers, ranging from 27-82 years, and even the existence of a small portion of carriers reaching their 90's without any disease symptoms. Moreover, studies have shown that penetrance is age-dependent, as it was minimal in carriers under the age of 35 years, increasing to 50% in 58 years of age and reaching full penetrance at 80 years of age (Majounie *et al.*, 2012, Murphy *et al.*, 2017). Whilst the vast majority of C9-FTD/ALS patients carry one expanded *C9orf72* allele, a rare case of a patient homozygous for the *C9orf72* HRE has also been described (Fratta *et al.*, 2013). Interestingly, the homozygous patient developed early-onset FTD with severe clinicopathological presentation but within the usual range seen in heterozygote *C9orf72* HRE carriers, suggesting that loss of *C9orf72* may not be the main driver of neurodegeneration in C9-FTD/ALS.

The frequency of *C9orf72* HRE mutations is significantly higher in patients with family history of disease compared to non-familial, sporadic cases (Majounie *et al.*, 2012). The *C9orf72* HRE accounts for up to 50% of familial ALS, 29% of familial FTD and 88% of combined familial FTD/ALS in Caucasians (Renton *et al.*, 2011, Majounie *et al.*, 2012, Lindquist *et al.*, 2013, Smith *et al.*, 2013). According to a recent meta-analysis, the overall mutation frequency of *C9orf72* is 20% for familial FTD, 16% for familial ALS and around 6%–8% for sporadic ALS and FTD (Marogianni *et al.*, 2019). The highest frequencies have been reported in familial ALS patients from Scandinavian populations; 50% in Denmark (Lindquist *et al.*, 2013), 46% in Sweden (Smith *et al.*, 2013) and 28% in Finland (Renton *et al.*, 2011, Majounie *et al.*, 2012).

Despite the high mutation frequencies in white individuals of Europe and North America, the *C9orf72* HRE is rare in Asian and Middle Eastern populations (Majounie *et al.*, 2012). Based on the geographical distribution of the *C9orf72* HRE, its prevalence in Caucasians of Europe and North America and a common 20 SNP haplotype shared by all expansion carriers, some studies have suggested that the *C9orf72* HRE derived from a single founder (Laaksovirta *et al.*, 2010, Renton *et al.*, 2011, Majounie *et al.*, 2012, Mok *et al.*, 2012, Ratti *et al.*, 2012, Smith *et al.*, 2013). However, the hypothesis of a common ancestor in *C9orf72* HRE has not been supported by other studies (Beck *et al.*, 2013, Chiang *et al.*, 2017).

C9orf72 HRE carriers exhibit significant heterogeneity in their clinical presentation. *C9orf72* HRE carriers may present with FTD, ALS or FTD and ALS. Importantly, the clinical heterogeneity is also present within families, with some family members exhibiting FTD whereas others ALS or both FTD and ALS (Simon-Sanchez *et al.*, 2012). ALS is a fatal, progressive neurodegenerative disease affecting both upper and lower motor neurons in motor cortex, brain stem and spinal cord of patients (Kiernan *et al.*, 2011). The neuronal loss leads to gradual loss of muscle control, weakness, spasticity, paralysis and eventually death due to respiratory failure usually within 2-5 years of onset (Rowland and Shneider, 2001).

The majority of C9-FTD patients present with the behavioural variant (bvFTD) and combination of symptoms including disinhibition, apathy, loss of empathy, compulsivity, overeating, and executive dysfunction (Rohrer *et al.*, 2015). For C9-ALS, a bulbar onset symptom is more frequent compared to non-C9 ALS (Chio *et al.*, 2012, Majounie *et al.*, 2012). Clinical heterogeneity is further enhanced by the wide range in the age of onset and disease duration. The age of onset range for C9-FTD/ALS is 27-83 years and the disease duration range is 1-22 years (Boeve *et al.*, 2012, Hsiung *et al.*, 2012, Mahoney *et al.*, 2012, Majounie *et al.*, 2012, Sabatelli *et al.*, 2012, Simon-Sanchez *et al.*, 2012, Woollacott and Mead, 2014).

Interestingly, *C9orf72* HRE carriers display a higher rate of psychiatric phenotypes compared to non-carriers (Devenney *et al.*, 2018, Silverman *et al.*, 2019). Among these phenotypes, *C9orf72* HRE carriers are in higher risk of developing psychosis, schizophrenia and suicide. A higher incidence of psychosis has been observed in *C9orf72* HRE carriers with FTD or FTD/ALS, with 21-56% of C9-FTD cases exhibiting psychotic symptoms compared to 3.5% in non-C9 bvFTD (Dobson-Stone *et al.*, 2012, Snowden *et al.*, 2012, Kaivorinne *et al.*, 2013). In many cases, due to the early onset of these psychotic symptoms years before FTD diagnosis, a misdiagnosis of psychiatric disorder may occur (Kaivorinne *et al.*, 2013). Unlike C9-FTD, C9-ALS is not associated with higher incidence of psychiatric symptoms compared to non-C9 ALS.

The *C9orf72* HRE has also been found as a rare cause of other neurodegenerative diseases, including Alzheimer's disease (Beck *et al.*, 2013, Cacace *et al.*, 2013, Harms *et al.*, 2013, Kohli *et al.*, 2013), Parkinson's disease (Xi *et al.*, 2012, Lesage *et al.*, 2013), corticobasal syndrome (Lindquist *et al.*, 2013), ataxia (Lindquist *et al.*, 2013), Huntington's disease-like syndrome (Beck *et al.*, 2013) and prion disease (Beck *et al.*, 2013).

No established association of the *C9orf72* HRE with anticipation currently exists. An inverse correlation between generation number and age of onset in FTD/ALS families, suggestive of genetic anticipation, has been shown by several studies (Boeve *et al.*, 2012, Chio *et al.*, 2012, Stewart *et al.*, 2012, Van Langenhove *et al.*, 2013). Evidence for anticipation was found in *C9orf72* pedigrees where younger generations exhibited earlier age of onset and longer repeat sizes (Gijselinck *et al.*, 2016, Van Mossevelde *et al.*, 2017b). However, others reported contraction of the *C9orf72* HRE in successive generations, arguing against the existence of anticipation (Jackson *et al.*, 2020). Interestingly, in contrast to other repeat expansions, the observed anticipation in some *C9orf72* pedigrees does not lead to more severe disease phenotypes in

younger generations which may be explained by additional disease-modifying factors. In conclusion, whether or not anticipation exists in *C9orf72* families is still a matter of debate as studying anticipation has been challenging due to the HRE size variability across tissues.

1.2.6. Disease modifiers in C9-FTD/ALS

The existence of disease modifiers in *C9orf72*-linked FTD/ALS was evident early on following the discovery of the *C9orf72* HRE due to the substantial clinical heterogeneity among *C9orf72* expansion carriers, including disease onset age and duration. *TMEM106B* has been described as a genetic modifier of C9-FTD (Gallagher *et al.*, 2014, van Blitterswijk *et al.*, 2014b). Homozygosity for the rs3173615 minor allele (G) was found to be protective against the development of FTD, but not ALS, therefore supporting a role for *TMEM106B* as a genetic modifier of disease presentation in *C9orf72* expansion carriers (van Blitterswijk *et al.*, 2014b). The major allele (A) of *TMEM106B* rs1990622 associated with later age of onset and death in *C9orf72* expansion carriers, and with an increased risk for developing FTLD-TDP (Gallagher *et al.*, 2014). Interestingly, *TMEM106B* has also been described as a genetic modifier of FTLD-TDP caused by *GRN* mutations, with the major allele (A) of rs1990622 associated with earlier age of onset in FTD patients with the *GRN* mutation (Cruchaga *et al.*, 2011). Intermediate ataxin-2 (ATXN2) repeats (27-33 repeats) have also been described as a genetic modifier in *C9orf72* expansion carriers by increasing their risk of developing ALS (van Blitterswijk *et al.*, 2014a).

Interestingly, *C9orf72* repeat expansions have been detected in patients carrying known pathogenic mutations in either *GRN* (p.C466LfsX46, p.R493X, p.C31LfsX35) or *MAPT* (p.P301L), suggesting that co-occurrence of two pathogenic mutations may

contribute to the pleiotropy seen in *C9orf72* repeat expansion carriers (van Blitterswijk *et al.*, 2013a).

Hypermethylation in the *C9orf72* locus has been suggested as another disease modifier (Russ *et al.*, 2014). *C9orf72* hypermethylation in blood and brain was associated with later age at death in FTD whereas in blood it was associated with longer disease duration in FTD (Russ *et al.*, 2014).

Potential effects of the global DNA methylation on C9-FTD/ALS have also been investigated. Combined analyses of epigenetic and genetic data show that CpG-SNPs within a 124.7 kb linkage disequilibrium block covering two overlapping genes *C6orf10* and *LOC101929163* may modify age of onset in *C9orf72* HRE carriers by controlling DNA methylation and gene expression (Zhang *et al.*, 2018b). For *C9orf72* HRE carriers carrying the AA genotype of the top-significant SNP rs9357140, the median age of onset was 6 years later compared to those carrying the GG genotype. Moreover, acceleration of ageing, assessed via analysis of genome-wide blood DNA methylation in *C9orf72* HRE carriers, significantly correlated with a shorter disease duration and earlier age of onset (Zhang *et al.*, 2017).

Finally, the *C9orf72* repeat expansion length has been described as a genetic modifier affecting disease age at onset and severity. Specifically, in the frontal cortex of FTD patients, repeat length correlated with age at onset, whereas in cerebellum, longer repeat sizes were associated with a survival disadvantage (van Blitterswijk *et al.*, 2013b). The *C9orf72* HRE size was also found to correlate with onset of disease (Beck *et al.*, 2013, Gijselinck *et al.*, 2016) and 5’ CpG DNA methylation (Gijselinck *et al.*, 2016). In contrast, other studies have reported no association between repeat expansion length and age at onset (Suh *et al.*, 2015). Moreover, an association between *C9orf72* repeat expansion size and shorter disease duration was found in individuals with FTD, but not ALS (Suh *et al.*, 2015). However, due to (i) the somatic

mosaicism of the HRE and repeat expansion length variability within individuals and (ii) the very long repeat expansions in many *C9orf72* carriers, the investigation of such associations is extremely difficult.

1.2.7. *C9orf72* promoter hypermethylation

Repeat-associated hypermethylation near *C9orf72* has been described in both FTD and ALS cases and is associated with the downregulation in *C9orf72* expression (Belzil *et al.*, 2013, Xi *et al.*, 2013, Liu *et al.*, 2014, Gijselinck *et al.*, 2016). Epigenetic silencing has been associated with other G/C rich repeat expansions, the most prominent example being the DNA methylation of the trinucleotide CGG repeat expansion in the promoter region of *FMR1* in Fragile X syndrome, leading to pathogenic loss of function (Sutcliffe *et al.*, 1992, Brykczynska *et al.*, 2016, Zhou *et al.*, 2016).

The *C9orf72* HRE is flanked by two CpG islands, generating a methylation-hypersensitive region that extends over the promoter until the second intron of the gene, covering a region of ~1kb (Figure 1.5). The HRE is located upstream of exon 1b, in the promoter region of V1 transcript whilst the 5' CpG island covers exon 1a and extends to the promoter region of V2 and V3 transcripts. The 3' CpG island is mostly downstream of exon 1b, in the second intron of the gene.

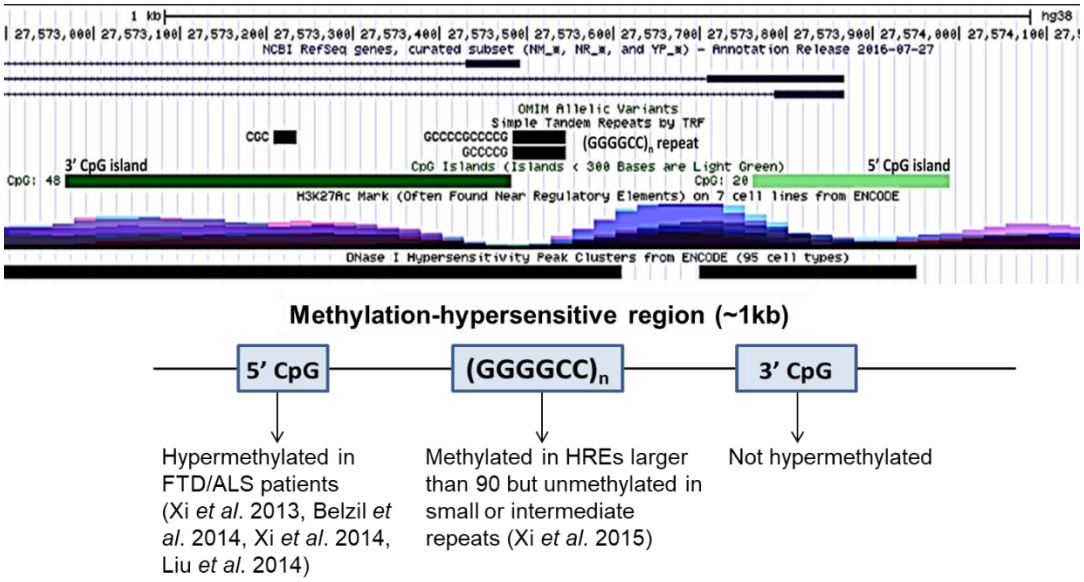


Figure 1.5 The *C9orf72* HRE is flanked by two CpG islands.
The *C9orf72* promoter contains a ~1kb methylation-sensitive region which consists of the HRE flanked by two CpG islands. The HRE is located in the promoter region of V1 whereas the 5' CpG island extends to the promoter region of V2 and V3 transcripts. Image acquired from UCSC Genome Browser on Human Dec. 2013 (GRCh38/hg38) Assembly.

The *C9orf72* HRE itself is methylated in 97% of *C9orf72* HRE carriers, reaching 100% in carriers of more than 90 repeats, whereas it is unmethylated in carriers of small (<22) and intermediate (22-90) repeats (Xiao *et al.*, 2015). Therefore, it has been suggested that HRE methylation state could be useful as a predictor of large versus intermediate repeat expansions (Xiao *et al.*, 2015). Moreover, 17% of C9-FTD and 36% of C9-ALS cases exhibit 5' CpG island hypermethylation which occurs exclusively in *C9orf72* HRE carriers (Xi *et al.*, 2013, Belzil *et al.*, 2014, Liu *et al.*, 2014, Xi *et al.*, 2014a, Jackson *et al.*, 2020). 5' CpG hypermethylation is associated with increased HRE length (Gijselinck *et al.*, 2016) and reduced *C9orf72* transcription (Belzil *et al.*, 2013, Xi *et al.*, 2013, Gijselinck *et al.*, 2016). In contrast, no hypermethylation has been observed at the 3' CpG island in *C9orf72* HRE carriers (Xi *et al.*, 2013, Liu *et al.*, 2014). The observations that (i) the majority of *C9orf72* HRE carriers exhibit a methylated HRE while a fraction of those also exhibit 5 'CpG island

hypermethylation and (ii) the presence of methylated HRE in the absence of hypermethylation at the 5'CpG island in C9orf72 HRE carriers (Xi *et al.*, 2014a), led to the hypothesis that HRE methylation may spread to the upstream 5' CpG island but not vice versa (Xiao *et al.*, 2015). Of note, another study reported 5' CpG island hypermethylation in the absence of a methylated HRE, arguing against the above hypothesis (Liu *et al.*, 2014).

Apart from DNA methylation, other epigenetic modifications such as histone trimethylation, may also be responsible for the reduction in *C9orf72* transcription in C9orf72 HRE carriers. Indeed, trimethylation of histones H3 and H4 at several lysine residues was detected in the C9orf72 promoter of *C9orf72* expansion carriers and was associated with reduced expression of *C9orf72* (Belzil *et al.*, 2013). Interestingly, treatment of *C9orf72* patient fibroblasts with a DNA/histone demethylating agent, 5′-aza-2′-deoxycytydine, resulted in increased *C9orf72* expression (Belzil *et al.*, 2013).

All these repressive epigenetic modifications of DNA and histones have been associated with reduced *C9orf72* expression, providing one mechanism that may lead to reduced *C9orf72* expression in C9-FTD/ALS patients. Some studies have observed that *C9orf72* hypermethylation and subsequent *C9orf72* transcript reduction are associated with lower RNA foci- and DPR-related toxicity, suggesting that hypermethylation is a disease modifier protecting against toxicity in C9-FTD/ALS (Liu *et al.*, 2014, Russ *et al.*, 2014, Bauer, 2016).

Insight from other GC-rich non-coding repeat expansion disorders shows similar effects of DNA methylation in gene expression. In Fragile X syndrome (FXS), a CGG trinucleotide repeat in the promoter region of *fragile X mental retardation 1* (*FMR1*) gene is responsible for silencing the expression of the gene consistent with a loss of function disease mechanism (Oberle *et al.*, 1991, Verkerk *et al.*, 1991). Hypermethylation of the expanded CGG repeat in the 5′UTR region of *FMR1* has

been described as the main mechanism by which the expansion can silence the expression of the gene (Bell *et al.*, 1991, Pieretti *et al.*, 1991, Sutcliffe *et al.*, 1992). Insight into the epigenetic silencing mechanism was provided by ESC models of mutant *FMR1*, where silencing was caused by direct binding of the CGG repeat mRNA on its complementary sequence in the DNA, a phenomenon that happens in neurons only after 45 days in culture and correlates with reduction of *FMR1* expression (Colak *et al.*, 2014). Fragile X E (FRAXE) mental retardation is caused by a CCG expansion in the 5'UTR of *Fragile X mental retardation 2* (*FMR2*) gene (Gecz *et al.*, 1996). Similar to FXS, a CCG expansion above 200 repeats leads to hypermethylation of the CpG island upstream of *FMR2* and transcriptional gene silencing (Knight *et al.*, 1993, Gu *et al.*, 1996). The aforementioned syndromes are examples of trinucleotide expansions in the promoter of genes leading to hypermethylation and transcriptional silencing of the downstream genes. However, these syndromes have a loss of function disease aetiology, whereas in C9-FTD/ALS, apart from *C9orf72* loss of function, RNA foci and DPR gain of function also have major roles in disease pathogenesis.

1.2.8. *C9orf72* HRE disease mechanisms

Evidence from studies in post-mortem patient tissue as well as *in vivo* and *in vitro* experimental models show that three well-described, non-mutually exclusive pathogenic mechanisms can contribute to *C9orf72* HRE-mediated neurodegeneration; a loss of C9orf72 protein function, a toxic gain of expanded RNA function and a toxic gain of DPR function. Soon after the discovery of the HRE potential disease mechanisms, a debate over which is the main driver of neurodegeneration led to a plethora of studies aiming to dissect these mechanisms. The most recent evidence suggest that toxicity and loss of function mechanisms act synergistically to cause the neurodegenerative phenotypes seen in C9-FTD/ALS.

However the molecular pathways and interactions that lead to disease, and in particular the significant heterogeneity in clinical presentation in C9-FTD/ALS carriers, is a subject of extensive research.

1.2.8.1. Toxic gain of function

C9-FTD/ALS toxicity pathogenic mechanisms arise from the transcription and translation of the intronic HRE to RNA foci and DPRs respectively. Normally, in unaffected individuals the intron 1 of *C9orf72* is spliced out of *C9orf72* mRNA transcripts and degraded. However, the presence of the HRE can interfere with normal splicing resulting in intron 1-retaining *C9orf72* pre-mRNA transcripts (Niblock *et al.*, 2016). Indeed, several studies have shown an increase in intron 1-retaining transcript levels in *C9orf72* HRE carriers compared to controls (Cohen-Hadad *et al.*, 2016, Jackson *et al.*, 2020). Moreover, additional stable expanded RNA species have been detected either in the form of intronic RNA or aborted RNA transcripts (Haeusler *et al.*, 2014, Tran *et al.*, 2015, van Blitterswijk *et al.*, 2015).

RNA foci consisting of expanded sense and antisense HRE-containing RNA are commonly found in the brain tissue of *C9orf72* HRE carriers (DeJesus-Hernandez *et al.*, 2011, Mizielinska *et al.*, 2013). These aberrant HRE-containing RNA transcripts can form G-quadruplexes and RNA hairpin secondary structures, and in addition to that, HRE-containing RNA can form hybrids with HRE-containing DNA, called R-loops (Fratta *et al.*, 2012, Reddy *et al.*, 2013, Haeusler *et al.*, 2014). All of these HRE-induced higher-order structures could potentially cause neurotoxicity by impeding vital cellular processes such as replication and transcription (Haeusler *et al.*, 2014). G-quadruplex RNA structures sequester nucleolin, the major nucleolar protein, leading to a translocation from the nucleoli to the nucleus, nucleolar stress and impairments in rRNA processing and maturation of ribosomes (Haeusler *et al.*, 2014).

HRE-containing RNA repeats sequester multiple RNA binding proteins, including splicing factors heterogeneous nuclear ribonucleoprotein (hnRNP) A1 (Sareen *et al.*, 2013, Cooper-Knock *et al.*, 2014), hnRNP A3 (Mori *et al.*, 2013b), hnRNP H (Lee *et al.*, 2013, Cooper-Knock *et al.*, 2014) and serine/arginine-rich splicing factor 2 (SRSF2) (Cooper-Knock *et al.*, 2014) as well as other proteins Pur-α (Sareen *et al.*, 2013), adenosine deaminase RNA specific B2 (ADARB2) (Donnelly *et al.*, 2013) and Aly/REF export factor (ALYREF) (Cooper-Knock *et al.*, 2014). Sequestration and loss of function of these regulatory proteins may result in altered RNA splicing pathways and defects in nucleocytoplasmic transport that could potentially contribute to C9-FTD/ALS disease pathogenesis. Interestingly, altered RNA splicing is characteristically observed in ALS due to defects in RNA-binding proteins such as TDP-43 and FUS that affect global splicing machinery (Butti and Patten, 2019).

Unconventional RAN translation of the accumulating HRE-containing RNA transcripts results in the production of DPRs, which are another pathological hallmark of C9-FTD/ALS, detected in cytoplasmic aggregates in patient brain tissue (Ash *et al.*, 2013, Gendron *et al.*, 2013, Mori *et al.*, 2013a, Mori *et al.*, 2013c, Zu *et al.*, 2013).

According to several *in vitro* and *in vivo* studies, the arginine-rich DPRs, GR and PR, are the most toxic DPR species (Mizielinska *et al.*, 2014, Wen *et al.*, 2014, Freibaum *et al.*, 2015, Jovicic *et al.*, 2015). Neurotoxic effects for poly-GA have also been reported (Zhang *et al.*, 2014). Poly-GR and poly-PR were shown to bind to hnRNPA2 hydrogels non-reversibly and cause toxicity and cell death via nucleolar stress and inhibition of RNA biogenesis in cultured cells (Kwon *et al.*, 2014). Evidence from two Drosophila models argued against a major pathogenic role for expanded RNA, instead suggested that DPRs were the mediators of toxicity in these *in vivo* models (Mizielinska *et al.*, 2014, Tran *et al.*, 2015). Arginine-rich DPRs were particularly neurotoxic in Drosophila models (Mizielinska *et al.*, 2014, Moens *et al.*, 2019) and it has been demonstrated that they are able to bind ribosomal proteins and induce a

toxic translational arrest (Kanekura *et al.*, 2016, Zhang *et al.*, 2018c, Moens *et al.*, 2019). Arginine-rich DPRs were also toxic to *S. cerevisiae* mainly by disrupting nucleocytoplasmic transport (Jovicic *et al.*, 2015). Nucleocytoplasmic transport was also affected by HRE-mediated toxicity in two Drosophila models, but whether the phenotypes are attributed to expanded RNA or DPR toxicity, or both, remains unclear (Freibaum *et al.*, 2015, Zhang *et al.*, 2015). In human cells transient overexpression of poly-GR and poly-PR was able to induce stress granule assembly and localization of nucleocytoplasmic transport factors into the stress granules, leading to nucleocytoplasmic transport defects (Zhang *et al.*, 2018a). Additional cellular pathways affected by DPR-mediated toxicity include the ubiquitin proteasome system (UPS) (Zhang *et al.*, 2014) and stress granules (Lee *et al.*, 2016, Boeynaems *et al.*, 2017, Zhang *et al.*, 2018c).

In mice, several GOF models have been described, however the majority failed to recapitulate robust C9-FTD/ALS pathology and neurodegeneration (Chew *et al.*, 2015, O'Rourke *et al.*, 2015, Peters *et al.*, 2015, Jiang *et al.*, 2016, Liu *et al.*, 2016, Zhang *et al.*, 2016, Herranz-Martin *et al.*, 2017, Schludi *et al.*, 2017, Zhang *et al.*, 2018c, Hao *et al.*, 2019). Among these, a bacterial artificial chromosome (BAC) transgenic mouse model showed the most robust phenotype, including RNA foci, poly-GA aggregates, TDP-43 inclusions, rapidly progressive disease with motor deficits and neurodegeneration (Liu *et al.*, 2016). Similarly, a mouse model expressing 66 GGGGCC repeats via adeno-associated virus-mediated delivery exhibited RNA foci, poly-GA and poly-GP inclusions, TDP43 pathology, behavioural and motor deficits and neuronal loss (Chew *et al.*, 2015). Two more recent mouse models overexpressing GFP-GR_{100} or GFP-PR_{28} displayed neurodegeneration and behavioural and/or motor deficits but without TDP43 pathology (Zhang *et al.*, 2018c, Hao *et al.*, 2019).

Altogether these models have shed light on the downstream effects of the DPR-mediated toxicity, however, the majority rely on exogenous overexpression. Human patient-derived iPSC models are necessary to assess endogenous levels of DPRs.

1.2.8.2. Loss of function

A frequent finding in C9-FTD/ALS is the reduced expression of *C9orf72* leading to haploinsufficiency of C9orf72 protein. Several studies have shown that *C9orf72* HRE carriers exhibit reduced levels of *C9orf72* transcripts V1 and V2 in blood and CNS compared to controls (DeJesus-Hernandez *et al.*, 2011, Gijselinck *et al.*, 2012, Belzil *et al.*, 2013, Fratta *et al.*, 2013, Xi *et al.*, 2013, Waite *et al.*, 2014, van Blitterswijk *et al.*, 2015, Jackson *et al.*, 2020). These findings have also been supported by *C9orf72* iPSC studies which will be discussed in detail in 1.3.5.2. The presence of the HRE may lead to a reduced *C9orf72* expression via several potential mechanisms. Firstly, the HRE can form G-quadruplex DNA or DNA:RNA R-loop structures, impeding RNA polymerase transcription and resulting in paused or abortive transcription of *C9orf72* (Haeusler *et al.*, 2014). Moreover, C9orf72 downregulation may be caused epigenetically, via HRE-mediated DNA hypermethylation of the *C9orf72* promoter or repressive histone modifications (Belzil *et al.*, 2013, Xi *et al.*, 2013, Liu *et al.*, 2014, Xi *et al.*, 2014a, Xi *et al.*, 2015b). Apart from reduced transcript levels, reduced levels of C9orf72 protein have also been detected in the frontal and temporal cortices and cerebellum of *C9orf72* HRE carriers (Waite *et al.*, 2014, Xiao *et al.*, 2015, Frick *et al.*, 2018). However, the description of a patient homozygous for the *C9orf72* HRE without more severe clinical and pathological presentation compared to heterozygous patients, created an argument against the loss of function as the main disease mechanism in C9-FTD/ALS (Fratta *et al.*, 2013).

Increasing evidence supports a role for C9orf72 protein in endosomal trafficking, autophagy and lysosomal biogenesis and function (Farg *et al.*, 2014, Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Webster *et al.*, 2016, Yang *et al.*, 2016, Shi *et al.*, 2018). Interestingly, this is in line with the roles of several other genes of the FTD-ALS spectrum, such as *CHMP2B* and *GRN* in the endosomal-lysosomal pathway, or *TBK1*, *OPTN*, *SQSTM1*, *UBQLN2* and *VCP* in autophagy. Therefore, loss of *C9orf72* could contribute to the phenotypes of autophagy dysregulation seen in FTD/ALS.

Loss of the *C9orf72* orthologue in *C. elegans* and zebrafish resulted in motor deficits, whereas overexpression of human *C9orf72* mRNA in zebrafish rescued the loss of function phenotypes (Ciura *et al.*, 2013, Therrien *et al.*, 2013). Human *C9orf72* gene has higher homology with the mouse orthologue compared to the *C. elegans* and zebrafish orthologues (Suzuki *et al.*, 2013). Several *C9orf72* loss of function mouse models have been generated, however, none of them has shown neurodegeneration phenotypes (Lagier-Tourenne *et al.*, 2013, Koppers *et al.*, 2015, Atanasio *et al.*, 2016, Burberry *et al.*, 2016, Jiang *et al.*, 2016, O'Rourke *et al.*, 2016, Sudria-Lopez *et al.*, 2016, Sullivan *et al.*, 2016, Ugolino *et al.*, 2016, Burberry *et al.*, 2020). Instead, *C9orf72* knockout mice were mostly characterised by immune system dysregulation phenotypes, including lymphadenopathy, splenomegaly, inflammation and autoimmunity (Atanasio *et al.*, 2016, Burberry *et al.*, 2016, Jiang *et al.*, 2016, O'Rourke *et al.*, 2016, Sudria-Lopez *et al.*, 2016, Sullivan *et al.*, 2016, Ugolino *et al.*, 2016, Burberry *et al.*, 2020). Some studies also reported cognitive decline (Jiang *et al.*, 2016), motor deficits (Atanasio *et al.*, 2016, Jiang *et al.*, 2016, Burberry *et al.*, 2020) and reduced survival in *C9orf72* knockout mice (Atanasio *et al.*, 2016, Burberry *et al.*, 2016, Jiang *et al.*, 2016, Sudria-Lopez *et al.*, 2016, Ugolino *et al.*, 2016, Burberry *et al.*, 2020). However, more recent studies of crossed *C9orf72*$^{+/-}$ or *C9orf72*$^{-/-}$ mice with BAC transgenic mice (GOF) have highlighted the contribution of *C9orf72* loss of function in the development of FTD/ALS (described in 1.2.8.3.).

Loss of function mouse models have supported a role for C9orf72 in immune system function and have highlighted its importance for proper macrophage and microglial function (O'Rourke *et al.*, 2016). Interestingly, *C9orf72* HRE carriers are significantly more likely to have been diagnosed with autoimmune disease prior to the neurological diagnosis (Miller *et al.*, 2016, Fredi *et al.*, 2019). Moreover, an environment deprived of immune-stimulating bacteria was protective against the reduced mortality, inflammation and autoimmunity phenotypes in *C9orf72* knockout mice, supporting a role for C9orf72 in regulating inflammatory responses to innate immunity activation (Burberry *et al.*, 2020). Specifically, loss of C9orf72 function resulted in infiltration of peripheral immune cells to the spinal cord, activation of resident microglia and neuroinflammation in response to immune-stimulating bacteria, whereas treatment with antibiotics rescued these phenotypes (Burberry *et al.*, 2020). This highlights the importance of gene x environment interactions in neurodegeneration and confirms the link between gut microbiota composition and neuroinflammation that has also been described for other neurodegenerative diseases such as Parkinson's disease (Sampson *et al.*, 2016). It has also been recently shown that myeloid cells, and specifically dendritic cells (DCs), lacking C9orf72 protein are hyperresponsive to activators of the stimulator of interferon genes (STING) protein, which is a regulator of innate immune response to cytosolic DNA, as STING cannot be degraded via autophagy in these cells (McCauley *et al.*, 2020). *C9orf72*$^{-/-}$ mice were more susceptible to experimental autoimmune encephalitis (EAE) and had increased antitumour immunity compared to wild-type mice. Blocking of STING was able to suppress hyperactive type I interferon responses in *C9orf72*$^{-/-}$ immune cells as well as splenomegaly and infammation in *C9orf72*$^{-/-}$ mice. Interestingly, elevated levels of type I interferon signature was detected in blood-derived macrophages, whole blood and brain tissue from patients with C9-ALS/FTD (McCauley *et al.*, 2020). These findings support the role of C9orf72 in the regulation of inflammatory responses and

suggest that *C9orf72* HRE carriers have altered immunophenotypes due to partial loss of C9orf72.

Loss of *TBK1*, another gene linked to FTD and ALS with a role in autophagy, results in similar phenotypes to the *C9orf72* loss. Whilst TBK1-deficient mice are embryonically lethal (Bonnard *et al.*, 2000), conditional loss of *TBK1* in T cells leads to T cell activation and infiltration of spleen and lymph nodes in T-cell-conditional TBK1-knockout mice (Yu *et al.*, 2015), whereas in $TBK1^{\Delta/\Delta}$ mice encoding a mutant TBK1 protein lacking catalytic activity which is expressed at exceedingly low levels, loss of TBK1 results in infiltration of immune cells to many organs and tissues and susceptibility to LPS-induced mortality (Marchlik *et al.*, 2010). Interestingly, TBK1 is recruited and activated by STING in response to cyclic GMP-AMP (cGAMP), produced by innate immune sensing of cytosolic double-stranded DNA, and stimulates the induction of type I interferons through the STING–TBK1–IRF-3 signalling axis (Tanaka and Chen, 2012, Zhao *et al.*, 2019a). Conditional knockout of *TBK1* in DCs in mice generates a similar immunophenotype to the *C9orf72* myeloid-conditional knockout, whereby mice develop systemic interferon-driven inflammation, increased propensity for autoimmune disease and cancer resistance (Xiao *et al.*, 2017). Moreover TBK1 has been shown to have a central role in inflammatory responses mediated by macrophages (Yu *et al.*, 2012) and be closely related to inflammatory diseases such as rheumatoid arthritis (RA) (Hammaker *et al.*, 2011) and herpes simplex encephalitis of childhood (Herman *et al.*, 2012). In conclusion, a link between autophagy and neuroinflammation is supported by both *C9orf72* and *TBK1* ALS/FTD genes and recent evidence suggests a convergence of these gene pathways that requires further investigation. Importantly, the above studies highlight the contribution of non-cell-autonomous disease mechanisms in C9-FTD/ALS that also need to be further investigated.

1.2.8.3. Synergy between gain and loss of function

A plethora of LOF and GOF dissection mouse models have been published to date, however, the vast majority of them fail to demonstrate neurodegenerative phenotypes typical of FTD/ALS, suggesting that these mechanisms do not act in isolation, but instead the pathogenic mechanism underlying C9-FTD/ALS is more complex.

Increasing amount of evidence suggest that *C9orf72* HRE-mediated loss and gain of function mechanisms act in synergy to drive neurodegeneration in C9-FTD/ALS. Loss of C9orf72 was shown to increase the vulnerability of human iPSC-motor neurons to neurodegenerative stimuli such as glutamate-induced excitotoxicity and DPRs (Shi *et al.*, 2018). Moreover, restoration of vesicle trafficking could rescue neurodegenerative phenotypes, demonstrating that enrichment of C9orf72 function has great therapeutic benefit in C9-FTD/ALS (Shi *et al.*, 2018). Additional *in vitro* models for the study of GOF and LOF mechanisms were provided by Abo-Rady *et al.* who generated human isogenic iPSC models from *C9orf72* HRE patient-derived iPSCs via CRISPR/Cas9-mediated HRE excision (gene-corrected) or *C9orf72* knockout (C9 HRE + KO) (Abo-Rady *et al.*, 2020). Loss of C9orf72 in patient motor neurons exacerbated axonal trafficking phenotypes and apoptosis.

Recently, the generation of combined LOF and GOF mouse models has provided the first *in vivo* evidence for a combinatorial effect between these two disease mechanisms (Shao *et al.*, 2019, Zhu *et al.*, 2020). Shao *et al.* generated *C9orf72*$^{+/-}$*;C9-BAC* and *C9orf72*$^{-/-}$*;C9-BAC* mice by crossing *C9orf72*$^{+/-}$ mice with *C9-BAC* mice expressing the human *C9orf72* sequence under the control of the endogenous regulatory elements (Shao *et al.*, 2019). Loss of *C9orf72* exacerbated motor deficits in a dose-dependent manner supporting a dual effect pathogenic mechanism. Zhu *et al.* studied the combined LOF and GOF effects in two different mouse models. In the first model, mice with neither, one or both inactivated endogenous *C9orf72* alleles were injected with AAV expressing 66 GGGGCC repeats (*C9orf72*$^{+/+}$*;66R, C9orf72*$^{+/-}$

;66R, C9orf72$^{-/-}$*;66R)* or 2 repeats as control (*C9orf72*$^{+/+}$*;2R, C9orf72*$^{+/-}$*;2R, C9orf72*$^{-/-}$*;2R)* (Zhu *et al.*, 2020). Reduction or loss of C9orf72 in mice expressing 66 repeats induced or accelerated, respectively, early death and resulted in increased DPR accumulation in a dose-dependent manner. In the second model, transgenic mice carrying a BAC that contained a human *C9orf72* gene expressing 450 repeats at a physiologically relevant level (*C9*^{450C}) were crossed twice with mice lacking one or both endogenous *C9orf72* alleles (*C9*^{450C}, *C9*^{450C}*;C9orf72*$^{+/-}$ and *C9*^{450C}*;C9orf72*$^{-/-}$) (Zhu *et al.*, 2020). The reduction or loss of endogenous C9orf72 in *C9*^{450C} mice led to an exacerbation of age-dependent cognitive abnormalities and motor deficits. Moreover, loss of C9orf72 resulted in degeneration of hippocampal neurons, glial activation, inhibition of autophagy and increased accumulation of DPRs in *C9*^{450C} mice.

1.2.9. C9orf72 function

Great progress has been made in elucidating aspects of the function of C9orf72 which include a role in vesicle trafficking such as endolysosomal trafficking and autophagy. C9orf72 has high homology to Differentially Expressed in Normal and Neoplasia (DENN) related proteins, which act as GDP/GTP exchange factors (GEFs) that activate Rab-GTPases which are important regulators of vesicle trafficking (Zhang *et al.*, 2012, Levine *et al.*, 2013). Indeed, several studies have reported interaction of C9orf72 with Rab-GTPases, including Rab1a (Webster *et al.*, 2016), Rab7 (Farg *et al.*, 2014), Rab7L1 (Aoki *et al.*, 2017), Rab8 (Corbier and Sellier, 2016, Sellier *et al.*, 2016, Tang *et al.*, 2020), Rab11 (Farg *et al.*, 2014, Tang *et al.*, 2020) and Rab39b (Corbier and Sellier, 2016, Sellier *et al.*, 2016, Yang *et al.*, 2016) all of which have roles in endolysosomal and autophagy pathways. Moreover, *C9orf72* knockdown in SH-SY5Y cells resulted in impaired endocytosis and autophagy, supporting a role for C9orf72 in regulating these pathways (Farg *et al.*, 2014). Similarly, *C9orf72*

knockdown in primary cortical neurons and human cell lines resulted in inhibition of autophagy initiation (Sellier *et al.*, 2016, Webster *et al.*, 2016). It has been demonstrated that C9orf72 forms a protein complex with guanine nucleotide exchange protein Smith-Magenis Chromosome Region 8 (SMCR8) and WD repeat-containing protein 41 (WDR41) (Amick *et al.*, 2016, Blokhuis *et al.*, 2016, Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Ugolino *et al.*, 2016, Yang *et al.*, 2016, Tang *et al.*, 2020). The C9orf72-SMCR8 complex regulates autophagy via interaction with the ULK1 complex (consisting of ULK1, FIP200, ATG13, and ATG101) which controls autophagy initiation (Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Webster *et al.*, 2016, Yang *et al.*, 2016). Furthermore, through its role in regulating ULK1-mediated autophagy, C9orf72 has a cell-autonomous role in neuronal and dendritic morphogenesis (Ho *et al.*, 2019).

Other distinct functions have also been described for C9orf72 under nutrient stress or cellular stress conditions. For instance, the C9orf72-SMCR8 complex may regulate autophagy in response to nutrient availability via the mTOR pathway (Ugolino *et al.*, 2016). Indeed, loss of *C9orf72* resulted in reduced phosphorylation of the mTOR substrate S6K1 and upregulation of the transcription factor EB (TFEB), suggesting C9orf72 is also a negative regulator of autophagy (Ugolino *et al.*, 2016). Decreased mTOR signalling was also observed in *C9orf72* knockout HeLa cells (Amick *et al.*, 2016). Moreover, C9orf72 was found to localise in lysosomes under starvation conditions (Amick *et al.*, 2016). Under conditions of glucose starvation, C9orf72 was found to regulate autophagy and lipid metabolism via coactivator-associated arginine methyltransferase 1 (CARM1), an epigenetic activator of autophagy and fatty acid synthesis genes (Liu *et al.*, 2018). Finally, it has also been shown that C9orf72 may have a role in the regulation of stress granule formation (Maharjan *et al.*, 2017) as well as degradation by p62-C9orf72-mediated selective autophagy (Chitiprolu *et al.*, 2018).

C9orf72 function is important for lysosomal biogenesis and function (Amick *et al.*, 2016, Shi *et al.*, 2018). *C9orf72* KO cells were characterised by swollen and clustered lysosomes, similar to the accumulation of enlarged lysosomes in macrophages and microglia from *C9orf72* KO mice (Amick *et al.*, 2016, O'Rourke *et al.*, 2016). C9orf72 function is required for normal macrophage and microglial function in mice, whilst loss of C9orf72 leads to severe inflammatory and autoimmunity phenotypes (O'Rourke *et al.*, 2016). Interestingly, the proposed functions of C9orf72 in endolysosomal and autophagy pathways are linked to the inflammatory phenotypes observed in *C9orf72* LOF mouse models, whereby proper lysosomal and autophagy function are indispensable for macrophages and microglia mediating the phagocytosis of cellular debris. Recent evidence also suggest a role for C9orf72 in the prevention of pathological inflammatory responses to the gut immune-stimulating bacteria (Burberry *et al.*, 2020). Furthermore, C9orf72 function is essential for maintaining immune homeostasis in myeloid cells, as it has been shown that loss of C9orf72 leads to hyperactive production of type I interferons and adaptive immune activation (McCauley *et al.*, 2020).

The distinct functions of the C9-L and C9-S protein isoforms of *C9orf72* are not yet fully understood, mainly due to the unavailability of well-characterised antibodies against the C9orf72 protein isoforms. Studies have shown that C9-L is the predominantly expressed protein isoform in the human brain, whereas C9-S is undetectable (Saberi *et al.*, 2017, Frick *et al.*, 2018). C9-L is expressed in cytoplasmic puncta co-localising with lysosomes (Xiao *et al.*, 2015, Frick *et al.*, 2018). Moreover, Frick *et al.* showed that C9-L was detected at the presynapses and interacted with all members of the Rab3 family, suggesting a role in synaptic vesicle regulation (Frick *et al.*, 2018). The C9orf72 role in autophagy is probably mediated by the C9-L isoform, as this isoform can form a functional complex with SMCR8 and WDR41, and C9-L but not C9-S, could rescue autophagy dysfunction *in vitro* (Sellier *et al.*, 2016, Ho *et*

al., 2019). Another potential role for the C9-L isoform has been proposed in the regulation of stress granule assembly (Maharjan *et al.*, 2017). Xiao *et al.* reported that C9-S is localised specifically at the nuclear membrane in neurons, and is translocated to the plasma membrane in C9-ALS motor neurons (Xiao *et al.*, 2015). Forced expression of both C9-L and C9-S isoforms could restore lysosomal biogenesis impairments in both *C9orf72* iPSC-MNs and *C9orf72* deficient (*C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$) iPSC-MNs, suggesting a potential role for both isoforms in the lysosomal biogenesis pathway (Shi *et al.*, 2018).

1.3. iPSC modelling of C9-FTD/ALS

1.3.1. Patient-derived iPSCs in disease modelling

Human iPSCs, first described over a decade ago, were the result of cellular reprogramming of human fibroblasts via a minimum set of four *bona fide* reprogramming factors; Octamer-binding transcription factor-3/4 (Oct3/4), Sex-determining region Y-box 2 (Sox2), Kruppel Like Factor-4 (Klf4), and c-Myc (Takahashi *et al.*, 2007). The discovery of iPSCs was the result of the seminal work by Takahasi and Yamanaka on the genes responsible for the maintenance of pluripotency in embryonic stem cells (ESCs) which are stem cells derived from the inner cell mass (ICM) of a blastocyst of pre-implantation stage. ESCs are characterised by (i) pluripotency, their ability to differentiate into cell types from all three germ layers (endoderm, medoserm and ectoderm) and (ii) limitless self-renewal capacity. Takahasi and Yamanaka's work on mouse ESCs led to a list of 24 candidate genes which were subsequently screened in different combinations and through a series of elimination rounds. The result was a minimal set of four genes (*Oct3/4*, *Klf4*, *Sox2* and *c-Myc* termed OKSM) sufficient to reprogram MEFs into "induced" PSCs (iPSCs) which are indistinguishable from ESCs in their capacity for self-renewal and

differentiation to all the three germ layers (endoderm, mesoderm, ectoderm). This initial four factor overexpression is sufficient to induce the expression of pluripotency networks and other typical ESC-specific markers, such as the transcription factor NANOG and several cell surface markers such as SSEA-4, TRA-1-60 and TRA-1-81 and supress lineage-specific markers.

The use of human induced pluripotent stem cells (iPSCs) has revolutionised the field of neurological disease research in particular, as it enabled the study of live patient-derived neurons or glia and allowed observations of the molecular mechanisms underlying neurological diseases as they develop. This is particularly true in the field of *C9orf72* research where iPSC studies have accelerated research developments by shedding light to novel disease mechanisms as well as previously unknown functions of the C9orf72 protein. *C9orf72* patient-derived iPSC models offer the advantage of fully recapitulating the pathological repeat expansion size present in *C9orf72* HRE carriers, therefore, allowing faithful disease modelling without relying on overexpression of transgenes. This is crucial for the detection of physiological levels of RNA foci, DPRs and *C9orf72* transcripts in *C9orf72* iPSC-neurons and assessment of the variation in these levels between different *C9orf72* HRE carriers.

Nonetheless, as with most model systems, there are some challenges in the use of iPSCs to model C9-FTD/ALS. The *C9orf72* HRE is prone to genomic instability (Reddy *et al.*, 2014, Abu Diab *et al.*, 2018) leading to HRE mosaicism (Beck *et al.*, 2013) which is an additional source of genetic heterogeneity in iPSC cultures. Moreover all three potential disease mechanisms leading to neurodegenerative phenotypes co-exist in patient-derived iPSC models, and although this adds to their physiological relevance, the contribution of each to cellular phenotypes cannot be easily dissected. The use of genome editing tools such as CRISPR/Cas9 can be used to dissect GOF vs LOF mechanisms in isogenic iPSC models that share a common genetic background. Finally, whilst iPSC models offer excellent systems for

investigating developmental phenotypes, they are limited in their capacity to fully model phenotypes associated with ageing, as they are strictly tied to developmental cues. Indeed, most iPSC models of neurodegenerative diseases mostly capture early disease phenotypes, thus additional stressors or accelerated *in vitro* ageing are needed to induce further disease-related phenotypes.

1.3.2. Genome editing in iPSCs

During the past decade, several genome editing technologies have emerged that have been extensively used in the field of iPSC disease modelling as they offer programmable, highly efficient and precise genome editing. The generation of isogenic cell lines that are genetically engineered versions of the same parental line are very useful tools in disease modelling and have been widely used to study the effects of specific mutations or gene knockouts. The main advantages of using isogenic iPSC lines in disease modelling are (i) elimination of genetic heterogeneity-related noise, (ii) dissection and study of the net effect of a proposed pathogenic mutation, (iii) robust and reproducible phenotyping, and (iv) significant reduction of the number of experimental iPSC controls needed for studies.

The key in genome editing technologies are engineered nucleases which are fused to DNA-binding domains or recruited to the DNA target by small RNA sequences. The way these programmable nucleases mediate genome editing is by generating double stand breaks (DSBs) in the DNA. DSBs are highly recombinogenic and trigger two main cellular repair mechanisms; non-homologous end-joining (NHEJ) that generates random insertions/deletions (indels) at the DNA break site or homologous recombination (HR) also referred to as homology directed repair (HDR) that incorporates donor DNA with homologous sequences at the cleavage site (Liang *et al.*, 1998, Symington and Gautier, 2011) (Figure 1.6 A). When nucleases are

designed to target essential exons of genes, NHEJ can generate indels that may cause frameshift and lead to loss of protein function, an application widely used for the generation of gene knockout cell or animal models.

The main three programmable nuclease technologies are Zinc Finger Nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs) and Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 Nucleases systems (CRISPR/Cas9) (Kim and Kim, 2014). ZFNs have been the first and most extensively studied technology (Bibikova *et al.*, 2003, Urnov *et al.*, 2005). In ZFNs, FokI endonuclease is fused to Zing Finger DNA binding motifs (Figure 1.6 B). As FokI needs to dimerize in order to produce a DSB, ZFNs are designed in pairs in order to target a genomic region (Bitinaite *et al.*, 1998). TALENs are based on naturally occurring TAL effectors that are released by plant pathogenic bacteria of the genus *Xanthomonas* upon plant infection, translocate to the plant cell nucleus and directly bind DNA to modulate expression of genes that facilitate bacterial survival and colonisation (Gu *et al.*, 2005, Kay *et al.*, 2007, Römer *et al.*, 2007). The DNA binding specificity of TAL Effectors is set by the repeat-variable di-residues (RVDs), whereby the RVDs correspond directly to the nucleotides in their target sites, one RVD to one nucleotide (Boch *et al.*, 2009, Moscou and Bogdanove, 2009). In TALENs, TAL effector domains are fused to FokI catalytic domain, and similarly to ZFNs, they are used in pairs to target DNA sequences (Miller *et al.*, 2011, Zhang *et al.*, 2011) (Figure 1.6 C). The newest addition to the designer nuclease technologies is the CRISPR/Ca9 Class 2, Type II system which is a naturally occurring adaptive immunity defence mechanism in bacteria (Figure 1.6 D) (Barrangou *et al.*, 2007, Jinek *et al.*, 2012). Upon infection of bacteria by viruses, foreign DNA is cleaved into short sequences and stored in the bacterial genome in the form of CRISPRs. CRISPRs are transcribed to RNAs and act as memory to fight viruses upon re-infection by guiding Cas9 nuclease to cleave the foreign DNA. The natural system consists of 3 basic

components, Cas9 nuclease, CRISPR targeting RNA (crRNA) and trans-activating crRNA (tracrRNA) but in the engineered system crRNA and tracrRNA are fused to a single guide RNA (sgRNA). The Cas9-sgRNA complex recognise DNA target sites that consist of an NGG protospacer adjacent motif (PAM) sequence, recognised by Cas9, and an adjacent 20 nucleotide protospacer sequence complementary to the sgRNA (Deltcheva *et al.*, 2011, Jinek *et al.*, 2012, Sternberg *et al.*, 2014).

Genome editing studies using CRISPR/Cas9 system have reported a high incidence of off-target effects, which is a disadvantage when scarless engineering applications are required (Cradick *et al.*, 2013, Fu *et al.*, 2013). As a solution to this, engineered versions of S. pyogenes Cas9 were generated by single point mutations in either one of the enzyme's two catalytic domains, RuvC and HNH. Each domain is responsible for cleavage of one DNA strand when DSB by Cas9 occurs. The D10A mutation in HNH domain generates a RuvC-active/HNH-dead Cas9 whereas the H840A mutation in RuvC domain causes Cas9 to have only HNH catalytic activity (Mali *et al.*, 2013, Ran *et al.*, 2013a). These engineered Cas9 versions have lost the ability to generate DSBs but nick the DNA strands instead (nickase Cas9). By pairing two sgRNAs targeting sequences close enough to each other the nicks generated from nCas9 can be turned to a DSB (Mali *et al.*, 2013, Ran *et al.*, 2013a). This paired sgRNA strategy has significantly reduced off-target issues related to CRISPR technology (Cho *et al.*, 2014). Other strategies to reduce off-target effects and increase specificity of CRISPR editing include: truncated guide RNAs (Fu *et al.*, 2014), engineered Cas9 versions with altered PAM specificity (Kleinstiver *et al.*, 2015), dimeric fusions of catalytically inactive SpCas9 to FokI nucleases (Guilinger *et al.*, 2014, Tsai *et al.*, 2014, Wyvekens *et al.*, 2015) and a high-fidelity Cas9 variant, Cas9-HF1, designed to reduce non-specific DNA contacts (Kleinstiver *et al.*, 2016).

Rapid developments on the field have led to the characterisation of a novel Class 2, Type V CRISPR/Cpf1 system of the bacterium *Francisella novicida* using the RNA-

guided endonuclease Cpf1 instead of Cas9 (Zetsche *et al.*, 2015). The CRISPR/Cpf1 system has some key differences from CRISPR/Cas9, based on some unique features of Cpf1 that: (i) is a single-crRNA-guided endonuclease, and unlike Cas9, it does not require tracRNA to target DNA, (ii) can produce staggered cuts with 5′ overhangs as opposed to the blunt ends generated by Cas9, (iii) cleaves target DNA at the proximal end of the PAM, far away from the seed region, and unlike Cas9, it allows subsequent rounds of Cpf1 targeting, as the target site is preserved and not disrupted by NHEJ, and (iv) relies on T-rich PAM motifs as opposed to the G-rich PAMs recognised by Cas9.

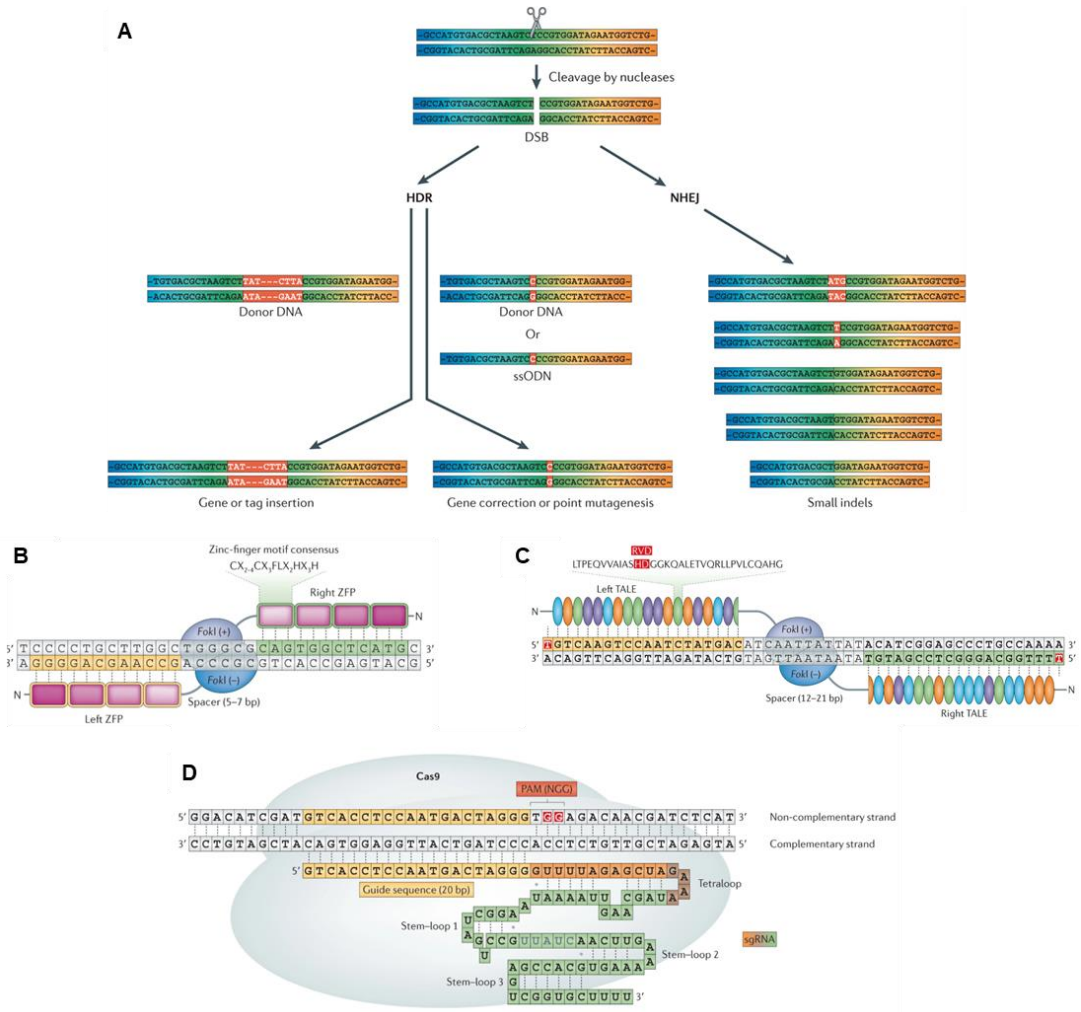


Figure 1.6 Overview of genome editing technologies.

A. Double strand breaks (DSBs) in DNA trigger two main cellular repair mechanisms; non-homologous end joining (NHEJ) that generates small indels at the DSB and homology directed repair (HDR) that inserts homologous donor DNA at the DSB site. B. Schematic representation of a ZFN pair directing FokI endonuclease at the target genomic locus. DNA recognition is mediated by zing-finger motifs. C. Schematic representation of a TALEN pair. Each TALEN consists of a range of up to 20 TAL effector domains fused to FokI endonuclease semi-domains. DNA specificity is mediated by RVDs in TAL effectors and DSBs only occur upon binding of both TALENs and dimerization of FokI. D. RNA-guided engineered nucleases (RGENs) are a different category of programmable nucleases where DNA specificity is mediated by small guide RNAs guiding Cas9 DNA endonuclease at the desired genomic locus. Modified from (Kim and Kim, 2014).

1.3.3. Cortical neuron development

1.3.3.1. Overview of the early human neurodevelopment

Early human neuronal development occurs via progressive restriction in fate potential in a tightly regulated manner. The earliest event in the development of the human nervous system is the neural induction, whereby a region of the embryonic ectoderm, one of the three germ layers generated during gastrulation, is specified to form the neural plate. Neural induction is achieved by molecules inhibiting bone morphogenetic protein (BMP) and Activin/Nodal signalling which are collectively members of the TGF-β signalling superfamily activating Smad signal transduction proteins. The activation of the fibroblast growth factor (FGF) and WNT signalling pathways has also been proposed to play a role in neural induction by further inhibiting BMP signalling (Baker *et al.*, 1999, Pera *et al.*, 2004), however their precise roles in neural induction remain controversial. Following neural induction, the neuroepithelium of the neural plate folds to form the neural tube, from which the brain and spinal cord develop, during a process termed neurulation. Specification of subtypes of neural progenitor cells (NPCs) along the dorsal-ventral (D-V) and anterior-posterior axes of the neural tube is achieved by the coordinated spatiotemporal gradient of morphogens during the neural patterning process. The specification of NPCs along the A-P axis is coordinated by gradients of Nodal, BMP4, RA (retinoic acid), and FGFs, whereby forebrain development is specified by BMP antagonism and lack of additional cues, whereas the most posterior parts are specified by highest concentrations of these morphogens (Zirra *et al.*, 2016). The morphogen gradient on the A-P axis is responsible for the activation of homeobox (HOX) genes that specify positional identity during development (Philippidou and Dasen, 2013). The specification of the NPCs along the D-V axis relies on opposing gradients of sonic hedgehog (SHH), secreted from the floor plate of the ventral neural tube, and Wnt and BMP, secreted from the roof plate of the dorsal neural tube (Zirra

et al., 2016). In order to generate the variety of neuronal and glial subtypes in the human cortex, NPCs undergo a series of temporally regulated cellular divisions. These divisions are initially symmetric, generating more NPCs to expand the stem cell pool. However, over time the potential of NPCs for proliferation changes, leading to temporally regulated asymmetric divisions generating first deep-layer neurons, then upper-layer neurons, and finally glial cells.

1.3.3.2. Human cortical neuron development

Early progenitor populations in the developing neocortex are generating the diverse cellular repertoire of neuronal and glial subpopulations in a tightly controlled spatiotemporal manner. In all mammals, the cortex consists of two main classes of neurons, excitatory glutamatergic projection neurons (80%) and inhibitory GABAergic interneurons (20%) (Hendry *et al.*, 1987, Beaulieu, 1993). A balance between the excitatory and inhibitory neurons of the cortex is critical for its normal function.

All excitatory neurons are generated from cortical neuron progenitors located in the dorsal telencephalon, within the anterior part of the neural tube. In early stages of corticogenesis, symmetrical divisions of the neural progenitors forming the pseudostratified neuroepithelium occur that enrich the pool of early progenitor neurons. As neurogenesis begins, the neuroepithelium gives rise to radial glial cells (RGCs) which proliferate in the ventricular zone (VZ) via interkinetic nuclear migration, in which the mitotic nuclei migrate between the basal and apical boundaries of the neuroepithelium and divide at the latter. As neurogenesis progresses, RGCs divide asymmetrically to generate cortical projection neurons (PNs), however, this generates ~10-20% of all PNs. The majority of PNs (~80% of all cortical PNs), are generated by intermediate cortical progenitors which are derived from RGCs' symmetrical divisions and proliferate away from the VZ, forming a new germinal zone,

the subventricular zone (SVZ). The intermediate precursors of the SVZ divide symmetrically to produce cortical PNs. The pool of cortical neurons is further enriched by additional progenitors called subapical progenitors (SAPs) (Pilz *et al.*, 2013). The vast majority of cortical interneurons are generated in the ganglionic eminences in the ventral telencephalon and migrate tangentially to reach the developing cortex (Zhu *et al.*, 1999, Hansen *et al.*, 2013).

Cortical interneurons and projection neurons are distributed into 6 different cortical layers (I-VI) of highly organised cytoarchitecture; the supragranular (layers I/II–III), the granular (layer IV), and the infragranular (layers V and VI). Each layer has a unique architecture, containing different neuronal subtypes with distinct gene expression and connectivity patterns where neurons form connections within and across cortical areas but also extra-cortical connections. The establishment of the 6 cortical layers is tied to a developmental process of temporal patterning, in which neurons of the deep layers are generated first, followed by neurons of the upper cortical layers.

1.3.3.3. In vitro cortical neuron differentiation

A prerequisite for faithful modelling of FTD and similar diseases that affect the frontal cortex is to understand human cortical development and employ cellular models that best mimic the developmental processes and capture the cellular complexity *in vitro*.

Interestingly, despite the huge complexity of the human cortex, *in vitro* neuronal differentiation methods are able to mimic the complex developmental and layering processes *in vitro*, capturing to a great extend its cellular diversity. The paradoxical simplicity by which such a complex neuronal culture can be achieved, even in the reductionist 2D form, lies in the simplicity of forebrain induction itself and in the self-organising properties of the cortical progenitor cells (Watanabe *et al.*, 2005, Eiraku *et*

al., 2008, Sasai, 2013). In the lack of any cues, iPSCs or ESCs spontaneously differentiate by default into neurons with mostly forebrain identity (Gaspard *et al.*, 2008).

Telencephalic identity is achieved by dual SMAD inhibition whereby two small molecule inhibitors Dorsomorphin and SB431542 applied on iPSCs or ESCs for 8-11 days act synergistically to induce a neuroepithelium with anterior positional identity (Chambers *et al.*, 2009, Shi *et al.*, 2012a). More specifically, SB431542 inhibits TGF-β-mediated activation of SMAD proteins and Dorsomorphin inhibits Bone morphogenetic protein (BMP)-mediated activation of SMADs. Replacement of noggin with Dorsomorphin in the dual SMAD inhibition protocol has been shown to reduce variation in cortical differentiation efficiency, with the protocol by Shi *et al.* featuring an efficiency of up to 100% (Kim *et al.*, 2010a, Shi *et al.*, 2012a). Several protocols have been described for the *in vitro* differentiation of cortical neurons which are summarised in Table 1.1.

The *in vitro* differentiation recapitulates human cortical development. Early in differentiation and following the initial cortical induction, polarised cortical rosette structures form, featuring interkinetic nuclear migration typical of the proliferation occurring in the ventricular zone of the developing cortex *in vivo* (Zhang *et al.*, 2001, Elkabetz *et al.*, 2008). According to the protocol described by Shi *et al*., which is employed in this Thesis, rosettes appear *in vitro* at ~day 12-14 and are positive for cortical primary progenitor markers such as paired box 6 (PAX6), forkhead box protein G1 (FOXG1) and orthodenticle homeobox 2 (OTX2) (Shi *et al.*, 2012a). The cycling progenitor cells within the rosettes are undergoing interkinetic nuclear migration. Consistent with the “inside-out” development of cortical layers, deep layer neurons are the first post-mitotic neurons to appear in the periphery of the rosettes as early as day 20 and can be identified by the expression of the deep layer markers TBR1 and CTIP2. Upper layer neurons appear later at day 70 and are positive for

BRN2 and SATB2 markers. Astrocytes expressing GFAP, GLAST and S100B markers are also generated in the cultures after day 50, but their proportion gradually increases over time *in vitro*. The whole cortical neuron differentiation takes approximately 80-100 days *in vitro*, closely resembling the cortical development *in utero* (Shi *et al.*, 2012a). Moreover, the *in vitro* generated cortical neurons form functional networks and are electrically active (Shi *et al.*, 2012a).

However, it is noteworthy that such differentiation protocols do not fully recapitulate the repertoire of neuronal classes since they consist primarily of glutamatergic projection neurons and lack GABAergic interneurons which originate from the ventral telencephalon and migrate tangentially to the neocortex during development (Zhu *et al.*, 1999, Hansen *et al.*, 2013).

Table 1.1 Published protocols for directed differentiation of cortical neurons (Modified from Zirra *et al.* 2016).

Study	**Cell Type**	**Compounds for neuronal induction**	**Duration (days)**
Watanabe *et al.* 2005	Cortical precursors	BMPRIA-Fc - BMP antagonist LeftyA - Activin/Nodal antagonist Dkk1 - Wnt antagonist	35
Eiraku *et al.* 2008	Cortical neurons	BMPRIA-Fc - BMP antagonist LeftyA - Activin/Nodal antagonist Dkk1 - Wnt antagonist	45-60
Chambers *et al.* 2009	Cortical neurons	Noggin - BMP antagonist SB431542 - Activin/Nodal antagonist	19
Li *et al.* 2009	Cortical neurons	Endogenous Wnt	30-35
Shi *et al.* 2012	Cortical neurons	Dorsomorphin - BMP antagonist SB431542 - Activin/Nodal antagonist	80-100

1.3.4. Insights from *C9orf72* iPSC models

1.3.4.1. Genomic instability and mosaicism of the C9orf72 HRE in iPSC models

The *C9orf72* HRE is characterised by genomic instability leading to mosaicism in carriers that is evident by the different HRE lengths not only between brain and blood, but also within different brain regions (Beck *et al.*, 2013, van Blitterswijk *et al.*, 2013b, Fratta *et al.*, 2015, Nordin *et al.*, 2015, Gijselinck *et al.*, 2016). Considering the HRE instability, an important step in the generation of iPSC-neurons for faithful disease modelling of C9-FTD/ALS, is the characterisation of the HRE length in cell cultures. Most published studies to date have carried out a characterisation of the HRE length, mainly via Southern blotting. Among these studies, several have reported HRE instability, evident as an expansion or contraction in HRE size during various differentiation stages that may result in mosaic cultures. In one of these, Almeida *et al.* differentiated iPSC-CNs from two *C9orf72* HRE carrier lines (2 clones per individual, expansion sizes: ~1000 and ~ 1.600/730/650 repeats) (Almeida *et al.*, 2013). Characterisation of the *C9orf72* HRE via Southern blotting revealed genomic instability of the expanded repeat, with both expansion and contraction in iPSCs following reprogramming, and contraction during differentiation of both lines to iPSC-CNs. In one of the lines, mosaicism was evident in all cell types, from parental fibroblasts, to iPSCs and differentiated CNs. Sareen *et al.* differentiated four ALS *C9orf72* HRE iPSC lines into MNs (expansion sizes: 3 lines ~800, 1 line ~70 repeats) (Sareen *et al.*, 2013). The *C9orf72* HRE was characterised by genomic instability, specifically somatic expansion, during reprogramming of two fibroblast lines to iPSCs and contraction of one line during differentiation to MNs. Dafinca *et al.* generated iPSC-CNs and iPSC-MNs from three *C9orf72* HRE patients (3 lines, 2 clones each, repeat expansion sizes: for line 1 ~510/690 and ~420–640, for line 2 ~1,210 and ~1,380, and for line 3 ~1,000 repeats) (Dafinca *et al.*, 2016). Similar to previous studies, mosaicism for different *C9orf72* HRE repeat lengths was observed via

Southern blotting in MNs derived from some of the patient lines. In another study by Esanov *et al.*, iPSCs from one *C9orf72* HRE patient were differentiated to MNs (Esanov *et al.*, 2016). *C9orf72* HRE repeat instability and subsequent mosaicism were evident across all stages from the initial mosaic pool of immortalised lymphocytes, to a contraction in iPSCs, an expansion in neuron precursors and finally further expansion in iPSC-MNs. Finally, Bardelli *et al.* also reported that HRE mosaicism was present in fibroblasts and peripheral blood cells from a *C9orf72* HRE carrier as well as in the generated iPSCs from these tissues (Bardelli *et al.*, 2020).

1.3.4.2. Patient-derived iPSC-neurons recapitulate pathological hallmarks of C9-FTD/ALS

Several *C9orf72* iPSC studies have emerged showing that iPSC-neurons from patients with the *C9orf72* HRE recapitulate major pathological hallmarks of C9-FTD/ALS, including RNA foci and DPRs. Sense and/or antisense RNA foci have been detected in the nuclei of *C9orf72* patient-derived CNs (Almeida *et al.*, 2013, Simone *et al.*, 2018, Yuva-Aydemir *et al.*, 2019), mixed neurons (Donnelly *et al.*, 2013), MNs (Sareen *et al.*, 2013, Dafinca *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Selvaraj *et al.*, 2018, Simone *et al.*, 2018, Ababneh *et al.*, 2020) and astrocytes (Zhao *et al.*, 2019b). Furthermore, it has been demonstrated that the RNA foci sequester several RNA-binding proteins in C9-FTD/ALS (Lee *et al.*, 2013, Mori *et al.*, 2013b, Cooper-Knock *et al.*, 2014). Several iPSC studies have also investigated the expanded RNA-mediated toxicity and have shown that RNA foci can sequester RNA-binding proteins including RNA-editing deaminase-2 (ADARB2) (Donnelly *et al.*, 2013 heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1) and purine-rich binding protein-α (Pur-α) (Sareen *et al.*, 2013) as well as other proteins such as Ran GTPase activating Protein 1 (RANGAP1) (Zhang *et al.*, 2015). It is not known whether this expanded RNA-mediated toxicity is the major driver of neurodegeneration in C9-FTD/ALS similar to what is known for the sequestration of muscleblind proteins by the CUG repeats in myotonic dystrophy (Jiang *et al.*, 2004).

Similarly, several DPRs, mainly the sense-derived poly-GP, poly-GR and poly-GA, have been identified in iPSC-MNs, iPSC-CNs and iPSC-astrocytes (Almeida *et al.*, 2013, Donnelly *et al.*, 2013, Dafinca *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Westergard *et al.*, 2016, Selvaraj *et al.*, 2018, Simone *et al.*, 2018, Almeida *et al.*, 2019, Yuva-Aydemir *et al.*, 2019, Zhao *et al.*, 2019b, Ababneh *et al.*, 2020, Abo-Rady *et al.*, 2020). Initial difficulties in DPR detection using conventional methods such as dot blot, western blot or immunocytochemistry were overcome by the gradual development of better-characterised anti-DPR antibodies and the use of more sensitive detection methods such as Meso Scale Discovery (MSD) immunoassay. Interestingly, an iPSC study has provided evidence for cell-to-cell transmission of DPRs, by demonstrating transmission of poly-GA and poly-GR to control iPSC-MNs using co-cultures and conditioned media from *C9orf72* iPSC-MNs (Westergard *et al.*, 2016).

Whilst RNA foci and DPRs have been a common pathological finding among iPSC studies, other major neuropathological hallmarks of C9-FTD/ALS, such as p62 inclusions co-localising with DPRs or cytoplasmic TDP-43 aggregates, have been less commonly reported. For instance, a few iPSC studies have found elevated p62 levels in *C9orf72* iPSC-neurons (Almeida *et al.*, 2013, Dafinca *et al.*, 2016), whereas p62 cytoplasmic inclusions have been detected under basal conditions (Dafinca *et al.*, 2016) and upon chronic sodium arsenite stress in *C9orf72* iPSC-MNs (Ratti *et al.*, 2020). Similarly, for TDP-43 no robust cytoplasmic inclusion pathology, typical of what is seen in C9-FTD/ALS patients, has been reported under basal conditions in *C9orf72* iPSC-neurons, only early-stage TDP-43 nucleocytoplasmic mislocalisation (Zhang *et al.*, 2015, Yuva-Aydemir *et al.*, 2019). However, upon chronic, mild oxidative stress insult by sodium arsenite treatment, *C9orf72* iPSC-MNs exhibited cytoplasmic TDP-43 aggregates that co-localised with stress granule markers as well as distinct cytoplasmic aggregates of phosphorylated TDP-43 (Ratti *et al.*, 2020).

In conclusion, HRE-related RNA foci and DPR pathology has been widely reported by iPSC studies, however, the typical TDP-43 pathology in C9-FTD/ALS has not been yet robustly recapitulated *in vitro*. This may reflect the relative immaturity of iPSC model systems that are tied to a strict developmental schedule and may not develop robust TDP-43 pathology without the application of additional stressors or extended neuronal maturation *in vitro* to account for the lack of ageing factor. The above findings from iPSC studies also confirm that RNA foci and DPR pathologies precede that of TPD-43 (Baborie *et al.*, 2014).

1.3.4.3. C9orf72 haploinsufficiency and DNA methylation in C9orf72 iPSC models

C9orf72 haploinsufficiency is evident via a reduction in *C9orf72* transcript levels in patient brain tissue and may cause C9-FTD/ALS via a loss of function mechanism. Several iPSC studies have reported a reduction in total *C9orf72* mRNA (Ababneh *et al.*, 2020) as well as in individual *C9orf72* transcript levels, mainly in transcript V2 but also in V1 (Almeida *et al.*, 2013, Donnelly *et al.*, 2013, Shi *et al.*, 2018, Ababneh *et al.*, 2020), in *C9orf72* iPSC-neurons compared to controls. However, others did not detect any reduction in *C9orf72* mRNA expression in *C9orf72* iPSC-neurons (Sareen *et al.*, 2013). Interestingly, some reports have also demonstrated increased expression of intron 1-retaining transcripts, containing the HRE, in *C9orf72* iPSC-neurons compared to controls (Cohen-Hadad *et al.*, 2016, Ababneh *et al.*, 2020). Furthermore, some studies have reported a reduction in C9orf72 protein levels, mainly detected as the C9-L protein isoform, in *C9orf72* iPSC-neurons compared to control neurons (Sivadasan *et al.*, 2016, Aoki *et al.*, 2017, Shi *et al.*, 2018), while other studies did not observe any differences (Sareen *et al.*, 2013, Ababneh *et al.*, 2020).

DNA hypermethylation of the *C9orf72* promoter region, consisting of the *C9orf72* HRE and its flanking CpG islands, is a well described mechanism that can contribute to the *C9orf72* haploinsufficiency via epigenetic silencing of the gene. A hypermethylated *C9orf72* promoter is frequently detected in patient brain and blood (Xi *et al.*, 2013,

Belzil *et al.*, 2014, Xi *et al.*, 2014a, Xi *et al.*, 2015b) and is a potential disease modifier in C9-ALS/FTD (Russ *et al.*, 2014).

The *C9orf72* promoter methylation status has also been assessed in *C9orf72* patient-derived iPSC models (Cohen-Hadad *et al.*, 2016, Esanov *et al.*, 2016, Ababneh *et al.*, 2020). In one of these studies, assessment of the *C9orf72* promoter DNA methylation and hydroxymethylation revealed a reduction in 5-methylcytosine (5mC) levels and a concomitant increase in 5-hydroxymethylcytosine (5hmC) levels following reprogramming (Esanov *et al.*, 2016). 5mC levels were re-acquired upon neuronal specification whereas 5hmC levels were elevated in iPSCs and MNs. However, some limitations of this study include the use of only one patient line and the considerable heterogeneity due to different HRE sizes within and between cellular populations of different developmental stages (as described in 1.3.4.1.). DNA methylation at the *C9orf72* promoter region, including the C9orf72 HRE and the upstream 5'CpG island, was also assessed in differentiated neural precursors and teratomas derived from *C9orf72* HRE human embryonic stem cell (hESC) (~270 repeats) and iPSC lines (haploidentical line from mother ~700 repeats and unrelated donor ~2,700 repeats)(Cohen-Hadad *et al.*, 2016). The 5' CpG and the *C9orf72* HRE were found to be hypermethylated in the *C9orf72* HRE iPSC-NPCs and teratomas. It was therefore suggested that reprogramming leads to hypermethylation of the *C9orf72* promoter region in iPSCs that can lead to milder phenotypes in iPSCs compared to ESCs. This is supported by the finding of increased intron 1-retaining *C9orf72* transcripts in NPCs and teratomas differentiated from *C9orf72* HRE hESCs compared to iPSCs. However, the considerable difference in the HRE size between the iPSC and hESC lines is another important factor that could influence the DNA methylation in *C9orf72* promoter. Finally, another more recent study reported a hypermethylated 5'CpG island in iPSC-MNs differentiated from two iPSC lines derived from one *C9orf72* HRE

carrier (Ababneh *et al.*, 2020). The 5'CpG hypermethylation was abolished in isogenic CRISPR/Cas9 HRE-corrected iPSC-MNs.

1.3.4.4. Novel disease mechanisms emerging from C9orf72 iPSC models

Importantly, apart from recapitulating C9-FTD/ALS pathology, several iPSC studies have shed light on novel disease mechanisms linked to *C9orf72* HRE, in some cases serving as a means to confirm the relevance of findings from *in vivo* studies in a human non-overexpression model.

Compromised autophagy

One of the first identified phenotypes in *C9orf72* iPSC-neurons was compromised autophagy (Almeida *et al.*, 2013). *C9orf72* iPSC-CNs were more susceptible to cell death induced by autophagy inhibitors and exhibited elevated levels of the autophagy adaptor p62 (Almeida *et al.*, 2013), whereas basal levels of autophagy were reduced in *C9orf72* induced neurons (Webster *et al.*, 2016). Elevated p62 levels as well as p62 aggregates, a common pathological finding in C9-FTD/ALS, were detected under basal conditions in *C9orf72* CNs and MNs, suggesting autophagy impairment (Dafinca *et al.*, 2016). Finally, p62 aggregation was also detected in *C9orf72* iPSC-MNs upon chronic stress induction by sodium arsenite treatment (Ratti *et al.*, 2020).

Impaired nucleocytoplasmic transport

Nucleocytoplasmic transport has been found by several independent groups to be impaired in C9-FTD/ALS (Freibaum *et al.*, 2015, Jovicic *et al.*, 2015, Zhang *et al.*, 2015, Cheng *et al.*, 2019, Moore *et al.*, 2019). Ran GTPase-activating protein 1 (RanGAP1), involved in the regulation of nucleocytoplasmic transport, was found to interact with the HRE RNA and co-localise with RNA foci and other protein

components of the nuclear pore complex (NPC), such as nucleoporin 205 (Nup205), a NPC scaffold protein in *C9orf72* iPSCNs (Zhang *et al.*, 2015). Loss of RanGAP1 function was linked to disrupted nuclear-cytoplasmic pattern of Ran protein, leading to higher cytoplasmic levels of the protein in *C9orf72* iPSCNs compared to controls, which was rescued by RanGAP1 overexpression. Impaired nucleocytoplasmic transport was responsible for the abnormal nuclear/cytoplasmic ratios of Ran and TDP-43 which were restored after ASO treatment in *C9orf72* iPSCNs. Nucleocytoplasmic transport phenotypes were also reported by Freibaum *et al.* who found a 35% increase in the nuclear/cytoplasmic ratio of RNAs in *C9orf72* iPSC-CNs compared to controls (Freibaum *et al.*, 2015). Moreover, Jovicic et al. investigated the integrity of nucleocytoplasmic transport in human induced neurons (iNs) directly converted from fibroblasts of *C9orf72* HRE carriers and controls (Jovicic *et al.*, 2015). They reported a disrupted nuclear/cytoplasmic localization of the Ran-GEF RCC1 with decreased nuclear localisation of RCC1 in *C9orf72* iNs compared to controls. Furthermore, the RNA editing enzyme adenosine deaminase acting on RNA 2 (ADAR2) was found to be mislocalised to the cytoplasm of *C9orf72* iPSC-MNs (Moore *et al.*, 2019). Several factors of the NXF1-NXT1 mRNA nuclear export machinery were identified as genetic modifiers of *C9orf72* HRE-mediated toxicity by Cheng *et al.* (Cheng *et al.*, 2019). Specifically, the RNA helicase DDX3X was identified as a repressor of *C9orf72* HRE RAN translation and elevated DDX3X levels led to a decrease in DPRs, rescuing *C9orf72* HRE phenotypes of glutamate-induced excitotoxicity and disrupted nucleocytoplasmic transport in *C9orf72* iPSCNs. Finally, another way that the *C9orf72* HRE may lead to nucleocytoplasmic transport disruption is via DRP-/mutant TDP-43-induced cellular stress and the formation of stress granules that co-localise with nucleocytoplasmic transport factors (Zhang *et al.*, 2018a). Treatment with stress granule inhibitors GSK, ISRIB, and Ataxin-2 ASOs were able to suppress nucleocytoplasmic transport defects and neurodegeneration in *C9orf72* iPSCNs.

DNA damage, oxidative stress, ER stress and mitochondrial dysfunction

Several reports have revealed dysregulated cellular pathways in *C9orf72* iPSC-neurons such as increased DNA damage, oxidative and ER stress, and mitochondrial dysfunction (Dafinca *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Andrade *et al.*, 2020). Specifically, it has been shown that *C9orf72* iPSC-MNs exhibit elevated ER calcium levels, reduced mitochondrial membrane potential and reduced levels of the antiapoptotic protein Bcl-2 compared to control MNs (Dafinca *et al.*, 2016). In addition to this, *C9orf72* iPSC-CN and iPSC-MNs also displayed increased susceptibility to apoptosis, elevated p62 levels, abnormal protein aggregation and stress granule formation compared to control neurons (Dafinca *et al.*, 2016). Poly-GP was found to cause mitochondrial dysfunction, age-dependent increase in oxidative stress and DNA damage, as indicated by an increase in DNA damage marker γH2AX, in *C9orf72* iPSC-MNs (Lopez-Gonzalez *et al.*, 2016). Reduction of oxidative stress partially reduced DNA damage in *C9orf72* iPSC-MNs, suggesting that oxidative stress could play an important role in the disease pathogenesis and its reduction has therapeutic potential in C9-FTD/ALS. Finally, increased DNA damage was also reported by Andrade *et al.* who reported increased levels of DNA damage marker γH2AX as well as RAD5, a component of the SSA repair machinery, and phosphorylated RAD52, *C9orf72* iPSC-MNs (Andrade *et al.*, 2020). Upon CRISPR/Cas9-mediated excision of the HRE, RAD52 hyperactivation was reduced in isogenic HRE-corrected MNs. These findings suggest HRE-mediated DNA damage in *C9orf72* iPSC-MNs, leads to deficits in homology-directed DNA double strand break (DSB) repair pathways.

Altered excitability and excitotoxicity

Neuronal excitability impairments are commonly found in *C9orf72* HRE carriers (Williams *et al.*, 2013). Several iPSC studies have also reported altered excitability in *C9orf72* iPSC-neurons. Specifically, two studies showed that *C9orf72* iPSC-MNs exhibited hyperexcitability compared to control MNs at 2 to 4 weeks post

differentiation (Wainger *et al.*, 2014, Devlin *et al.*, 2015). However, Sareen *et al.* reported loss of excitability in 2-month-old *C9orf72* iPSC-MNs as well as altered expression of genes involved in membrane excitability, including the delayed rectifier potassium channel (KCNQ3), which is consistent with hypoexcitability (Sareen *et al.*, 2013). These conflicting results may be attributed to the different developmental stage of the MNs, as revealed by temporal analysis of *C9orf72* iPSC-MN excitability (Devlin *et al.*, 2015). Specifically, Devlin *et al.* showed that *C9orf72* iPSC-MNs were characterised by intrinsic hyperexcitability at early time points (3-4 weeks) *in vitro*, followed by progressive loss of action potential output and synaptic activity in MNs reaching 9-10 weeks *in vitro*. Of note, the loss of excitability manifested at a similar timepoint during *in vitro* MN differentiation as in the study of Sareen *et al.* (Sareen *et al.*, 2013). In contrast to the above studies, Selvaraj *et al.* reported no differences in the excitability of *C9orf72* iPSC-MNs compared to control MNs and isogenic CRISPR/Cas9 *C9orf72* HRE-corrected MNs (Selvaraj *et al.*, 2018). The lack of changes in excitability was attributed to the purity of MN cultures, compared to other studies using mixed cultures of MNs and glia, supporting non-cell-autonomous effects of glia on MN function. Indeed, in a follow-up study, Zhao *et al.* demonstrated that *C9orf72* iPSC-astrocytes induced progressive loss of action potential output in control iPSC-MNs caused by an underlying loss of voltage activated Na+ and K+ currents (Zhao *et al.*, 2019b).

Excessive exposure to the neurotransmitter glutamate is a mechanism of excitotoxicity leading to neuronal death. Glutamate-induced excitotoxicity is a major contributor to the neurodegeneration in ALS and emerging evidence support impaired glutamatergic transmission in FTD pathogenesis (Starr and Sattler, 2018, Benussi *et al.*, 2019). Donelly *et al.* showed that *C9orf72* iPSCNs were highly susceptible to glutamate-mediated excitotoxicity and knockdown of ADARB2, the RNA binding protein sequestered by RNA foci, resulted in increased susceptibility to glutamate-

induced excitotoxicity in control iPSCNs (Donnelly *et al.*, 2013). Interestingly, these phenotypes were rescued upon treatment of *C9orf72* iPSCNs with antisense oligonucleotides (ASO) against the *C9orf72* transcript or HRE. Similarly, in another study by Selvaraj *et al*., iPSC-MNs, but not iPSC-CNs, exhibited vulnerability to excitotoxicity due to increased expression of GluA1 AMPA receptor (AMPAR) subunit leading to increased Ca^{2+}-permeable AMPAR expression (Selvaraj *et al.*, 2018). These deficits were abolished following CRISPR/Cas9-mediated correction of the HRE in isogenic MNs. Interestingly, loss of C9orf72 protein has also been shown to result in accumulation of glutamate receptors and vulnerability of iPSC-MNs to glutamate excitotoxicity (Shi *et al.*, 2018). Specifically, *C9orf72* iPSC-MNs as well as *C9orf72* +/- and *C9orf72* +/- iPSC-MNs exhibited reduced survival in response to excess glutamate that was attributed to elevated levels of NMDA (NR1) and AMPA (GLUR1) receptors in these neurons compared to their respective controls. Moreover, overexpression of either of the two C9orf72 protein isoforms was able to rescue *C9orf72* iPSC-MN survival in response to glutamate treatment. Finally, glutamate-induced excititoxicity was also reported in *C9orf72* iPSCNs by Cheng *et al.* who also demonstrated that overexpression of the RNA helicase DDX3X, identified as a repressor of *C9orf72* HRE RAN translation, could rescue the glutamate-induced neuron death and improved survival of *C9orf72* iPSCNs (Cheng *et al.*, 2019).

Additional findings

Several other iPSC studies have identified additional cellular pathways affected by the *C9orf72* HRE. Among these, compromised extracellular vesicle secretion and endosome formation as well as dysfunctional trans-Golgi network was reported in *C9orf72* iPSC-MNs (Aoki *et al.*, 2017). Other affected pathways in *C9orf72* iPSCNs include axonal trafficking, axonal degeneration and synaptic vesicle recycling (Coyne *et al.*, 2017, Yuva-Aydemir *et al.*, 2019, Abo-Rady *et al.*, 2020). Coyne *et al.* showed

that synaptic vesicle cycling was impaired in *C9orf72* iPSC-MNs due to posttranscriptional reduction in the levels of the Hsc70-4/HSPA8 chaperone (Coyne *et al.*, 2017). Axonal degeneration and partial TDP-43 translocation to the cytoplasm, reminiscent of the TDP-43 pathology observed in C9-FTD/ALS patients, was observed in *C9orf72* iPSCNs (Yuva-Aydemir *et al.*, 2019). It was also demonstrated that transcription elongation factor AFF2/FMR2 regulates the transcription of the HRE. CRISPR-Cas9-mediated knockout of AFF2/FMR2 resulted in decreased expression of the mutant *C9orf72* allele containing the HRE and reversal of axonal degeneration and TDP-43 mislocalisation. Finally, *C9orf72* iPSC-MNs exhibited axonal transport defects, as indicated by lysosomal track displacement in distal and proximal axons, compared to control MNs (Abo-Rady *et al.*, 2020). This was accompanied by reduced levels of ubiquitously expressed chaperone HSP70 and altered stress granule formation compared to control MNs. Interestingly, all these phenotypes were exacerbated in isogenic *C9orf72* iPSC-MNs that contained the HRE as well as a *C9orf72* knockout, supporting a combination of gain and loss of function mechanisms in C9-ALS/FTD pathogenesis.

The first *C9orf72* three-dimensional neuronal model revealed re-engagement of cell cycle-associated proteins and a senescence-associated secretory phenotype in *C9orf72* iPSCNs (Porterfield *et al.*, 2020). Specifically, *C9orf72* iPSCNs grown on Alvetex scaffold spontaneously re-expressed cyclin D1 12 weeks post-differentiation, suggesting cell cycle re-engagement. *C9orf72* iPSCNs exhibited increased expression of senescence-associated genes including CXCL8, a chemokine overexpressed by senescent cells. In addition to this, increased levels of components of the senescence-associated secretory phenotype were present in media from *C9orf72* iPSCNs compared to controls.

1.3.4.5. *iPSCNs elucidate C9orf72 function*

C9orf72 has a well-described role in endolysosomal trafficking and autophagy from several *in vitro* studies (see 1.2.9). Indeed, compromised autophagy has also been detected in *C9orf72* iPSCNs (Almeida *et al.*, 2013, Dafinca *et al.*, 2016, Webster *et al.*, 2016), suggesting this could be due to loss of C9orf72 function.

The C9orf72 protein was found to localise in early endosomes and was required for normal vesicle trafficking and lysosomal biogenesis in iPSC-MNs (Shi *et al.*, 2018). *C9orf72* haploinsufficiency could trigger neurodegeneration by causing accumulation of glutamate receptors, leading to MN excitotoxicity, and hypersensitivity of MNs to neurotoxic DPRs by impairing their clearance (Shi *et al.*, 2018). This could be rescued by restoring C9orf72 levels or treatment with small molecule modulators of vesicle trafficking.

Finally, Sivadasan *et al.* provided evidence that C9orf72 regulates axonal actin dynamics via regulation of the GTPase activity of Arf6 and the phosphorylation of cofilin, a ubiquitous actin-binding factor required for the reorganization of actin filaments (Sivadasan *et al.*, 2016).

1.3.4.6. *Non-cell-autonomous toxicity of C9orf72 patient-derived iPSC-glia*

Mouse LOF studies have shown that *C9orf72* is expressed at higher levels in myeloid cells compared to neurons (O'Rourke *et al.*, 2016, Rizzu *et al.*, 2016), it is important for proper macrophage and microglia function (O'Rourke *et al.*, 2016), and has a role in the regulation of inflammatory responses and autoimmunity (Burberry *et al.*, 2016, O'Rourke *et al.*, 2016, Burberry *et al.*, 2020). Therefore, iPSC studies investigating the cell-autonomous and non-cell-autonomous mechanisms by which glia and microglia may contribute to neurodegeneration in C9-FTD/ALS are necessary.

Several iPSC studies have explored the non-cell-autonomous effects of *C9orf72* patient-derived glia and demonstrated a toxic effect of *C9orf72* astrocytes and oligodendrocytes to MNs and other cell types either in co-cultures or via conditioned media. *C9orf72* transdifferentiated human astrocytes were toxic to co-cultured mouse MNs (Meyer *et al.*, 2014), whereas *C9orf72* iPSC-astrocyte conditioned medium was toxic to control and patient iPSC-MNs (Madill *et al.*, 2017). *C9orf72* patient-derived oligodendrocytes were also able to induce MN death both in co-cultures and via oligodendrocyte conditioned media (Ferraiuolo *et al.*, 2016). Moreover, a non-cell-autonomous effect of *C9orf72* iPSC-astrocytes on the modulation of the autophagy pathway has been demonstrated, whereby treatment with *C9orf72* iPSC-astrocyte conditioned medium resulted in reduced expression of the autophagosomal marker LC3-II, with a concomitant accumulation of p62 puncta and increased SOD1 expression in HEK293T cells (Madill *et al.*, 2017). As well as revealing the non-cell-autonomous mechanisms of *C9orf72* iPSC-astrocytes, studies have also demonstrated cell-autonomous disease mechanisms in these cells. *C9orf72* iPSC-astrocytes were shown to exhibit increased oxidative stress and *C9orf72* iPSC-astrocyte conditioned medium was found to be toxic to control MNs by inducing oxidative stress (Birger *et al.*, 2019). In another study by Varcianna *et al.*, *C9orf72* induced astrocytes exhibited impaired extracellular vesicle (EV) biogenesis and EVs with dysregulated miRNAs responsible for the regulation of axonal maintenance (Varcianna *et al.*, 2019). *C9orf72* astrocyte conditioned medium and astrocyte-derived EVs were both toxic to control mouse MNs, due to the miRNA dysregulation, in particular miR-494-3p, a negative regulator of the axon guidance protein semaphorin 3A (SEMA3A). Restoration of miR-494-3p levels was able to restore Sema3A levels in control mouse MNs and increase MN survival (Varcianna *et al.*, 2019). Finally, Zhao *et al.* showed that *C9orf72* iPSC-astrocytes recapitulated key pathological features of C9-ALS and caused a progressive loss of action potential output in co-cultured control MNs which was reversed upon CRISPR/Cas-9-mediated

correction of the HRE (Zhao *et al.*, 2019b). Importantly, these phenotypes were only present in control MNs co-cultured with *C9orf72* iPSC-astrocytes and not in *C9orf72* iPSC-MN-enriched cultures alone, highlighting the importance of non-cell-autonomous toxicity mechanisms in neurodegeneration.

1.3.4.7. Novel therapies emerging from iPSC studies

iPSC studies have been particularly useful in elucidating C9-FTD/ALS pathogenic mechanisms as well as in providing evidence for the therapeutic potential of certain treatments. Moreover, human iPSC-neurons offer ideal platforms for drug-screening. One example of promising C9-FTD/ALS therapies as demonstrated by two independent iPSC studies, is the use of antisense oligonucleotides (ASOs) to target HRE-mediated toxicity in C9-FTD/ALS which is currently in clinical trials (https://clinicaltrials.gov/ct2/show/study/NCT03626012) (Donnelly *et al.*, 2013, Sareen *et al.*, 2013). In one study, Donnelly *et al.* reported sequestration of RNA editing regulator ADARB2 by the *C9orf72* RNA foci, aberrant gene expression and susceptibility to glutamate excitotoxicity in *C9orf72* iPSC-neurons compared to control neurons (Donnelly *et al.*, 2013). In the second study by Sareen *et al.* the *C9orf72* iPSC-MNs exhibited RNA foci that co-localised with RNA-binding proteins hnRNPA1 and Pur-α, as well as aberrant gene expression, and reduced excitability compared to control MNs (Sareen *et al.*, 2013). The use of ASOs targeting the *C9orf72* transcript resulted in reversal of the toxicity phenotypes in *C9orf72* iPSC-neurons in both studies. Furthermore, as the ASO-mediated *C9orf72* knockdown had no adverse effect on *C9orf72* iPSC-neurons, the authors argued against a loss of function mechanism as the major pathogenic cause of C9-FTD/ALS. However, recent studies have demonstrated an important function of C9orf72 protein in autophagy and microglial function and provided evidence for a contribution of *C9orf72* haploinsufficiency in FTD/ALS pathogenesis, cautioning against the use of ASOs that

reduce *C9orf72* expression (O'Rourke *et al.*, 2016, Sellier *et al.*, 2016, Webster *et al.*, 2016, Shi *et al.*, 2018, Shao *et al.*, 2019, Abo-Rady *et al.*, 2020, Zhu *et al.*, 2020). Interestingly, the use of ASOs also restored the impaired nucleocytoplasmic transport phenotype that was responsible for the abnormal nuclear/cytoplasmic ratios of Ran and TDP-43 in *C9orf72* iPSCNs (Zhang *et al.*, 2015). Collectively, the ASO intervention studies in human *C9orf72* iPSC-neurons have shown that specific targeting of the *C9orf72* transcript can rescue gain of function toxicity and has therapeutic value.

Apart from ASOs, several studies have shown the therapeutic potential of compounds in C9-FTD/ALS. The use of retigabine, an activator of Kv7 potassium channels, was able to block the hyperexcitability of *C9orf72* iPSC-MNs (Wainger *et al.*, 2014). In another study, Trolox, a water-soluble antioxidant and vitamin E analog, was shown to rescue the poly-GR-induced increase in mitochondrial reactive oxygen species (ROS) and DNA damage in *C9orf72* iPSC-MNs (Lopez-Gonzalez *et al.*, 2016). In a phenotypic screen to repurpose existing drugs, Imamura *et al.* identified the Src/c-Abl pathway as a novel potential therapeutic target in ALS, and Bosutinib, a Src/c-Abl inhibitor, was shown to increase survival of *C9orf72* iPSC-MNs (Imamura *et al.*, 2017). The small molecule modulators of vesicle trafficking, YM201636 and Apilimod, which are inhibitors of the lipid kinase PIKFYVE, could rescue neurodegeneration in *C9orf72* iPSC-MNs (Shi *et al.*, 2018). Finally, using small molecules that bind and specifically stabilise the *C9orf72* HRE G-quadruplex RNA, Simone *et al.* demonstrated a reduction in RNA foci burden in both *C9orf72* iPSC-CNs and iPSC-MNs as well as a reduction in the levels of poly-GP in *C9orf72* iPSC-MNs, providing proof of principle that targeting the *C9orf72* HRE G-quadruplex structure has therapeutic potential (Simone *et al.*, 2018).

1.4. Mitophagy

1.4.1. PINK1/Parkin-dependent mitophagy

Mitochondria are indispensable for the maintenance of neurons as they supply ATP required for neuronal growth, survival, and function. Neurons are particularly vulnerable to mitochondrial dysfunction due to their high energy requirements as well as their post-mitotic nature and long life. Therefore, the regulation of mitochondrial health and homeostasis by quality control mechanisms is crucial for neuronal health. The mitochondria are dynamic organelles, forming the mitochondrial network which is maintained via several regulated processes such as mitochondrial fusion and fission, elimination of damaged mitochondria via non-selective or selective autophagic degradation and *de novo* mitochondrial biogenesis, all tuned to ensure an equilibrium between mitochondrial degradation and biogenesis.

The clearance of damaged mitochondria via selective autophagy, known as mitophagy, is highly specialised and involves recognition of the damaged mitochondria, engulfment by the autophagic membrane and fusion of the autophagosome with the lysosome for degradation (Pickles *et al.*, 2018). In the classical mitophagy pathway, the PTEN-induced putative kinase 1 (PINK1) and the cytosolic E3 ubiquitin ligase Parkin are the mediators of mitochondrial targeting for degradation. In healthy mitochondria, PINK1 kinase which contains a mitochondrial targeting sequence, is constitutively imported to the mitochondria through the translocase of the outer membrane (TOM) and translocase of the inner membrane (TIM23) complexes. Then, PINK1 is proteolytically cleaved in the transmembrane segment by the inner membrane PINK1/PGAM5-associated rhomboid-like (PARL) protease (Deas *et al.*, 2010, Jin *et al.*, 2010, Meissner *et al.*, 2011). Finally, processed PINK1 translocates to the cytoplasm for N-end recognition and proteasomal degradation (Yamano and Youle, 2013). However, in damaged mitochondria, protein

import is compromised and therefore PINK1 accumulates in the outer mitochondrial membrane (OMM) of the damaged mitochondria and phosphorylates ubiquitin (Ub) chains bound to OMM protein substrates on Ser65, triggering the recruitment of cytoplasmic Parkin (Koyano *et al.*, 2014, Shiba-Fukushima *et al.*, 2014). PINK1 then phosphorylates Parkin on Ser65 enhancing its E3 ligase activity (Kondapalli *et al.*, 2012, Shiba-Fukushima *et al.*, 2012). Activation of Parkin results in the ubiquitination of several OMM protein substrates, leading to further phosphorylation of ubiquitin by PINK1, and therefore triggering further Parkin recruitment and activation via a feed-forward mechanism amplifying phospho-ubiquitin (Ordureau *et al.*, 2015). Mitofusins 1 and 2 (Mfn1, Mfn2) are OMM proteins essential for mitochondrial fusion which are ubiquitinated by Parkin (Gegg *et al.*, 2010, Poole *et al.*, 2010). The ubiquitination of mitofusins results in inhibition of mitochondrial fusion promoting mitochondrial fragmentation and mitophagy. Ultimately, excessive ubiquitination of OMM proteins flags the damaged mitochondria for degradation by macroautophagy, as ubiquitinated chains are bound by autophagy receptor proteins that connect them to LC3 on autophagosomes. So far, five different autophagy receptor proteins have been identified; sequestosome 1 (SQSTM1) or p62, next to BRCA1 gene 1 (NBR1), optineurin (OPTN), nuclear dot protein 52 (NDP52) and Tax1 Binding Protein 1 (TAX1BP1) (Lazarou *et al.*, 2015). Among these, OPTN and NDP52 have been described as the preferential adaptors to induce mitophagy (Lazarou *et al.*, 2015). Mitochondrial damage activates TBK1 kinase which increases mitochondrial recruitment of OPTN, NDP52 and p62, and phosphorylates OPTN, therefore increasing its affinity for ubiquitin and phospho-ubiquitin chains (Heo *et al.*, 2015, Richter *et al.*, 2016).

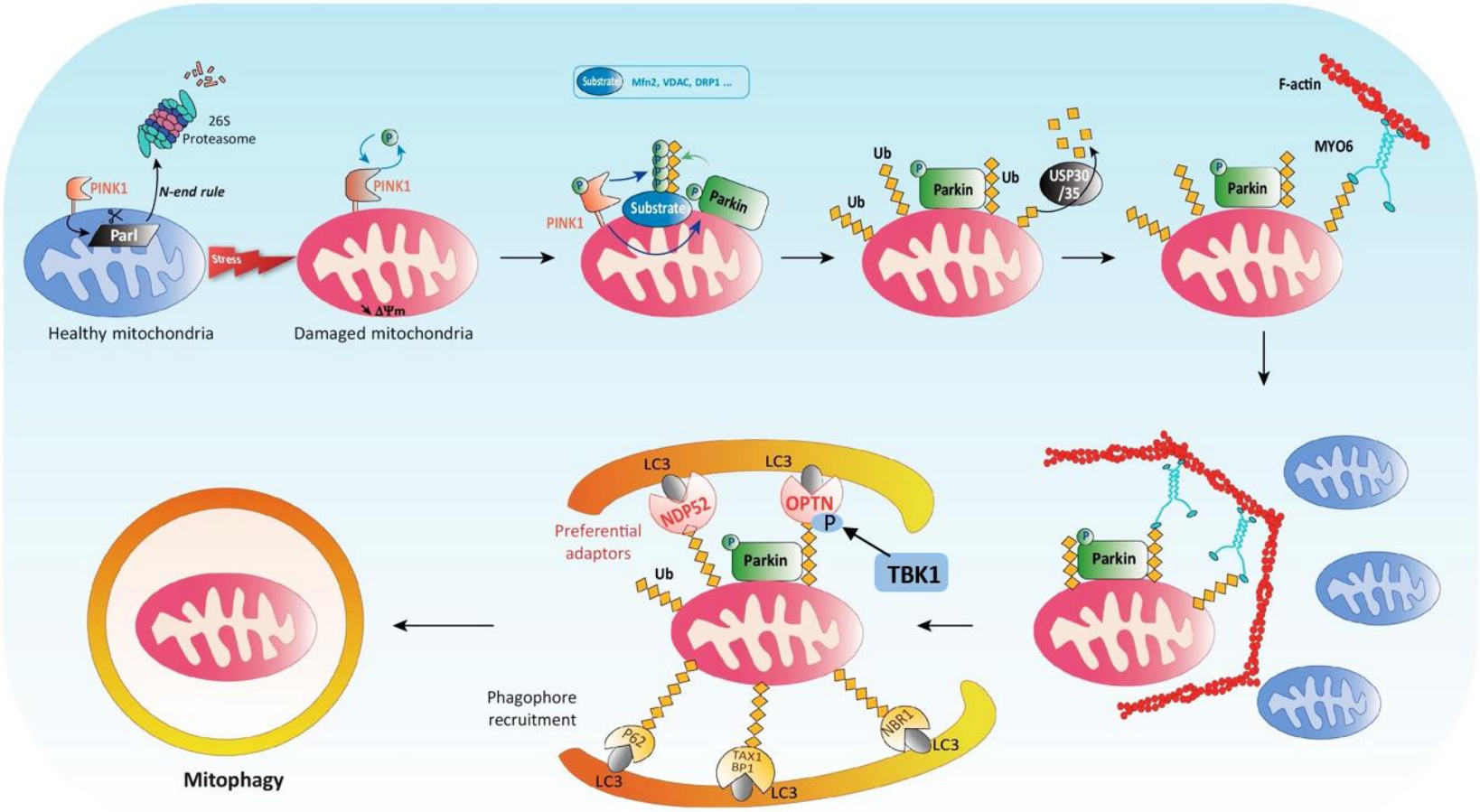


Figure 1.7 Summary of Parkin/PINK1-dependent mitophagy.

Under normal conditions, PINK1 is constitutively imported in mitochondria, cleaved by PARL protease in the transmembrane segment, and translocated to the cytoplasm for N-end recognition and proteasomal degradation. Upon mitochondrial damage, PINK1 accumulates in the outer

(Continued Figure legend) mitochondrial membrane (OMM) of the damaged mitochondria and phosphorylates ubiquitin chains bound to OMM protein substrates on Ser65, triggering the recruitment of cytoplasmic Parkin. PINK1 phosphorylates Parkin on Ser65 enhancing its E3 ligase activity. Parkin ubiquitinates many OMM substrates with K48- and K63-linked ubiquitin chains, including the mitofusins 1 and 2. K63-chains recruit autophagy receptors such as sequestosome 1 (SQSTM1) or p62, next to BRCA1 gene 1 (NBR1), optineurin (OPTN), nuclear dot protein 52 (NDP52) and Tax1 Binding Protein 1 (TAX1BP1) that connect the damaged mitochondria to LC3 on autophagosomes for degradation. TBK1 kinase increases mitochondrial recruitment of OPTN, NDP52 and p62, and phosphorylates OPTN, therefore increasing its affinity for ubiquitin and phospho-ubiquitin chains.

1.4.2. FTD/ALS-related genes in autophagy and mitophagy pathways

Several FTD/ALS-associated genes are involved in the autophagy and mitophagy processes. The *C9orf72* HRE is the most common cause of FTD/ALS and C9orf72 protein has been shown to have important roles in the regulation of autophagy initiation via interaction with ULK1, as well as in lysosomal function (Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Webster *et al.*, 2016, Yang *et al.*, 2016). Moreover, a recent report of an interaction between C9orf72 and mitochondrial fission protein fis1 suggest a potential role in mitophagy that requires futher investigation (Chai *et al.*, 2019). The FTD/ALS-associated proteins OPTN and p62 are selective autophagy/mitophagy receptor proteins (Maruyama *et al.*, 2010, Fecto *et al.*, 2011, Rubino *et al.*, 2012, Le Ber *et al.*, 2013, Pottier *et al.*, 2015). Haploinsufficiency of *TBK1* causes FTD/ALS (Cirulli *et al.*, 2015, Freischmidt *et al.*, 2015, Pottier *et al.*, 2015). TBK1 kinase phosphorylates OPTN and p62 to facilitate selective mitophagy, as well as SMCR8, the binding partner of C9orf72, to regulate autophagy initiation (Heo *et al.*, 2015, Richter *et al.*, 2016, Sellier *et al.*, 2016). Another FTD/ALS-linked gene, *VCP*, encodes for valosin-containing protein or p97 which has important roles in the promotion of mitophagy by mediating the proteasomal degradation of mitofusins (Tanaka *et al.*, 2010), autophagosome maturation (Tresse *et al.*, 2010), clearance of damaged lysosomes via macroautophagy (Papadopoulos *et al.*, 2017) and ER-associated degradation (ERAD) (Houck *et al.*, 2014). Mutations in *UBQLN2,* encoding ubiquilin 2, are a rare cause of FTD/ALS (Deng *et al.*, 2011). Ubiquilin 2 regulates proteostasis by facilitating the clearance of misfolded proteins via the proteasome and autophagy degradation pathways (Hjerpe *et al.*, 2016, Wu *et al.*, 2020). *CHMP2B* is a rare cause of FTD/ALS encoding CHMP2B, a subunit of the endosomal

sorting complex required for transport III (ESCRT-III) involved in multivesicular body (MVB) formation and endosome-lysosome fusion in the autophagy pathway.

Haploinsufficiency of *GRN* is associated with FTD whereas depletion of *GRN* causes neuronal ceroid lipofuscinosis (NCL), a lysosomal storage disease (Smith *et al.*, 2012). PGRN is localised in lysosomes and its function is important for lysosomal function. Moreover, TMEM106B which is a major risk factor for FTD has a role in lysosomal morphology and function (Brady *et al.*, 2013, Stagi *et al.*, 2014).

Vesicle-associated membrane protein-associated protein B (VAPB), an ALS-associated protein, is an ER-tethering protein interacting with the protein tyrosine phosphatase interacting protein-51 (PTPIP51) of the OMM and therefore recruiting regions of ER to the mitochondria (De Vos *et al.*, 2011, Stoica *et al.*, 2014). It has been shown that VAPB-PTPIP51 tethers regulate calcium homeostasis and autophagy (De Vos *et al.*, 2011, Gomez-Suaga *et al.*, 2017). Mutations in *ALS2* gene, encoding alsin, cause a rare juvenile-onset, autosomal recessive form of ALS via loss of protein function (Hadano *et al.*, 2001, Yang *et al.*, 2001). Alsin has a role in endosomal vesicle trafficking and fusion within the autophagy cascade by acting as a GEF for the Ra5 GTPase family (Otomo *et al.*, 2003, Devon *et al.*, 2006). Other rare causes of ALS are mutations in dynactin1 (*DCTN1)* (Münch *et al.*, 2004) and factor-induced gene 4 (*FIG4)* (Chow *et al.*, 2009) genes encoding dynactin1 and fig4 respectively. Dynactin1 (DCTN1) is the largest subunit of the dynactin complex, which interacts with motor protein dynein to enable the retrograde transport of autophagy substrates such as authophagosomes and lysosomes (Laird *et al.*, 2008, Ishikawa *et al.*, 2019). Phosphatase fig4 has a role in the autophagy/lysosomal pathway which is mediated via regulation

of the levels of the signaling lipid molecule phosphatidylinositol-3,5,-bisphosphate (PI(3,5)P2) involved in endosomal trafficking (Ferguson *et al.*, 2009, Kyotani *et al.*, 2015).

1.5. Research aims

We hypothesise that *C9orf72* haploinsuffiency plays a key role in disease pathogenesis. The aim of this thesis is to establish human iPSC-CN models of C9-FTD/ALS together with *C9orf72* knockout iPSC, in order to investigate the presence of the main pathological hallmarks of C9-FTD/ALS and understand the impact of *C9orf72* haploinsufficency on neuronal health.

Specifically, the aims of this thesis as discussed in individual chapters are:

- To characterise *C9orf72* iPSC-CNs for pathologies relevant to FTD, specifically RNA foci and DPRs.
- To investigate *C9orf72* mRNA expression both temporally during corticogenesis in control iPSC-CNs but also in *C9orf72* iPSC-CNs compared to control iPSC-CNs.
- To assess the temporal DNA methylation in *C9orf72* iPSC-CNs compared to control iPSC-CNs and investigate how it influences *C9orf72* mRNA expression and RNA foci/DPR burden in *C9orf72* iPSC-CNs.
- To characterise the levels of C9orf72 protein both temporally during corticogenesis in control iPSC-CNs but also in *C9orf72* iPSC-CNs compared to control iPSC-CNs.
- To generate and characterise *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs.
- To investigate mitophagy as well as expression of autophagy genes in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ CNs.

Chapter 2 Materials and Methods

2.1. Ethics

Informed consent was obtained from all participants in this study. The study protocol, patient information sheets and consent forms were reviewed and approved by the joint UCL and National Hospital for Neurology and Neurosurgery joint research ethics committee, study reference number 09/0272.

2.2. Derivation of fibroblast lines from skin biopsies

Fibroblast lines were generated from 4 mm skin punch biopsies taken under local anaesthetic following informed consent. Biopsies were dissected into ~1-2 mm pieces and plated on 5 cm^2 petri dishes in Dulbecco's modified Eagle's medium (DMEM) GlutaMAX (Thermo Fisher Scientific) supplemented with 10% (v/v) heat-inactivated fetal bovine serum (FBS) (Thermo Fisher Scientific), 50 U ml^{-1} penicillin and 50 µg ml^{-1} streptomycin (Thermo Fisher Scientific) at 37ºC, 5% CO_2, humidified. Fibroblast outgrowth was observed in the periphery of the explants 1-2 weeks post dissection. When confluent, fibroblasts were gently passaged by washing with Dulbecco's Phosphate Buffered Saline (DPBS) (Thermo Fisher Scientific) and incubating in 2 ml 0.05% trypsin-EDTA (Thermo Fischer Scientific) for 5 min at 37ºC, 5% CO_2, humidified. Cells were then collected and centrifuged at 200 $g_{(av)}$ for 5 min to remove trypsin. The cell pellet was resuspended in fresh DMEM-Glutamax with 10% (v/v) FBS, 50 U ml^{-1} penicillin and 50 µg ml^{-1} streptomycin and transferred to a T25 cell culture flask (Thermo Fisher Scientific) to expand. Upon reaching ~80% confluency, fibroblasts were passaged again with trypsin as described and transferred to T75 cell culture flasks in a 1:4 split ratio to allow further expansion. Fibroblasts were frozen at low passages in FBS containing 10% (v/v) Dimethyl sulfoxide (DMSO) (Sigma) in 2

ml cryovials (Thermo Fisher Scientific). For long-term cryopreservation, fibroblasts were stored in liquid nitrogen.

2.3. Pluripotent stem cell lines used in this study

Six pluripotent stem cell lines were used in this study, one embryonic stem cell (ESC) line and five induced pluripotent stem cell lines from three patients carrying the GGGGCC expansion in *C9orf72* and two unaffected individuals. iPSC lines had been previously generated using the retroviral Yamanaka reprogramming factors OCT4, KLF4, SOX2, and c-MYC (Takahashi *et al.*, 2007). The pluripotent stem cell lines used in this study and their characteristics are summarised in Table 2.1. Control iPSC line 1 (CTRL-1) was a kind gift from Dr Tilo Kunath, control iPSC line 2 (CTRL-2) has been obtained from Coriell repository, SHEF6 ESCs (CTRL-3) were obtained from the UK Stem Cell Bank, *C9orf72* patient-derived iPSC lines 1 and 2 (C9-1 and C9-2) were generated in house and finally *C9orf72* iPSC line 3 (C9-3) was a kind gift from Prof Chris Shaw.

Table 2.1 Summary of pluripotent stem cell lines used in this study.

Line ID	Status	Disease	Gender	Age	AAO	Repeat	Karyotype	Reprogramming	Clones
CTRL-1	Unaffected	N/A	M	78	N/A	N/A	46XY, normal	Retrovirus	1
CTRL-2	Unaffected (ESCs)	N/A	F	N/A	N/A	N/A	46XX, normal	N/A	1
CTRL-3	Unaffected	N/A	M	64	N/A	N/A	Normal	Retrovirus	1
C9-1	Affected	ALS	M	50	39	~241	46XY, normal	Retrovirus	1
C9-2	Affected	ALS	M	58	52	~638	46XY, normal	Retrovirus	1
C9-3	Affected	ALS	M	74	72	~602	46XY, normal	Retrovirus	1

Abbreviations; ESCs: embryonic stem cells, ALS: amyotrophic lateral sclerosis, M: Male, F: Female, AAO: age at onset, N/A: non applicable.

2.4. iPSC cultures

Human iPSCs were grown on 6-well cell culture plates (Thermo Fisher Scientific) pre-coated with Geltrex LDEV-Free, hESC-Qualified, Reduced Growth Factor Basement Membrane Matrix (Thermo Fisher Scientific). Geltrex was allowed to thaw on ice overnight, then aliquoted on ice and stored at -20^{o}C. For the preparation of coated plates, Geltrex was diluted 1:100 in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) medium (Thermo Fisher Scientific). The solution was mixed gently and used immediately to coat 6-well plates (1 ml of the solution per well of a 6-well plate).

iPSCs were maintained in Essential 8 medium (Thermo Fisher Scientific) at 37^{o}C, 5% CO_2, humidified. Medium was replaced daily and cultures were passaged every 4-5 days, using 0.5 mM Ethylenediaminetetraacetic acid (EDTA) (Thermo Fisher Scientific). For passaging iPSCs, all medium was removed and 1 ml of 0.5 mM EDTA was added on the well, followed by 5 min incubation at 37^{o}C, 5% CO_2, humidified. At the end of the incubation, plates were inspected under the microscope to ensure partial detachment of the iPSCs. EDTA was aspirated off the well and 1 ml fresh Essential 8 medium was added to resuspend the iPSCs and transfer to a new Geltrex coated plate using a 1:6 split ratio. iPSCs were frozen down in ice-cold Essential 8 medium containing 10% DMSO and stored in liquid nitrogen. The passage numbers for all pluripotent lines used in this study are; P32-P40 for CTRL-1, P40-P48 for CTRL2, P80-P88 for CTRL-3, P35-P43 for C9-1, P44-P52 for C9-2 and P70-P78 for C9-3. iPSC cultures were tested for mycoplasma contamination every two weeks using the MycoAlert Mycoplasma Detection Kit (Lonza).

2.5. Cortical neuron differentiation

Differentiation of iPSCs to CNs was performed following the protocol previously described by Shi *et al.*, that recapitulates human corticogenesis *in vitro* (Shi *et al.*, 2012a). iPSCs were collected using EDTA as described in 2.4 and pooled into 12-well plates (Thermo Fisher Scientific). Once the cells had reached 100% confluency the media was changed to neural induction media, which is neural maintenance media (NMM) containing two SMAD inhibitors. Neural maintainence media is a 1:1 mix of N2 (DMEM/F-12 GlutaMAX, 1x N2, 5 µg ml^{-1} insulin, 1 mM l-glutamine, 100 µm nonessential amino acids, 100 µM 2-mercaptoethanol, 50 U ml^{-1} penicillin and 50 mg ml^{-1} streptomycin) and B27 (Neurobasal, 1x B27, 200 mM l-glutamine, 50 U ml^{-1} penicillin and 50 µg ml^{-1} streptomycin) (all Thermo Fisher Scientific). For neural induction, NMM was supplemented with 1 µM Dorsomorphin, a selective inhibitor of Bone morphogenetic protein (BMP) signaling and 10 µM SB431542, a selective inhibitor of transforming growth factor-β (TGF-β) superfamily type I activin receptor-like kinase (ALK) receptors. Cells were maintained in neural induction media for 10 days with daily feeding until a uniform neuroepithelium layer had formed.

The 100% confluent neuroepithelium layer was replated using dispase (Thermo Fisher Scientific). Stock dispase was dissolved in DPBS to 10 mg ml^{-1}, incubated at 37°C for 30 min, sterilised through a 0.22 µM filter (Millipore) and stored in aliquots at -20°C. Dispase was added directly into media at 1 mg ml^{-1} final concentration and incubated at 37°C for 15-20 min. The detached neuroepithelium was washed in DPBS followed by sedimentation by gravity for a total of 3 washes. The neuroepithelial pieces were then plated onto laminin (Sigma) coated 12-well plates (1 mg ml^{-1} laminin stock was diluted 1:50 with DPBS to a final concentration of 20 µg ml^{-1} and incubated on 12-well plates for 4 hours or overnight) and grown in neural maintenance medium. Neural rosettes could be observed 2-5 days following replating. Rosettes were expanded and passaged onto fresh laminin coated plates with dispase between days

12-25 as described above, until a substantial amount of neurogenesis had occurred. Once a high number of post-mitotic neurons were observed in the cultures, neurons were replated as single cells using StemPro Accutase (Invitrogen). Briefly, medium was removed from wells and neurons were incubated with 400 µl Accutase per well of the 12-well plate for 5 min at 37°C, 5% CO_2, humidified. Neurons were then gently mixed, transferred to a 15 ml tube with NMM and centrifuged at 200 $g_{(av)}$ for 6 min. The neuronal pellet was gently resuspended in NMM and plated in laminin coated 12-well plates at a 1:3 ratio. Between passages 25 and 35, neuronal precursor cells (NPCs) were frozen down in ice-cold NMM containing 10% DMSO and stored in liquid nitrogen. Neurons were replated for the final time between day 35-40 on poly-L-ornithine (Sigma) and laminin double coated plates. Plates were first coated with 0.01% (w/v) poly-L-ornithine for 4 h at 37°C, then poly-L-ornithine was removed and plates were coated with 10 µg ml^{-1} laminin for 4 h at 37°C. All neuronal cultures were maintained at 37°C, 5% CO_2, humidified.

2.6. HEK293 cell cultures

HEK293 cells were maintained in DMEM Glutamax medium supplemented with 10% (v/v) FBS in T75 flasks at 37°C, 5% CO_2, humidified. When ~80% confluent, cells were washed with DPBS and incubated with 2 ml 0.05% trypsin for 5 min at 37°C, 5% CO_2, humidified. The cells were collected in a 15 falcon tube and centrifuged at 270 $g_{(av)}$ for 5 min to remove trypsin. Fresh DMEM Glutamax media was added to the cell pellet and cells were transferred in 6-well plates at a 0.3 x 10^6 seeding density, ready for transfections. HEK cells were frozen down in FBS containing 10% DMSO and stored in liquid nitrogen.

2.7. Genome engineering

2.7.1. TALENs

TALENs were generated using TALE-Toolbox from the Zhang lab (Addgene, Kit #1000000019) following the published protocol of Sanjana *et al*. based on hierarchical ligation assembly Figure 2.1 (Sanjana *et al.*, 2012). TALEN Targeter online software (Cornell University) was used to find suitable TALEN pairs against exon 2 of *C9orf72* gene. For the construction of TALENs, a TALEN monomer library has been generated by PCR amplifying all monomer plasmids using primers with added ligation arms and restriction enzyme sites. TALEN sequences were divided in hexamers which were generated by assembling monomers in a 10 µl Golden Gate digestion-ligation reaction consisting of 1 µl of each monomer (six monomers in total), 1 µl Tango buffer (10x), 1 µl DTT (10 mM), 1 µl ATP (10 mM), 0.75 µl BsmBI and 0.25 µl T7 ligase (3,000 U μl^{-1}) (all from New England Biolabs). The GoldenGate reaction was carried out in a thermocycler (37ºC, 5 min then 20ºC, 5 min for a total of 15 cycles). Hexamers were treated with exonuclease PlasmidSafe (Epicentre) to degrade non-circular ligation products and then PCR amplified in a 50 µl PCR reaction using 1 µl of the PlasmidSafe-treated hexamer, 10 µl Herculase Buffer 10x, 0.5 µl dNTP mix 100 mM, 1 µl Hex-F (5'- CTTAAACCGGCCAACATACC -3') and Hex-R (5'- AGTCTGTCTTTCCCCTTTCC -3') primers (10 µM each), 0.5 µl Herculase II Fusion polymerase (Agilent) and 37 µl ddH_2O. The PCR program used was initial denaturation 95ºC, 2 min, 35 cycles 95ºC, 20 sec, 60ºC, 20 sec, 72ºC, 30 sec and final extension 72ºC, 3 min. Amplified hexamer products were ran on a 2% (w/v) agarose gel and the right size hexamers were extracted from the gel using the MinElute gel extraction kit (Qiagen) as described in 2.20. Purified hexamers were quantified by semi-quantitative gel electrophoresis and integrated intensity was quantified using the LiCOR Image Studio Lite 3.1. The concentration of each hexamer was normalised to 20 ng μl^{-1}. Finally, the hexamers were mixed in a second and final

Golden Gate digestion-ligation reaction for assembly of the final multimer in the TALEN cloning vector. The 10 µl Golden Gate reaction contained 1 µl of each purified and normalised hexamer (20 ng μl^{-1}), 1 µl TALE backbone vector (100 ng μl^{-1}), 0.75 µl BsaI-HF (20 U μl^{-1}), 1 µl NEBuffer 4 (10x), 1 µl bovine serum albumin (BSA) (10x), 1 µl adenosine triphosphate (ATP) (10 mM), 0.25 µl T7 ligase (3,000 U μl^{-1}) and 1 µl ddH_2O. A Negative Control reaction that contained ddH_2O instead of hexamers was prepared in parallel. Digestion-ligation reactions were carried out at 20 cycles 37°C for 5 min, 20°C for 5 min and final 80°C for 20 min. TALEN vectors were used to transform competent *E. coli* as described in 2.9. Colonies emerging on the agar plate were screened for the correct insert by colony PCR. Briefly, 10 colonies were selected from the agar plate. Each colony was streaked onto a new LB agar (Sigma) plate containing 100 µg ml^{-1} ampicillin (Sigma) and incubated at 37°C to save the colony, and the rest was dissolved in 100 µl distilled H_2O. Colony PCR was performed using 1 µl of each colony suspension, 0.25 µl dNTP (100mM), 2.5 µl Taq-B polymerase buffer (10x), 0.25 µl of each TALE-Seq-F1 (5'- CCAGTTGCTGAAGATCGCGAAGC -3') and TALE-Seq-R1 (5'- TGCCACTCGATGTGATGTCCTC -3') primers (10 µM each), 0.1 µl Taq-B polymerase (5 U μl^{-1}) and 20.9 ddH_2O. The PCR was performed in a Veriti 96-Well Fast Thermal Cycler with the following cycling parameters: 94°C for 3 min, 31 cycles 94°C for 30 sec, 60°C for 30 sec and 68°C for 2 sec, and 68°C for 5 min with final hold at 4°C. PCR products were resolved by agarose gel electrophoresis as described in 2.19. Bacterial colonies with correctly assembled TALENs were further grown (as described in 2.10) and plasmid DNA was extracted (as described in 2.10.1) and sequenced to verify TALENs sequence (as described in 2.18). For sequencing, three primers were used to verify the entire sequence as shown in Figure 2.1; TALE-Seq-F1 (5'- CCAGTTGCTGAAGATCGCGAAGC -3') annealing just before the first monomer, TALE-Seq-F2 (5'- ACTTACACCCGAACAAGTCG -3') annealing at the beginning of the seventh

monomer, and TALE-Seq-R1 (5'- TGCCACTCGATGTGATGTCCTC -3') annealing after the final 0.5 monomer. Sanger Sequencing was perfomed as described in 2.18.

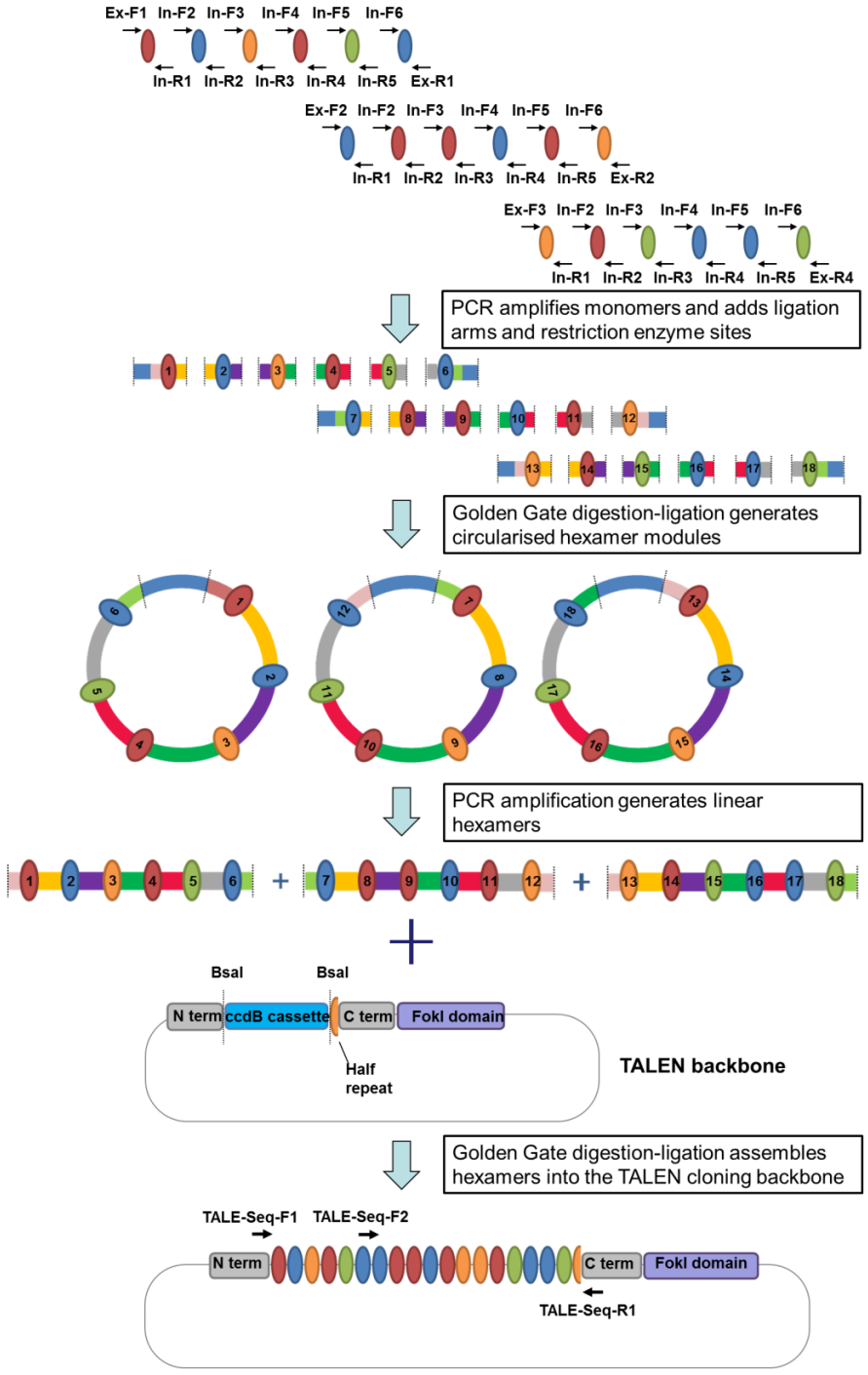


Figure 2.1 Schematic overview of TALEN hierarchical modular assembly protocol (Sanjana *et al.*, 2012).

Initially, each TALE monomer was amplified using specific primers that add appropriate ligation adaptors. Hexameric tandem repeats were then assembled in a

(Continued Figure legend) Golden Gate digestion-ligation reaction that generates circularised hexamer modules. Linear DNA was removed by PlasmidSafe exonuclease treatment, leaving only the circular hexamers. Each hexamer was amplified by PCR and purified. Hexamers were then ligated into the TALEN cloning backbone in a Golden Gate digestion-ligation reaction. The assembled TALEN plasmid was transformed into competent *E. coli* and isolated colonies were amplified by colony PCR. Sanger sequencing of the colonies was performed to verify the correct assembly using primers: TALE-Seq-F1 annealing just before the first monomer, TALE-Seq-F2 annealing at the beginning of the seventh monomer, and TALE-Seq-R1 annealing after the final 0.5 monomer.

2.7.2. CRISPR/Cas9n

For the generation of CRISPR/Cas9n targeting exon 2 of *C9orf72*, the protocol by Ran *et al.* was adapted (Ran *et al.*, 2013b). CRISPR Finder tool from the Wellcome Trust Sanger Institute was used (http://www.sanger.ac.uk/htgt/wge/) to find suitable sgRNAs in *C9orf72* exon 2. The CRISPR plasmids used have been developed by Feng Zhang lab and contain the Cas9 D10A nickase pSpCas9n(BB)-2A-GFP (PX461) (Addgene Plasmid #48140) and pSpCas9n(BB)-2A-Puro (PX462) (Addgene Plasmid #48141), vector maps shown in Figure 2.2. For CRISPR pairs 0 and 17, the oligos shown in Table 2.2 were ordered from Sigma after adding overhangs at the ends of the sgRNA sequences. The overhang 5'- CACCG -3' was added at the 5' end of the top strand whereas the overhang 5'- AAAC -3' was added at the 5' end and a C at the 3' end of the bottom strand sgRNA.

Oligos from each pair (top and bottom strands) were phosphorylated with T4 PNK kinase and annealed by incubation at 37^{o}C for 30 min followed by 95^{o}C for 5 min and then ramp down to 25^{o}C at 5^{o}C/min. A total reaction volume of 10 µl was used consisting of 1 µl of each 100 µM oligo (top and bottom), 1 µl 10x T4 Ligation Buffer (New England Biolabs) and 1 µl T4 polynucleotide kinase (PNK) (New England Biolabs) and 6 µl ddH_2O.

Table 2.2 Oligo sequences used for the generation of CRISPR/Cas9n constructs for the pair 0 and pair 17.

Oligo	Sequence (5'-3')
0_LEFT_TOP	CACCGATCCTTCGAAATGCAGAGAG
0_LEFT_BOTTOM	AAACCTCTCTGCATTTCGAAGGATC
0_RIGHT_TOP	CACCGTTCTCCATTTAGAGTGTGGT
0_RIGHT_BOTTOM	AAACACCACACTCTAAATGGAGAAC
17_LEFT_TOP	CACCGAGCAGCTACTTTTGCTTACT
17_LEFT_BOTTOM	AAACAGTAAGCAAAAGTAGCTGCTC
17_RIGHT_TOP	CACCGACTTAAAGCAATCTCTGTCT
17_RIGHT_BOTTOM	AAACAGACAGAGATTGCTTTAAGTC

Paired oligos were diluted 1:200 and cloned into the backbone in a single digestion-ligation step. Briefly, 100 ng of backbone vector were mixed with 2 µl of phosphorylated and annealed oligo duplex diluted 1:200, 2 µl Tango buffer (10x), 1 µl DTT (10 mM), 1 µl ATP (10 mM), 1 µl FastDigest *Bbs*I (Fermentas) and 0.5 µl T7 ligase (New England Biolabs) and ddH_2O up to 20 µl total reaction volume. Ligation reaction was incubated at 37 °C for 5 min followed by 21 °C for 5 min, for a total of 6 cycles. Ligation products were treated with PlasmidSafe exonuclease to digest any residual linearized DNA. Briefly, 11 µl of the ligation reaction were incubated with 1.5 µl PlasmidSafe buffer (10x), 1.5 µl ATP (10 mM), and 1 µl PlasmidSafe exonuclease at 37 °C for 30 min, followed by 70 °C for 30 min. The PlasmidSafe-treated product was used to transform competent *E. coli* as described in 2.9. Plasmid DNA was extracted from colonies as described in 2.10.1 and sequenced for verification of the correct sgRNA sequence using the sequencing primer 5'-GAGGGCCTATTTCCCATGATTCC -3' (as described in 2.18). The sequencing results were referenced against the pSpCas9(BB) cloning vector sequence to check that the 20 nt guide sequence is inserted between the U6 promoter and the remainder of the sgRNA scaffold.

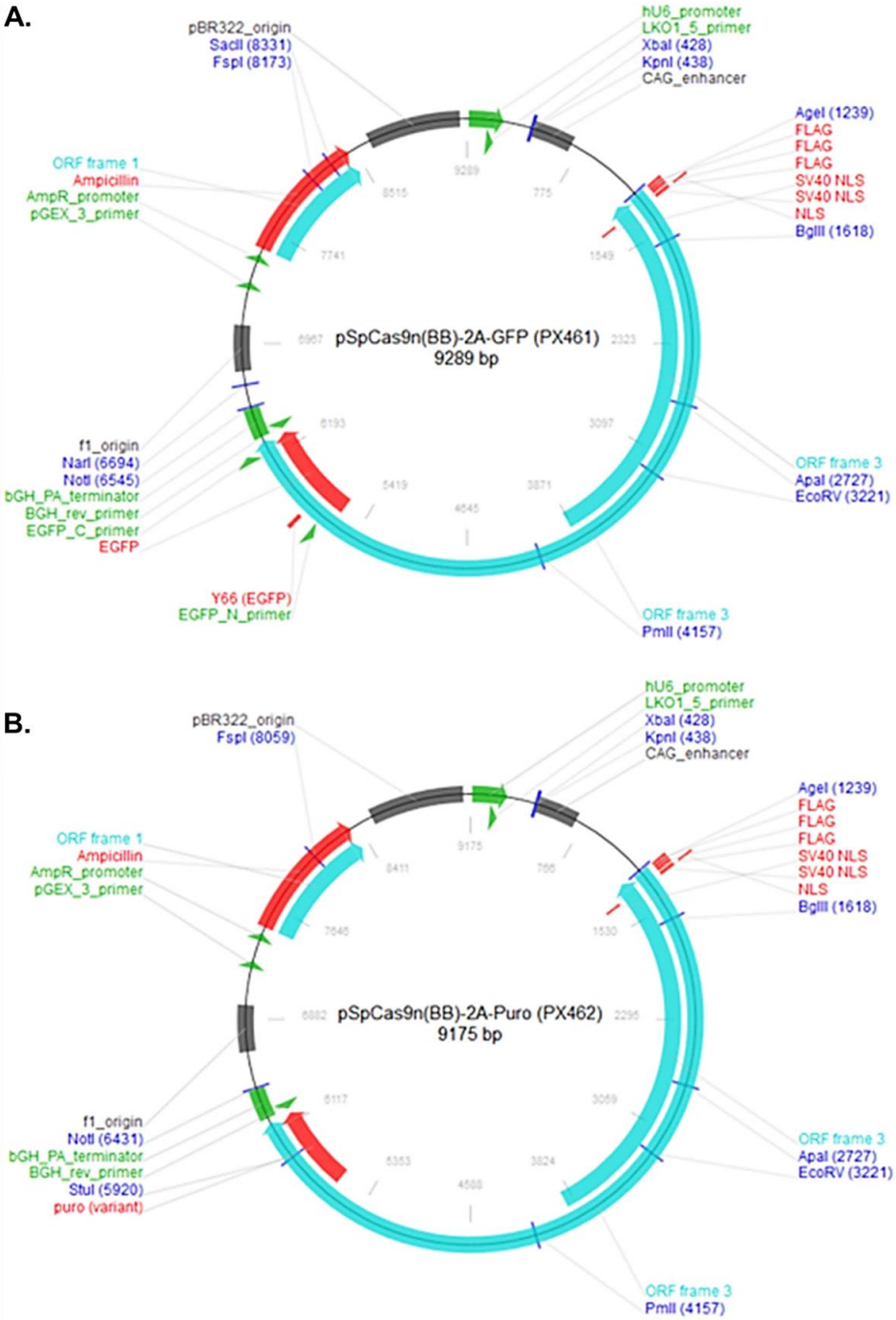


Figure 2.2 Vector maps for the CRISPR/Cas9 D10A nickase plasmids.

The CRISPR/Cas9 D10A nickase plasmids pSpCas9n(BB)-2A-GFP (PX461) (Addgene Plasmid #48140) (A) and pSpCas9n(BB)-2A-Puro (PX462) (Addgene Plasmid #48141) (B). Main plasmid features include CAG enhancer, hybrid form of the CBA promoter (CBh), nuclear localisation signal (NLS), 3x FLAG and ampicillin resistance.

2.8. Preparation of competent *E. Coli*

The Inoue method was followed for the preparation of competent *E. Coli* (Inoue *et al.*, 1990). Lysogeny Broth (LB) agar (Sigma) plates were prepared by dissolving 35 g of LB agar (15 g/L Agar, 10 g/L Tryptone, 5 g/L Yeast Extract, 5 g/L NaCl) in 1 L H_2O which was autoclaved for 15 min at 121°C to sterilise, poured onto 10 cm dishes (Thermo Fisher Scientific) and allowed to set for 1 h. *E. Coli* bacteria (*Stbl3* strain, kind gift of Dr Suran Nethisinghe) were then streaked onto the LB plates and incubated for 16-20 h at 37°C. Lysogeny Broth (LB) medium (Sigma) was prepared by dissolving 20 g LB powder (10 g/L Tryptone, 5 g/L Yeast Extract, 5 g/L NaCl) in 1 L H_2O and autoclaving the solution for 15 min at 121°C. A single bacterial colony was picked from the previously inoculated plate with *E. Coli* and transferred in 25 ml LB medium in a 250 ml flask and incubated 6-8 hours at 37°C with vigorous shaking. 10 ml, 4 ml and 2 ml of this culture were used to inoculate flasks containing 250 ml LB each. All flasks were incubated for 14 hours at 18°C with moderate shaking. OD_{600} was regularly monitored and when reached 0.55 the culture was incubated on ice for 10 min. Cells were harvested by centrifugation at 2,500 $g_{(av)}$ for 10 min at 4°C. Medium was discarded and bacteria were resuspended gently in 80 ml ice-cold Inoue transformation buffer (55 mM $MnCl_2{\cdot}4H_2O$, 15 mM $CaCl_2{\cdot}2H_2O$, 250 mM KCl, 10 mM PIPES (0.5M, pH 6.7) in pure H_2O up to final volume 1 L, filter sterilised through a 0.45 µm filter (Nalgene)). Cells were centrifuged at 2,500 $g_{(av)}$ for 10 min at 4°C, supernatant discarded and pellet resuspended gently in 20 ml ice-cold Inoue transformation buffer. 1.5 ml of DMSO was added to the bacteria suspension, mixed by swirling and left in ice for 10 min. Competent bacteria were aliquoted in eppendorfs, snap-frozen and stored in -80°C.

2.9. Transformation of competent *E. Coli*

For the transformation of competent *E. Coli*, frozen competent bacteria were thawed on ice. 5 µl of ligation product was added to 50 µl of bacteria and incubated for 5 min on ice. The mixture was then heat shocked at 42°C for 45 sec and subsequently incubated for 5 min on ice. 250 µl of Super Optimal broth with Catabolite repression medium (New England Biolabs, 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) was added and incubated at 37°C for 1 h in an orbital shaker at 300 rpm. Finally, 100 µl of transformation product was streaked on LB agar (Sigma) plates containing 100 µg ml^{-1} ampicillin (Sigma) and incubated overnight at 37°C.

2.10. Cloning

Individual bacterial colonies from the transformed *E. coli* with TALEN or CRISPR/Cas9n plasmids were subcloned in starter cultures of 5 ml LB medium containing 100 µg ml^{-1} ampicillin (Sigma) and grown overnight at 31°C, vigorous shaking. For selected colonies, small scale purification of plasmid DNA was carried out as described in 2.10.1. The plasmid sequence in all selected colonies was verified by Sanger Sequencing as described in detail in 2.7.1 and 2.7.2. The starter cultures of the plasmids with the correct sequence were further grown in large scale cultures in 250 ml LB with 100 µg ml^{-1} ampicillin, overnight at 31°C, vigorous shaking. Purified plasmid DNA was obtained by following the large scale purification protocol described in 2.10.2. Finally, plasmid concentration was measured using NanoDrop Spectrophotometer.

2.10.1. Small scale plasmid DNA purification

Purification of plasmid DNA was performed using the QIAprep spin miniprep kit (Qiagen) following manufacturer's instructions. Briefly, the bacterial cultures were pelleted by centrifugation at 6,800 $g_{(av)}$ and pellets were resuspended in 250 µl Buffer P1 (50 mM Tris-Cl, pH 8.0, 10 mM EDTA, 100 µg/ml RNase A) containing 1:1,000 LyseBlue reagent and transferred to eppendorf tubes. Bacteria were lysed under alkaline conditions, with the addition of 250 µl Buffer P2 (200 mM NaOH, 1% SDS (w/v)) followed by mixing by inverting the tubes 4–6 times until the solution turned blue in the presence of LyseBlue indicator. The lysates were subsequently neutralised and adjusted to high-salt binding conditions by addition of 350 µl Buffer N3 (proprietary) and mixed by inverting the tubes 4–6 times until the solution turned colourless. Samples were centrifuged for 10 min at 16,000 $g_{(av)}$ at RT and 800 µl of the supernatant was loaded on QIAprep 2.0 spin column. Samples were centrifuged for 1 min at 16,000 $g_{(av)}$, RT, the flow-through was discarded and the column was washed with 500 µl Buffer PB (proprietary) followed by 1 min centrifugation at 16,000 $g_{(av)}$, RT. Flow-through was discarded and the spin column was washed with 750 µl Buffer PE (proprietary) and centrifuged for 1 min. The spin column was centrifuged empty for 1 min to dry from any residual buffer, placed on a clean eppendorf tube for DNA elution and incubated with 25 µl Buffer EB (10 mM Tris-Cl, pH 8.5) for 1 min before a final centrifugation at 16,000 $g_{(av)}$, RT for 1 min.

2.10.2. Large scale plasmid DNA purification

For large scale plasmid purification, plasmid DNA was isolated using the EndoFree Plasmid Maxi Kit (Qiagen) according to manufacturer's instructions. Briefly, bacterial cultures were centrifuged at 6,000 $g_{(av)}$ for 15 min at 4°C, then the bacterial pellets were resuspended in 10 ml Buffer P1 (50 mM Tris-Cl, pH 8.0, 10 mM EDTA, 100 µg/ml RNase A) containing 1:1,000 LyseBlue reagent and mixed thoroughly. 10 ml Buffer P2 (200 mM NaOH, 1% SDS (w/v)) were added to the resuspended cells for alkaline lysis, mixed thoroughly by inverting the tube 4–6 times until the solution turned blue in the presence of LyseBlue indicator and incubated at RT for 5 min. 10 ml chilled Buffer P3 (3.0 M potassium acetate, pH 5.5) were added to precipitate the lysate and mixed thoroughly inverting 4–6 times. The lysate was transferred onto the QIAfilter Cartridge, incubated at RT for 10 min and then filtered into a 50 ml tube. For the removal of endotoxins, 2.5 ml Buffer ER (proprietary) were added to the filtered lysate, mixed by inverting the tube 10 times and incubated on ice for 30 min. Then the lysate was transferred onto a QIAGEN-tip 500 pre-equilibrated with 10 ml Buffer QBT (750mM NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol (v/v), 0.15 % Triton X-100 (v/v)) and allowed to pass through the resin by gravity flow. The column was then washed 2x with 30 ml Buffer QC (1.0 M NaCl, 50 mM MOPS, pH 7.0, 15% isopropanol (v/v)) and DNA was eluted with 15 ml Buffer QN (1.6 M NaCl, 50 mM MOPS, pH 7.0, 15 % isopropanol (v/v)). DNA was precipitated with 10.5 ml RT isopropanol and centrifugation at 5,000 $g_{(av)}$ for 60 min at 4°C. The DNA pellet was washed with endotoxin-free 70% EtOH and centrifuged at 5,000 $g_{(av)}$ for 60 min at 4°C. The DNA pellet was air-dried for 10 min and resuspended in 500 µl endotoxin-free Buffer TE (10 mM Tris-Cl, pH 8.0, 1 mM EDTA).

2.11. HEK293 transfections

For transfections with TALEN and CRISPR/Cas9n nucleases, HEK293 cells were plated 24h prior to transfection (as described in 2.6) at a density of 5 x 10^5 cells per well of a 6-well plate in 3 ml of DMEM Glutamax with 10% FBS and maintained at 37^{o}C, 5% CO_2, humidified. Transfections were performed after 24 h using FuGENE HD reagent (Roche) at a FuGENE HD:DNA ratio of 3:1. Briefly, for each reaction 3.3 µg of plasmid DNA were added in 155 µl OptiMEM reduced serum medium (Thermo Fisher Scientific), for a final concentration of 0.02 µg $µl^{-1}$ of DNA. Then, 9.9 µl FuGENE HD were added to the DNA solution and mixed gently by pipetting. The complex was incubated for 10 min at RT. Finally, 150 µl of the complex were added per well of HEK293 cells. HEK293 cells were also transfected with EGFP as a positive control for transfection. Transfected cells were visualised by light and fluorescence microscopy in a Nikon Eclipse TS100 microscope equipped with a Leica DC 500 digital camera.

2.12. T7 endonuclease I assay

T7 endonuclease I assay was used for the functional testing of the generated CRISPR and TALEN nucleases. The principle of the enzyme cleavage assay using T7 endonuclease is summarised in Figure 2.3. When a random mixture of edited and non-edited PCR products is melted and then gradually re-annealed at a slow rate, DNA heteroduplexes form. T7 endonuclease I incubation cleaves the heteroduplexes at the site of the mismatch which are then visualised by gel electrophoresis. The presence of cleaved PCR products indicates the presence of mutations and thus successful genome editing and functional nucleases.

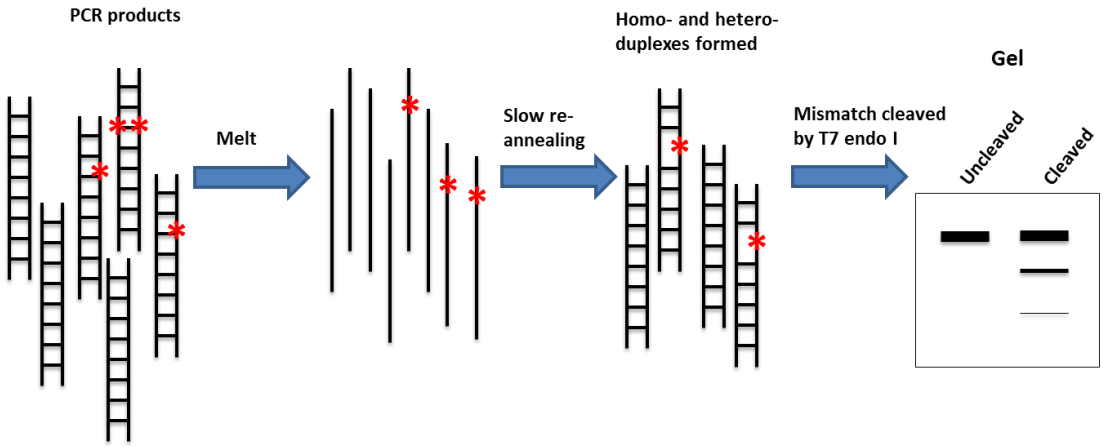


Figure 2.3 Overview of the T7 endonuclease I mutation detection assay for the functional testing of CRISPR/Cas9n and TALEN nucleases.
Cross-hybridisation of modified and non-modified PCR products results in the production of heteroduplexes which are cleaved at the site of the mutation due to mismatch in the DNA strands by T7 endonuclease I. Cleaved products can be detected by gel electrophoresis and indicate the presence of mutation and successful genome editing mediated by the nucleases.

HEK293 cells were harvested for T7 endonuclease I assay at 48 h or 72 h post-transfection with CRISPR/Cas9n or TALEN constructs. Genomic DNA was extracted using QuickExtract™ DNA extraction solution (Epicentre). HEK293 cell pellets were obtained after incubation with trypsin as described in 2.6. Pellets were resuspended in 500 µl DPBS, transferred into eppendorf tubes and centrifuged at 270 $g_{(av)}$ for 5 min. Supernatant was aspirated and the cell pellet was resuspended in 100 µl QuickExtract™ solution and mixed well. Genomic DNA (gDNA) was extracted by incubation at 68 °C for 15 min and 95 °C for 8 min on a thermal block.

Target sites were PCR amplified using a touchdown PCR program: Tm from 65°C to 55°C at a rate of −1°C/cycle and fixed at 55°C for 20 cycles; 3 min extension for each cycle, using primers forward 5'- TCTCCAGTCATACTAGAATTGCACA -3' and reverse 5'- CTACCTCCCACTTCATAGAGTGTGT -3' that generate a 805 bp amplicon containing all nuclease cleavage sites. 200 ng of PCR product were cross-hybridised and digested with T7 endonuclease I at 37°C for 15 min. Digestion

products were then analysed on a 1.5% TBE agarose gel as described in 2.19. For the calculation of the modification efficiency, the integrated intensity of each band was quantified using the LiCOR Image Studio Lite 3.1. For each lane, the fraction of the PCR product cleaved (f_{cut}) was calculated using the formula: $f_{cut} = a / (a + b)$, where a = the integrated intensity of both of the cleavage product bands and b = the integrated intensity of uncleaved PCR product band. The percentage of gene modification was estimated using the formula: $100 \times (1- (1 - f_{cut})^{1/2})$. This formula is based on the assumptions that (i) strand reassortment during the duplex formation is random, (ii) there is a negligible probability of the identical mutations reannealing during duplex formation and (iii) the nuclease digestion is complete.

2.13. Pluripotent media comparison for optimal single-cell survival post nucleofection

CTRL-1 iPSCs were routinely maintained in Essential 8 and grown as described in 2.4. For the purpose of this experiment, CTRL-1 iPSCs were also adapted to mTeSR1 (StemCell Technologies) and TeSR2 (StemCell Technologies) pluripotent cell culture media. The iPSCs were grown for at least two passages on their respective media (Essential 8, mTeSR1 and TeSR2) under feeder-free confitions as described in 2.4, before nucleofections. CTRL-1 iPSCs grown in each of the three media were incubated in 10 µM Y-27632-containing media for 1 hr prior to nucleofection with 1.5 µg of each left and right pair 0 pSpCas9n(BB)-2A-GFP vectors using the 4D Lonza nucleofection system (2.14). At 24 hr post nucleofection, single GFP-positive iPSCs were sorted by FACS into 96-well plates containing the respective growth medium and plates were inspected daily for colony detection.

2.14. Nucleofection of iPSCs

IPSCs were incubated with 10μM Rho-associated, coiled-coil containing protein kinase (ROCK) inhibitor Y-27632 (Stemcell) for at least 1 h prior to the nucleofection. iPSC were dissociated into single cells by incubation with accutase for 5 min at 37°C. 8 x 10^5 cells were used per nucleofection. Nucleofections were performed using the 4D Lonza nucleofection system and the P3 Primary Cell Kit (Lonza) as per manufacturer's instructions. Briefly, after centrifugation at 190 $g_{(av)}$ for 3 min the cell pellet was resuspended in 82 μl Nucleofection Solution P3 and 18 μl Supplement and loaded on the provided nucleofection cuvette. iPSCs were nucleofected using the pre-set program for H9 Stem cells using CB-150 pulse (proprietary). Cells were plated onto geltrex coated 6-well plates in mTeSR1 medium containing 10μM Y-27632. 24 h post-transfection, medium was replaced with fresh mTeSR1 and iPSCs were visualised under a Nikon Eclipse TS100 microscope equipped with a Leica DC 500 digital camera to ensure successful transfection and survival.

2.15. Cell sorting and clonal expansion of iPSCs

Control iPSCs that were nucleofected with the GFP vector pSpCas9n(BB)-2A-GFP were maintained in mTeSR1 medium containing 10 μM Y-27632 for at least 1 h prior to cell sorting. Cells were then dissociated to single cells by incubation with accutase for 5 min at 37°C, 5% CO_2, humidified. Accutase was removed by centrifugation at 190 $g_{(av)}$ for 3 min and the cell pellet was resuspended in 500 μl cell sorting medium (mTeSR1 containing 10 μM Y-27632, 2 mM EDTA, 50 U ml^{-1} penicillin and 50 μg ml^{-1} streptomycin) and kept on ice until sorting. Single GFP+ cells were sorted by fluorescence-activated cell sorting (FACS) on 96-well geltrex coated plates that contained 1:1 mTeSR1:Conditioned mTeSR1 supplemented with 10μM Y-27632 and

1x penicillin/streptomycin (recovery medium) using Beckman Coulter MoFlo XDP with a CyCLONE (with the following parameters; nozzle: 100 µm, sheath pressure: 30 psi, laser: 488nm (blue) 150mW, filter: 530/40, and sample temperature: 4-6^{o}C). The 96-well plates were briefly centrifuged at 190 $g_{(av)}$ for 3 min to help iPSC attachment to the matrix and increase survival. Sorted GFP+ cells were also pooled and seeded on 10 cm geltrex coated dishes. After 2 days the medium was replaced with fresh recovery medium without Y-27632 and medium was replaced daily after that. When colonies started appearing in wells of 96-well plates they were split 1:3 to new 96-well plates. Colonies were also manually picked under the microscope from the 10 cm dishes to wells of 96-well plates, when circular colonies appeared before they started merging. One 96-well "copy" plate was used for DNA extraction (described in 2.16.1) and screening for engineered clones (described in 2.17). Following establishment of edited iPSC clones, all iPSC lines were adapted to Essential 8 medium.

2.16. Genomic DNA extraction

2.16.1. Genomic DNA extraction from CRISPR/Cas9n iPSC lines (96-well format)

Extraction of gDNA from CRISPR/Cas9n nucleofected iPSC lines was performed in 96-well format for subsequent PCR screening for CRISPR/Cas9n-induced indels (described in 2.17). Briefly, CRISPR/Cas9n nucleofected iPSC lines were EDTA passaged as described in 2.4, transferred to MicroAmp Optical 96-well Reaction Plates 0.2 ml (Applied Biosystems) and covered with adhesive PCR plate seals (Thermo Fisher Scientific). The plates were centrifuged at 200 $g_{(av)}$ to remove the media and cell pellets were resuspended in 20 µl QuickExtract^{TM} solution and mixed well. The gDNA was extracted in a Veriti 96-Well Fast Thermal Cycler (Applied Biosystems) by incubation at 68 °C for 15 min and 95 °C for 8 min.

2.16.2. Genomic DNA extraction using DNeasy Blood & Tissue Kit (QIAGEN)

Genomic DNA was extracted from iPSCs using the DNeasy Blood & Tissue Kit (QIAGEN). Briefly, iPSCs were detached using EDTA as described in 2.4 and centrifuged for 5 min at 300 $g_{(av)}$. The iPSC pellet was resuspended in 200 µl PBS, then 20 µl Proteinase K were added followed by 200 µl Buffer AL. The sample was mixed thoroughly by vortexing, and incubated at 56°C for 10 min. 200 µl ultrapure RNAse-free ethanol (VWR) were added to the sample, followed by vortexing. The mixture was loades onto a DNeasy Mini spin column placed in a 2 ml collection tube and centrifuged at 6000 $g_{(av)}$ for 1 min. Flowthrough was discarded and 500 µl Buffer AW1 were added onto the spin column and centrifuged for 1 min at 6000 $g_{(av)}$. Flowthrough was discarded, then 500 µl Buffer AW2 were added onto the spin column and centrifuged for 3 min at 20,000 $g_{(av)}$ to dry the DNeasy membrane. The column was placed in a clean 1.5 ml Eppendorf tube and 200 µl Buffer AE were added directly onto the DNeasy membrane. The column was incubated at room temperature for 1 min, and then centrifuged for 1 min at 6000 $g_{(av)}$ to elute the genomic DNA. Genomic DNA concentration was measured using a NanoDrop Spectrophotometer.

2.17. PCR screening for CRISPR/Cas9n edited iPSC lines

Screening for genetically engineered iPSC lines was performed by PCR using primers that amplify 100-200bp regions surrounding the nuclease DNA cleavage site. For CRISPR pair 0, the forward and reverse primers 5'- GGGCTCCAAAGACAGAACAG -3' and 5'- ATCCCCATTCCAGTTTCCAT -3' respectively, were designed using Primer 3, to generate an amplicon of 185 bp flanking the CRISPR/Cas9n target site in exon 2 of *C9orf72.*

For the PCR, 2 µl of the extracted gDNA from 2.16.1 were used in a 25 µl total volume PCR reaction containing 0.25 µl dNTPs, 5 µl Herculase Buffer (Agilent), 0.25 µl

forward primer 5'- GGGCTCCAAAGACAGAACAG -3', 0.25 µl reverse primer 5'- ATCCCCATTCCAGTTTCCAT -3', 0.25 µl Herculase II (Agilent) and 17 µl H_2O. The reactions were plated on MicroAmp Optical 96-well Reaction Plates (0.2 mL) and the PCR was performed in a Veriti 96-Well Fast Thermal Cycler with the following cycling parameters: 95°C for 3 min, 35 cycles 95°C for 30 sec, 58°C for 15 sec and 72°C for 30 sec, and 72°C for 7 min with final hold at 4°C.

Screening iPSC clones for the presence of indels was carried out by resolving 5 µl of the 185 bp region amplified products pre-mixed with 1 µl of 6x orange G, on a 3% agarose TBE gel pre-stained with GelRed dye. An additional lane of 2 µl of hyperladder 100 bp (Bioline) was included. High resolution agarose gels were run at 50 V for 30 min and then 100V for 2 h in 1x TBE buffer to allow good separation and detection of products containing indels under UV light. PCR products were gel-extracted as described in 2.20 and the extracted DNA bands were sequenced using the CRISPR/Cas9n pair 0, forward (5'- GGGCTCCAAAGACAGAACAG -3') primer as described in 2.18.

2.18. Sanger sequencing

Prior to sequencing, PCR Clean-up was performed by incubation of 15 µl of PCR product with 6 µl Exonuclease I/FastAP™ Thermosensitive Alkaline Phosphatase mix (Thermo Scientific) (mix: 50 µl Exonuclease I, 200 µl FastAP Thermosensitive Alkaline Phosphatase and 750 µl H_2O) at 37°C for 15 min, followed by inactivation at 85°C for 15 min. Purified products were then used in sequencing reactions with the BigDye Terminator V.3.1 Cycle Sequencing Kit (Applied Biosystems). Briefly, 5 µl of the purified PCR product were mixed with 0.5 µl BigDye (Applied Biosystems), 2 µl sequencing buffer (Applied Biosystems), 1 µl sequencing primer and 2.5 µl ddH_2O. The sequencing reaction was performed in a Veriti 96-Well Fast Thermal Cycler with

the following cycling parameters: 94°C for 1 min, 25 cycles of 94°C for 30 sec, 50°C for 15 sec, 60°C for 4 min and final hold at 4°C. BigDye Terminator removal was done by centrifugation of sequencing reactions through Sephadex (Sigma) containing filter plates (Corning). Briefly, 2.9 g Sephadex G-50 (Sigma) were mixed with 40 ml autoclaved distilled H_2O and left to hydrate at RT on a tube roller (VWR) for 30 min. 350 µl of the hydrated Sephadex were added on each well of a Corning™ FiltrEX 96-well filter plate (Sigma) which was then centrifuged for 3 min at 750 $g_{(av)}$ to remove the water. The entire sequencing reaction (10 µl) was diluted 1:2 in distilled H_2O and loaded on the Sephadex plate. The sequencing reactions were filtered through Sephadex onto a clean 96-well non-skirted PCR plate (VWR) via centrifugation for 5 min at 910 $g_{(av)}$. Sequencing was performed using Applied Biosystems 3730XI Sequencer and analysis of results was carried out using the Sequencher software (Gene Codes Corporation).

2.19. Agarose gel electrophoresis

5 µl of PCR product were mixed with 1 µl 6x orange G (final 1x concentration 3.3% v/v glycerol and 3.33% w/v Orange G dye in ddH_2O) loading dye and separated on a 1.5% agarose (Sigma) Tris-Borate EDTA (TBE) (89 mM Tris, 89 mM boric acid, 2 mM EDTA) (Thermo Fisher Scientific) gel pre-stained with GelRed (Biotium) dye. An additional lane of 2 µl of hyperladder 100 bp (Bioline) was also included. Agarose gels were run at 100 V for 30-45 min in 1x TBE buffer and visualised under UV light.

2.20. Gel extraction

PCR products were extracted from agarose gels using the QIAquick Gel Extraction Kit (Qiagen). Briefly, the DNA bands were excised from the gel under brief UV light exposure using sterile scalpels and placed into eppendorf tubes. Then, all DNA slices were weighed and for every 100 mg of gel (1 gel volume), 3 volumes Buffer QG (proprietary) were added into the tubes (400 mg maximum amount of gel per spin column). The samples were incubated at 50°C for 10 min with frequent vortexing until the gel slice had dissolved. Due to optimal DNA adsorption to the membrane at pH ≤7.5, the yellow colour of Buffer QG was used as pH indicator in all samples. In case of orange or violet colour indication 10 µl of 3 M sodium acetate, pH 5.0 were added to the samples. Then, 1 gel volume of isopropanol was added to the sample and mixed. The samples were transferred on QIAquick spin columns placed onto 2 ml collection tubes. The columns were centrifuged at 16,000 $g_{(av)}$ for 1 min. The flow-through was discarded and QIAquick columns were placed onto the 2 ml collection tubes. 500 µl Buffer QG were added to each QIAquick column and centrifuged at 16,000 $g_{(av)}$ for 1 min. The flow-through was discarded and QIAquick columns were placed onto the 2 ml collection tubes. 750 µl of Buffer PE (10 mM Tris-HCl pH 7.5, 80% ethanol) were added to each QIAquick column to wash. The columns were then centrifuged at 16,000 $g_{(av)}$ for 1 min and placed back onto the 2 ml collection tubes. A final centrifugation step at 16,000 $g_{(av)}$ for 1 min was applied to remove any residual ethanol from the wash Buffer PE. Finally, QIAquick columns were placed onto a clean 1.5 ml eppendorf tube and 20 µl of elution Buffer EB (10 mM Tris-HCl, pH 8.5) were added to the centre of the QIAquick membrane and centrifuged at 16,000 $g_{(av)}$ for 1 min.

2.21. RT-qPCR genetic analysis for the detection of karyotypic abnormalities in iPSCs

C9orf72$^{+/-}$ and *C9orf72*$^{-/-}$ CRISPR/Cas9n edited iPSC lines were characterised for the presence of common karyotypic abnormalities reported in iPSCs/ESCs using the qPCR hPSC Genetic Analysis Kit (StemCell Technologies). Genomic DNA was extracted from all iPSC lines as described in 2.16.2. 1.5 µL of Genomic DNA Control were added to 56.5 µL of nuclease-free water (final concentration 5 ng/µL) in a 0.65 mL microcentrifuge tube, followed by vortexing on high for 3 - 5 sec. For each iPSC line, 290 ng of genomic DNA were mixed with nuclease-free water up to 58 µL (final concentration 5 ng/µL) in 0.65 mL microcentrifuge tubes. 145 µL of pre-mixed qPCR Master Mix (2X) and ROX Reference Dye were added in each genomic DNA sample and the genomic DNA control, followed by gentle pipetting to mix the solutions. The primer-probe stock mixes were prepared by resuspending each genetic assay in 33 µL of TE Buffer. For each assay, 12 µL of stock primer-probe mix were mixed with 60 µL ddH_2O. The RT-qPCR reaction consisted of 7 µL gDNA from each sample (in duplicate reactions) and 3 µL of each primer-probe mix, in separate reactions for each assay. The RT-qPCR was performed on a Stratagene mx3000p System using the following conditions: 95^{o}C for 3 min, 40 cycles 95^{o}C for 15 sec, 60^{o}C for 1 min. Chr 4p genetic assay was used as reference for the calculation of ΔCt values and genomic DNA control sample was used as reference for the calculation of ΔΔCt values. Copy number was calculated using the formula: Copy number = ($2^{-\Delta\Delta Ct}$) x 2.

2.22. Bisulfite sequencing

Bisulfite sequencing was performed by collaborator Dr Ming Zhang in Pror Ekaterina Rogaeva's laboratory on gDNA extracted from fibroblasts, iPSCs and temporal iPSC-CN samples from controls and *C9orf72* HRE carriers. Briefly, 100 ng gDNA were used for bisulfite treatment using the Imprint DNA Modification kit (Sigma) as per manufacturer's instructions. Following the bisulfite conversion, the 5' region of the $(GGGGCC)_n$ repeat (5' CpG island) was amplified by semi-nested PCR. The CpG sites amplified are shown in Figure 2.4. 2 µl of bisulfite-treated DNA were amplified using FastStart PCR Master Mix (Roche) in a touchdown PCR (for primers BSP_1F, BSP_1R: Tm from 68°C to 58°C at a rate of −1°C/cycle and fixed at 58°C for 30 cycles; for primers BSP_2F, BSP_2R: Tm from 67°C to 57°C at a rate of −1°C/cycle and fixed at 57°C for 20 cycles; 3 min extension for each cycle). Primer sequences are shown in Table 2.3 and their binding sites in Figure 2.4. For detection of methylation, inspection of sequencing electropherograms and an integrated tool within Mutation Surveyor version 4.0 (Softgenetics) were used. Only samples with >95% conversion rate of non-CpG C were included in the analysis. For each sample, the total number of methylated CpG sites was estimated based on the following scoring system: unmethylated CpG (T peak) or methylated CpG (T/C double peaks). C peak indicating 100% methylation was not observed in any of the samples.

Table 2.3 Bisulfite sequencing primers.

Primer	Sequence (5'-3')
BSP_1F	TTTATTAGGGTTTGTAGTGGAGTTTT
BSP_1R	AAATCTTTTCTTATTCACCCTCAAC
BSP_2F	TATTAGGGTTTGTAGTGGAGTTTT
BSP_2R	CCACACCTACTCTTACTAAACCC

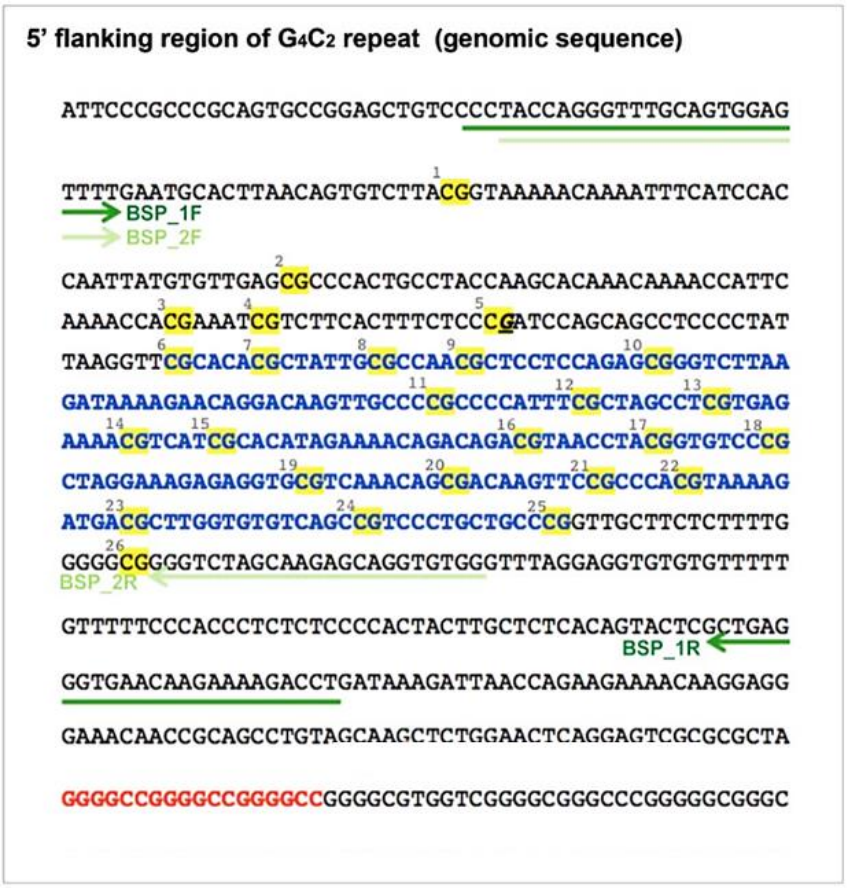


Figure 2.4 Schematic illustration of the 5' CpG island upstream of the *C9orf72* HRE. The genomic sequence (5'-3') of the region upstream of the *C9orf72* HRE is shown with highlighted sequences of the 5' CpG island in blue, the *C9orf72* HRE in red, and the CpG dinucleotides in yellow. The positions of the primers for the bisulfite sequencing are shown by colored arrows. The investigated CpG sites are numbered at the upper left corner. Figure adapted from (Xi *et al.*, 2013).

2.23. $(G_4C_2)_n$-methylation assay

The $(G_4C_2)_n$-methylation assay used to assess the methylation of the *C9orf72* HRE (Xi *et al.*, 2015b) . Figure 2.5 provides an overview of the assay which is a combination of the methylation specific PCR (MSP) and the repeat primed PCR (rp-PCR) methods. The first step of the assay is an MSP, where genomic DNA is modified by sodium bisulfite to convert all unmethylated, but not methylated, cytosines to uracils. Then the PCR results in the replacement of uracils by thymines. Therefore, the $(GGGGCC)_n$ repeat can be converted either to $(GGGGTC)_n$ when methylated, or to $(GGGGTT)_n$ when unmethylated. The two states can be discriminated in a rp-PCR, which is also used to detect the HRE, with the use of specific labelled primers; FAM for methylated DNA amplification (blue channel) and HEX for unmethylated DNA amplification (green channel).

If the G_4C_2-repeat is methylated in all DNA copies, only the blue channel is expected to have products; and if the G_4C_2-repeat is unmethylated in all DNA copies, only the green channel is expected to have products. However, if the G_4C_2-repeat is methylated in some copies and unmethylated in others, both channels will have products.

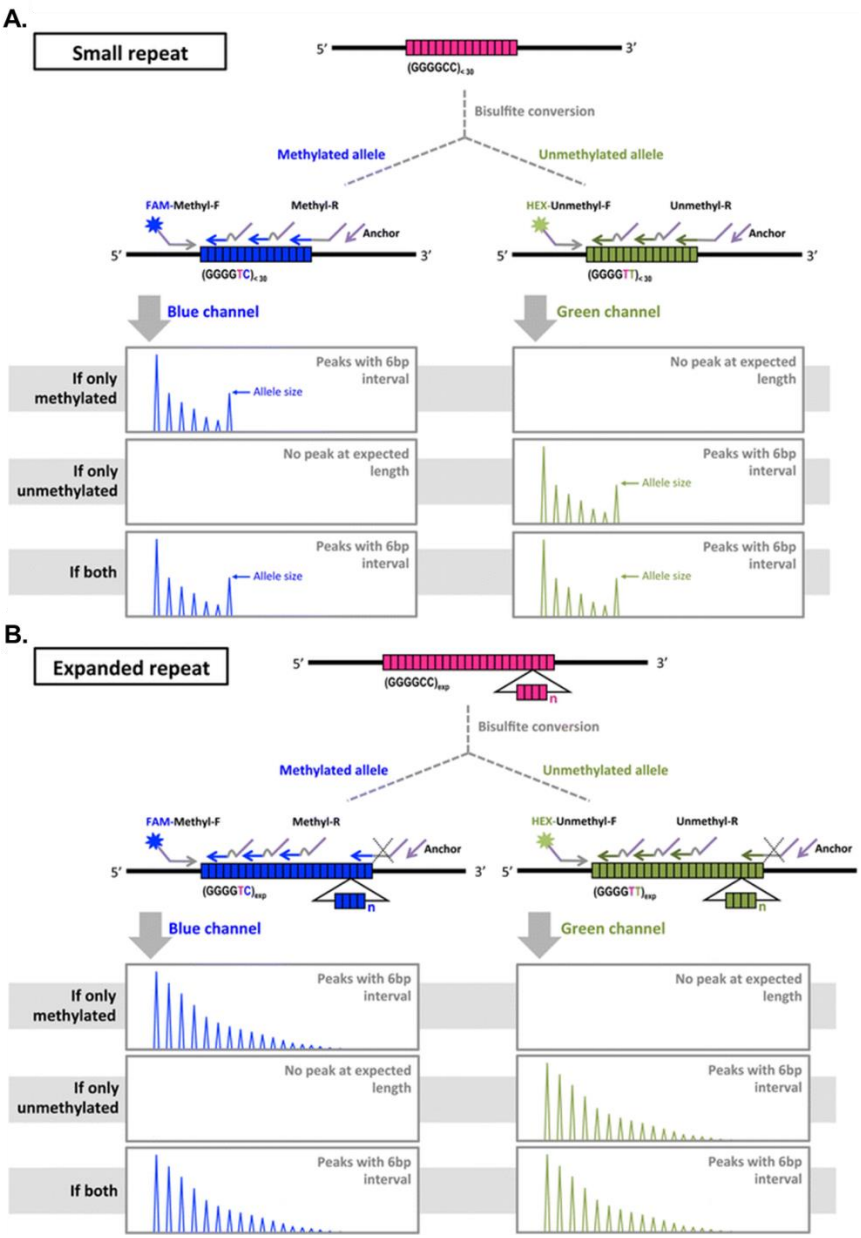


Figure 2.5 Schematic overview of the $(G_4C_2)_n$-methylation assay.

The $(G_4C_2)_n$-methylation assay combines the repeat-primed PCR (rpPCR) method that detects the presence of the expanded repeats with a methylation-specific PCR (MSP) that detects the presence of methylated expanded repeats in *C9orf72*. Briefly, bisulfite-treated DNA is amplified by rpPCR using either methylation-specific FAM labelled $(GGGGTC)_n$ or unmethylation-specific HEX labelled $(GGGGTT)_n$ probes for the detection of methylated and unmethylated HRE respectively. The expected results for small repeats are shown in A, whereas for large repeat expansions in B. Blue channel detects methylated repeats whereas green detects unmethylated repeats.

The following $(G_4C_2)_n$-methylation assay was performed by Dr Ming Zhang in Pror Ekaterina Rogaeva's laboratory on gDNA extracted from fibroblasts, iPSCs and temporal iPSC-CN samples from controls and *C9orf72* HRE carriers (as described in 2.16.2).

Methyl-specific (M-primers) and unmethyl-specific (U-primers) primers were designed using MethPrimer software and their sequences are provided in Table 2.4. 0.5–1.0 µg of DNA was used for bisulfite treatment using the EZ DNA Methylation-Lightning™ Kit (Zymo) according to the manufacturer's protocol. Bisulfite treated DNA was eluted in 10 µl elution buffer. 5 µl of the treated DNA were used as template for the methyl-specific (M) PCR and the other 5 µl were used as template for the unmethyl-specific (U) PCR reaction. MSP reactions were performed in 50 µl final volume using FastStart PCR Master Mix (Roche) with 800 nM of 7-Deaza-dGTP (Roche). A known methylated *C9orf72* HRE carrier gDNA sample was used as a positive control for the methylation detection. The PCR conditions were: initial denaturation at 96.5 °C for 13 min, followed by one stage where annealing temperature was decreased from 70 to 63 °C at a rate of −1 °C/2 cycles and then from 61 to 57 °C at a rate of −2 °C/2 cycles and a second stage which had annealing temperature at 55 °C for 35 cycles. For each cycle, 1 min was used for the denaturation and annealing steps, and 3 min was used for the extension step. Ramp rate was 3.2 °C/s for the denaturation and extension steps and 1.3 °C/s for the annealing step. Final extension was at 72 °C for 10 min. Finally, 2.5 µl of PCR product were mixed with HiDi Formamide (Applied Biosystems) and GeneScan 400HD Rox size standard (Applied Biosystems) and resolved on an ABI 3100 DNA Analyzer. Data were visualized by Genotyper software (version 3.6, Applied Biosystem).

Table 2.4 $(G_4C_2)_n$-methylation assay primers.

Primer sets	Primer name	Primer sequences (5'-3')
Methyl-specific primers	**FAM-Methyl-F**	**FAM-**TGTAAAACGACGGCCAGTAGTTTTGGAATTTAGGAGT***C***G***C***
	Methyl-R1	CAGGAAACAGCTATGACC***G***AACCC***G***CCCC***G***ACCAC***G***CCCC***G***ACCCC***G***ACCCC***G***
Unmethyl-specific primers	**HEX-Unmethyl-F**	**HEX-**TGTAAAACGACGGCCAGTAGTAAGTTTTGGAATTTAGGAGT***T***G***T***G
	Unmethyl-R1	CAGGAAACAGCTATGACC***A***AACCC***A***CCCC***A***ACCAC***A***CCCC***A***ACCCC***A***ACCCC***A***A
Anchor primer		CAGGAAACAGCTATGACC

The non-CpG sites covered by the MSP primers are underlined. The CpG sites covered by MSP primers are in bold Italic font. The additionally attached linker sequences are double underlined.

2.24. RNA fluorescence *in situ* hybridisation (FISH)

RNA fluorescence *in situ* hybridisation (FISH) was performed by collaborators in iPSCs and iPSC-CNs using the following protocol. Cells were fixed 4% methanol-free formaldehyde (Pierce) for 10 min at RT, then dehydrated in a graded series of alcohols, air-dried and rehydrated in PBS. Then cells were permeabilised for 10 min in 0.1% Triton X-100 in PBS, briefly washed in PBS and incubated for 30 min in pre-hybridisation solution (40% formamide, 2x saline-sodium citrate (SSC), 1 mg/ml tRNA, 1 mg/ml salmon sperm DNA, 0.2% BSA, 10% dextran sulphate, 2 mM ribonucleoside vanadyl complex) at 67°C. Hybridisation solution (40% formamide, 2x SSC, 1 mg/ml tRNA, 1 mg/ml salmon sperm DNA, 0.2% BSA, 10 % dextran sulphate, 2 mM ribonucleoside vanadyl complex, 0.2 ng/ll (C4G2)4 LNA probe, 50 TYE563-labelled, Exiqon) was incubated with the cells for 2 h at 67°C. Cells were washed for 30 min at 67°C with high-stringency buffer (2x SSC, 0.1% Triton X-100) and then for 20 min each, in 0.2x SSC buffer. Nuclei were counterstained with 1 µg ml^{-1} 4′,6-diamidino-2-phenylindol (DAPI). Coverslips were then dehydrated in 70% and 100% EtOH and mounted onto slides in Vectashield. Images were acquired using an LSM710 confocal microscope (Zeiss) using a plan-apochromat 40x/1.4 NA oil immersion objective. RNA foci and cell counting was performed in Fiji-ImageJ (Schindelin *et al.*, 2012). Five to ten z-stacks were acquired for each field and at least 4 independent fields imaged. Maximum intensity projections of each z-stack were analyzed in Fiji-ImageJ and nuclear RNA foci automatically counted by using the analyze particles function to identify nuclei and the find maxima function to identify RNA foci in each nucleus. At least 180 neurons were counted for each independent differentiation.

2.25. Southern blot

Southern blotting was performed by collaborators in gDNA extracted from iPSCs and iPSC-CNs. Genomic DNA was extracted from cells using the Qiagen Puregene kit following manufacturer's instructions. 5 µg of genomic DNA were digested with EcoRI and BamHI and run on a 0.8% agarose gel at 60 V overnight. The gel was agitated for 30 min in 1L of denaturation buffer (National Diagnostics) followed by 1L of neutralisation buffer (National Diagnostics). DNA was transferred onto a positively charged nylon membrane (Roche Applied Science) via capillary action overnight using 20x SSC transfer buffer. Following transfer, DNA was baked onto the nylon membrane at 80 ºC for 2 hours, then briefly wetted in water and incubated in DIG Easy Hyb buffer (Roche) at 46 ºC for 3 hours. It was then transferred into fresh DIG Easy Hyb buffer with 100 µg/ml salmon sperm DNA (Thermo) and 1 kb DIG labelled probe and left overnight at 46 ºC. The membrane was washed for 15 minutes in 0.5x SSC with 0.1% SDS at a rising temperature from 46ºC to 65ºC, before being washed twice more at 65ºC. Following washes, the membrane was processed using the Roche DIG wash and block buffer set following manufacturer's instructions and using the AP-conjugated anti-digoxigenin (11093274910, Roche) at 1:10,000. The membrane was incubated in CDP-star ready to use (Roche) sealed in acetate and exposed to film (Roche Lumi-Film chemiluminescent detection film) for >1 hour before being developed using an X4 automatic processor (XOgraph).

2.26. Meso Scale Discovery (MSD) immunoassay

Two Meso Scale Discovery (MSD) immunoassays, a poly(GP) and a poly(GR) assay, were established by Dr Adrian Isaac's group using a previously generated rabbit anti-poly(GP) antibody (Mizielinska *et al*., 2014) and newly generated affinity purified rabbit anti-(GR)7 antibody (Eurogentec) respectively. Briefly, iPSC-CNs were lysed

in RIPA buffer with protease inhibitors (Roche complete mini EDTA-free) and then sonicated using a Soniprep 150 (Renaissance Scientific Limited) probe sonicator. Lysates were centrifuged at 16,000 $g_{(av)}$, the soluble fraction was collected and protein concentration was determined by DC protein assay (Bio-Rad) following manufacturer's instructions and equal amounts of protein used in the immunoassay. Capture was performed with either unlabelled anti-poly(GP) or anti-poly(GR) antibodies. Detection on an MSD sector imager utilised either sulfo-tagged anti-poly(GP) antibody or biotinylated anti-poly(GR) antibody followed by sulfo-tagged streptavidin. Prior to analysis, the average reading from a calibrator containing no peptide was subtracted from each reading. A four-parameter logistic regression curve was fit to the values obtained using the peptide calibrators using GraphPad Prism, and concentrations interpolated. Specificity was confirmed with a peptide cross-reactivity assay using (GP)7, (PR)7 or (GR)7 synthetic peptides (Biogenes) at a concentration of 100 ng/ml. The lower limit of detection was calculated after fitting a fourparameter logistic regression curve using the MSD workbench 4.0 software.

2.27. Immunocytochemistry

Cells were fixed in 4% (v/v) paraformaldehyde (4 ml 10% formaldehyde Ultra pure, methanol free (Polysciences), 5 ml H_2O and 1 ml 10x phosphate-buffered saline (PBS) (Thermo Fisher Scientific)) for 15 min at RT followed by permeabilisation with 0.25% (v/v) Triton X-100 in PBS for 10 min at RT and blocking with 1% BSA/PBST for 30 min at RT. They were then incubated in primary antibody diluted in 1% BSA/PBST either overnight at 4°C or for 2 h at RT. Primary antibody was removed and cells were washed 3 times in PBS. Secondary antibodies were diluted in 1% BSA/PBST and applied for 1 h at RT in the dark. A list of primary and secondary antibodies used is shown in Table 2.5. Cells were washed 5 times with PBS, nuclei

were counterstained using 1 µg ml^{-1} 4′,6-diamidino-2-phenylindol (DAPI) (Thermo Fisher Scientific) for 5 min and mounted with ProLong® Gold antifade reagent (Thermo Fisher Scientific). Images were aquired using a Leica DM5500B fluorescent microscope.

Table 2.5 List of primary and secondary antibodies used for immunocytochemistry.

Antibody	Species	Dilution	Company	Product code
Primary antibodies:				
SSEA4	Mouse	1:300	Biolegend	MC-813-70
OCT4	Goat	1:400	Santa Cruz	sc-8629
TRA-1-81	Mouse	1:300	Biolegend	330702
NANOG	Rabbit	1:400	Cell Signaling	4903
KI67	Mouse	1:500	BD Transduction Laboratories	550609
FOXG1	Rabbit	1:300	Abcam	ab18259
TBR1	Rabbit	1:300	Abcam	ab31940
CTIP2	Rat	1:300	Abcam	ab18465
SATB2	Mouse	1:100	Abcam	ab51502
BRN2	Goat	1:100	Santa Cruz	sc-6029
BIII Tubulin (TUJ1)	Mouse	1:1,000	BioLegend	801202
BIII Tubulin (TUJ1)	Rabbit	1:1,000	BioLegend	802001
Secondary antibodies:				
AlexaFluor 568 anti-goat IgG	Donkey	1:500	Thermo Fisher Scientific	A-11057
AlexaFluor 488 anti-mouse IgG	Donkey	1:500	Thermo Fisher Scientific	A-21202
AlexaFluor 568 anti-mouse IgG	Goat	1:500	Thermo Fisher Scientific	A-11001
AlexaFluor 488 anti-rabbit IgG	Goat	1:500	Thermo Fisher Scientific	A-11008
AlexaFluor 488 anti-rat IgG	Goat	1:500	Thermo Fisher Scientific	A-11006

2.28. Lentiviral shRNA transduction

For the generation of the stable *C9orf72* knock-down cell line, HEK293 cells were transduced with lentiviral short hairpin RNAs (shRNAs) against *C9orf72* V3LHS_303669 and V3LHS_303670 (Open Biosystems). The mature antisense targeting sequences of the shRNAs are 5'- TTCACTAGAGTGTCTCTGT -3' and 5'-ATGCATCCATATTCTTCCT -3' respectively. Mock transduction with non-targeting shRNA was performed as a control. $5x10^4$ HEK293 cells were seeded per well of a 24-well plate and maintained in DMEM Glutamax (Thermo Fisher Scientific) supplemented with 10% (v/v) FBS (Thermo Fisher Scientific). 24h post-seeding, cells were transduced with 250 µl serum-free DMEM containing 4 µg ml^{-1} polybrene (Sigma) and $7x10^5$ viral particles (Multiplicity of infection (MOI) 10). Cells were then spinfected at 620 $g_{(av)}$ for 45 min and maintained at 37°C, 5% CO_2, humidified. After 6 hours, 500 µl of serum-containing DMEM medium was added to each well. 24 h post-transduction cells were transferred in 5 cm dishes. For the generation of stable lines cells were maintained in DMEM Glutamax containing 10% FBS and 2 µg ml^{-1} puromycin (Thermo Fisher Scientific) for 3 days which was then switched to DMEM Glutamax containing 10% FBS and 0.5 µg ml^{-1} puromycin for further 5 days. At the end of the puromycin selection, the surviving HEK293 cells were expanded further in DMEM Glutamax with 10% FBS as described in 2.6.

2.29. Western blot

Cells were lysed in RIPA buffer (10 mM Tris-Cl, 1 mM EDTA, 0.5 mM EGTA, 1% Triton X-100, 0.1% sodium deoxycholate, 0.1% SDS and 140 mM NaCl) supplemented with Complete protease inhibitor cocktail (Roche) and 1% (v/v) phenylmethylsulfonyl fluoride (PMSF) (Sigma) for 30 min on ice. Cell lysates were centrifuged for 10 min at 16,000 $g_{(av)}$ at 4°C and protein concentrations in the soluble

fractions were estimated using the BioRad DC Protein Assay Kit and absorbance values were measured in a Tecan Spark plate reader. Equal amounts of protein and 3 µl of Precision Plus Protein™ All Blue Standards molecular weight marker (BioRad) were electrophoresed on NuPAGE 10% Bis-Tris Gels (Invitrogen) for 1 h 30 min at 150V using 1x NuPAGE MES SDS Running Buffer (Life Technologies) and a XCell SureLock™ Mini-Cell Electrophoresis System (Thermo Fisher Scientific). Proteins were then transferred on nitrocellulose membranes (Whatman) using the XCell SureLock™ Mini-Cell Electrophoresis System and ice-cold transfer buffer (25 mM Tris, 190 mM glycine, 20% methanol (vol/vol). pH 8.3) for 1 h at 30V on ice. Membranes were blocked with 5% BSA/PBST for 1 h at RT with agitation on a tube roller (VWR). Primary antibodies were diluted in 5% milk in PBST (PBST-M) and applied overnight at 4ºC with agitation on a tube roller. Primary antibody solution was removed and the membranes were washed in PBST 3 times, 5 min each with agitation on a tube roller. Membranes were incubated with secondary antibodies for 1 h at room temperature in a 1:5,000 dilution in 5% milk in PBST-M with agitation on a tube roller. Secondary antibody solution was removed and membranes were washed in PBS 5 times, 5 min each with agitation on a tube roller. Table 2.6 shows the list of primary and secondary antibodies used in this thesis. Western blots were visualised using an Odyssey Infrared imaging system (LI-COR Biosciences) and image analysis was performed using the LiCOR Image Studio Lite 5.2 (LI-COR Biosciences).

Table 2.6 List of primary and secondary antibodies used in western blots.

Antibody	Species	Dilution	Company	Product code
Primary antibodies:				
C9orf72	Rabbit	1:1,000	Proteintech	22637-1-AP
C9orf72	Rabbit	1:1,000	Sigma	HPA023873
GLAST	Rabbit	1:1,000	Abcam	AB416
GFAP	Rabbit	1:1,000	Dako	Z0334
PINK1	Rabbit	1:1,000	Cell Signalling	6946
MFN2	Rabbit	1:1,000	Cell Signalling	9482
TIM23	Mouse	1:1,000	BD Transduction Laboratories	611223
Phospho-ubiquitin (Ser65)	Rabbit	1:1,000	Cell Signalling	70973
GAPDH	Mouse	1:10,000	Ambion	AM4300
Beta-Actin	Mouse	1:10,000	Sigma	A3853
Beta-III-Tubulin	Rabbit	1:10,000	BioLegend	802001
Beta-III-Tubulin	Mouse	1:10,000	BioLegend	801202
Secondary antibodies:				
AlexaFluor 680 anti-mouse IgG	Goat	1:5,000	Invitrogen	A-21058
IRDye 800 anti-rabbit IgG	Goat	1:5,000	Rockland Immunochemicals	926-32211

2.30. Mitophagy induction with Oligomycin/Antimycin A treatment

Briefly, mitochondrial damage was induced by combined treatment with Oligomycin, an inhibitor of ATP synthase, and Antimycin A, an inhibitor of cellular respiration, causing increased superoxide generation coupled with loss of mitochondrial membrane potential (ΔΨm). Control and *C9orf72* knockout iPSC-CNs at day 90 *in vitro* were treated with 1 µM Oligomycin/Antimycin A (both Sigma, dissolved in DMSO) diluted in N2B27 medium for 12 h, 24 h and 36 h or with DMSO for 36 h as control. The N2B27 medium in all iPSC-CN lines was changed 1 h prior to the start of the treatment. Oligomycin/Antimycin A treatment medium was changed every 12 h in all samples. Cells were lysed and prepared for western blot as described in 2.29.

2.31. RNA extraction

RNA extraction was performed using Trizol (Thermo Fisher Scientific) according to manufacturer's instructions. Briefly, frozen Trizol lysates were thawed on ice before 200 µl of chloroform was added per 1 ml of Trizol reagent. Samples were vortexed for 10 sec, incubated at RT for 2-3 min and centrifuged at 12,000 $g_{(av)}$ for 15 min at 4ºC. The aqueous phase was collected in RNAse-free tubes on ice and 0.5 µl glycogen (Thermo Fisher Scientific) were added per tube and mixed briefly. RNA was precipitated by adding 500 µl 100% isopropanol (VWR) for each ml of Trizol reagent used and inverting 5 times. Tubes were incubated at RT for 10min and then centrifuged at 12,000 $g_{(av)}$ for 30 min. The RNA pellet was washed 3 times with 1 ml 75% ethanol (75 ml ultrapure RNAse-free ethanol (VWR) and 25 ml RNAse-free H_2O (Ambion)) for each ml of Trizol reagent used followed by centrifugation at 7,500 $g_{(av)}$ for 5min at 4ºC. Pellets were left to air dry for 10 min and were then resuspended in 45 µl RNAse-free water (Ambion) on ice. RNA was then treated with DNAse for removal of gDNA by adding 10 U DNAse I Recombinant, RNAse-free (Roche) and 5 µl 10x incubation buffer (Roche) per sample and incubated at 37ºC for 30min. RNA was precipitated by adding 1/10 of the volume RNA grade sodium acetate (Thermo Fisher Scientific) and 2.5x 100% ethanol followed by incubation on dry ice for at least 15 min and centrifugation at 12,000 $g_{(av)}$ for 30 min. RNA pellet was washed 3 times with 500 µl 75% ethanol followed by centrifugation at 7,500 $g_{(av)}$ for 5 min at 4ºC. Finally the pellet was air-dried and resuspended in 25 µl RNAse free H_2O. RNA concentration was measured using NanoDrop Spectrophotometer.

2.32. RT-PCR

1 µg RNA was used for cDNA synthesis using the Superscript IV reverse transcription kit (Thermo Fisher Scientific) following manufacturer's instructions. Briefly, 1 µg RNA was mixed with 1 µl 50 µM random hexamers, 1 µl 10 mM dNTP mix and nuclease-free H_2O up to 13 µl final volume in nuclease-free PCR tube strips (VWR). Reagents were mixed and incubated at 65^{o}C for 5 min in a Techne PCR machine to anneal the primers to the RNA, followed by 1 min incubation on ice. In an Eppendorf tube, the RT reaction was prepared by mixing 4 µl 5x SSIV Buffer, 1 µl 100mM DTT, 1 µl RNaseOUT Recombinant RNase Inhibitor and 1 µl SuperScript IV Resevrse Transcriptase (200 U/µl). The RT reaction mix was added to the annealed RNA and the following RT incubation conditions were followed in a Techne PCR machine: 23^{o}C for 10 min, 50^{o}C for 10 min, 80^{o}C for 10 min and final hold at 4^{o}C. No reverse transcriptase control (NRT), no template control (NTC) and control human brain total RNA (Thermo Fisher Scientific) reactions were also included.

2.33. Reverse transcription quantitative PCR (RT-qPCR)

2.33.1. SYBR Green gene expression analysis

Primers were designed using Primer Blast NCBI tool https://www.ncbi.nlm.nih.gov/tools/primer-blast/ and ordered from Sigma. All primers were tested and optimised using serial dilutions of cDNA (1, 1:10, 1:100, 1:1,000, 1:10,000) in SYBR green reactions as described below. Single dissociation curves, linearity (R^2) of the standard curve and primer amplification efficiency were assessed for every primer pair.

cDNA was diluted 1:10 and RT-qPCR was performed in a total volume of 20 µl, consisting of 10 µl Power SYBR Green (Thermo Fisher Scientific), 1 µl forward primer

(10 µM), 1 µl reverse primer (10 µM), 2 µl cDNA and 6 µl ddH_2O per reaction. All reactions were performed in technical duplicate. The real-time cycling program was: 95°C for 10 min, 40 cycles of 95°C for 30 sec, 60°C for 1 min, 72°C for 1 min, and one last cycle at 95°C for 1 min, 55°C for 30 sec and 95°C for 30 sec. RT-qPCR was performed using a Stratagene mx3000p System. A list of the primers used is provided in Appendix A. Data were analysed using the ΔΔCt method by Livak, where $\Delta Ct_{S1} = Ct_{GOI(S1)} - Ct_{REF(S1)}$, $\Delta Ct_{S2} = Ct_{GOI(S2)} - Ct_{REF(S2)}$, $\Delta\Delta Ct = \Delta Ct_{S1} - \Delta Ct_{S2}$, and normalised relative expression = $2^{-\Delta\Delta Ct}$, S1 = Sample 1 and S2 = Sample 2 (Livak and Schmittgen, 2001). Statistical analysis was performed as described in 2.36.

2.33.2. Taqman assays for total *C9orf72* mRNA expression analysis

An inventoried TaqMan assay was used for the detection of total *C9orf72* mRNA (Table 2.7). For Taqman RT-PCR reactions, 2 µl cDNA (diluted 1:5) were mixed with 10 µl TaqMan Universal Master Mix (2x), 1 µl TaqMan Gene Expression Assay (20x) (Applied Biosystems) and 7 µl of H_2O. Standard Taqman cycling conditions were applied: 50°C for 2 min, 95°C for 10 min, 40 cycles of 95°C for 15 sec and 60°C for 1 min and finally 25°C for 5 min. All reactions were set up in technical duplicate and RT-qPCR was performed in a Stratagene mx3000p System. Details on the TaMan assays are shown in Table 2.7. Control human brain total cDNA was used as reference for the calculation of relative expression values.

Table 2.7 Information on the TaqMan assays used in this thesis.

Primer	Sequence/ID	Details	Coverage
C9orf72 Assay	Hs00376619_m1	FAM-labelled MGB probe	All transcripts
TBP Assay	Hs00427620_m1	FAM-labelled MGB probe	All transcripts

2.33.3. LNA probes and primers for *C9orf72* transcript analysis

Custom made primers and FAM-labelled LNA probes (IDT) were used for the assessment of the relative expression of different *C9orf72* variants V1, V2 and V3 by RT-qPCR, as previously described (Fratta *et al.*, 2013, Simone *et al.*, 2018). Details of primers used are given in Table 2.8. For the real-time PCR reaction, 1 µl forward primer (10 µM), 1 µl reverse primer (10 µM) and 1 µl LNA probe were mixed with 12.5 µl Gene expression Mastermix (2x) (Thermo Fisher Scientific), 3 µl cDNA (1:10 diluted) and H_2O up to 25 µl total volume. Standard Taqman cycling conditions were used as described. All reactions were performed in technical duplicate and RT-qPCR was performed in a Stratagene mx3000p System.

Table 2.8 Sequences of primers and FAM-labelled LNA probes for the detection of the three *C9orf72* transcripts.

Primer	Sequence/ID	Details	Coverage
C9orf72 V2F	GCGGTGGCGAGTGGATAT	Oligo	Variant 2
C9orf72 V3F	GAGCAGGTGTGGGTTTAGGAGA	Oligo	Variant 3
C9orf72 V2+3R	TGGGCAAAGAGTCGACATCA	Oligo	Variants 2 & 3
C9orf72 V2+3 probe	ATTTGGATAATGTGACAGTTGG	FAM-labelled LNA probe	Variants 2 & 3
C9orf72 V1F	TCATCTATGAAATCACACAGTGTTC	Oligo	Variant 1
C9orf72 V1R	GGTATCTGCTTCATCCAGCTT	Oligo	Variant 1
C9orf72 V1 probe	ATGATGATGATATTGGTGAC	FAM-labelled LNA probe	Variant 1

2.34. Semi-quantitative RT-PCR

A previously published protocol (Waite *et al.*, 2014) was employed with some modifications for the semi-quantitative assessment of the splice variants of *C9orf72.* The PCR reaction comprised of 1 µl cDNA (1:5 dilution), 0.25 µl dNTPs, 2.5 µl Taq buffer, 0.25 µl forward primer (10 µM), 0.25 µl reverse primer (10 µM), 0.1 µl Taq

polymerase and H_2O in a total volume 25 µl. Forward primers were located in exon 1a (5'- GTCAAACAGCGACAAGTTCCG -3') and exon 1b (5'- AGGCGCAGGCGGTGGCGAGTG -3') and the common reverse primer in exon 2 (5'- ACCTGTTCTGTCTTTGGAGCC -3'). TATA-binding protein (*TBP*) (forward ACGCCGAATATAATCCCAAG and reverse CTTCACTCTTGGCTCCTGTG) was used as housekeeping gene to normalise values across different samples. The cycling conditions for exon 1a and *TBP* were initial denaturation at 94°C, 30 cycles of denaturation at 94°C, annealing at 59°C and extension at 72°C, final extension at 72°C for 10 min and hold at 4°C. Cycling conditions for exon 1b were the same but the number of cycles was 25 instead of 30. Products were then separated on agarose gels as described in 2.19.

2.35. Allelic expression imbalance

Prior to allelic expression imbalance analysis, the genotypes of all control and *C9orf72* patient iPSC lines for rs17769294 and rs10122902 were determined by PCR followed by Sanger sequencing. The primers F1: 5'- AGGCACAGAGAGAATGGAAGA -3' and R1: 5'- AACAGCCACAGGTTTGCAAG -3', and F2: 5'- AGAAAATGCTCCAGGTTATGTGA -3' and R2: 5'- TGTCAGCTCGGATCTCATGT -3' were designed to amplify rs17769294 and rs10122902 respectively, using Primer3 (Koressaar and Remm, 2007, Untergasser *et al.*, 2012). For the PCR, 1 µl of the extracted gDNA (described in 2.16.2) from each line was used in a 25 µl total volume PCR reaction containing 0.25 µl dNTPs, 5 µl Herculase Buffer (Agilent), 0.25 µl forward primer, 0.25 µl reverse primer, 0.25 µl Herculase II (Agilent) and 18 µl ddH_2O. The reactions were plated on MicroAmp Optical 96-well Reaction Plates (0.2 mL) and the PCR was performed in a Veriti 96-Well Fast Thermal Cycler with the following cycling parameters: 95°C for 4 min, 35 cycles 95°C for 30 sec, 55°C for 15 sec and 72°C for 30 sec, and 72°C for 7 min with final hold at 4°C. Sequencing of the PCR products was performed as described in

2.18, using primers F1: 5'- AGGCACAGAGAGAATGGAAGA -3' for rs17769294 and F2: 5'- AGAAAATGCTCCAGGTTATGTGA -3' for rs10122902.

Allele discrimination for *C9orf72* was based on real-time genotyping of the SNP rs10122902 c.870C>T using a TaqMan SNP assay (Assay ID C___3137372_30). Genotyping was performed on gDNA and cDNA samples of each study line in triplicate. A total reaction volume of 20 µl contained 10 µl TaqMan Universal PCR Master Mix (2x), 0.5 µl 40x TaqMan assay mix and 40 µg gDNA or 2 µl cDNA. Cycling conditions were: 10 min at 95°C followed by 40 cycles of denaturing at 92°C for 15 sec and annealing/extension at 60°C for 1 min. Real-time genotyping was performed using a Stratagene mx3000p System. Calculation of allelic expression imbalance was performed using the formula:

AEI or rF (relative frequency) = $1 / (2^{\Delta Ct'} + 1)$, where:

$\Delta Ct' = \Delta Ct^{cDNA} - \Delta Ct^{gDNA}$ and $\Delta Ct^{cDNA} = Ct^{allele1} - Ct^{allele2}$, $\Delta Ct^{gDNA} = Ct^{allele1} - Ct^{allele2}$

2.36. Statistical analysis

Statistical analyses were performed on R Studio (R 3.6.0) using one-way ANOVA with post-hoc Tukey HSD multiple comparisons of means. Tukey HSD adjusted *p* values < 0.05 were considered as significant. All ANOVA and Tukey *p* values from the experiments in this thesis are summarised in Appendix B. Pearson correlations were also performed on R Studio (R 3.6.0). Plots were generated using the R package ggplot2.

Chapter 3 Generation and characterisation of cortical projection neurons from *C9orf72* HRE carriers and controls

3.1. Introduction

Recent studies have shown that the HRE in *C9orf72* may cause disease by three non-mutually exclusive mechanisms; haploinsufficiency of the *C9orf72* gene, toxicity of the expanded RNA repeats and toxicity of the DPRs (Ash *et al.*, 2013, Mori *et al.*, 2013c). All these mechanisms may contribute to disease pathogenesis, but the precise degree of involvement of each remains unclear. Further studies are needed to dissect the role of each mechanism as well as their interactive outcome which could explain the heterogeneous clinical presentation of *C9orf72* HRE carriers and elucidate the selective predisposition to FTD or ALS within the spectrum of these diseases.

Understanding the mechanisms of *C9orf72* mediated neurodegeneration has been greatly accelerated by the study of patient-derived iPSC models, which have allowed the investigation of the physiological expression of the *C9orf72* HRE in differentiated neuronal/glial subtypes of interest and within the patient's own genetic background. Studies of iPSCNs have recapitulated pathological hallmarks of the *C9orf72* HRE, most consistently RNA foci and to a lesser extend DPRs. Specifically, sense and/or antisense RNA foci detection has been reported in a wide-range of differentiated CNS cell-types, including CNs, MNs and astrocytes (Almeida *et al.*, 2013, Donnelly *et al.*, 2013, Sareen *et al.*, 2013, Devlin *et al.*, 2015, Zhang *et al.*, 2015, Dafinca *et al.*, 2016, Livesey *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Selvaraj *et al.*, 2018, Simone *et al.*, 2018, Yuva-Aydemir *et al.*, 2019, Zhao *et al.*, 2019b, Ababneh *et al.*, 2020). However, in contrast to the reproducible detection of RNA foci across studies, the detection of DPRs has been more variable, especially in the early studies. The development of new DPR antibodies and more sensitive methods such as the Meso Scale Discovery (MSD) immunoassay, has led to a more robust detection of DPRs in iPSC studies. To date, several DPR species, mainly the sense-derived poly-GP, poly-GR and poly-GA, have been detected in published studies (Almeida *et al.*, 2013, Donnelly *et al.*,

2013, Dafinca *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Westergard *et al.*, 2016, Selvaraj *et al.*, 2018, Simone *et al.*, 2018, Almeida *et al.*, 2019, Cheng *et al.*, 2019, Yuva-Aydemir *et al.*, 2019, Zhao *et al.*, 2019b, Ababneh *et al.*, 2020, Abo-Rady *et al.*, 2020). RNA foci and DPRs are some of the earliest pathological features detected in iPSC models of C9-FTD/ALS, however, their expression is not associated with reduced neuronal viability.

This chapter describes the generation and characterisation of iPSC-CNs from three *C9orf72* HRE carriers and the characterization of RNA foci and DPR in these neurons, evaluating their potential as a model of C9-FTD.

3.2. Aims

The specific aims of this chapter were:

- To generate and characterise human iPSCs from patients carrying the HRE in *C9orf72*.
- To differentiate control and *C9orf72* HRE carrier iPSC lines into CNs.
- To characterise *C9orf72* HRE iPSC-CNs for pathologies relevant to C9-FTD, specifically RNA foci and DPRs.

3.3. Results

3.3.1. Fibroblast derivation and characterisation

The first stage in the generation of patient-derived *in vitro* models for the study of the molecular mechanisms underlying the HRE in *C9orf72* was the establishment of fibroblast lines from two patients carrying the *C9orf72* HRE. Biopsies taken in clinic were dissected as described in 2.2. Fibroblasts started emerging in the dissected cultures approximately 1-2 weeks post-dissection and were further grown to generate lines. Phase contrast microscopy images of the generated *C9orf72* fibroblast lines are shown in Figure 3.1. Patient fibroblasts displayed typical fibroblast morphology similar to control fibroblasts (CTRL-1). Moreover, fibroblast identity was confirmed via positive immunofluorescence for the fibroblast marker S100 calcium-binding protein A4 (S100A4) (Figure 3.2) as well as via RT-qPCR for fibroblast marker genes *vimentin* (*VIM*) and *S100A4* (Figure 3.3).

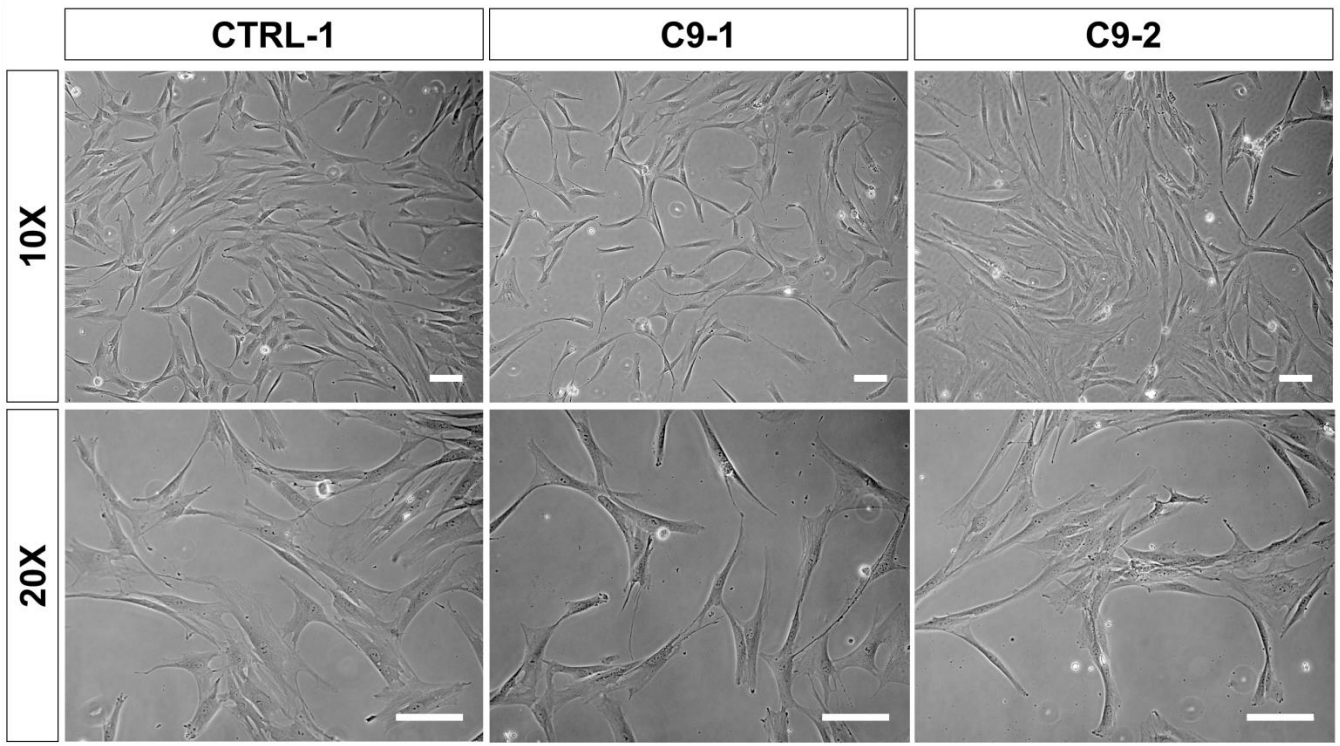


Figure 3.1 Phase contrast images of control and *C9orf72* HRE patient fibroblast lines.

The fibroblast lines C9-1 and C9-2 were derived from skin biopsies from patients carrying the *C9orf72* HRE. Patient fibroblasts (C9-1 and C9-2) display typical fibroblast morphology, spindle-shaped with an oval flat nucleus, similar to that of fibroblasts from an unaffected donor (CTRL-1). Scale bar 100 µm.

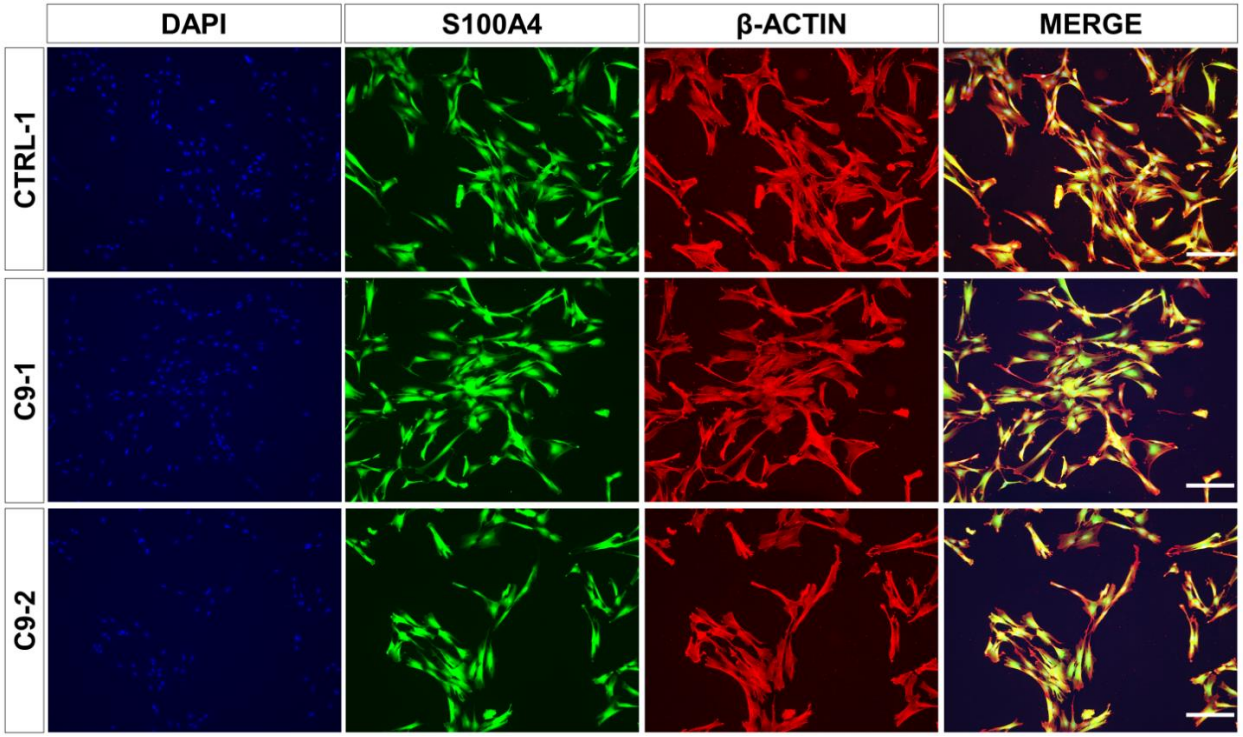


Figure 3.2 *C9orf72* HRE patient fibroblasts express fibroblast-specific marker S100 calcium-binding protein A4 (S100A4). Confirmation of fibroblast identity in C9-1 and C9-2 *C9orf72* HRE patient-derived fibroblast lines via immunofluorescence revealed that the lines were positive for S100A4 fibroblast marker, similar to control fibroblasts CTRL-1. Fibroblast cytoskeleton was stained with β-actin and nuclei were counterstained with DAPI. Scale bar 200 μM.

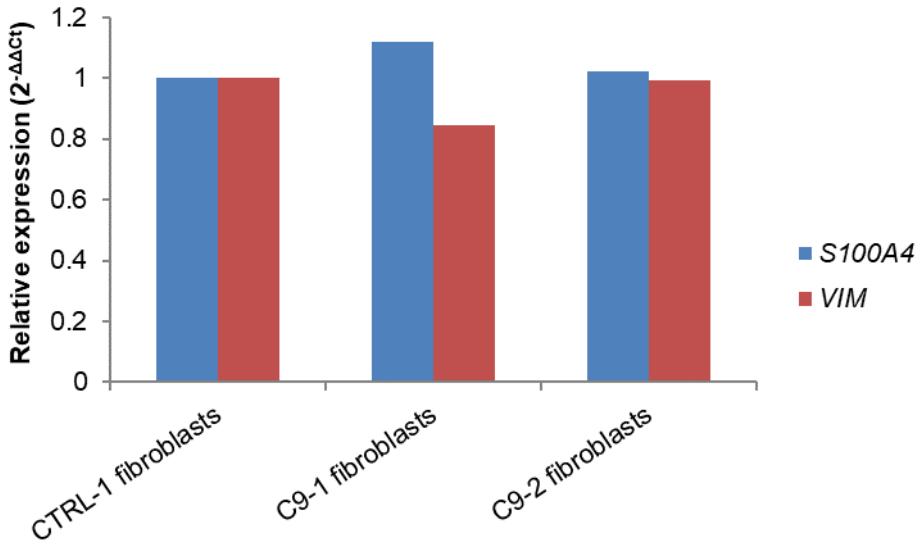


Figure 3.3 *C9orf72* fibroblast lines express fibroblast specific markers *S100A4* and *VIM*.

Real time quantitative PCR was used to measure the expression of two fibroblast marker genes *S100A4* and *VIM* in *C9orf72* HRE fibroblast lines C9-1 and C9-2 compared to control (CTRL-1) fibroblasts. Expression values were calculated relative to CTRL-1 and normalised to *TBP* housekeeping gene. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method (n=1 experiment, performed in technical duplicate).

3.3.2. Characterisation of control and patient-derived iPSC lines

In order to generate iPSCs, patient fibroblasts described in 3.3.1 were transduced with retrovirus expressing Oct4, Sox2, Klf4 and cMyc (Takahashi *et al.*, 2007). Approximately 30 days post-transduction, colonies with a similar appearance to hESCs started to emerge and were picked and expanded for further characterisation. The characteristics of the pluripotent stem cell lines used in this thesis are summarised in 2.3 and Table 2.1. Lines C9-1 and C9-2 were generated in house whereas C9-3 was obtained as a kind gift from Prof Chris Shaw.

Control and *C9orf72* patient iPSCs were further characterised before use in downstream experiments. All iPSC lines displayed a typical ESC-like morphology when observed by phase-contrast microscopy, with characteristic high nuclear-to-cytoplasmic ratio and well-defined borders (Figure 3.4). To further confirm successful reprogramming and pluripotency, immunofluorescence for a panel of pluripotency markers was performed. All iPSC lines were positive for core pluripotency transcription factor OCT4 and ESC-specific cell surface stage-specific embryonic antigen (SSEA-4) (Figure 3.5) as well as the pluripotency factor NANOG and cell surface pluripotency marker TRA-1-81 (Figure 3.6). Furthermore, all iPSCs were further characterised for the expression of pluripotency genes *POU5F1* (OCT4), *SOX2*, *NANOG* and *DNMT3B* and suppression of fibroblast lineage gene *VIM* and *S100A4* via RT-qPCR (Figure 3.7).

Aberrant karyotypic abnormalities have been described for pluripotent lines, mainly arising during the reprogramming process or prolonged cell culture (Mayshar *et al.*, 2010, Martins-Taylor *et al.*, 2011, Taapken *et al.*, 2011). Therefore, prior to using iPSC for downstream experiments G-band

karyotyping was performed externally (TDL Genetics) to ensure that all study lines carry a normal karyotype (Figure 3.8).

Together, these results confirm successful reprogramming of *C9orf72* repeat expansion fibroblasts, whereby four factor reprogramming establishes pluripotency networks by inducing the endogenous expression of *bona fide* pluripotency transcription factor NANOG and other typical ESC cell surface markers whilst suppressing fibroblast identity markers.

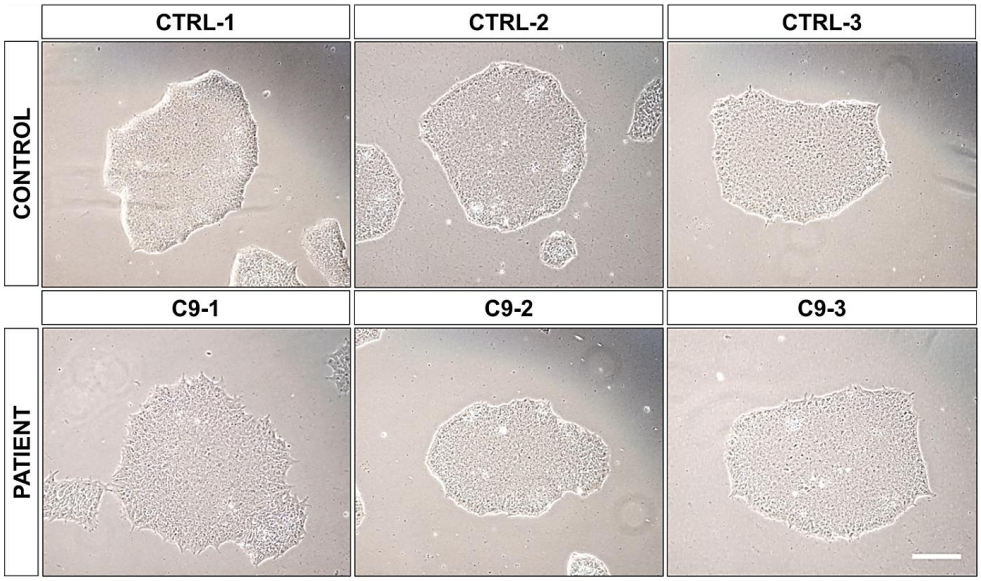


Figure 3.4 Phase contrast images of iPSC lines.
CTRL-1 and CTRL-2 are iPSC lines derived from neurologically normal individuals. CTRL-3 is a hESC line. C9-1, C9-2 and C9-3 are iPSC lines derived from individuals carrying the *C9orf72* repeat expansion. All lines displayed a typical ESC-like morphology. Scale bar 100 µm.

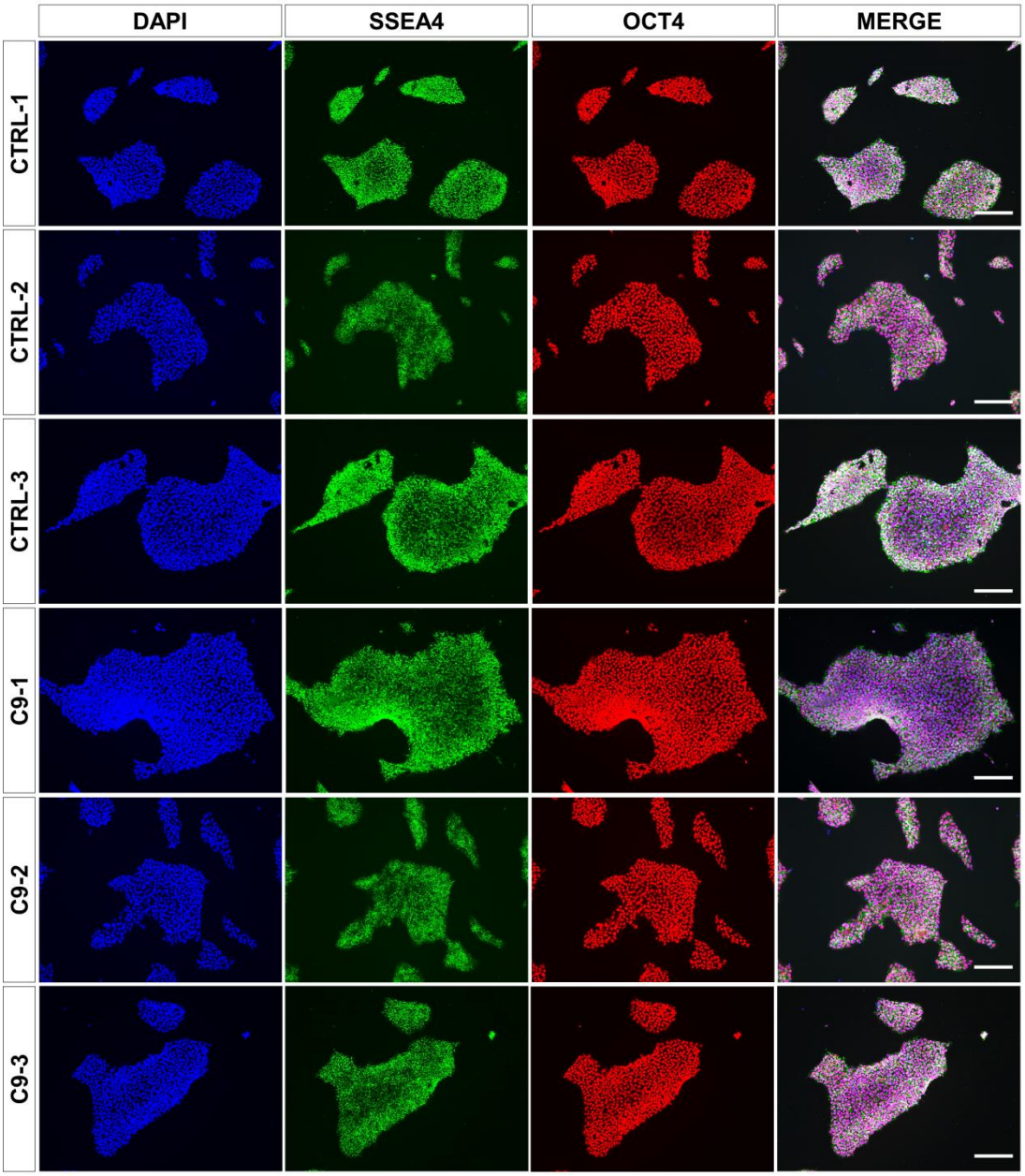


Figure 3.5 Expression of the pluripotency markers OCT4 and SSEA4 in iPSCs. Control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE (C9-1, C9-2 and C9-3) pluripotent lines were characterised by positive immunofluorescence for core pluripotency transcription factor OCT4 (red) and cell surface stage-specific embryonic antigen 4 (SSEA4) (green). Cell nuclei were counterstained with DAPI. Scale bar 200 µm.

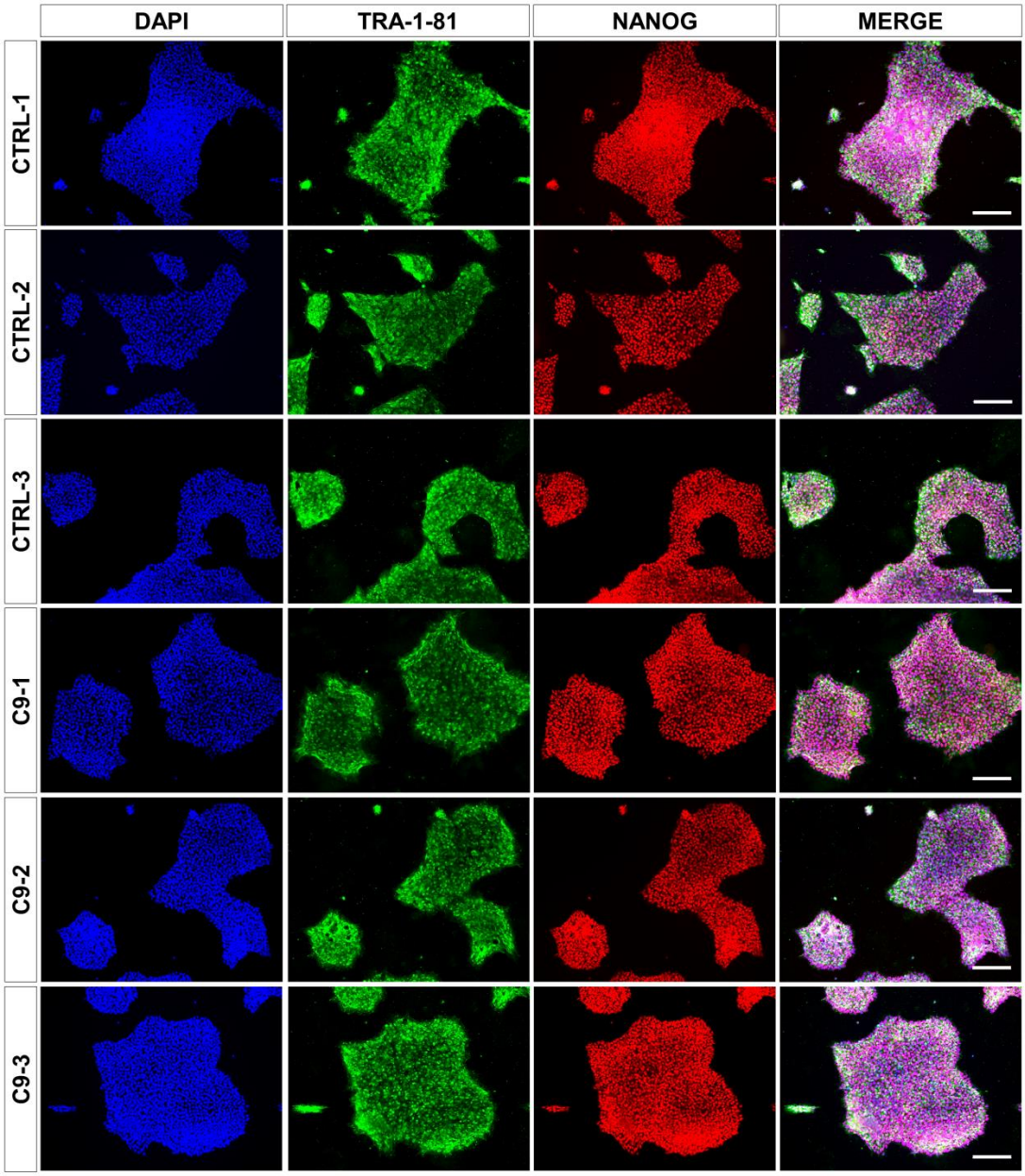


Figure 3.6 Expression of the pluripotency markers TRA-1-81 and NANOG in iPSCs. Control (CTRL-1, CTRL-2 and CTRL-3) and C9orf72 HRE (C9-1, C9-2 and C9-3) pluripotent lines were characterised by positive immunofluorescence for core pluripotency transcription factor NANOG (red) and cell surface pluripotency marker TRA-1-81 (green). Cell nuclei were counterstained with DAPI. Scale bar 200 μm.

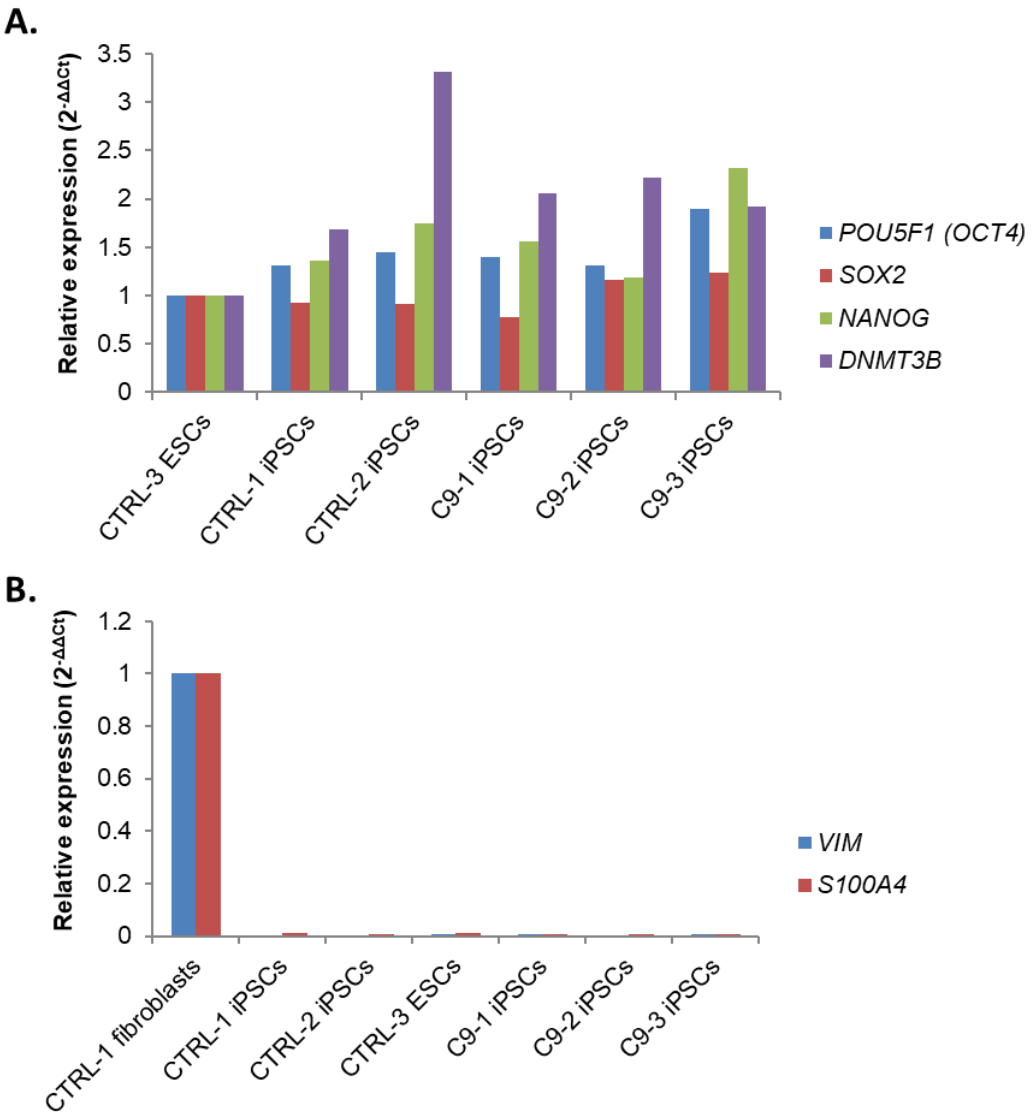


Figure 3.7 iPSC express pluripotency markers but not lineage-specific genes.
(A) RT-qPCR analysis was performed in all control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE (C9-1, C9-2 and C9-3) pluripotent lines to measure the expression of pluripotency marker genes *POU5F1* (OCT3/4), *SOX2*, *NANOG* and *DNMT3B*. Expression values were calculated relative to the ESC line CTRL-3 and normalised to *TBP* housekeeping gene. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method (n=1 experiment, performed in technical duplicates). (B) RT-qPCR was performed in all control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE (C9-1, C9-2 and C9-3) pluripotent lines to measure the expression of fibroblast identity markers *VIM* and *S100A4*. Expression values were calculated relative to the CTRL-1 fibroblast line and normalised to *TBP* housekeeping gene. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method (n=1 experiment, performed in technical duplicate).

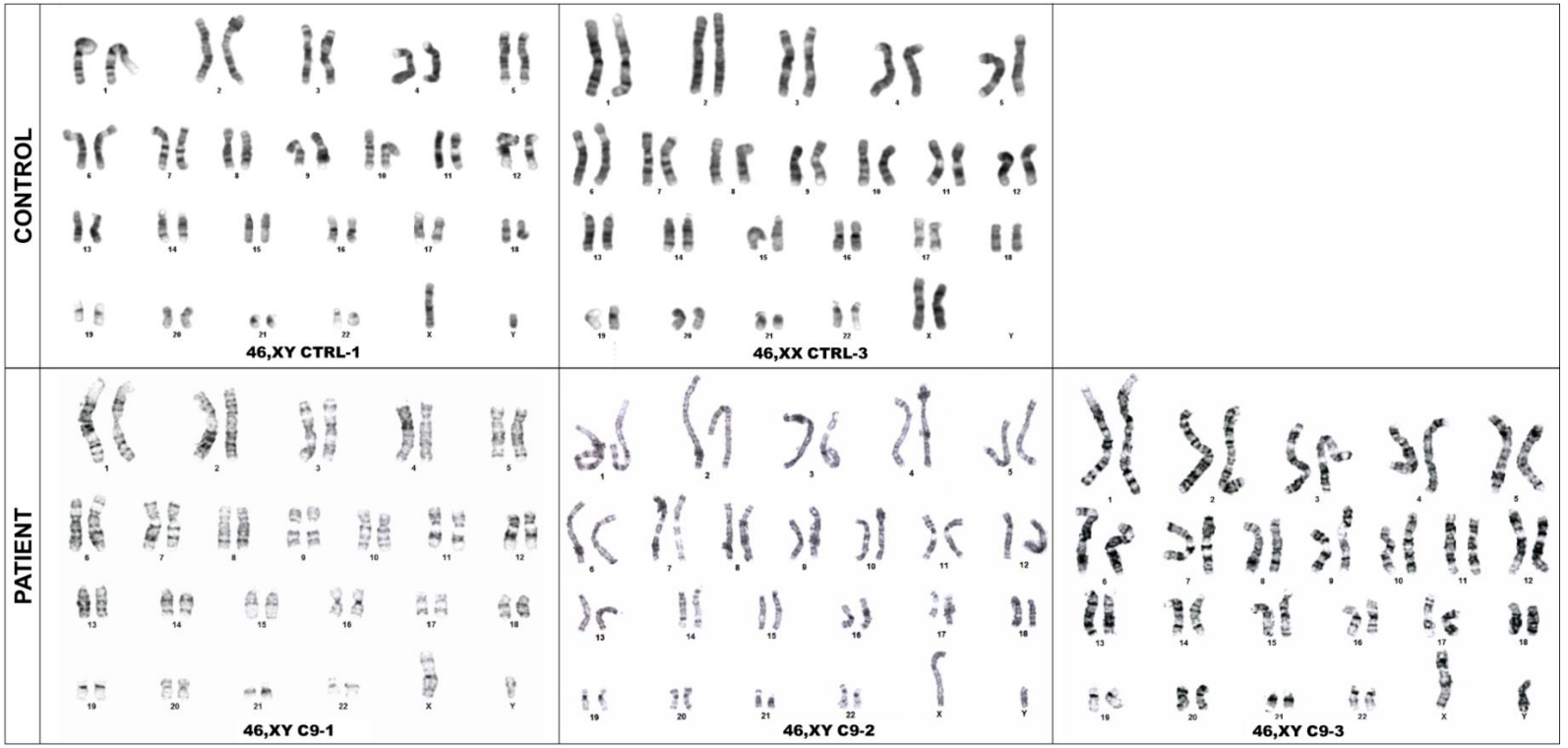


Figure 3.8 iPSC lines have normal karyotypes.

G-band karyotyping was performed externally (TDL Genetics) for two control (CTRL-2 and CTRL-3) and all three *C9orf72* HRE (C9-1, C9-2 and C9-3) pluripotent lines. All lines tested exhibited normal karyotypes. Control line CTRL-2 was obtained from Coriell repository and had a normal karyotype (not shown).

3.3.3. Differentiation and characterisation of control and *C9orf72* cortical neurons

The generation of CNs from human pluripotent stem cell lines was performed using a published protocol based on 10 day induction of the neuroepithelium using Dorsomorphin and SB431542 (Shi *et al.*, 2012a). The initial induction was followed by an extended period of *in vitro* neurogenesis that generates deep and upper layer neurons in a temporally controlled manner that closely resembles *in vivo* human corticogenesis (Shi *et al.*, 2012a) (Figure 3.9). Following replating of the neurepithelial layer, polarised rosette structures were observed by 15 DIV in all control and *C9orf72* HRE lines, confirming successful neural induction (Figure 3.10).

Confirmation of the cortical identity of the precursors generated following neuronal induction, immunofluorescence was performed for the early forebrain marker OTX1/2 and Ki-67, a marker of proliferating cells, at day 15 post-induction (Figure 3.11). By 20 DIV, single neurons had appeared in the periphery of the rosettes which were positive for the neuronal marker βIII-tubulin (TUJ1) (Figure 3.12). This population of early-born, deep layer neurons was subsequently enriched in culture by 30-40 DIV and expressed layer 6 marker TBR1 (data not shown). By 90 DIV, a large number of late-born, upper layer neurons had been generated which were positive for late progenitor and upper layer marker BRN2 (Figure 3.13). All control and *C9orf72* HRE pluripotent stem cell lines generated both deep and upper layer cortical neurons, demonstrating that the presence of the HRE did not impede neuronal differentiation.

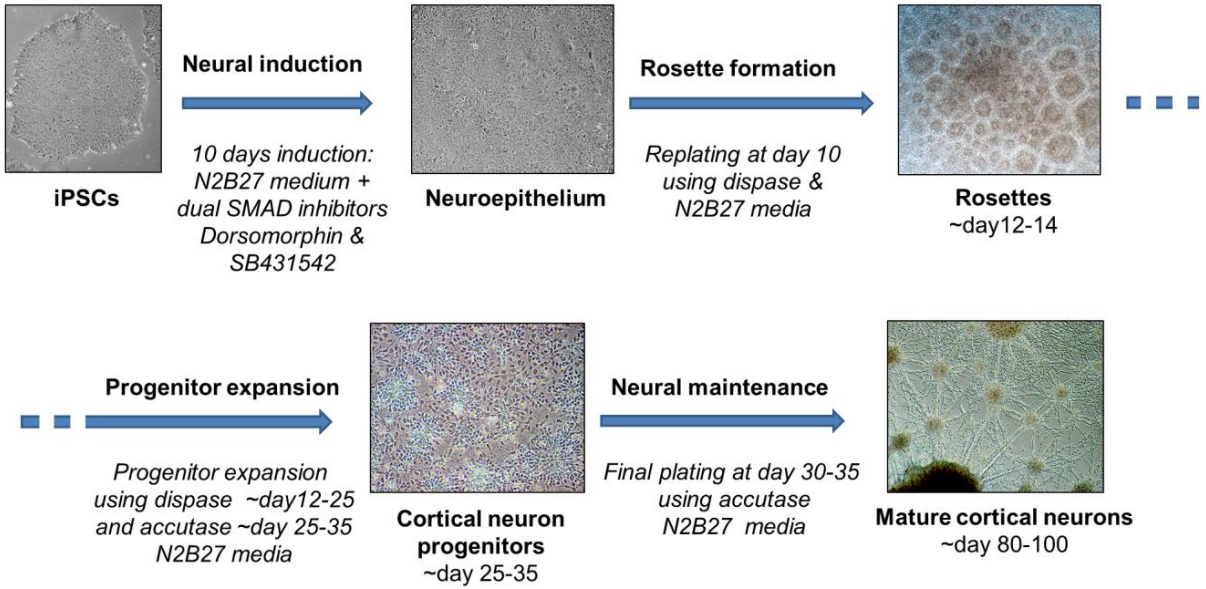


Figure 3.9 Overview of the Shi *et al*. 2012 CN differentiation protocol.

Briefly, 100% confluent iPSCs were induced towards formation of neuroepithelium by a 10 day induction in N2B27 media containing the dual SMAD inhibitors dorsomorphin and SB431542. The neuroepithelium was broken in smaller pieces by incubation in dispase, replated and maintained in N2B27 media. By ~ day 12 rosettes started appearing and expanding over time in culture. Early passages of rosettes at ~days 12-25 were done using dispase, but as significant neurogenesis had occurred in cultures by days ~25-35, neurons were passaged with accutase. For the final plating of neurons at ~day 35, accutase was used and neurons were left to mature until ~day 100 in N2B27 maintenance media.

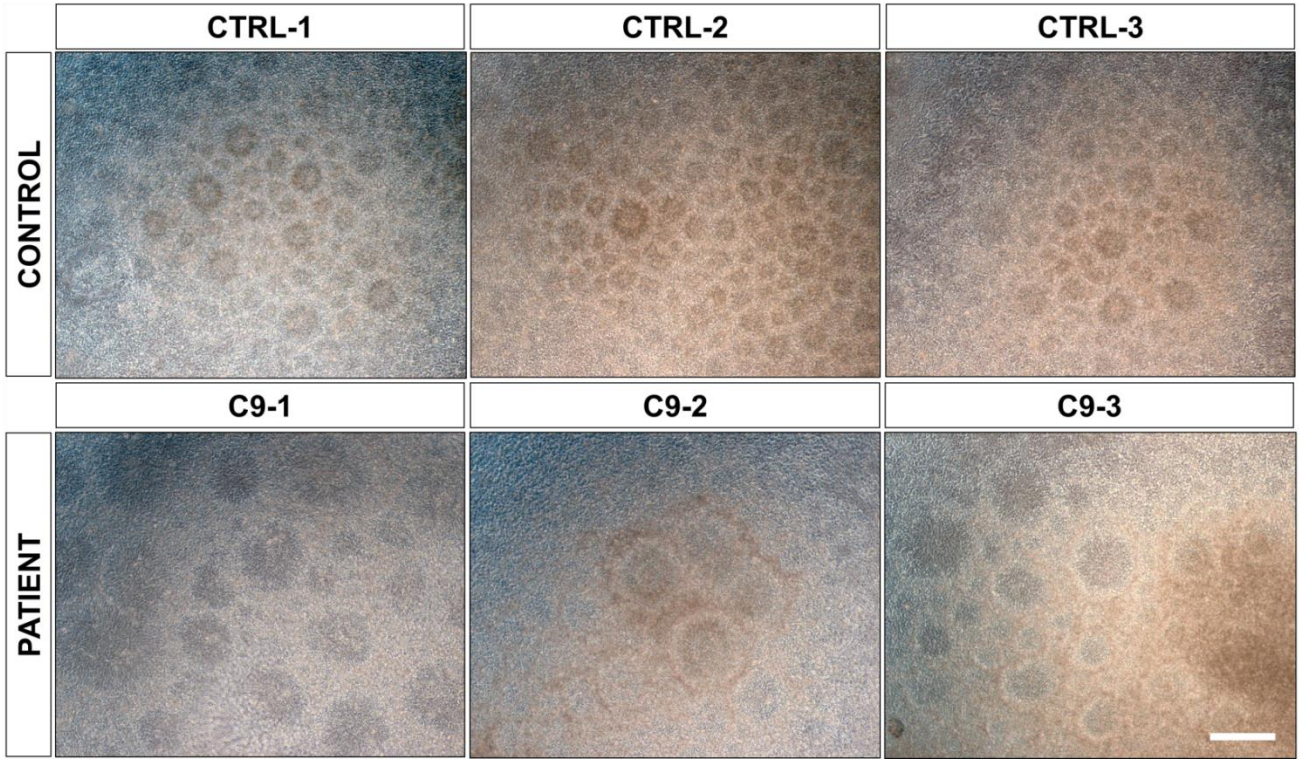


Figure 3.10 Control and *C9orf72* repeat expansion cortical rosettes at 15 DIV.

Phase contrast images of cortical rosettes from three control (CTRL-1, CTRL-2 and CTRL-3) and three *C9orf72* HRE (C9-1, C9-2 and C9-3) lines at 15 DIV of the cortical differentiation protocol. Scale bar 200 µm.

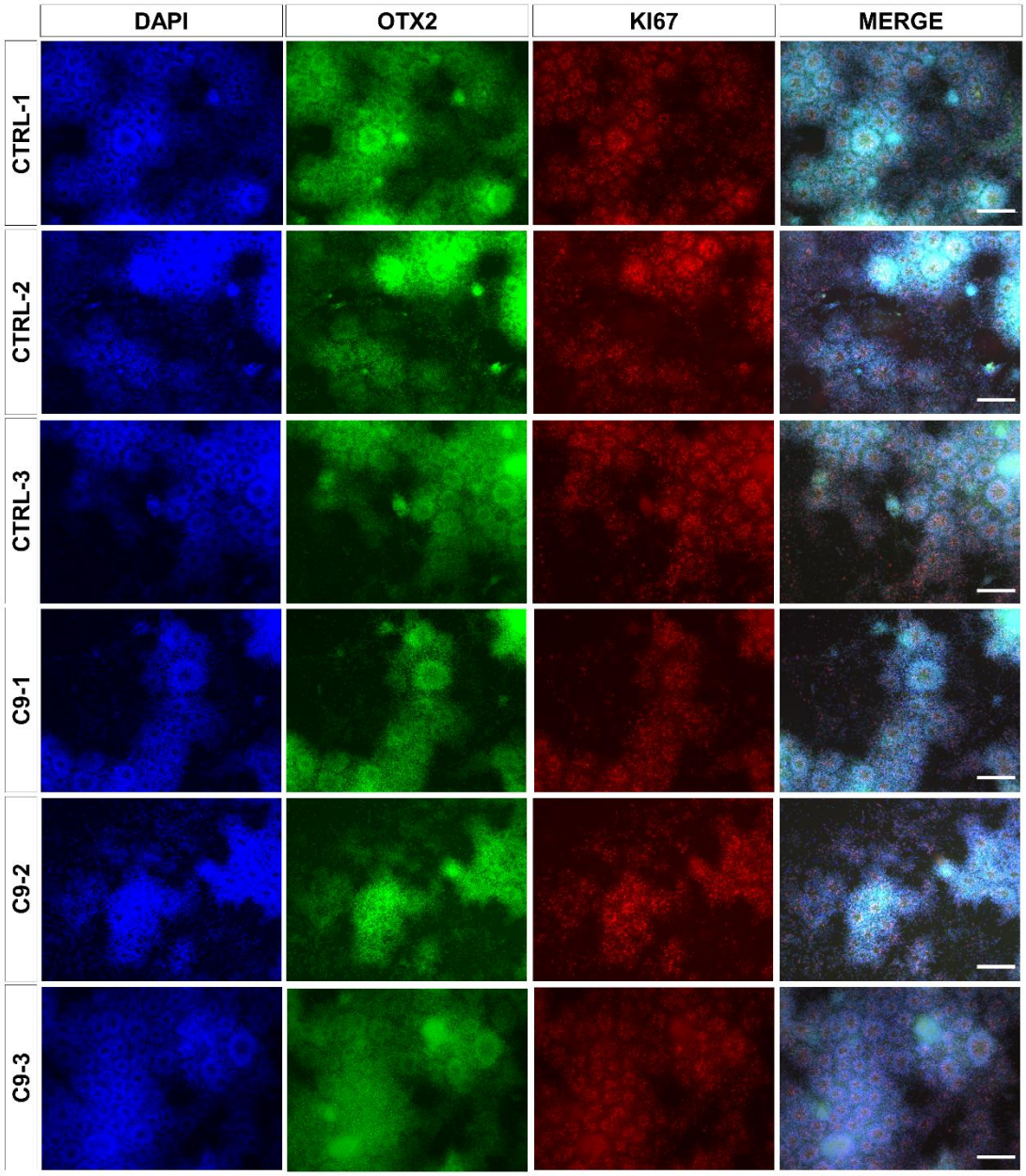


Figure 3.11 Cortical rosettes at day 15 express early forebrain marker OTX2. Positive immunofluorescence for the early forebrain marker OTX2 (green) confirmed the cortical identity of induced neural tissue in all control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE (C9-1, C9-2 and C9-3) differentiated lines. The cycling cortical progenitors forming the polarised rosettes were also positive for the marker of proliferating cells Ki-67 (red). Nuclei were counterstained with DAPI. Scale bar 200 µm.

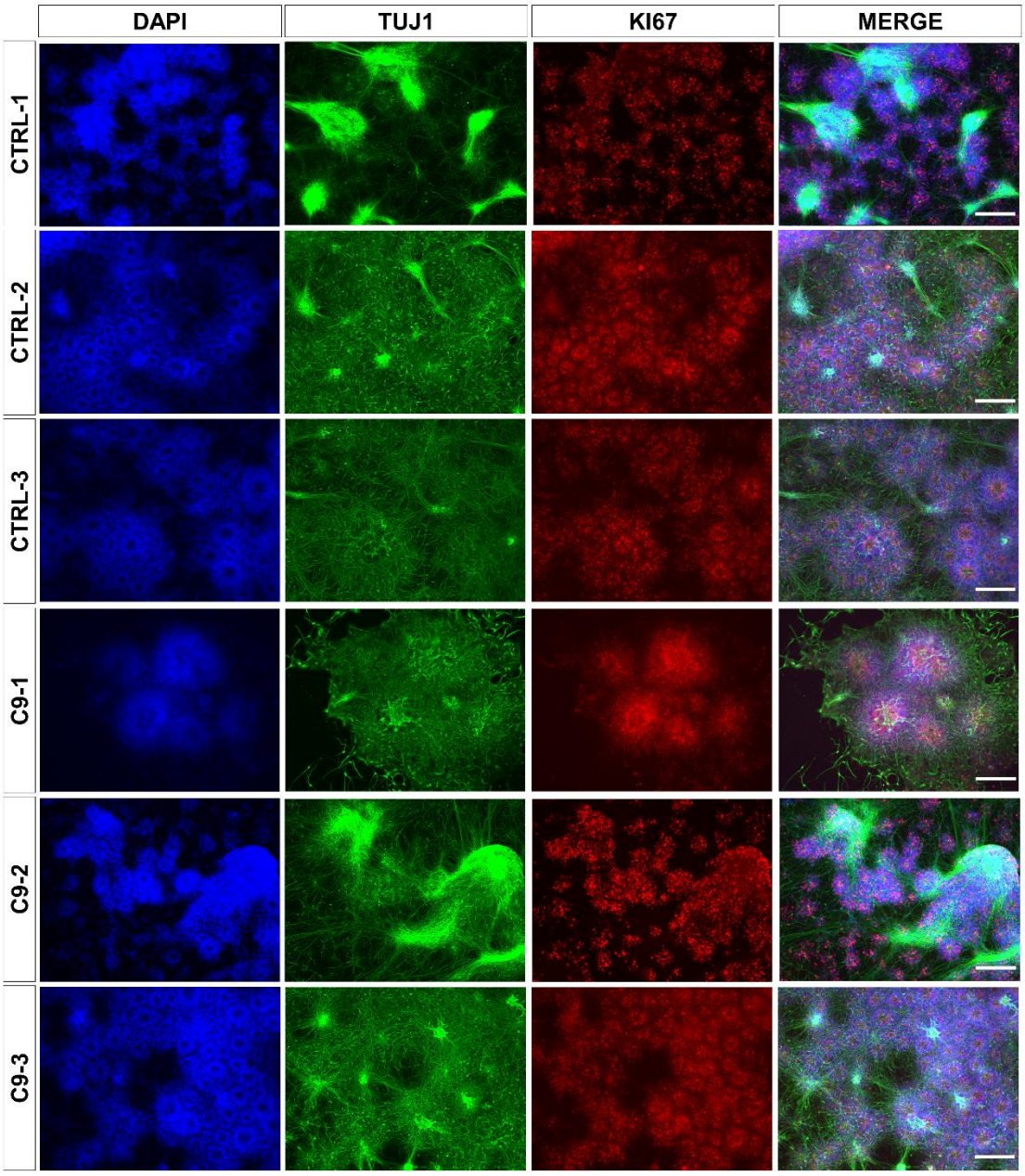


Figure 3.12 Early-born TUJ1 positive neurons emerge at the periphery of the cortical rosettes at 15-20 DIV.

Positive immunofluorescence for the neuronal marker βIII-tubulin (TUJ1) (green) in all control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE (C9-1, C9-2 and C9-3) differentiating cultures at 15-20 DIV revealed some early-born neurons generated from the cortical progenitors. The cycling progenitors forming the polarised rosettes were positive for the proliferation marker Ki-67 (red). Nuclei were counterstained with DAPI. Scale bar 200 μm.

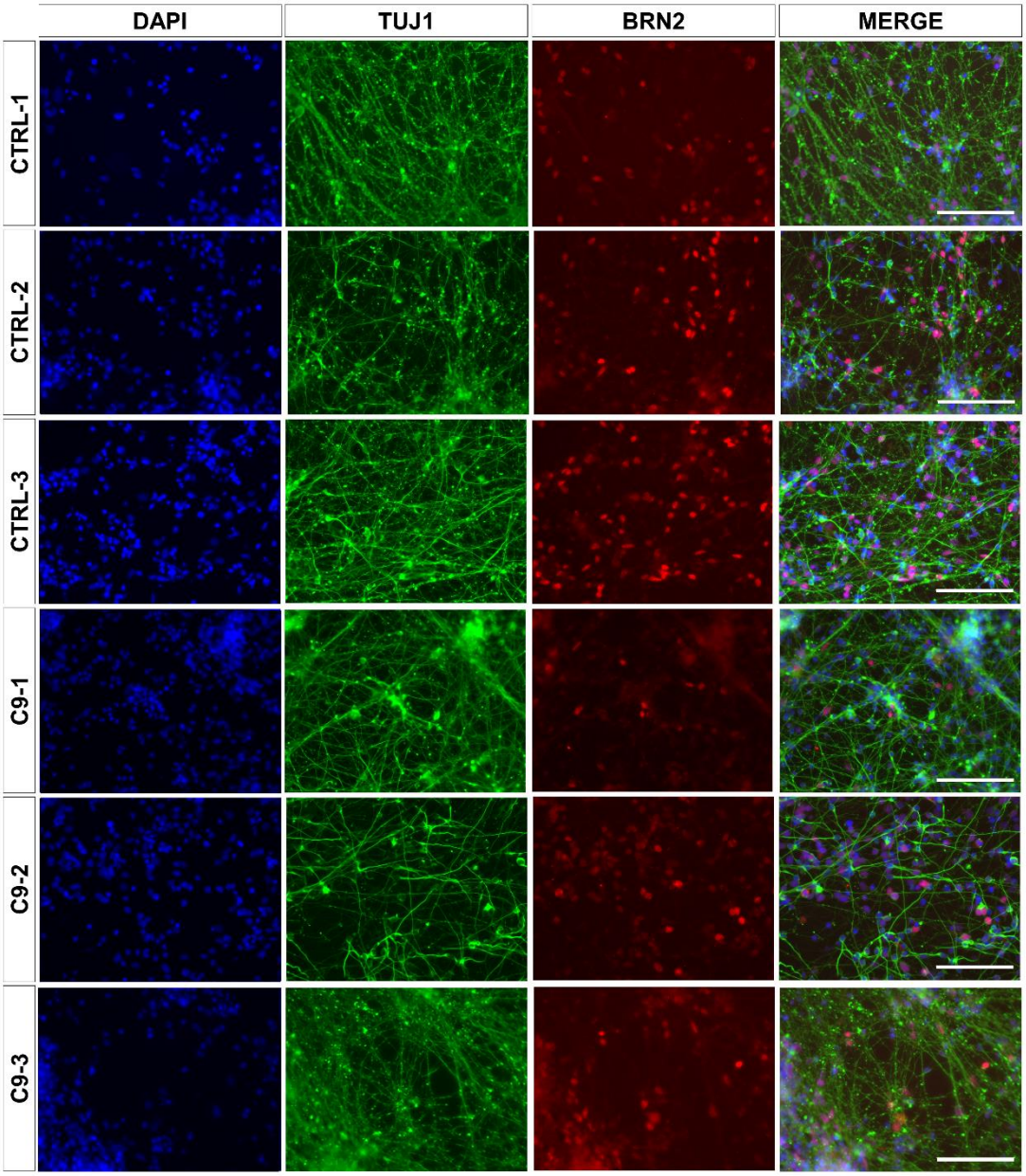


Figure 3.13 iPSC-CNs at 90 DIV express BRN2 marker of late cortical progenitors and upper layer (II-Vb) neurons.

Positive immunofluorescence for the late cortical progenitor and upper layer (II-Vb) marker BRN2 (red) in all control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE (C9-1, C9-2 and C9-3) lines at 90 DIV confirmed the presence of late-born, upper layer cortical neurons in all lines. Neurons were also stained with TUJ1 (green). Nuclei were counterstained with DAPI. Scale bar 100 µm.

3.3.4. Characterisation of CN cultures at 150 DIV

During i*n vitro* differentiation, neurogenesis precedes gliogenesis, similar to what happens during the development of the vertebrate CNS (Miller and Gauthier, 2007). In the current CN differentiation protocol, gliogenesis starts at ~60 DIV and numbers of astrocytes progressively increase until 100 DIV (Shi *et al.*, 2012b). Many of the experiments presented in chapters 3 and 4 were performed in iPSC-CNs cultured for extended periods *in vitro*, instead of the conventionally assessed 80-100 DIV timepoints. Therefore, 150 DIV control and *C9orf72* iPSC-CNs were further characterised for the expression of glial markers glutamate/aspartate transporter (GLAST), vimentin, and glial fibrillary acidic protein (GFAP). The levels of GLAST astrocytic marker were assessed by western blot in control and *C9orf72* 150 DIV iPSC-CN (n=3 independent inductions) (Figure 3.14). No differences in the levels of GLAST+ astrocytes was observed between control and *C9orf72* 150 DIV iPSC-CNs. Moreover the GLAST:TUJ1 ratios were similar between all iPSC-CN lines (Figure 3.14 C). The relative expression levels of two more astrocytic marker genes *VIM* and *GFAP*, as well as the neuronal marker *MAP2*, were assessed in control and *C9orf72* 150 DIV iPSC-CNs via RT-qPCR (Figure 3.15 A). Overall, no differences in *MAP2* and *VIM* mRNA levels were observed between control and C9orf72 150 DIV iPSC-CNs, with the exeption of a significant increase in *VIM* mRNA in C9-1 compared to CTRL-2 iPSC-CNs (p_{adj}<0.05). *GFAP* mRNA levels were significantly upregulated in C9-1 150 DIV iPSC-CNs compared to all controls and the rest of the *C9orf72* iPSC-CN lines (p_{adj}<0.0001, n=3 independent inductions). Western blot analysis confirmed the increase of the GFAP protein levels in C9-1 but further experiments are required (n=1 induction) (Figure 3.15 B). Increased GFAP may indicate the presence of reactive astrocytes (astrogliosis) in C9-1 150DIV iPSC-CNs, and further experiments should be done to confirm this.

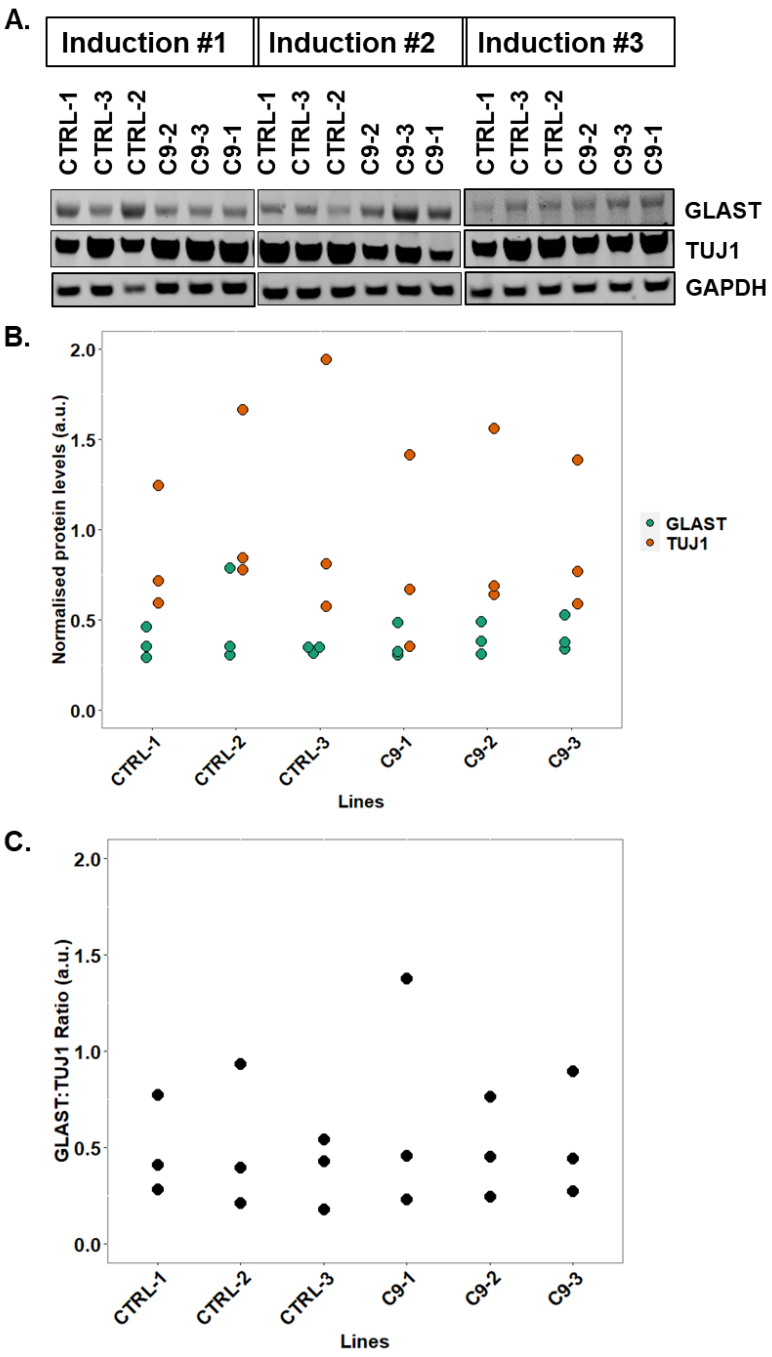


Figure 3.14 No difference in the levels of GLAST+ astrocytes and GLAST:TUJ1 ratios in 150 DIV iPSC-CNs of *C9orf72* patients and controls.

A. Western blot analysis of GLAST astrocytic marker in control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* (C9-1, C9-2 and C9-3) 150 DIV iPSC-CNs (n=3 independent inductions). TUJ1 was used as a neuronal marker and GAPDH as a normalising control. B. Densitometric quantification of western blot data shown in A. C. Ratios of GLAST:TUJ1 protein levels from A.

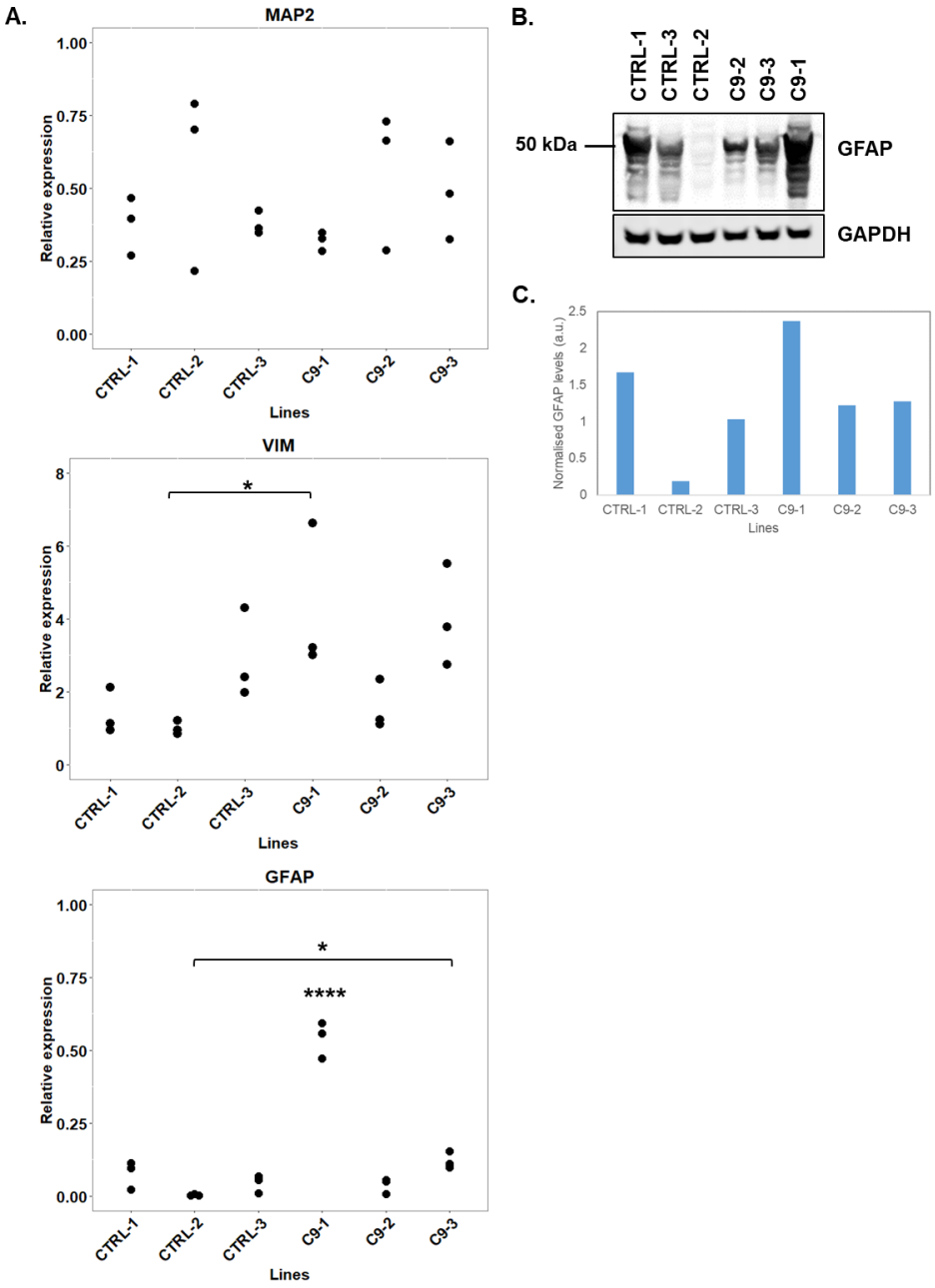


Figure 3.15 C9-1 150 DIV iPSC-CN exhibit significantly upregulated *GFAP* levels. A. Relative expression levels of *MAP2*, *VIM* and *GFAP* genes in control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* (C9-1, C9-2 and C9-3) 150 DIV iPSC-CNs as determined by RT-qPCR. Ct values of each gene were normalised to Ct values of housekeeping gene *ACTB*. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method and shown relative to control human brain. * represents p_{adj}<0.05, **** represent P_{adj}<0.0001 (n=3 independent inductions). B. Western blot analysis of GFAP astrocytic/astrogliosis marker in control and *C9orf72* 150 DIV iPSC-CNs (n=1 induction). GAPDH was used as a normalising control. C. Densitometric quantification of western blot data shown in B.

3.3.5. *C9orf72* iPSCs carry a stable repeat expansion during CN differentiation

The *C9orf72* HRE is characterised by genomic instability which ofter results in mosaicism of the HRE in different tissues and even within different CNS regions, in *C9orf72* HRE patients (Beck *et al.*, 2013, van Blitterswijk *et al.*, 2013b, Fratta *et al.*, 2015, Nordin *et al.*, 2015, Gijselinck *et al.*, 2016). Such genomic instability is a source of genetic heterogeneity in iPSC disease models. Thus, careful characterisation of *C9orf72* iPSC models is necessary for faithful, robust and reproducible disease modelling of C9-FTD/ALS.

To confirm the presence of the repeat expansion in *C9orf72* iPSC lines, and its absence from control iPSC lines, Southern blot analysis (performed by Dr Thomas Moens) was performed on genomic DNA extracted from the two control and three *C9orf72* expansion iPSC lines using a *C9orf72* hexanucleotide expansion-specific probe as described in 2.25. The Southern blot shown in Figure 3.16 confirmed the absence of the repeat expansion from the control iPSC lines CTRL-1 and CTRL-2 and revealed an expanded allele in all three patient-derived iPSC lines C9-1, C9-2 and C9-3. The size of the HRE varied among the patient iPSC lines but in all patient iPSCs the repeat expansion remained stable without signs of mosaicism in culture. The estimated repeat expansion sizes were approximately 241 repeats for C9-1, 638 repeats for C9-2 and 602 repeats for C9-3.

However, genomic instability in large repeat expansions may also arise during differentiation of iPSCs to neurons, mainly due to proliferation events in neuronal progenitor populations *in vitro* that increase the likelihood of errors during DNA replication at the *C9orf72* locus. Therefore, characterisation of the HRE in differentiated iPSC-CNs is another important step for the generation of *in vitro* neuronal models for *C9orf72* expansions. *C9orf72* iPSC-CNs maintained a stable HRE during CN differentiation as shown in the Southern blot in Figure 3.16.

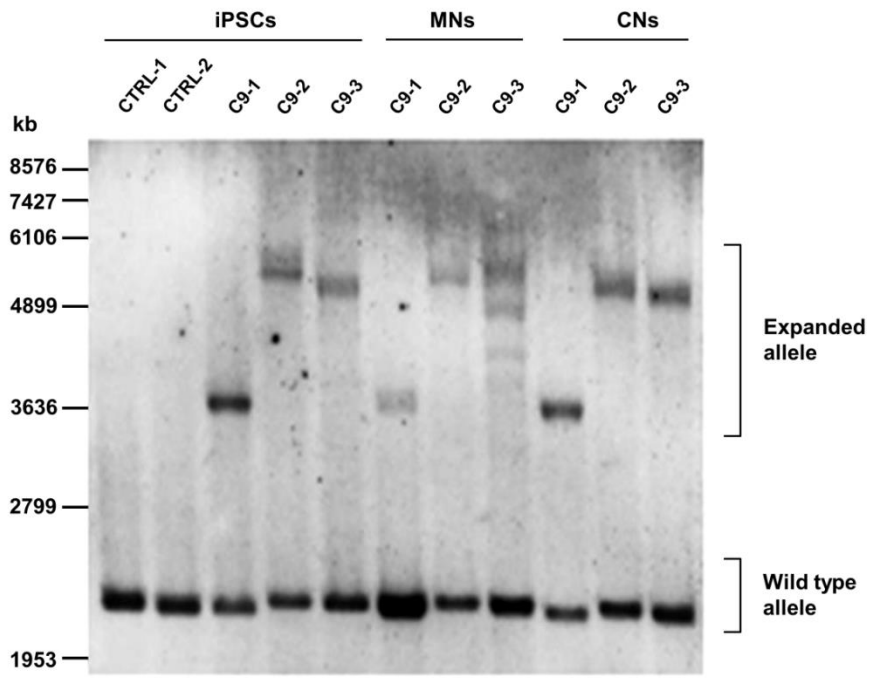


Figure 3.16 *C9orf72* HRE iPSCs and differentiated iPSC-CNs carry a stable HRE. Southern blot using a *C9orf72* HRE-specific probe revealed that the two control iPSC lines (CTRL-2 and CTRL-2) were negative for the HRE, whereas the three *C9orf72* iPSC lines (C9-1, C9-2 and C9-3) all carried different lengths of the HRE. Specifically, C9-1 had an estimated 241 repeats size, C9-2 had 638 and C9-3 had 602. All three expanded alleles were stable in iPSCs and remained stable during differentiation to CNs. MNs differentiated from the C9-1, C9-2 and C9-3 *C9orf72* expansion carrier lines also displayed a same HRE size as in iPSCs and genomic stability with the exception of C9-3 line which appeared to be unstable during MN differentiation, leading to mosaicism in culture. *Abbreviations; CNs: cortical neurons, MNs: motor neurons* (Modified from Simone *et al.* 2018).

3.3.6. *C9orf72* iPSC-CNs contain sense RNA foci

One major feature of the pathology associated with the repeat expansion in *C9orf72* is the presence of the RNA foci in FTD/ALS patient tissue (DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011). Therefore all control (CTRL-1, CTRL-2 and CTRL-3) and C9orf72 HRE carrier (C9-1, C9-2 and C9-3) iPSC-CNs were assessed for the presence of sense RNA foci via RNA Fluorescence *in situ* hybridization (FISH) (performed by Dr Roberto Simone) using a sense $(C_4G_2)_4$ probe as described in 2.24. Sense RNA foci were specifically detected in 53-63 DIV iPSC-CNs of all three *C9orf72* HRE lines (C9-1, C9-2 and C9-3) but not in any of the three controls (CTRL-1, CTRL-2 and CTRL-3) (Figure 3.17). Quantification of the RNA foci revealed variability in the percentage of foci between the three *C9orf72* HRE lines. Specifically, line C9-1 exhibited the highest percentage of sense RNA foci containing nuclei, followed by C9-3 line and C9-2 line had the lowest percentage (n=2 independent inductions, at least 180 neurons per induction were counted).

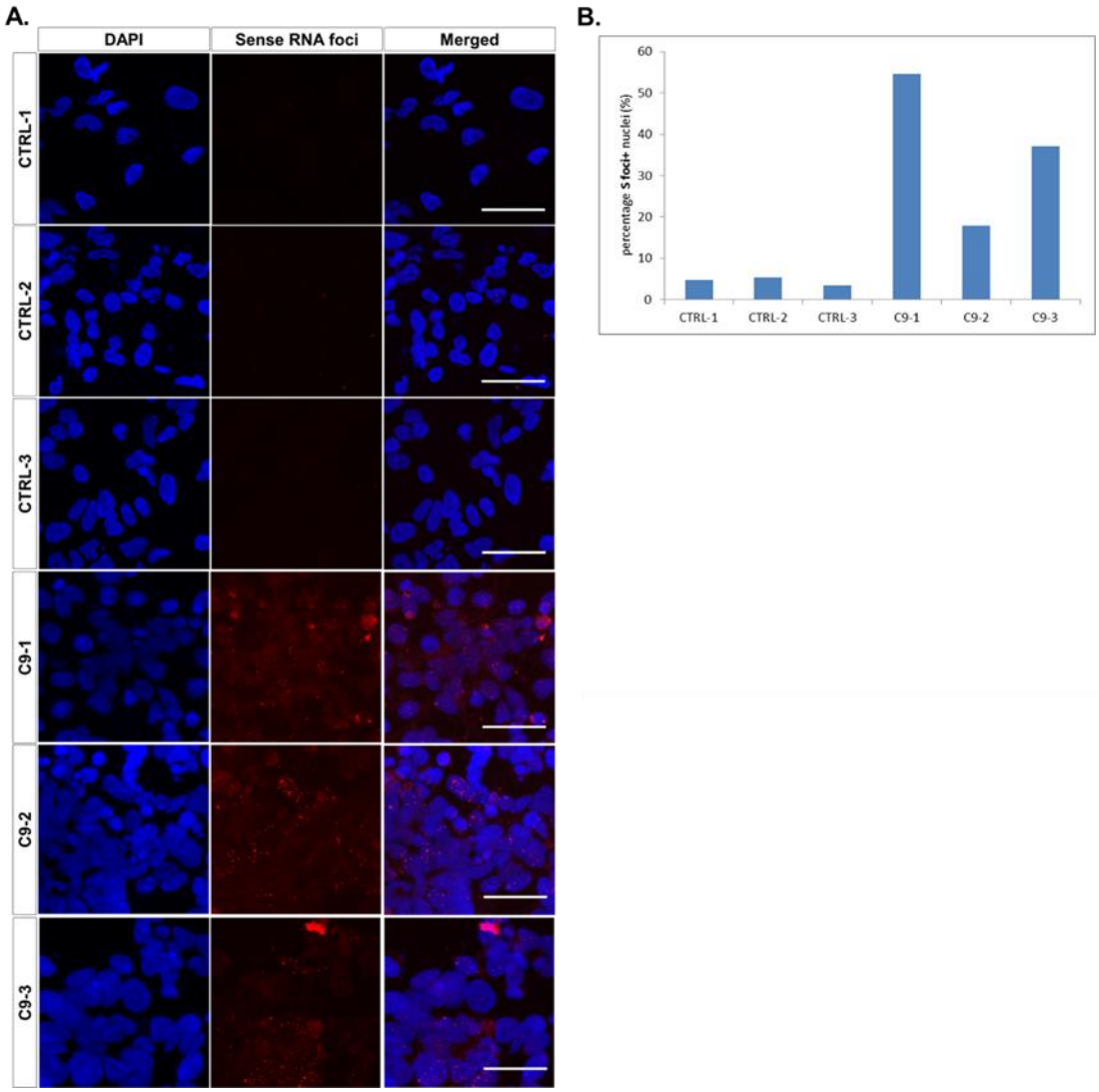


Figure 3.17 Sense RNA foci are detected in *C9orf72* iPSC-CNs.

A. Sense RNA foci were speficically detected in the nuclei of C9orf72 iPSC-CNs at 53-63 DIV via RNA FISH (n=2 independent inductions). B. Quantification of RNA FISH shown in A. The percentage of sense RNA positive nuclei is elevated in *C9orf72* iPSC-CNs compared to control iPSC-CNs (n=2 independent inductions, at least 180 neurons per induction were counted). Scale bars 50 µm.

3.3.7. *C9orf72* iPSC-CNs exhibit elevated poly-GP levels

Another hallmark of *C9orf72* pathology is the detection of dipeptide protein repeats (DPRs) produced in all reading frames from the HRE in patient tissue. Detection of different DPRs has been reported in published *in vivo* and *in vitro* models of *C9orf72* toxicity. However, in contrast to RNA foci detection, DPR detection remains largely variable.

Investigation of DPR pathology in 60-80 DIV *C9orf72* HRE and control iPSC-CNs via immunofluorescence did not produce any reliable and specific for *C9orf72* signal for all the assessed poly-dipeptides (poly-GP, poly-GR, poly-GA, and poly-PR) (data not shown). However, elevated levels of poly-GP were detected specifically in protein lysates of 150 DIV iPSC-CNs from all three *C9orf72* lines (C9-1, C9-2 and C9-3) but not in any of the control lines (CTRL-1, CTRL-2 and CTRL-3) using a poly-GP Meso Scale Discovery (MSD) immunoassay (performed by Javier Gilbert Jaramillo) (n=1 induction) (Figure 3.18 A). The assay was established at Prof. Adrian Isaac's laboratory, using a previously generated rabbit anti-poly-GP antibody (Mizielinska *et al*., 2014) as described in 2.26. Similar to RNA foci expression, poly-GP production was variable among the three *C9orf72* repeat expansion lines. Specifically, C9-3 line exhibited the highest levels of poly-GP, whereas C9-1 and C9-2 lines had similar levels and lower than C9-3. Poly-GP was below detection level in all control iPSC-CNs. Similar to poly-GP, a poly-GR MSD assay was also established using a newly generated affinity purified rabbit anti-(GR)7 antibody (Eurogentec) (performed by Mica Clarke) (n=1 induction). Elevated poly-GR levels were detected in C9-3 iPSC-CNs (Figure 3.18 B). However, further assays need to be performed for both poly-GP and poly-GR to establish the levels of endogenous DPR expression in *C9orf72* iPSC-CNs.

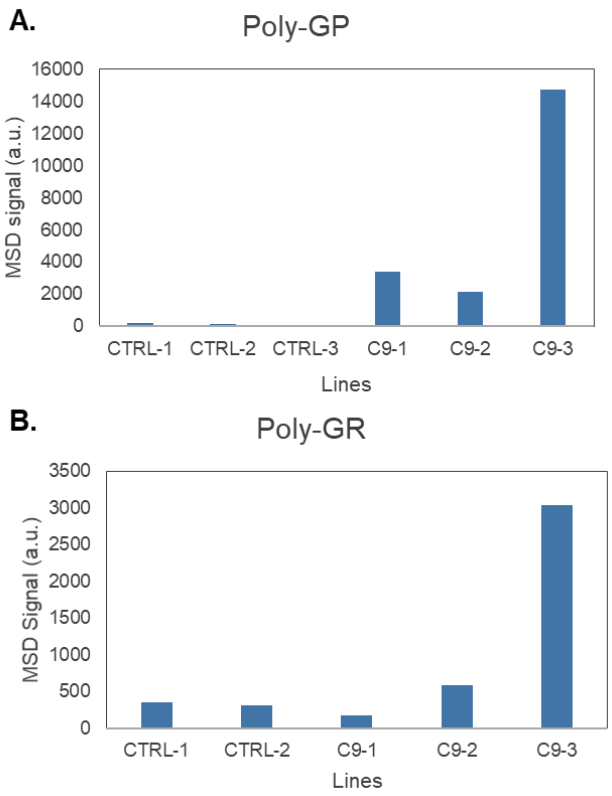


Figure 3.18 *C9orf72* iPSC-CNs exhibit elevated levels of poly-GP.

A. Poly-GP levels were measured by Meso Scale Discovery (MSD) immunoassay in 150 DIV iPSC-CNs from three control (CTRL-1, CTRL-2 and CTRL-3) and three *C9orf72* HRE (C9-1, C9-2 and C9-3) lines. Levels of poly-GP were elevated in iPSC-CNs from all three *C9orf72* HRE lines (C9-1, C9-2 and C9-3) compared to control lines (CTRL-1, CTRL-2 and CTRL-3). Data is shown as MSD counts, n=1 induction for each line, experiment performed in technical duplicate. B. Poly-GR levels were measured by MSD immunoassay in 120 DIV iPSC-CNs from two control (CTRL-1 and CTRL-2) and three *C9orf72* HRE (C9-1, C9-2 and C9-3) lines. Levels of poly-GR were elevated in iPSC-CNs from C9-3 line compared to control lines CTRL-1 and CTRL-2. Data is shown as MSD counts, n=1 induction for each line, experiment performed in technical duplicate.

3.4. Discussion

This chapter describes the generation and characterisation of fibroblast, iPSC and cortical glutamatergic neurons from patients carrying the *C9orf72* HRE.

C9-FTD/ALS is caused by long repeat expansions in *C9orf72* which are prone to genomic instability (Reddy *et al.*, 2014, Abu Diab *et al.*, 2018) and mosaicism with subsequent implications in diagnostics and disease modelling. Screening for *C9orf72* repeat expansions is carried out on blood samples which may exhibit different repeat lengths than the CNS (Beck *et al.*, 2013, Fratta *et al.*, 2015, Nordin *et al.*, 2015, Gijselinck *et al.*, 2016).. Similarly, in disease modelling, iPSCs are usually derived from peripheral tissue such as skin fibroblasts, or lymphoblasts. When the repeat expansion size in these tissues does not reflect the actual size in the brain, then this can lead to discrepancies in disease-related phenotypes. In addition to that, mosaicism may exist or arise during cell culture, leading to variability due to genetic heterogeneity between cellular subpopulations carrying different repeat expansion sizes. Therefore, characterisation of the *C9orf72* HRE length and stability is important for faithful disease modelling.

Here, characterisation of the *C9orf72* HRE size by southern blot in *C9orf72* iPSCs and iPSC-CNs revealed that the expanded repeat remains stable in iPSCs and during differentiation to CNs. No post-mortem tissue was available for examination of the *C9orf72* HRE length in the CNS of the *C9orf72* HRE carriers who participated in this study. In other published studies using iPSCs to model C9-FTD/ALS, some have reported a stable HRE in iPSCs and differentiated neurons (Donnelly *et al.*, 2013), whereas others observed instability of the repeat (Almeida *et al.*, 2013, Sareen *et al.*, 2013). In the study by Sareen *et al*., authors reported genomic instability, specifically somatic expansion, during reprogramming of two fibroblast lines to iPSCs and contraction of one line during differentiation to MNs (Sareen *et al.*, 2013). Almeida *et*

al. also reported repeat instability, both expansion and contraction, during reprogramming of fibroblasts to iPSCs and differentiation to neurons (Almeida *et al.*, 2013).

Moreover, this chapter describes the differentiation and characterisation of control and *C9orf72* repeat expansion carrier CNs, the main cell type affected in FTD. The CN differentiation protocol used in this study results in stepwise generation of both deep and upper layer cortical excitatory neurons and the cultures become progressively enriched with astrocytes offering a reductionist 2D model closely relevant to human cortex. As expected, corticogenesis was not found to be impaired in *C9orf72* repeat expansion carrier lines, since no differences in the ability of *C9orf72* iPSCs to form neuronal rosettes or deep/upper layer neurons were observed. This observation is consistent with no known effects of the *C9orf72* repeat expansion on development. Interestingly, 150 DIV iPSC-CN cultures from control and *C9orf72* lines had a similar ratio of astrocytes to neurons, however, *GFAP* levels were significantly upregulated in C9-1 150 DIV iPSC-CNs compared to the rest of the lines. This could be an indication of astrogliosis in this line that should be further investigated.

Even though the cellular repertoire in these cortical cultures is limited compared to human brain, the striking resemblance of *in vitro* cortical neuron generation to human embryonic corticogenesis and the presence all cortical layer neurons and astrocytes provide a physiologically relevant *in vitro* model to investigate *C9orf72*-linked FTD phenotypes. Other published studies of *C9orf72* iPSC models have also used iPSC-CNs (Almeida *et al.*, 2013, Dafinca *et al.*, 2016, Yuva-Aydemir *et al.*, 2019), mixed neuronal cultures (Donnelly *et al.*, 2013), or iPSC-MNs (Sareen *et al.*, 2013, Devlin *et al.*, 2015, Dafinca *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Sivadasan *et al.*, 2016, Westergard *et al.*, 2016, Aoki *et al.*, 2017, Selvaraj *et al.*, 2018, Simone *et al.*, 2018, Zhang *et al.*, 2018a, Cheng *et al.*, 2019, Ababneh *et al.*, 2020, Abo-Rady *et al.*, 2020) for the investigation of C9-FTD/ALS-specific phenotypes. Furthermore, studies have

also investigated C9-FTD/ALS pathologies as well as non-cell-autonomous toxicity mechanisms in iPSC-atsrocytes (Ferraiuolo *et al.*, 2016, Birger *et al.*, 2019, Zhao *et al.*, 2019b) and iPSC-oligodendrocytes (Livesey *et al.*, 2016), (see 1.3.4.6). Finally, some studies have used transdifferentiation protocols, whereby patient cells such as fibroblasts are directly reprogrammed to induced-neurons or induced-astrocytes (Meyer *et al.*, 2014, Jovicic *et al.*, 2015, Shi *et al.*, 2018, Bauer *et al.*, 2019).

The use of iPSC models to investigate complex, age-related neurodegenerative diseases such as FTD and ALS may be challenging due to the relative immaturity of the derived neurons. IPSC-derived neurons have been shown to follow a strict developmental route and despite *in vitro* maturation, they are still closer to embryonic than adult human neurons (Stein *et al.*, 2014, Burke *et al.*, 2020). Ageing signatures are absent from these neurons and need to be further induced by overexpression of progerin (Miller *et al.*, 2013) or by telomere shortening (Vera *et al.*, 2016). Instead, induced neurons, produced via transdifferentiation of fibroblasts directly into neurons, have been shown to retain the age-related signatures of the parental cells (Mertens *et al.*, 2015, Tang *et al.*, 2017). Therefore, induced neurons offer an alternative tool for exploring phenotypes in C9-FTD/ALS.

Recapitulation of the disease-related pathology seen in patients is a prerequisite for faithful disease modelling. In the present chapter, the *C9orf72* iPSC-CNs exhibited hallmarks of C9-FTD/ALS pathology that were absent from control iPSC-CNs. Specifically, *C9orf72* HRE iPSC-CNs displayed sense RNA foci, poly-GP and poly-GR with considerable variability between the three *C9orf72* HRE lines.

Quantification of the RNA foci revealed variability in the amount of nuclei containing RNA foci among the different patient lines, with C9-1 neurons having the highest and C9-2 the lowest percentage of foci positive nuclei. Interestingly, line C9-1 which had the highest number of foci positive nuclei and highest number of foci per nucleus, was

also the one line with the smallest size of ~241 repeats and the earliest age of disease onset at 39 years. Detection of RNA foci in iPSNs has also been consistently reported by other studies (Almeida *et al.*, 2013, Donnelly *et al.*, 2013, Sareen *et al.*, 2013, Dafinca *et al.*, 2016). In the study by Sareen *et al*. the iPSC line with the shortest size of ~70 repeats had the fewest RNA foci per nucleus which is opposite to what we observed in the case of C9-1 neurons, also supporting that repeat length on its own cannot explain the severity of foci pathology observed but other factors probably influence this. Indeed, it has been suggested that hypermethylation in the promoter region of *C9orf72* is protective against the toxicity of RNA foci and DPRs in human post-mortem tissue (Liu *et al.*, 2014). Epigenetics and differential methylation in the HRE region may explain the difference in the foci and this is a subject examined extensively in chapter 4.

Investigation of DPR pathology in *C9orf72* CNs by immunofluorescence did not reveal any specific and reliable signal for any of the anti-poly-dipeptide antibodies assessed (poly-GA, poly-GP, poly-GR, poly-PR). However, poly-GP, and to a lesser degree poly-GR, were detected in 150 and 120 DIV *C9orf72* iPSC-CN lysates respectively, by MSD immunoassay. Variability was evident in DPR expression between the different *C9orf72* lines. However, as these assays were performed only once, no conclusions can be drawn and more repeats should be performed to establish endogenous levels of DPRs and variation in *C9orf72* iPSC-CNs. Moreover, immunoassays for poly-GA, poly-PR and poly-AP DPRs should also be established and performed to complete the screening for DPR pathology in these cultures.

Difficulties in the detection of DPRs have been repeatedly reported for *in vitro* models, with some studies reporting no DPRs detected (Sareen *et al.*, 2013) while others detected DPRs in *C9orf72* iPSCNs using dot blot (Almeida *et al.*, 2013, Dafinca *et al.*, 2016, Lopez-Gonzalez *et al.*, 2016, Selvaraj *et al.*, 2018, Ababneh *et al.*, 2020), western blot (Zhao *et al.*, 2019b), immunofluorescence (Donnelly *et al.*, 2013,

Westergard *et al.*, 2016) or MSD immunoassay (Lopez-Gonzalez *et al.*, 2016, Simone *et al.*, 2018, Almeida *et al.*, 2019, Cheng *et al.*, 2019, Yuva-Aydemir *et al.*, 2019, Abo-Rady *et al.*, 2020). On the other hand, post-mortem detection has been more straightforward. Such difficulties in DPR detection in *in vitro* models may be a consequence of current technical limitations in detection of such DPR species, lack of specific/well-characterised antibodies or reflect the immaturity of the *in vitro* generated neurons. Indeed, in recent years, the development of novel highly-specific DPR antibodies combined with the use of high sensitivity immunoassays has resulted in a more robust and reproducible detection of DPRs *in vitro*. Moreover, as already discussed, iPSCNs are fairly immature and the introduction of *in vitro* ageing or transdifferentiation protocols could help accelerate the observation of age-specific phenotypes (Miller *et al.*, 2013, Mertens *et al.*, 2015). Finally, similar to RNA foci, variability in DPR production among the three *C9orf72* patient lines may be influenced by other factors such as methylation in *C9orf72* locus or additional modifiers.

Variability in iPSC models can be overcome by the use of genetically engineered isogenic lines that share a common genetic background. Ideally, isogenic lines carrying different *C9orf72* repeat expansion lengths could be used to delineate the threshold of repeat expansion size for detection of pathology and establish the relation between repeat expansion size and severity of pathology. However, variability in neuronal cultures is a natural consequence of the clinical heterogeneity that characterises patients carrying the *C9orf72* HRE. Thus, despite the challenges, the study of *C9orf72* iPSCNs can offer insights into the mechanisms influencing this heterogeneity.

To conclude, this chapter shows *C9orf72* iPSC-CNs are a useful *in vitro* model to study FTD for several reasons: (i) they are the main cell type affected in FTD, (ii) they contain the patient's *C9orf72* HRE and precise genome without overexpression artefacts, (iii) they offer a relevant model to study all three potential *C9orf72* FTD

disease mechanisms, (iv) they carry a stable HRE during differentiation and (v) they recapitulate *C9orf72*-linked FTD/ALS pathological hallmarks of RNA foci and DPRs. The next chapter describes the investigation of haploinsufficiency by assessing *C9orf72* expression, methylation and protein production in the established *C9orf72* HRE iPSC-CN models.

Chapter 4 Analysis of *C9orf72* temporal expression and methylation

4.1. Introduction

Haploinsufficiency of *C9orf72* and therefore loss of C9orf72 protein function is another mechanism by which the long hexanucleotide repeat expansion in *C9orf72* may lead to FTD/ALS pathogenesis. C9orf72 protein function is not yet fully characterised. However, *in silico* analysis has shown that it is a structural homolog of the DENN guanine nucleotide exchange factor (GEF) family proteins (Levine *et al.*, 2013). An increasing number of independent functional studies suggest that C9orf72 is involved in autophagy and lysosomal trafficking (Farg *et al.*, 2014, Amick *et al.*, 2016, O'Rourke *et al.*, 2016, Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Ugolino *et al.*, 2016, Webster *et al.*, 2016, Yang *et al.*, 2016, Aoki *et al.*, 2017, Shi *et al.*, 2018). However, there are several unknowns that still require further investigation, such as the significance of the two C9orf72 protein isoforms, and the link between loss of C9orf72 protein and neurodegeneration.

Reduced *C9orf72* mRNA transcript levels have been reported by several studies in patient tissue, blood and iPSCNs (DeJesus-Hernandez *et al.*, 2011, Gijselinck *et al.*, 2012, Almeida *et al.*, 2013, Belzil *et al.*, 2013, Donnelly *et al.*, 2013, Fratta *et al.*, 2013, Mori *et al.*, 2013c, Xi *et al.*, 2013, Waite *et al.*, 2014, Tran *et al.*, 2015, van Blitterswijk *et al.*, 2015, Shi *et al.*, 2018). One way by which the repeat expansion in *C9orf72* may lead to reduced expression of the gene is via epigenetic regulation. The long GC rich repeat can trigger epigenetic modifications such as DNA methylation or histone methylation in the promoter region of *C9orf72* and therefore silence its expression. Gene silencing due to hypermethylation of CpG-rich repeats has also been described in other repeat expansion diseases, such as Fragile X syndrome (FXS), whereby hypermethylation of the expanded CGG repeat in the 5′UTR region of *FMR1* leads to silencing of the gene (Bell *et al.*, 1991, Pieretti *et al.*, 1991, Sutcliffe *et al.*, 1992). Indeed, the *C9orf72* HRE itself has been found to be hypermethylated in patient tissue (Xiao *et al.*, 2015). Moreover, the HRE is also flanked by two CpG islands (see Figure

1.5), and the 5' CpG island has been found to be hypermethylated in *C9orf72* HRE FTD/ALS patients (Xi *et al.*, 2013, Belzil *et al.*, 2014, Xi *et al.*, 2014b).

In contrast to *C9orf72* transcript expression studies, studies investigating C9orf72 protein levels are limited. The main obstacle for investigating C9orf72 protein has been the lack of specific C9orf72 antibodies. For example, certain antibodies have been shown to cross-react with glial fibrillary acidic protein (GFAP) (Satoh *et al.*, 2014, Waite *et al.*, 2014). Indeed, a recent study characterising C9orf72 antibodies has highlighted the issues of antibody specificity and raised concerns over the use of non-specific anti-C9orf72 antibodies in published studies (Laflamme *et al.*, 2019). However, specific antibodies against the short (C9-S) and long (C9-L) isoforms of C9orf72 have also recently been developed (Xiao *et al.*, 2015, Xiao *et al.*, 2016, Frick *et al.*, 2018).

In chapter 3, the iPSC-CNs from controls and *C9orf72* HRE patients were characterised for the presence of CN identity markers as well as potentially toxic RNA foci and DPR species which may lead to neurodegeneration via a gain of function mechanism. The focus of this chapter is the investigation of *C9orf72* haploinsufficiency in *C9orf72* HRE iPSC-CNs, by the examination of epigenetic alterations, together with RNA and protein levels of C9orf72 throughout CN differentiation.

4.1. Aims

The aim of this chapter was to investigate haplosinsufficiency of *C9orf72* in patient-derived CNs. The specific aims of this chapter were:

- To assess the temporal DNA methylation in the *C9orf72* promoter region in control and *C9orf72* HRE iPSC-CNs during corticogenesis.
- To investigate the expression of the different *C9orf72* mRNA transcripts (V1, V2, V3 and total) in control and *C9orf72* HRE iPSC-CNs at 150 DIV.
- To characterise the levels of C9orf72 protein in control iPSC-CNs temporally during corticogenesis.
- To investigate the levels of C9orf72 protein in control and *C9orf72* HRE iPSC-CNs at 150 DIV.

4.2. Results

4.2.1. Methylation of the *C9orf72* HRE in fibroblasts

The *C9orf72* HRE has been found to be methylated in patient tissue (Xiao *et al.*, 2015). However, little is known about the methylation status of the *C9orf72* HRE in fibroblasts, which is the starting material for many iPSC studies. Some studies have reported that the *C9orf72* HRE is unmethylated in patient fibroblasts (Xi *et al.*, 2015a, Cohen-Hadad *et al.*, 2016), while in some cases, methylation of the HRE has also been reported (Xi *et al.*, 2015a).

The *C9orf72* HRE DNA methylation status was assessed in *C9orf72* HRE carrier fibroblasts via the $(G_4C_2)_n$-methylation assay (Xiao *et al.*, 2015). The $(G_4C_2)_n$-methylation assay combines the repeat-primed PCR (rpPCR) method that detects the presence of the expanded repeats with a methylation-specific PCR (MSP) that detects the presence of methylated expanded repeats in *C9orf72*. Briefly, bisulfite-treated DNA is amplified by rpPCR using either methylation-specific FAM labelled (GGGGTC)n or unmethylation-specific HEX labelled (GGGGTT)n probes for the detection of methylated and unmethylated HRE DNA respectively, in a qualitative manner (Xiao *et al.*, 2015) (see 2.23).

The $(G_4C_2)_n$-methylation assay was performed on bisulfite-treated DNA from 5 *C9orf72* HRE and 4 control fibroblast lines. Two independent methylation assay experiments were performed using fibroblasts of different passages per assay. The results of the HRE methylation detection in fibroblasts are shown in Figure 4.1. As shown in the unmethylation channel (green), the $(G_4C_2)_n$-methylation assay confirmed the presence of the HRE in all patient fibroblasts but not in controls, as expected (Figure 4.1). The methylation channel (blue) detected the presence or absence of DNA methylation at the HRE, which was unmethylated in all patient fibroblasts, with the exception of one patient line, C9-5, which had a methylated HRE (Figure 4.1).

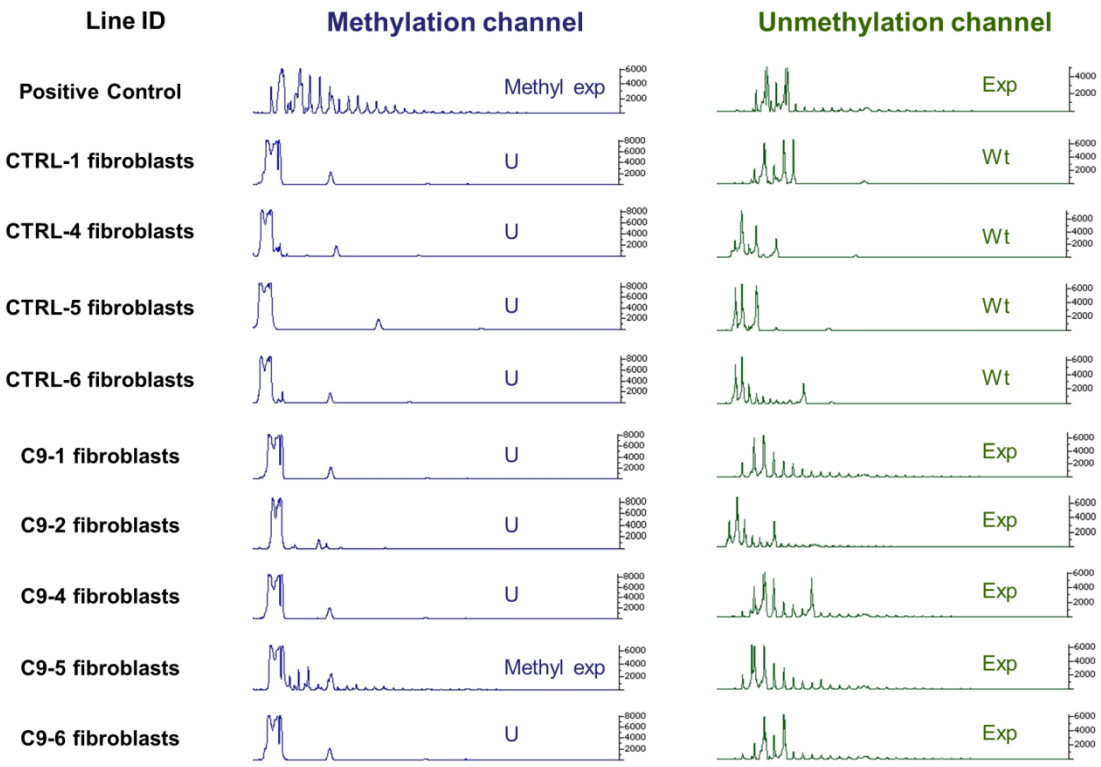


Figure 4.1 The *C9orf72* HRE is unmethylated in patient fibroblast lines.

The methylation status of the *C9orf72* HRE was examined using the $(G_4C_2)_n$-methylation assay in 4 control (CTRL-1, CTRL-2, CTRL-3 and CTRL-4) and 5 *C9orf72* HRE carrier (C9-1, C9-2, C9-4, C9-5 and C9-6) fibroblast lines. A known methylated *C9orf72* HRE carrier gDNA sample was used as a positive control for the assay. The assay was performed twice using genomic DNA from fibroblast samples of different passages in each experiment. The Unmethylation channel (green) detects the presence or absence of an HRE at *C9orf72* whereas the Methylation channel (blue) detects presence or absence of DNA methylation at the *C9orf72* HRE.The $(G_4C_2)_n$-methylation assay results confirmed the presence of the HRE in all *C9orf72* repeat expansion fibroblast lines and its absence from all controls (unmethylation channel). Despite a similar peak trace between CTRL-6 and C9-2 lines in the unmethylation channel, there are no extended small spikes beyond the peak spike for CTRL-6, indicating this line carries a small expansion within the non-pathogenic range. On the other hand, C9-2 has extended spikes beyong the preak spike, indicating the presence of a repeat expansion allele. The HRE of four *C9orf72* repeat expansion fibroblasts was unmethylated (methylation channel). The C9-5 fibroblast line had a methylated HRE (methylation channel). *Abbreviations:* Exp: Expansion in *C9orf72*, Wt: Wild-type, U: unmethylated, Methyl exp: methylated *C9orf72* expansion.

4.2.2. Methylation of the 5' and 3' CpG islands flanking the *C9orf72* HRE in fibroblasts

The *C9orf72* HRE is flanked by two CpG islands, one upstream of the HRE (5' CpG island) and one downstream of the HRE (3' CpG island) which are potential hypermethylation hotspots. The 5' CpG in particular has been found to be hypermethylated in *C9orf72* HRE patients (Xi *et al.*, 2013, Xi *et al.*, 2014b).

The DNA methylation status of these two CpG islands was investigated via bisulfite sequencing in control and *C9orf72* HRE patient fibroblasts. Figure 4.2 shows the DNA methylation status of 26 CpG (24/26 determined and 2/26 undetermined) sites at the 5' CpG island (A) and 22 CpG sites at the 3' CpG island (B) in 4 control and 5 patient fibroblast lines. In Figure 4.2 A, the 5' CpG island was fully unmethylated in all control fibroblasts and in 4 out of 5 C9orf72 HRE fibroblast lines at all the CpG sites assessed (24/24). However, hypermethylation was detected in the 5' CpG island of the C9-5 patient fibroblast line, where 22 out of the 24 CpG sites assessed were methylated. The observed 5'CpG hypermethylation in C9-5 is in line with the $(G_4C_2)_n$-methylation assay finding that the HRE itself is methylated (Figure 4.1), supporting an association between HRE methylation and 5' CpG hypermethylation. The 3' CpG was fully unmethylated (22/22 CpG sites) in all control and *C9orf72* HRE fibroblast lines as shown in Figure 4.2 B.

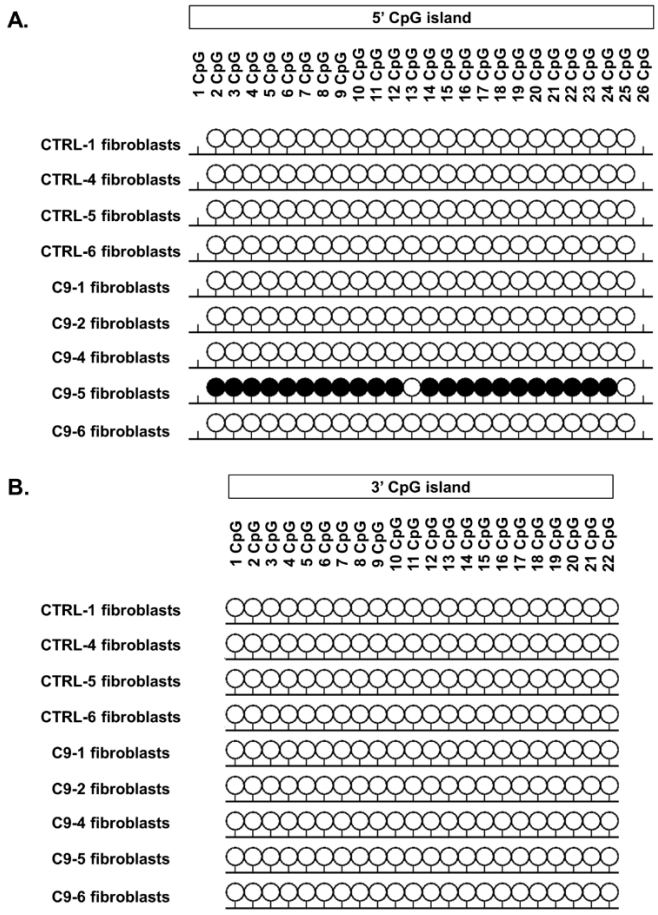


Figure 4.2 The 5' and 3' CpG islands flanking the *C9orf72* HRE are not hypermethylated in patient fibroblasts.

The DNA methylation status of the 5' (26 sites) and 3' CpG (22 sites) islands flanking the *C9orf72* HRE was assessed using a Bisulfite Sequencing PCR (BSP) assay in 4 control (CTRL-1, CTRL-2, CTRL-3 and CTRL-4) and 5 *C9orf72* repeat expansion carrier (C9-1, C9-2, C9-4, C9-5 and C9-6) fibroblast lines. (A) All the assessed CpG residues of the 5' CpG island (24/26) were found to be unmethylated in control and *C9orf72* HRE patient fibroblast lines with the exception of C9-5 line which was hypermethylated in 22/24 sites of the 5' CpG island. For all samples, 24/26 CpG sites were assessed and 2/26 were undetermined. (B) At the 3' CpG island, all of the sites assessed via BSP were unmethylated in both control and *C9orf72* HRE patient fibroblasts. White lollipop circles represent non-methylated CpG sites, black lollipop circles represent methylated CpG sites and sticks without circles represent sites with non-determined methylation status. The lollipop diagram shows equally spaced CpGs across the target regions, not reflecting true genomic positions. Diagrams show the DNA methylation results for one *C9orf72* allele, representing 50% of methylation status (the other allele is wt and unmethylated in all samples).

4.2.3. Temporal methylation of the *C9orf72* HRE during CN differentiation

Cortical neurons are primarily affected in FTD, therefore assessment of DNA methylation at the *C9orf72* promoter region was performed in control and *C9orf72* HRE iPSC-CNs. Specifically, promoter region methylation assessment included assessment of DNA methylation status of the *C9orf72* HRE and the two CpG islands flanking the HRE, similar to the experiments performed in fibroblast lines in 4.2.1 and 4.2.2. Moreover, the temporal establishment of DNA methylation during differentiation of iPSCs to CNs was investigated by assessing the methylation status of a series of samples taken at different timepoints throughout the differentiation protocol.

In order to assess the temporal methylation of the *C9orf72* HRE during CN differentiation, 3 control (1 ESC and 2 iPSC) and 3 patient iPSC lines were differentiated into CNs as described in 2.5. For all lines, iPSCs and CNs at 0, 10, 30, 50, 70, 80, 110 and 200 DIV were collected for DNA extraction and methylation analysis (see 2.22 and 2.23). The temporal DNA methylation of the HRE was assessed via two independent $(G_4C_2)_n$-methylation assays using CNs derived from two independent inductions. The representative chromatograms of the $(G_4C_2)_n$-methylation assay corresponding to the temporal DNA methylation of the *C9orf72* HRE are shown in Figure 4.3. As expected, no HRE was detected in control iPSC-CNs in any of the lines, at any of the timepoints examined (Figure 4.3, A, B and C, unmethylation channel). In contrast, the HRE was present in all *C9orf72* iPSC-CNs (Figure 4.3, D, E and F, unmethylation channel). Moreover, the HRE was found to be methylated in all *C9orf72* iPSC-CN lines (Figure 4.3, D, E and F, methylation channel), in contrast to the unmethylated HRE of the corresponding *C9orf72* fibroblasts (Figure 4.1). Analysis of the temporal *C9orf72* HRE methylation revealed a stably methylated HRE across different timepoints during CN differentiation in C9-2 and C9-3 patient lines (Figure 4.3, E and F). Specifically, the HRE methylation at these two lines was established as early as at the DIV 0 iPSC stage and maintained

until the most mature CNs that were examined. The third patient line, C9-1, however, was characterised by methylation instability of the HRE (Figure 4.3, D). Specifically, the HRE was unmethylated in C9-1 iPSCs, DIV 10, 50, 70 and 110 CNs but methylated in DIV 30, 80 and 200 CNs. Moreover, it is noteworthy that no CN methylation data are available on the additional control (CTRL-4, CTRL-5 and CTRL-6) and *C9orf72* HRE patient (C9-4, C9-5 and C9-6) fibroblast lines shown in Figure 4.1 due to the unavailability of reprogrammed iPSC lines.

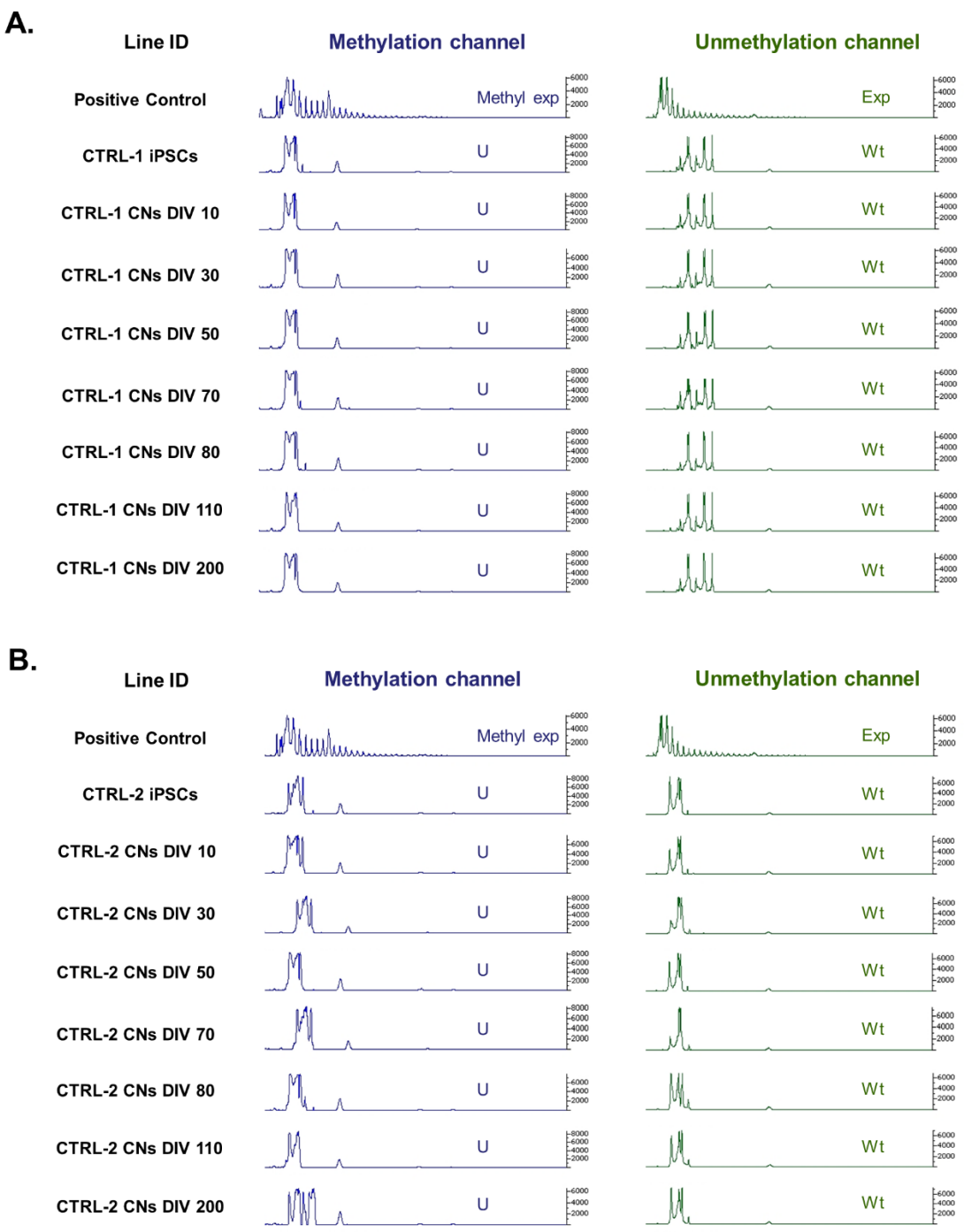
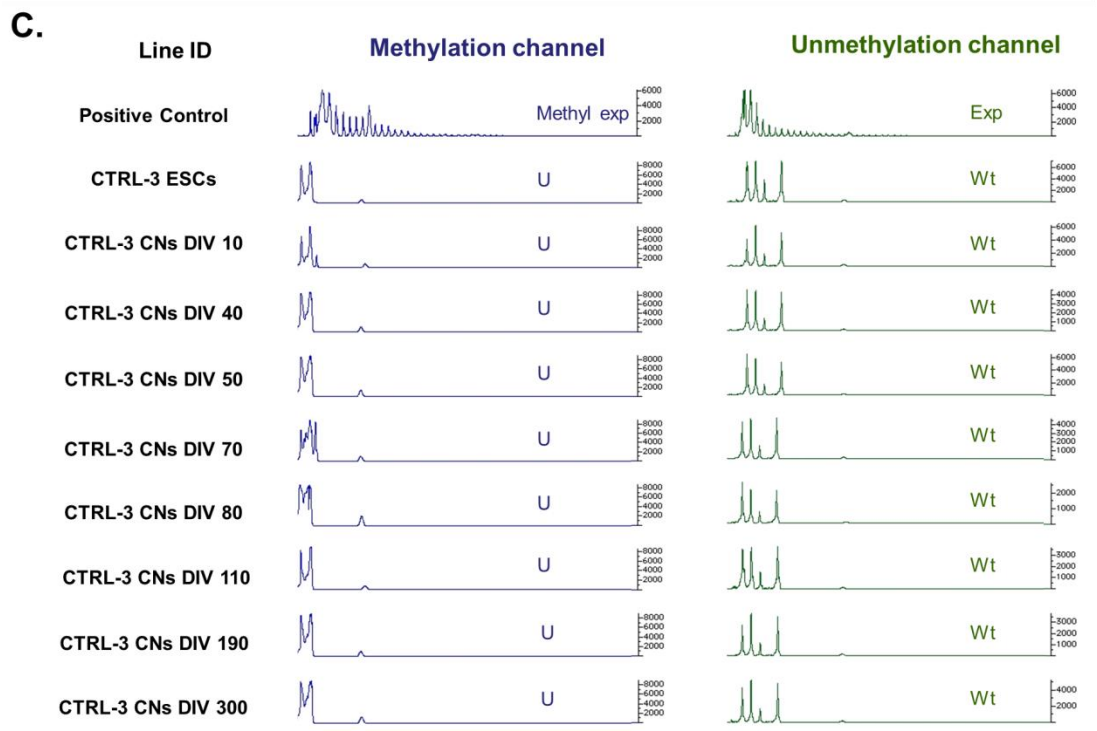
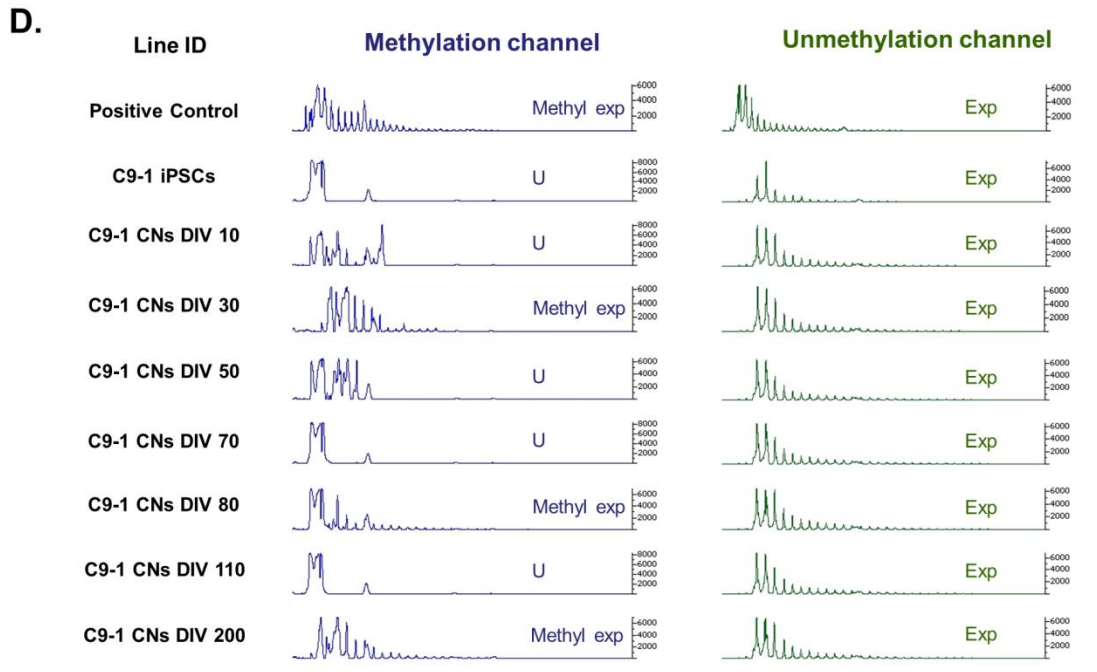
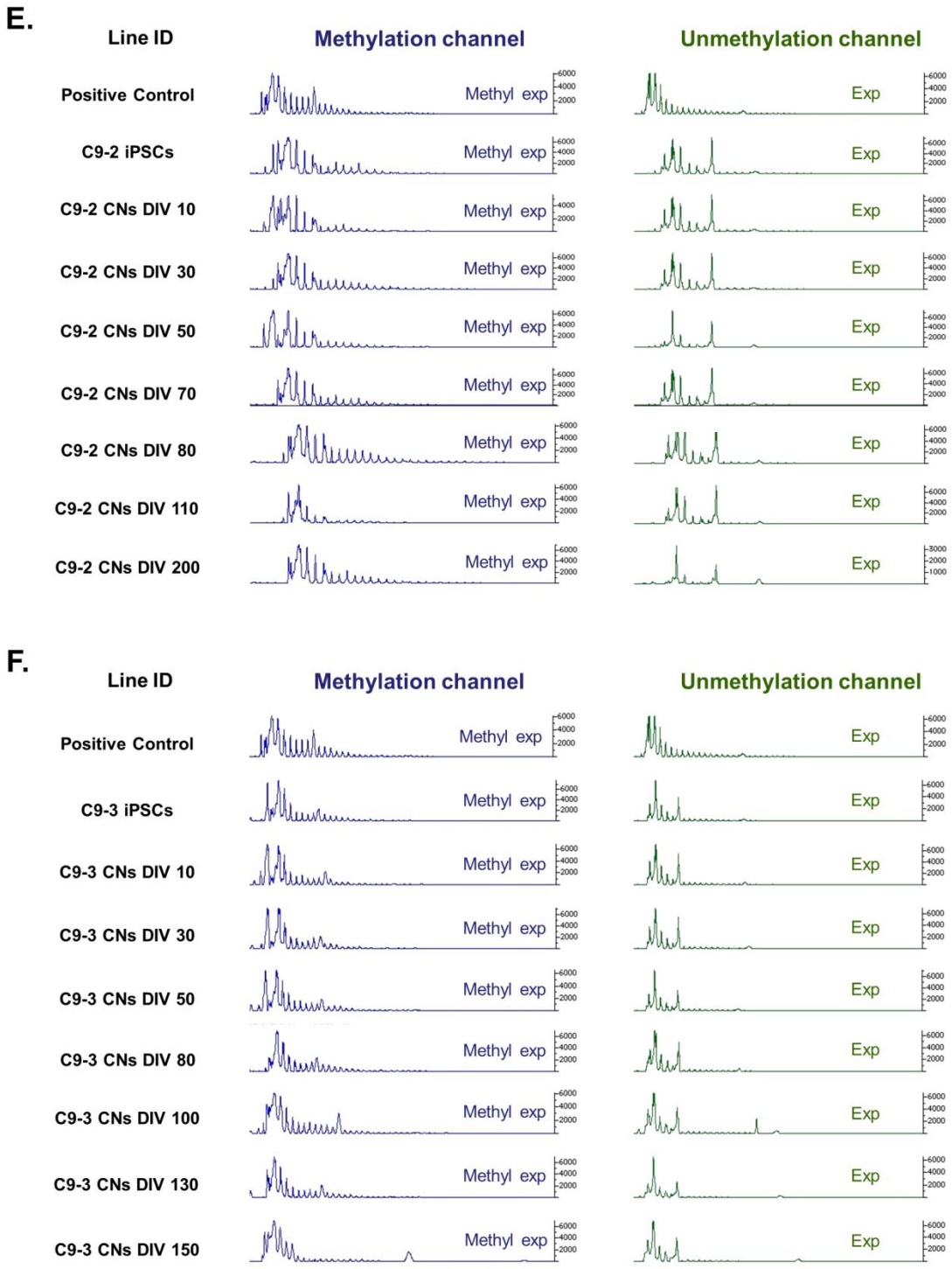


Figure 4.3 The *C9orf72* HRE is methylated in patient iPSC-CNs.





(Continued) Figure 4.3 The *C9orf72* HRE is methylated in patient iPSC-CNs.



(Continued) Figure 4.3 The *C9orf72* HRE is methylated in patient iPSC-CNs.

The temporal methylation of the *C9orf72* HRE at different timepoints during CN differentiation was assessed via $(G_4C_2)_n$-methylation assay in 3 control (CTRL-1, CTRL-2 and CTRL-3) and 3 *C9orf72* HRE patient (C9-1, C9-2 and C9-3) lines. Briefly, control and *C9orf72* HRE patient iPSC lines were differentiated into CNs and samples were collected at fixed timepoints during cortical differentiation (DIV 0, 10, 30, 50, 70, 80, 110 and 200) for the $(G_4C_2)_n$-methylation assay. Representative chromatograms

(Continued Figure legend) of the assay consist of two detection channels. Unmethylation channel (Green) is detecting the presence or absence of the HRE whereas the Methylation channel (Blue) is detecting presence or absence of methylation of the HRE. As expected, no HRE or methylation was detected in control iPSCs/ESCs and CNs across all timepoints assessed shown in A, B and C. All patient iPSCs and CN samples carried the expanded *C9orf72* allele (unmethylation channel) across all timepoints assessed (D, E and F). C9-2 (E) and C9-3 (F) patient lines had a methylated HRE in all the timepoints assessed (E and F), however, C9-1 patient line (D) showed methylation instability of the HRE during CN differentiation (methylation channel). The Unmethylation channel (green) detects the presence or absence of an HRE at *C9orf72* whereas the Methylation channel (blue) detects presence or absence of DNA methylation at the *C9orf72* HRE. *Abbreviations:* Exp: Expansion in *C9orf72*, Wt: Wild-type, U: unmethylated, Methyl exp: methylated expansion.

4.2.4. Temporal methylation of the 5' and 3' CpG islands flanking the *C9orf72* HRE during cortical differentiation

For the investigation of the DNA methylation at the *C9orf72* promoter region in CNs, the temporal DNA methylation at the 5' and 3' CpG islands flanking the *C9orf72* HRE was also assessed via BSP. Similar to 4.2.3, samples were collected at different timepoints throughout the CN differentiation at 0, 10, 30, 50, 70, 80, 110 and 200 DIV. Figure 4.4 shows the temporal DNA methylation at the 5' CpG island in control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE patient (C9-1, C9-2 and C9-3) lines. The 5' CpG island was unmethylated at all stages during CN differentiation and at all the CpG sites (26/26) assessed in all the control lines (Figure 4.4, A, B and C). In *C9orf72* HRE patient CNs, however, the 5' CpG island was methylated, with variable levels of methylation between patient lines and stages of CN differentiation. In C9-1, no methylation was detected at early stages of differentiation, whereas from DIV 70 onwards, a low degree of methylation in 2-4 CpG sites out of 26, mostly clustered at the 5' end of the island, was observed (Figure 4.4, D). C9-2 was characterised by a hypermethylated 5' CpG island, with all CpG residues assessed methylated in CNs (Figure 4.4, E). C9-3 also displayed a hypermethylated 5' CpG island, however, unlike C9-2, not all CpG sites were found methylated (Figure 4.4, F). A trend towards a complete hypermethylation pattern at later timepoints was observed and by 150 DIV the 5' CpG island was almost fully methylated at all CpG sites (21/22) (Figure 4.4, F).

The DNA methylation assessment results for the 3' CpG island are shown in Figure 4.5. Specifically, the 3' CpG island was fully unmethylated in all 22 CpG sites assessed in control CNs (Figure 4.5, A, B and C). In contrast to the 5' CpG island, the 3' CpG was not hypermethylated in patient CNs (Figure 4.5). Patient lines C9-1 and C9-2 only showed a low degree of methylation at no more than 4 CpG sites out of 22 in total (Figure 4.5, D and E), whereas C9-3 displayed a fully unmethylated 3' CpG island (Figure 4.5, F).

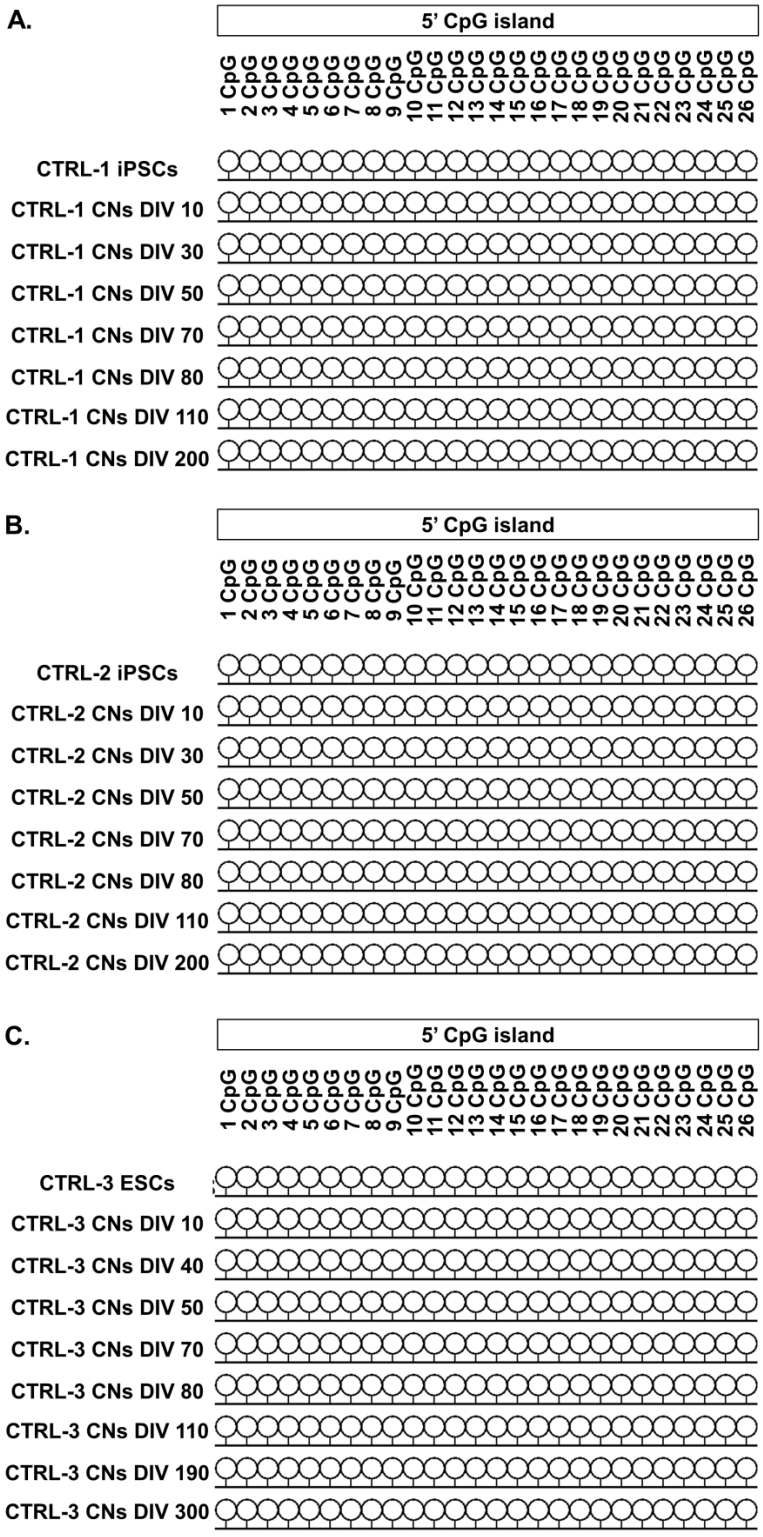
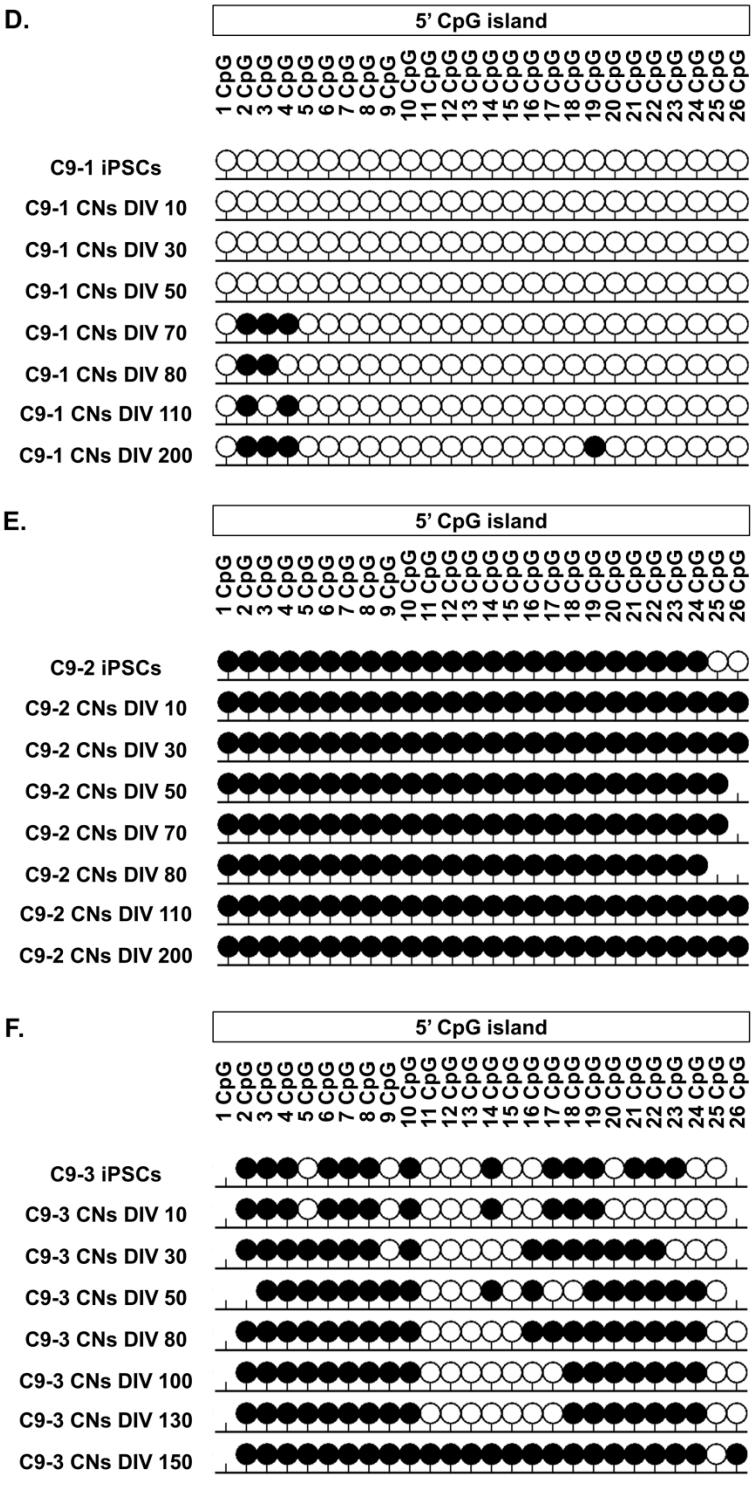


Figure 4.4 The 5' CpG island upstream of *C9orf72* is hypermethylated in *C9orf72* HRE patient iPSC-CNs.



(Continued) Figure 4.4 The 5' CpG island upstream of *C9orf72* is hypermethylated in *C9orf72* HRE patient iPSC-CNs.

The temporal DNA methylation of 26 CpG sites of the 5' CpG island upstream of the *C9orf72* HRE was assessed via BSP in different timepoints during CN differentiation for control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE patient (C9-1, C9-2 and C9-3) lines. All 26 CpG sites of the 5'CpG island were unmethylated in all the control

(Continued Figure legend) lines CTRL-1 (A), CTRL-2 (B) and CTRL-3 (C) at all the timepoints assessed. The 5' CpG islands of patient lines C9-2 (E) and C9-3 (F) were hypermethylated. The third patient line, C9-1 (D), was unmethylated in iPSCs and young cortical progenitors and showed hypomethylation in higher timepoints starting from DIV 70-80 which did not exceed 4/26 residues. White lollipop circles represent non-methylated CpG sites, black lollipop circles represent methylated CpG sites and sticks without circles represent sites with non-determined methylation status. Lollipop diagram is showing equally spaced CpGs across the target regions, not reflecting true genomic positions. Diagrams show the DNA methylation results for one *C9orf72* allele, representing 50% of methylation status (the other allele is wt and unmethylated in all samples).

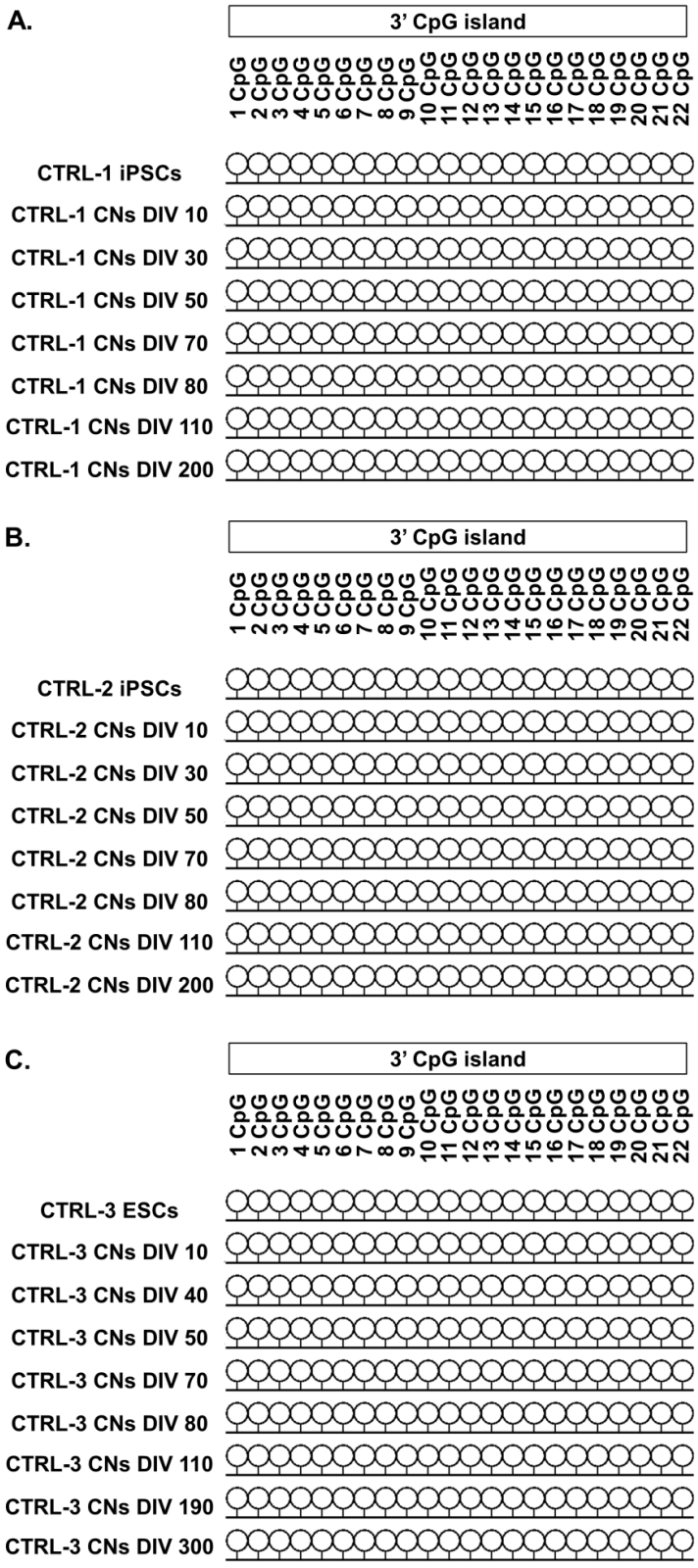
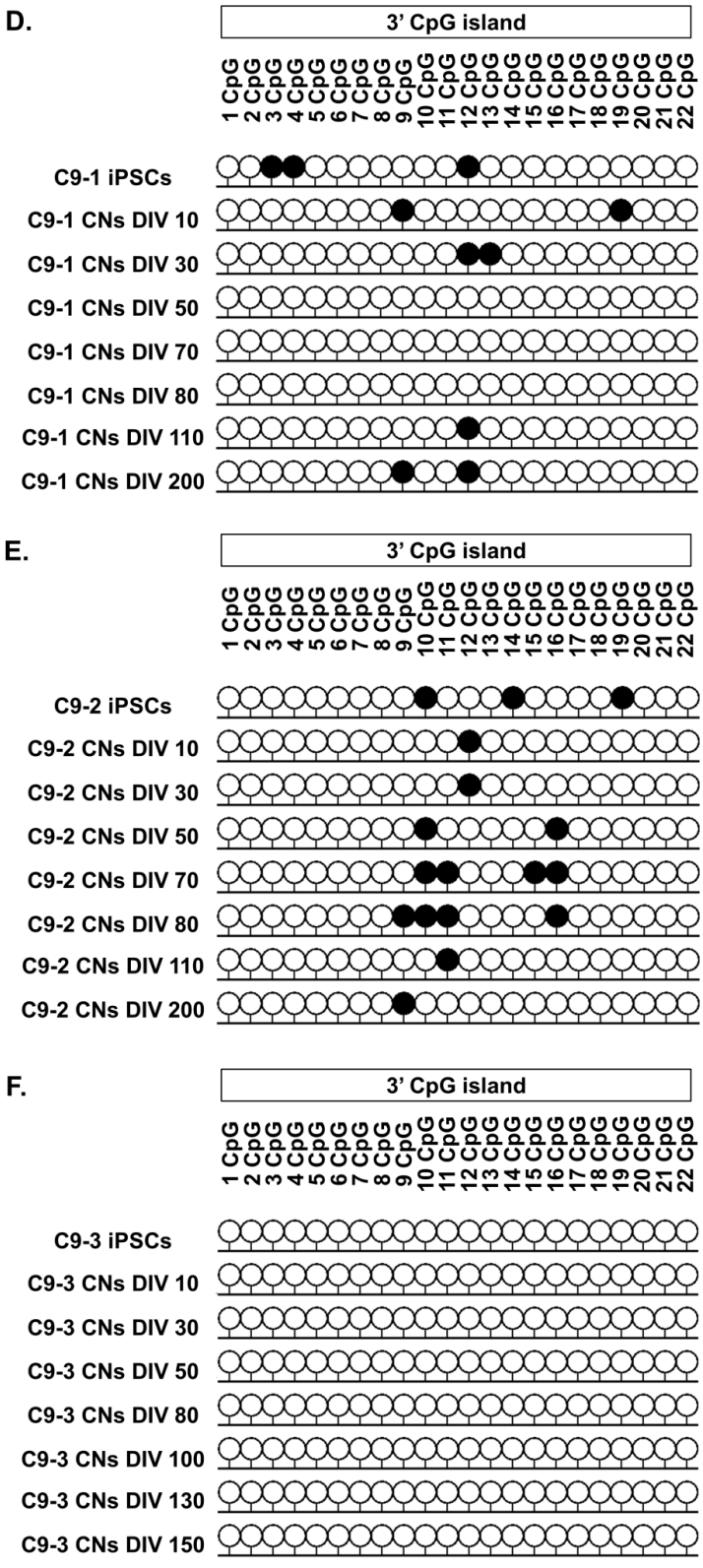


Figure 4.5 The 3' CpG island downstream of *C9orf72* is not hypermethylated in *C9orf72* HRE patient iPSC-CNs.



(Continued) Figure 4.5 The 3' CpG island downstream of *C9orf72* is not hypermethylated in *C9orf72* HRE patient iPSC-CNs.

The temporal DNA methylation of 22 CpG sites of the 3' CpG island downstream of the *C9orf72* HRE was assessed via BSP in different timepoints during cortical

(Continued Figure legend) neuron differentiation for control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE patient (C9-1, C9-2 and C9-3) lines. All 22 CpG residues of the 3' CpG island were unmethylated in all the control lines CTRL-1 (A), CTRL-2 (B) and CTRL-3 (C) and at all timepoints assessed. C9orf72 HRE patient line C9-3 (F) also displayed a fully unmethylated 3' CpG islands at all the CpG sites and timepoints assessed. Only a small amount of methylation in no more than 4 CpG sites out of the 22 was detected in samples from the C9-1 (D) and C9-2 (E) patient lines. The above data suggest that the 3' CpG island, unlike the 5' CpG island, is hypomethylated in *C9orf72* HRE patient CNs. White lollipop circles represent non-methylated CpG sites, black lollipop circles represent methylated CpG sites and sticks without circles represent sites with non-determined methylation status. Lollipop diagram is showing equally spaced CpGs across the target regions, not reflecting true genomic positions. Diagrams show the DNA methylation results for one *C9orf72* allele, representing 50% of methylation status (the other allele is wt and unmethylated in all samples).

4.2.5. Assessment of *C9orf72* RNA expression

4.2.5.1. Expression of total C9orf72 mRNA

So far, the results have suggested differential DNA methylation in control and *C9orf72* HRE iPSC-CNs. In order to determine whether the *C9orf72* mRNA is reduced in the *C9orf72* HRE iPSC-CNs and whether this reduction correlates with the presence of DNA methylation in *C9orf72* promoter region, *C9orf72* mRNA expression analysis was performed.

Total *C9orf72* mRNA levels were assessed via RT-qPCR using a TaqMan Assay detecting all three *C9orf72* mRNA transcripts as described in 2.33.2, in control and *C9orf72* iPSC-CNs at 150 DIV. The rationale for assessing *C9orf72* expression at 150 DIV was that by that timepoint the patient lines C9-2 and C9-3 had established a fully methylated 5' CpG island pattern according to the BSP data (Figure 4.4, E and

F). Total *C9orf72* transcript Ct values were normalised to the housekeeping gene *TBP* Ct values. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) where ΔΔCt values were calculated relative to control human brain tissue *C9orf72* expression. RT-qPCR was performed in technical duplicate using RNA samples from three independent inductions per line (n=3).

Figure 4.6 shows the total *C9orf72* mRNA relative expression in control and *C9orf72* 150 DIV iPSC-CNs. A one way analysis of variance (ANOVA) revealed a statistically significant difference in the total *C9orf72* mRNA expression between the lines (F (5, 13) = 22.58, p = 5.48e-06). Tukey post hoc multiple comparison of means testing showed that C9-2 and C9-3 CNs had significantly reduced *C9orf72* mRNA expression compared to all other lines at the 0.05 significance level. Specifically, *C9orf72* mRNA expression in C9-2 CNs was significantly reduced compared to control lines CTRL-1 (p_{adj} = 0.0000370), CTRL-2 (p_{adj} = 0.0000756) and CTRL-3 (p_{adj} = 0.0000685). Similarly, C9-3 CNs had significantly lower *C9orf72* total mRNA levels compared to controls CTRL-1 (p_{adj} = 0.0007694), CTRL-2 (p_{adj} = 0.0016843) and CTRL-3 (p_{adj} = 0.0015114). C9-1 total *C9orf72* levels were not reduced compared to controls and both C9-2 and C9-3 displayed significantly lower *C9orf72* mRNA compared to C9-1 (p_{adj} = 0.0002605 and p_{adj} = 0.0063814 respectively).

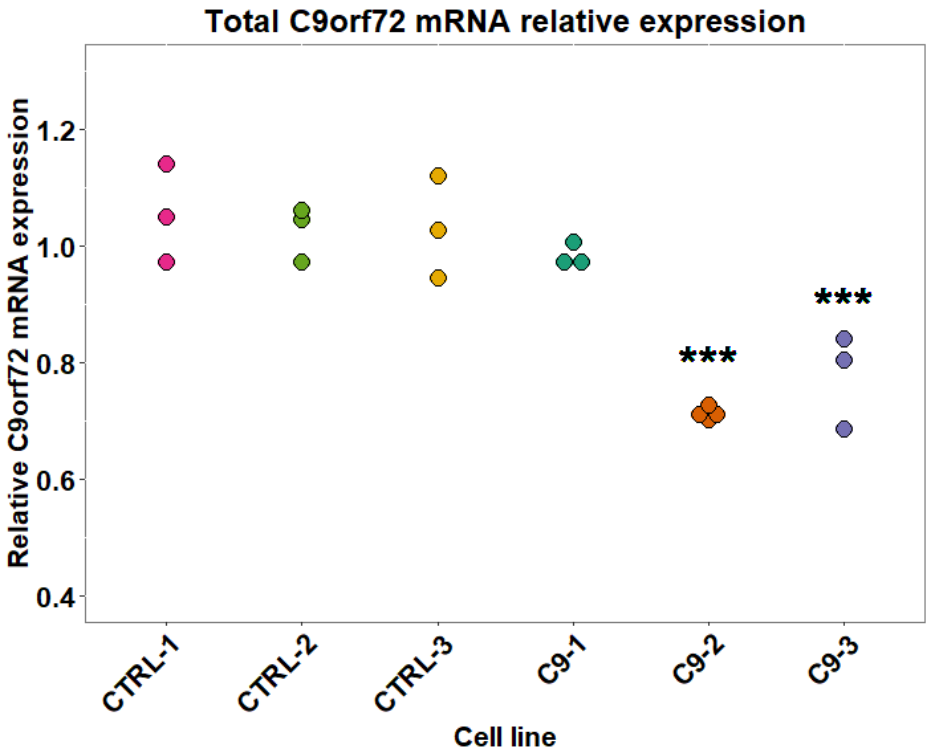


Figure 4.6 *C9orf72* total mRNA expression is reduced in C9-2 and C9-3 iPSC-CNs. The total *C9orf72* mRNA relative expression in control and patient 150 DIV CNs was assessed by RT-qPCR using a TaqMan Assay. Expression values were normalised to *TBP* gene. Relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method and shown relative to control human brain. A significant reduction in total *C9orf72* mRNA was observed in patients C9-2 and C9-3 (n=3-4 independent inductions, performed in technical duplicate). *** represent p_{adj}<0.001.

4.2.5.2. *Transcript-specific expression of C9orf72*

C9orf72 is alternatively spliced to produce 3 main transcript variants (Figure 4.7). Studies have shown that the presence of the HRE can lead to transcript-specific expression alterations. Therefore the expression of the three *C9orf72* RNA transcripts V1 (transcript NM_145005), V2 (transcript NM_018325) and V3 (transcript NM_001256054) was investigated with semi-quantitative RT-PCR in control and *C9orf72* 150 DIV iPSC-CNs (2.34). For this experiment, a modified version of a previously published RT-PCR assay that detects all three *C9orf72* transcripts in a semi-quantitative manner was utilised (Waite *et al.*, 2014). Specifically, the detection of all three *C9orf72* transcripts was based on the use of two sets of primers, with a common reverse primer 2R binding on exon 2 and two forward primers binding on either of the non-coding exons 1a or 1b (Figure 4.7). RT-PCR products were resolved on a 1.5% agarose gel and discriminated based on their amplicon size. The RT-PCR reaction for V2 transcript produced a 221 bp long amplicon, whereas the reaction for V1 and V3 transcripts produced a 242 bp amplicon corresponding to V1 and a 340 bp amplicon corresponding to V3 (Figure 4.7).

The RT-PCR gel of *C9orf72* transcripts in control and *C9orf72* 150 DIV iPSC-CNs is shown in Figure 4.8 A. The experiment was performed for n=3 independent CN inductions from each line, in addition to control human brain (Figure 4.8 A). Expression values were calculated by normalisation of the densitometric values of each transcript to the values of the housekeeping gene *TBP*. The normalised expression of the three *C9orf72* transcripts, expressed in arbitrary units (a.u.), is shown in Figure 4.8 B. No significant differences in the expression of V2 transcript were observed in *C9orf72* iPSC-CNs. However, a significant upregulation of V1 was found in C9-1 patient iPSC-CNs compared to controls (p_{adj}<0.0001 and p_{adj}<0.001) and the other two *C9orf72* lines (p_{adj}<0.001 and p_{adj}<0.01). Similarly, V3 transcript

levels were also significantly upregulated in C9-1 patient iPSC-CNs compared to controls (p_{adj}<0.0001 and p_{adj}<0.001) and the other two *C9orf72* lines (p_{adj}<0.001).

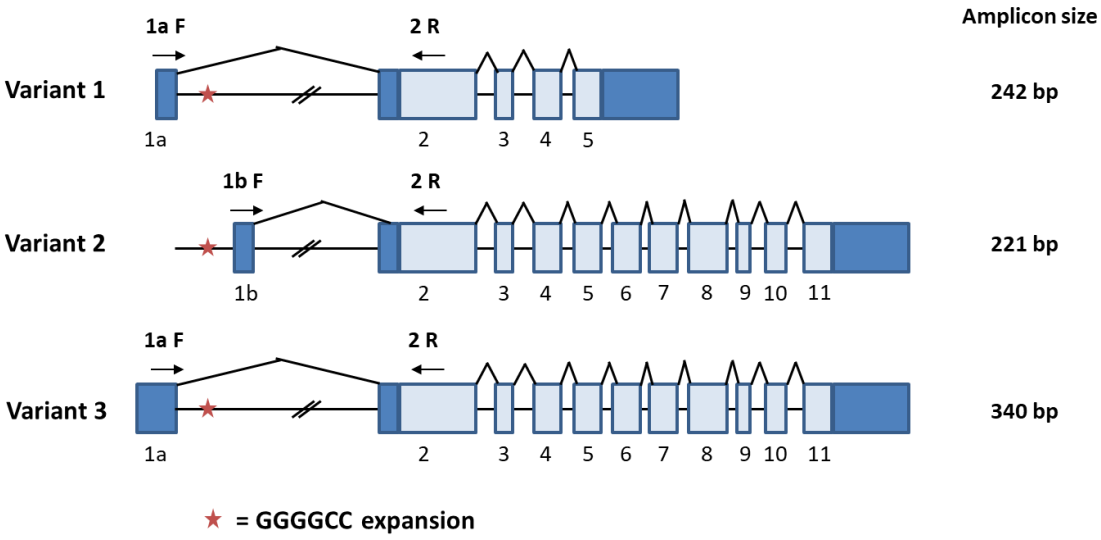


Figure 4.7 Schematic illustration of the semi-quantitative RT-PCR primer binding sites for the detection of the three *C9orf72* transcript variants.

A modified version of a previously published protocol was used for the assessment of expression of the three *C9orf72* transcript variants V1 (transcript NM_145005), V2 (transcript NM_018325) and V3 (transcript NM_001256054) in a semi-quantitative manner (method modified from Waite *et al.*, 2014). The three transcripts can be discriminated by their different amplicon sizes, 221 bp for V2 and 242 bp/340 bp for V1/V3, after being resolved on an agarose gel.

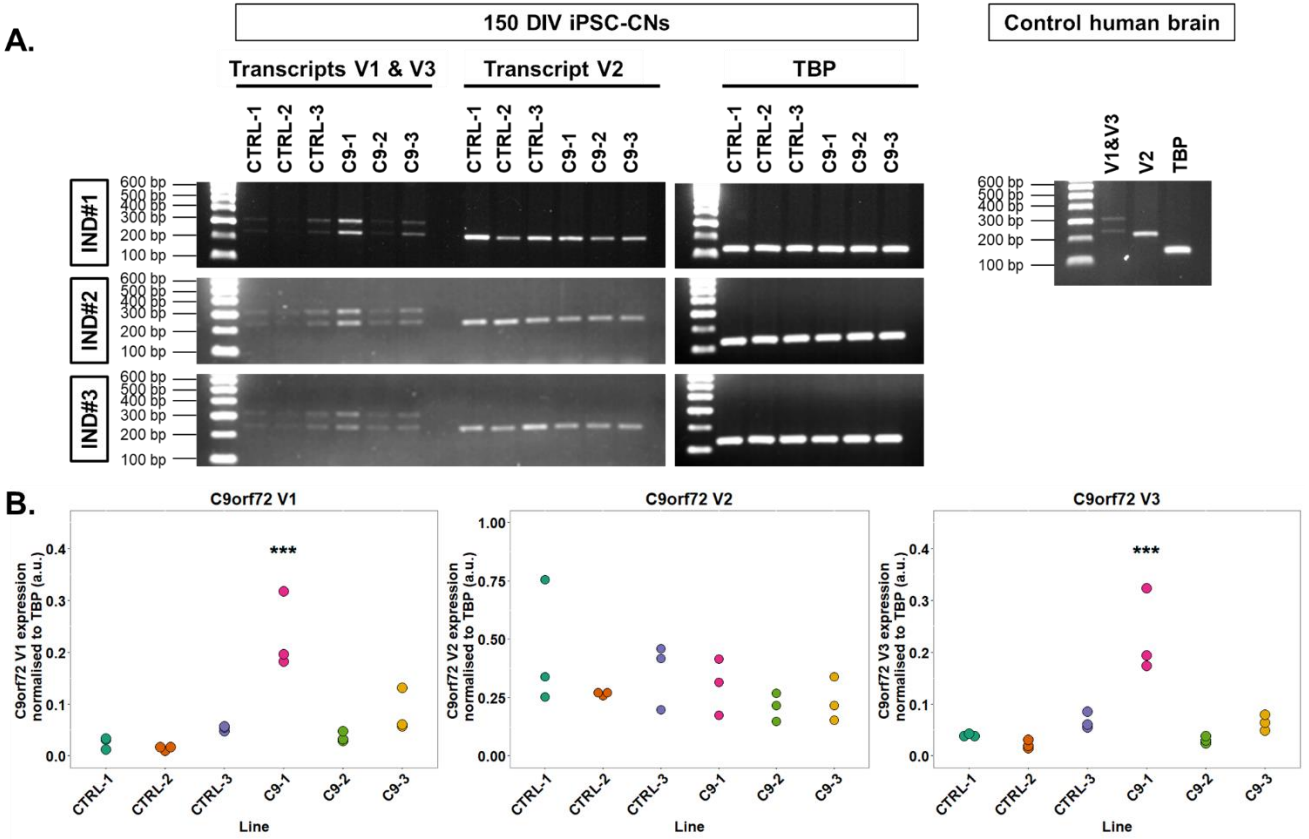


Figure 4.8 *C9orf72* transcripts V1 and V3 are significantly upregulated in C9-1 patient 150 DIV iPSC-CNs compared to control and other patient CNs.

A. Control and *C9orf72* HRE patient iPSCs were differentiated to CNs and harvested at 150 DIV for RNA extraction and semi-quantitative analysis of *C9orf72* mRNA transcripts by RT-PCR. RT-PCR was performed using previously published primers (Waite *et al.*, 2014) detecting the three *C9orf72* transcripts V1 (transcript NM_145005), V2 (transcript NM_018325) and V3 (transcript NM_001256054) (Figure 4.7). Additional RT-PCR

(Continued Figure legend) reactions for housekeeping gene *TBP* were performed for normalisation of expression values. All RT-PCRs were also performed on control human brain cDNA. B. Gel densitometry quantification of RT-PCR data revealed V1 and V3 *C9orf72* transcripts were significantly upregulated in C9-1 patient CNs.

In addition to RT-PCR, the relative expression of all three *C9orf72* transcripts was also assessed by RT-qPCR using previously published, custom-made, transcript-specific locked nucleid acid (LNA) probes in the 150 DIV iPSC-CN samples of control and *C9orf72* lines (Fratta *et al.*, 2013, Simone *et al.*, 2018) (2.33.3). Figure 4.9 shows the relative expression of V1, V2 and V3 *C9orf72* transcripts normalised to the housekeeping gene *TBP* for the three control and three patient study lines. No significant differences in the expression levels of C9orf72 transcripts V1 and V2 were observed. However, transcript V3 was found to be significantly upregulated in C9-1 iPSC-CNs compared to controls (p_{adj}=0.001) and the other two *C9orf72* lines (p_{adj}=0.001 and p_{adj}=0.01).

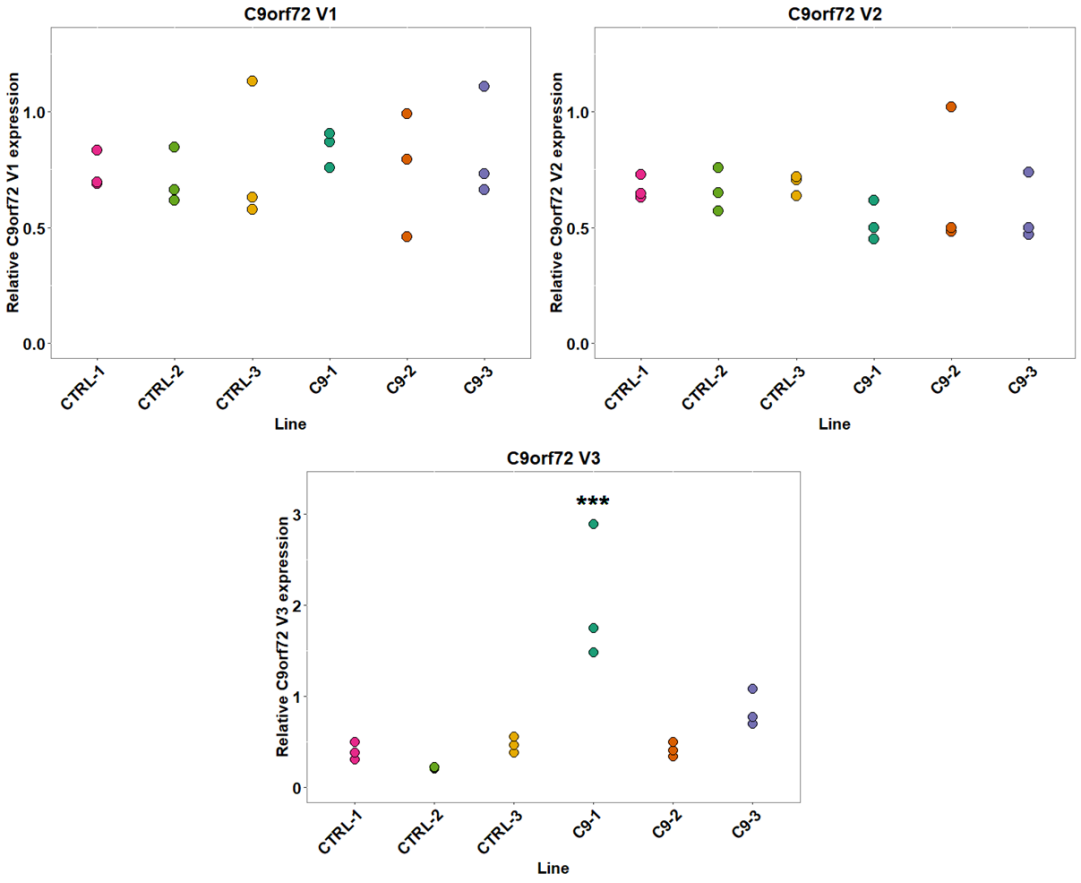


Figure 4.9 *C9orf72* V3 is significantly upgregulated in C9-1 150 DIV iPSC-CNs. Relative expression of V1 (transcript NM_145005), V2 (transcript NM_018325) and V3 (transcript NM_001256054) *C9orf72* transcripts in DIV 150 iPSC-CNs from three control (CTRL-1, CTRL-2 and CTRL-3) and three *C9orf72* (C9-1, C9-2 and C9-3) lines. Expression values were normalised to *TBP* gene and relative expression was calculated using the $2^{-\Delta\Delta Ct}$ method and shown relative to control human brain expression (n=3 independent inductions, performed in technical duplicate). *** represent p_{adj}<0.001.

4.2.6. Allele expression imbalance analysis of *C9orf72* in hypermethylated *C9orf72* iPSC-CNs

In order to assess whether methylation of the HRE or 5' CpG hypermethylation have any negative effect on the expression of the expanded *C9orf72* allele, allele expression imbalance analysis was performed. For this purpose, the *C9orf72* coding region was searched for common (minor allele frequency (MAF)>1%) SNPs. The SNPs with the highest MAF that were selected for sequencing and downstream allele expression imbalance analysis were rs17769294, c.620A>G, MAF=0.1043 (ExAC) and rs10122902, c.870C>T, MAF=0.2362 (ExAC). All control and *C9orf72* HRE lines C9-2 and C9-3 were sequenced for both SNPs.

Figure 4.10 shows the electropherograms of the sequencing for the rs17769294 (c.620A>G) and rs10122902 (c.870C>T) in control and hypermethylated *C9orf72* patient lines. C9-3 was heterozygous for rs17769294 whereas CTRL-3 and C9-3 lines were both heterozygous for rs10122902.

For the allele expression imbalance analysis, rs10122902 was genotyped real-time in gDNA and cDNA from 150 DIV iPSC-CN from CTRL-3 and C9-3 lines using a TaqMan genotyping assay (Assay ID C___3137372_30) as described in 2.35. Experiments were performed on 150 DIV iPSC-CNs from 3 independent CN inductions (n=3). As shown in Figure 4.11, no significant allele expression imbalance is present for C9orf72 in C9-3 patient-derived 150 DIV CNs, as the relative frequency (rf) was calculated between the threshold values $0.4 < rF < 0.6$.

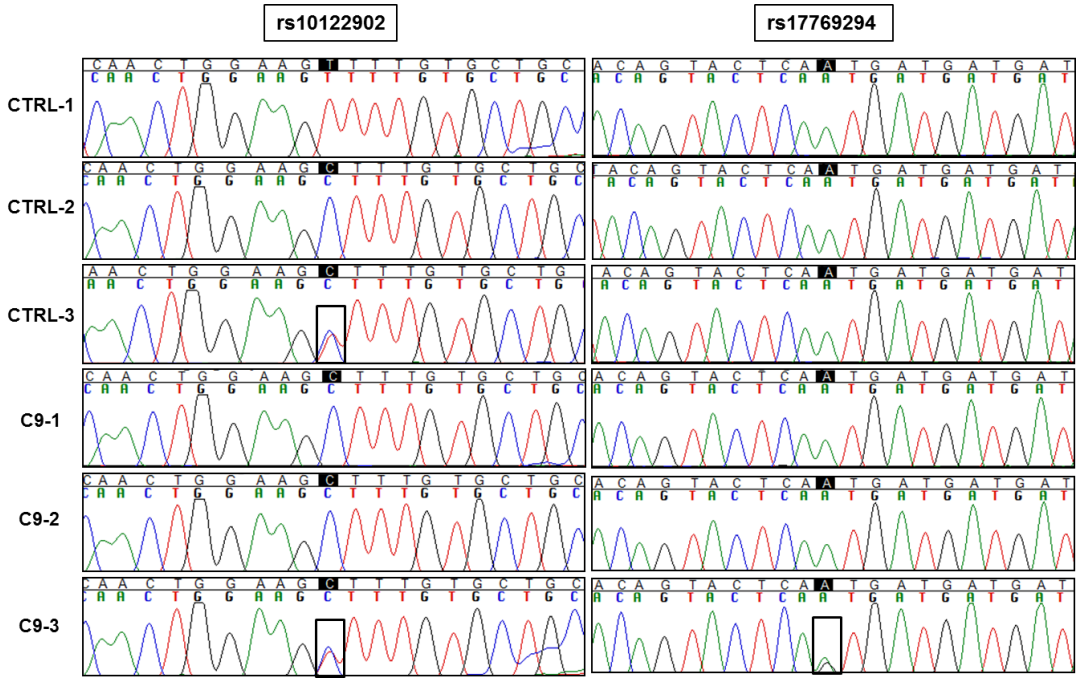


Figure 4.10 CTRL-3 and C9-3 lines are heterozygous for rs10122902 and C9-3 line is heterozygous for rs17769294.

Sanger sequencing electropherograms of rs10122902 c.870C>T and rs17769294 (c.620A>G) in control (CTRL-1, CTRL-2 and CTRL-3) and *C9orf72* HRE patient (C9-1, C9-2 and C9-3) lines. Lines CTRL-3 and C9-3 were heterozygous for rs10122902 whereas line C9-3 was also heterozygous for rs17769294. Black boxes indicate the heterozygous lines for both SNPs. cDNA sequence is shown.

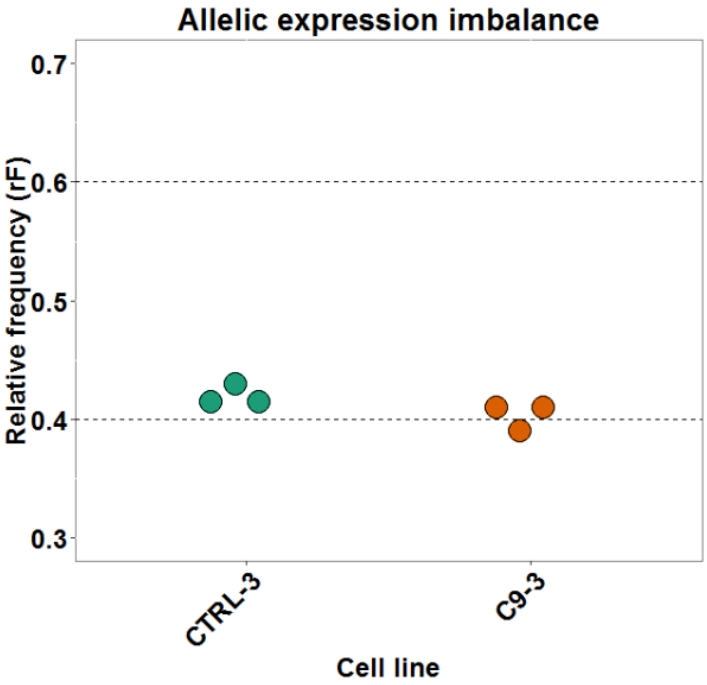


Figure 4.11 No significant *C9orf72* allelic expression imbalance is observed for hypermethylated C9-3 line 150 DIV iPSC-CNs.

Allele expression imbalance analysis for *C9orf72* was performed by real-time genotyping of rs10122902 gDNA and cDNA from CTRL-3 and C9-3 150 DIV iPSC-CN. N=3 independent inductions. Horizontal dashed lines represent 0.4 and 0.6 thresholds, whereby 0.4 > rF > 0.6 is considered evidence for allelic expression imbalance.

4.2.7. Characterisation of commercial antibodies for C9orf72 detection by western blot

The detection of a novel, uncharacterised protein such as C9orf72 via immunoblotting is challenging and limited by the availability of specific commercial antibodies. Here, the specificity of commercially available C9orf72 antibodies was tested in western blot. For this purpose, stable *C9orf72* knockdown HEK293 cell lines were generated via lentiviral shRNA transduction (2.28) for use as negative controls for C9orf72 protein detection. Figure 4.12 shows representative blots from two antibodies, Sigma HPA023873 and Proteintech 22637-1-AP, that detected specifically the C9orf72 long isoform (C9-L) as a single band at 50 kDa which was present in wild-type and mock-

transduced HEK cells but reduced in both *C9orf72* knockdown HEK293 cell lines (#669 and #670). The short isoform of C9orf72 (C9-S) was not detected by any of these antibodies. Additional non-specific bands were present for both antibodies. However, some of these could be attributed to non-specific binding of the anti-rabbit secondary antibody.

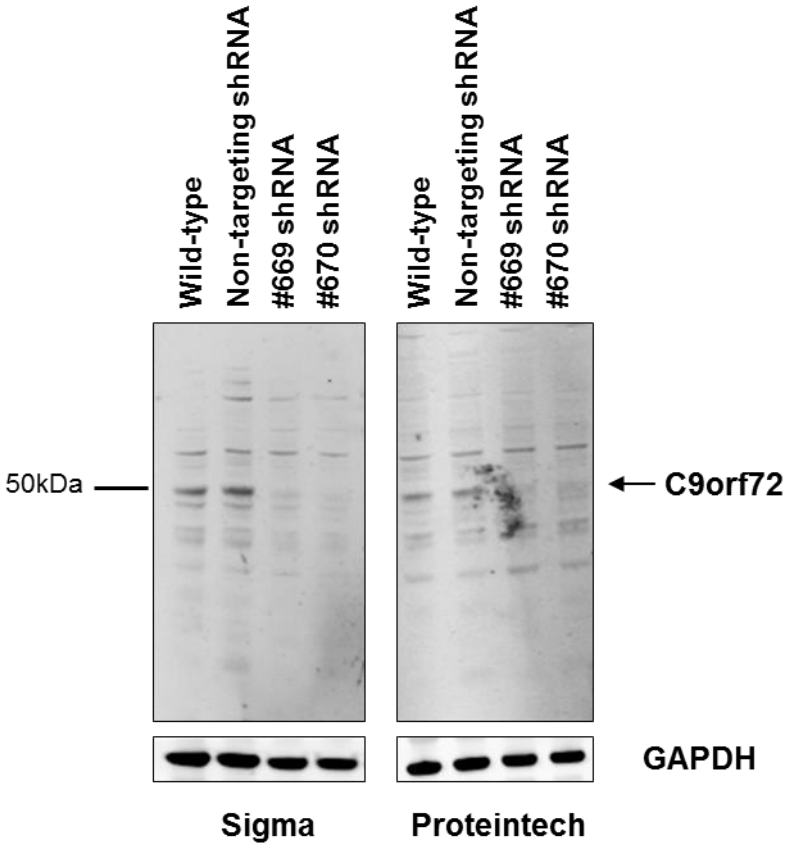


Figure 4.12. C9-L isoform is detected as a single band at 50 kDa via western blot. The specificity of commercially available antibodies against C9orf72 was validated by western blot using stable C9orf72 knockdown HEK293 cell lines (#669 and #670). Two rabbit polyclonal antibodies from Sigma (HPA023873) and Proteintech (22637-1-AP) detected specifically the long isoform of C9orf72 at 50 kDa in wild-type and mock-transduced HEK293 cells and not in the knockdown lines (arrow). GAPDH was used as loading control.

4.2.8. Investigation of temporal C9orf72 protein levels in control iPSC-CNs at extended timepoints

The temporal expression of C9orf72 in iPSC-CNs has not been investigated. Therefore assessment of C9orf72 protein levels in control iPSC-CNs at different timepoints *in vitro* was carried out to investigate its developmental regulation. Moreover, the analysis of C9orf72 expression levels was assessed in iPSC-CNs of extended timepoints, up to 365 DIV, which are not conventionally assessed in published studies.

CTRL-1 iPSC-CNs collected at different timepoints during CN differentiation and maturation were analysed by western blot. Figure 4.13 shows the western blot results for C9orf72 temporal expression in n=4 independent inductions of the CTRL-1 line. The timepoints assessed ranged from day 0 (iPSCs) until 260 DIV, and with one sample at 365 DIV (n=1). The main protein isoform detected was the C9-L protein isoform at ~50 kDa. Gel densitometry shown in Figure 4.13 B shows levels of C9orf72 remained stable throughout cortical differentiation and maturation (n=4 independent inductions, normalised to β-actin). Wild-type and stable *C9orf72* knockdown HEK293 cell lines were used as positive and negative controls respectively for the detection of C9orf72.

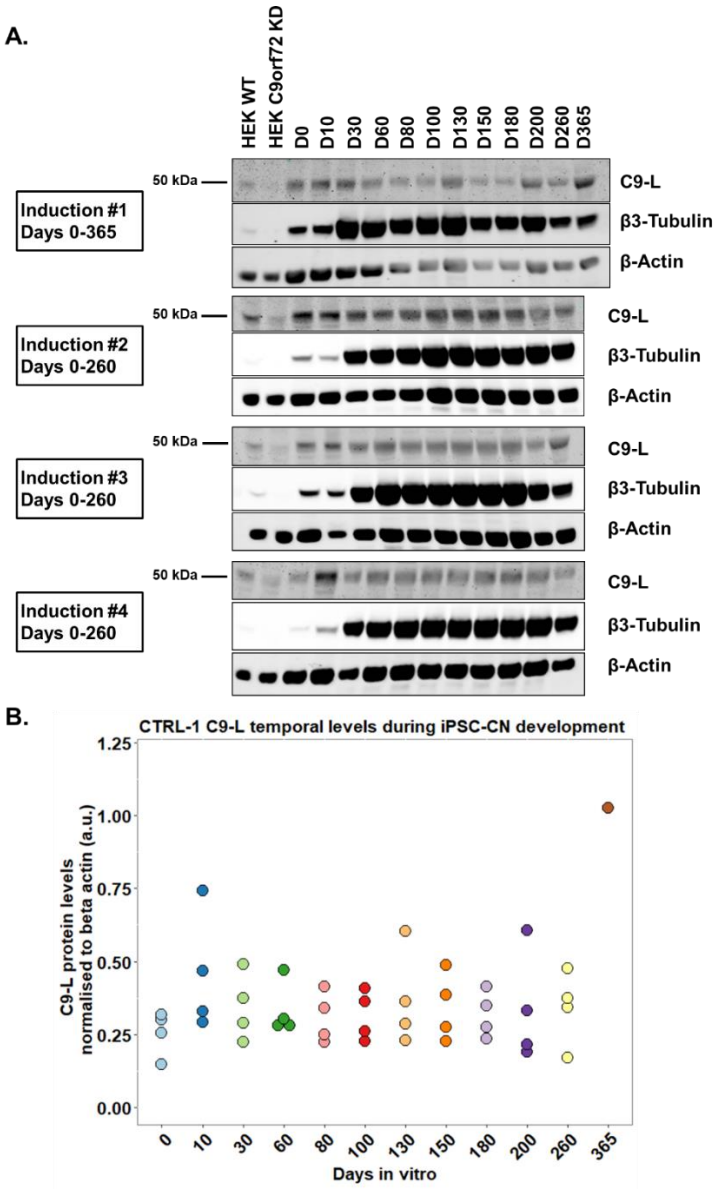


Figure 4.13. C9-L is stably expressed throughout *in vitro* cortical development.
A. Western blot analysis of C9orf72 protein levels was performed in neuronal lysates collected at different timepoints during *in vitro* cortical neuron differentiation of the control line CTRL-1. The timepoints ranged from 0 days (iPSC stage) to extended timepoints of 260 days (n=4) and 365 days (n=1) *in vitro*. C9-L was detected at 50 kDa. Wild-type and *C9orf72* knock down HEK lysates were included in each blot as positive and negative controls respectively, for the detection of C9orf72 protein. All blots were reprobed with neuronal β3 tubulin and β-Actin housekeeping proteins. Each blot represents one independent induction (n=4). B. Quantification of western blots shown in (A) by densitometry revealed that C9-L is stably expressed at similar levels across all timepoints of *in vitro* cortical development assessed (n=4 independent inductions, normalised to β-actin).

4.2.9. C9orf72 protein levels in control and *C9orf72* iPSC-CNs

We next assessed C9orf72 protein levels in control and *C9orf72* iPSC-CNs at selected timepoints during cortical differentiation, to determine if there is haploinsufficiency of C9orf72 protein in patient cells. The C9orf72 protein analysis was performed on iPSC-CNs at extended timepoints of 150 and 260 DIV. The rationale for choosing these extended timepoints for C9orf72 protein analysis is that hypermethylation was established at 150 DIV in C9-3 patient line. Therefore the analysis at extended timepoints is more likely to allow the detection of any potential reduction in C9orf72 protein.

C9orf72 expression analysis at 150 DIV iPSC-CNs from patients and controls is shown in Figure 4.14. The C9-L isoform was detected at 50 kDa in both control and *C9orf72* 150 DIV iPSC-CNs by western blot using the C9orf72 rabbit polyclonal antibody 22637-1-AP (Proteintech). As revealed by densitrometric analysis, no evidence of haploinsufficiency in C9-L levels was observed in C9orf72 iPSC-CNs at 150 DIV (n=4 independent inductions, normalised to β3-tubulin). Even though this C9orf72 antibody has been raised against a 1-169 aa C9orf72 peptide sequence and can potentially recognise both isoforms, no C9-S isoform could be detected in iPSC-CN lysates.

The levels of C9-L were also assessed in control and *C9orf72* 260 DIV iPSC-CNs by western blot. Figure 4.15 shows the three blots corresponding to three independent CN inductions. Similar to what was observed for *C9orf72* 150 DIV iPSC-CNs, *C9orf72* 260 DIV iPSC-CNs did not display reduced C9orf72 protein levels compared to controls.

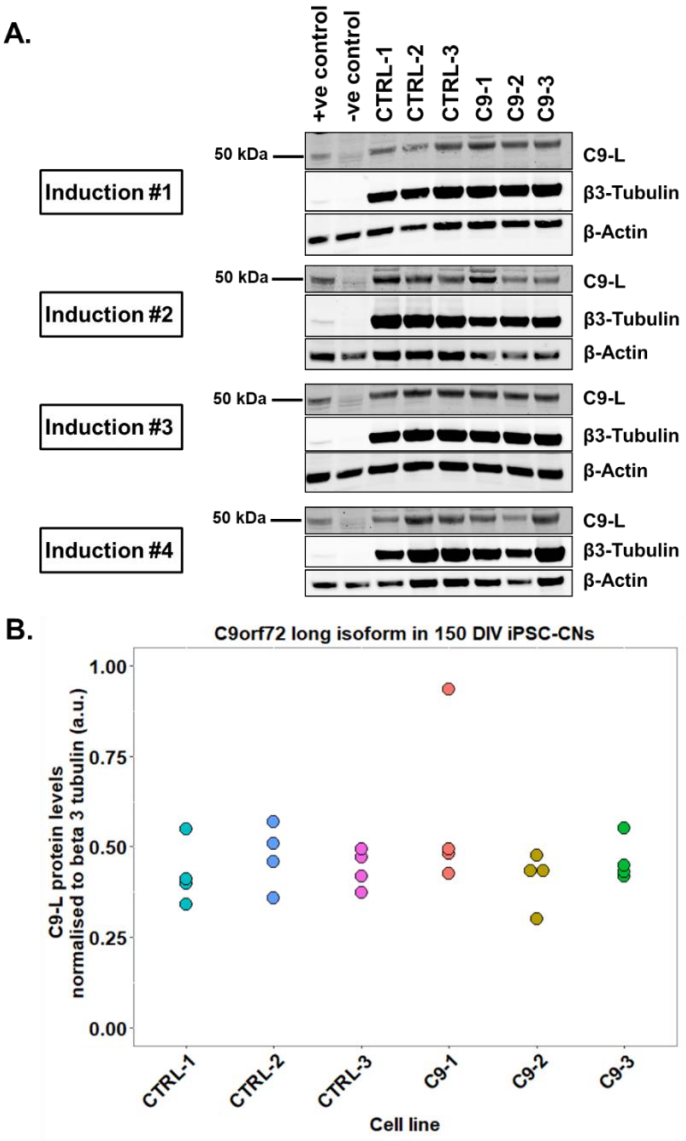


Figure 4.14 No evidence of haploinsufficiency in *C9orf72* 150 DIV iPSC-CNs compared to controls.

A. C9orf72 protein levels were assessed by western blot in 150 DIV iPSC-CNs from control lines (CTRL-1, CTRL-2 and CTRL-3) and C9orf72 HRE lines (C9-1, C9-2 and C9-3. Wild-type and *C9orf72* knock down HEK293 lysates were included in each gel as positive and negative controls respectively, for the detection of C9orf72 protein. The long isoform of C9orf72 (C9-L) was detected at 50 kDa, but not the short (C9-S, ~24kDa). All blots were reprobed with neuronal β3 tubulin and β-Actin housekeeping proteins. Each blot represents one independent induction (n=4). B. Densitometric quantification of western blot data shown in A revealed no difference in C9-L levels between C9orf72 and control 150 DIV iPSC-CNs (n=4 independent inductions, normalised to β3 tubulin).

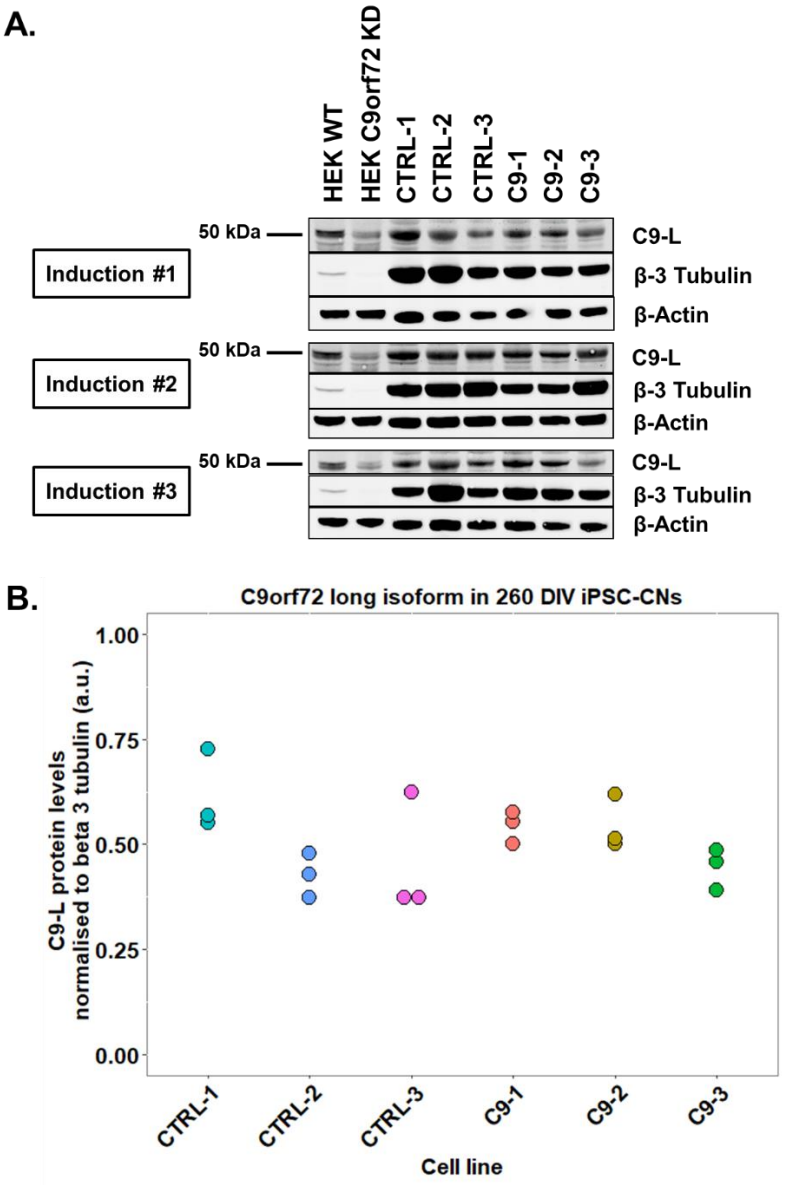


Figure 4.15. No evidence of haploinsufficiency in *C9orf72* 260 DIV iPSC-CNs compared to controls.

A. C9orf72 protein levels were assessed by western blot in 260 DIV iPSC-CNs from control lines (CTRL-1, CTRL-2 and CTRL-3) and C9orf72 HRE lines (C9-1, C9-2 and C9-3. Wild-type and *C9orf72* knock down HEK293 lysates were included in each gel as positive and negative controls respectively, for the detection of C9orf72 protein. C9-L was detected at 50 kDa. All blots were reprobed with neuronal β3 Tubulin and β-Actin housekeeping proteins. Each blot represents one independent induction (n=3). B. Densitometric quantification of western blot data shown in A revealed no difference in C9-L levels between C9orf72 and control 260 DIV iPSC-CNs (n=4 independent inductions, normalised to β3 tubulin).

4.2.10. Relationships between *C9orf72* promoter methylation, RNA foci, total *C9orf72* mRNA and HRE length

Studies have shown an association of 5' CpG hypermethylation in *C9orf72* promoter with increased HRE length (Gijselinck *et al.*, 2016) and reduced *C9orf72* transcription (Belzil *et al.*, 2013, Xi *et al.*, 2013, Liu *et al.*, 2014, Gijselinck *et al.*, 2016) as well as with lower RNA foci- and DPR-related toxicity, suggesting that hypermethylation protects against toxicity in C9-FTD/ALS (Liu *et al.*, 2014, Russ *et al.*, 2014, Bauer, 2016). Some of the above associations have also been observed in the present study, however, with the limitation of the very small sample size. Despite this caviat, Pearson correlation analysis has revealed an inverse correlation between 5' CpG methylation and percentage of RNA foci (r=-0.9699) and 5' CpG methylation and *C9orf72* total mRNA expression (r=-0.9966) (Figure 4.16). Moreover, 5' CpG methylation was correlated with *C9orf72* HRE size (r=0.9746) (Figure 4.16).

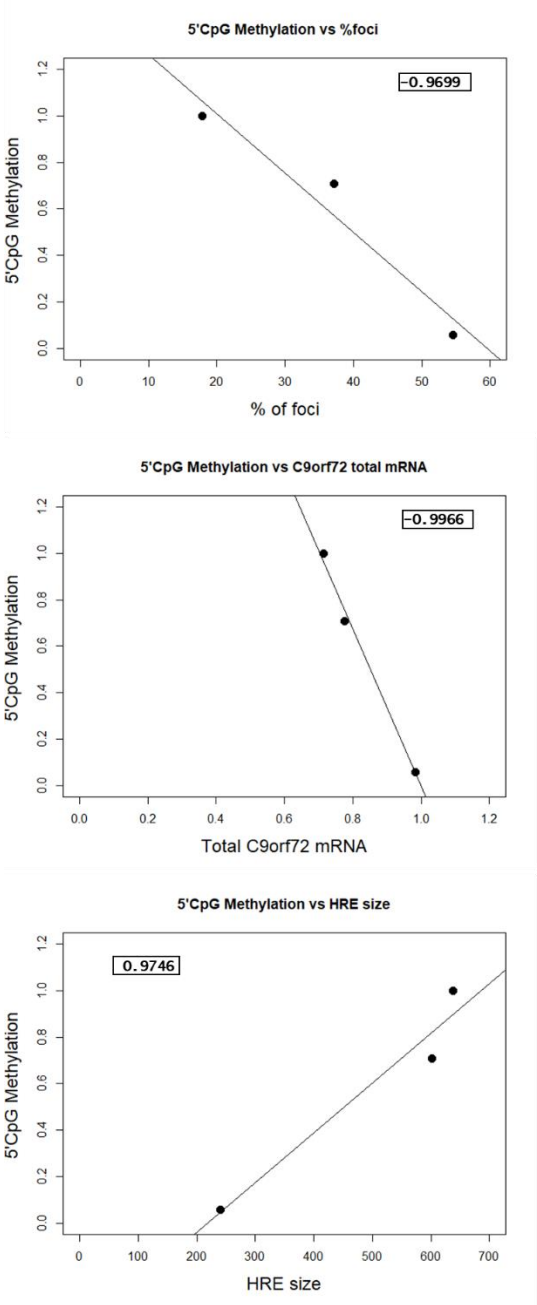


Figure 4.16 5' CpG methylation is inversely correlated with percentage of RNA foci and *C9orf72* total mRNA levels and positively correlated with the *C9orf72* HRE size. Pearson correlation analysis has revealed an inverse correlation between 5' CpG methylation and percentage of RNA foci (r=-0.9699) and 5' CpG methylation and *C9orf72* total mRNA expression (r=-0.9966). Moreover, 5' CpG methylation was correlated with *C9orf72* HRE size (r=0.9746).

4.3. Discussion

The work presented in this chapter shows the investigation of *C9orf72* haploinsufficiency iPSC-CNs derived from *C9orf72* HRE carriers, which was carried out as a comprehensive analysis of (i) DNA methylation in *C9orf72* promoter, (ii) *C9orf72* mRNA expression and (iiii) C9orf72 protein levels.

We first of all investigated methylation of the *C9orf72* HRE and promoter region. The *C9orf72* promoter was found to be hypermethylated in *C9orf72* patient-derived CNs, however with considerable variability between the patient lines. In two *C9orf72* HRE carrier lines, C9-2 and C9-3, the HRE was methylated in all samples throughout *in vitro* CN development, whereas in another line, C9-1, was characterised by HRE methylation instability during *in vitro* cortical development. Due to the qualitative nature of the $(G_4C_2)_n$-methylation detection assay, the levels of *C9orf72* HRE methylation accurate cannot be determined quantitatively. In another study, HRE methylation was associated with large expansions and was undetected in intermediate ones (Xiao *et al.*, 2015). This is in accordance with the observations of the present study, suggesting the existence of a possible HRE methylation threshold starting at intermediate repeat sizes than the pathogenicity threshold of ~30 repeats. This study contained a low number of HRE carrier lines, so further studies with more lines or ideally, isogenic lines with different, defined HRE sizes, would permit a more complete investigation of the link between HRE length and DNA methylation.

Similar to the HRE, the 5'CpG island was found to be hypermethylated in *C9orf72* iPSC-CNs of C9-2 and C9-3 lines, but unmethylated in iPSC-CNs of C9-1. This observed correlation between HRE methylation and 5' CpG island hypermethylation is further supported by other studies (Xiao *et al.*, 2015, Cohen-Hadad *et al.*, 2016). In fact, it has been suggested that HRE methylation can spread to the 5′-CpG island region but not the other way around (Xiao *et al.*, 2015). This is also supported by the

findings of the present study, where *C9orf72* promoter hypermethylation is HRE-specific and observed in *C9orf72* HRE carrier lines that exhibit high repeat lengths. In the only *C9orf72* HRE line which did not exhibit promoter hypermethylation, C9-1, HRE methylation may be detected, however, it fails to spread to the proximal 5' CpG because it apparently fails to establish at the HRE first, due to methylation instability. This highlights the importance of temporal assessment of DNA methylation in the present study. Finally, consistently with other reports, the 3' CpG island was not hypermethylated in *C9orf72* HRE lines at any of the timepoints assessed (Xi *et al.*, 2013).

The temporal assessment of methylation in parental fibroblasts, iPSCs and different stages during cortical differentiation has proven crucial for gaining temporal information on when DNA methylation is established in *C9orf72* HRE lines. Interestingly, *C9orf72* promoter hypermethylation was present in iPSCs of patient lines C9-2 and C9-3. Moreover, in the case of C9-2 patient line, HRE and 5' CpG were unmethylated in fibroblasts, but following reprogramming the resulting iPSCs displayed a methylated HRE and hypermethylated 5'CpG. Reprogramming has been known to act as an eraser of global epigenetic signatures in the genome, however, low passage iPSCs have been shown to retain some epigenetic memory from the parental cells (Kim *et al.*, 2010c, Leitch *et al.*, 2013). The hypermethylation in *C9orf72* iPSCs is suggestive of *de novo* methyltransferase activity setting the DNA methylation pattern early during reprogramming to iPSCs. One hypothesis is that the HRE triggers hypermethylation in *C9orf72* promoter of iPSCs by acting as a methylation hotspot. This has also been suggested by another study of *C9orf72* hypermethylation in iPSCs (Cohen-Hadad *et al.*, 2016). Interestingly, that study also showed that in contrast to *C9orf72* HRE iPSCs, the HRE and 5' CpG island of a *C9orf72* ESC line were unmethylated. The authors concluded this difference in DNA methylation between iPSCs and ESCs is down to the reprogramming process. *C9orf72* promoter

methylation has not been described in *C9orf72* HRE human ESCs by other studies and it is indeed possible that DNA methylation is absent from ESCs derived from a pre-implantation embryo compared to reprogrammed iPSCs. However, despite the two lines being haploidentical (iPSCs were generated from mother of embryo), a major difference between the two was the different HRE size of the *C9orf72* allele. The ESC line exhibited a much smaller HRE of ~270 repeats compared to the ~700 repeat HRE of the mother's iPSCs. Moreover, these HRE sizes are similar to the sizes used in the present study, which showed that iPSCs of ~240 repeats (C9-1) do not exhibit hypermethylation compared to larger HRE-bearing lines ~638/602 (C9-2 and C9-3) which are hypermethylated. Thus, although reprogramming may be permissive towards hypermethylation in iPSCs compared to *bona fide* ESCs due to epigenetic memory, a minimum HRE size threshold may also be required to observe hypermethylation in iPSCs, but larger studies are required to address this hypothesis. Moreover, the present study was limited by the use of three C9orf72 lines and one clone per line, so no conclusions can be made from the above observations.

In contrast to our findings, another study using one *C9orf72* HRE patient line reported 5' CpG demethylation in patient lymphoblasts during reprogramming to iPSCs and re-aquisition of methylation upon neural induction (Esanov *et al.*, 2016). The authors concluded that demethylation in iPSCs is due to global DNA methylation reduction during reprogramming and that neuronal induction promotes an increase in methylation. Arguably, the demethylation during reprogramming to iPSCs in this study is more likely to be reflective of the HRE instability, as southern blot analysis showed an unstable repeat in lymphoblasts which contracted in iPSCs, possibly below methylation threshold, and expanded again in differentiated iPSC-MNs.

Finally, further insight into methylation mechanisms can be gained from research on other GC-rich repeat expansion diseases such as Fragile X and Friedrich's ataxia where hypermethylation is a common feature (Castaldo *et al.*, 2008, Urbach *et al.*,

2010). It is known from Fragile X iPSC models that the methylated CGG repeat expansions in *FMR1* gene lead to gene silencing and they are retained during reprogramming to iPSCs. Moreover, even in rare cases of unmethylated patient somatic cells, reprogramming has been shown to generate iPSCs with methylated and silenced *FMR1,* leading to the notion that the generation of unmethylated FMR1 mutant iPSC is not possible (de Esch *et al.*, 2014). Interestingly, in a report of unmethylated full *FMR1* mutation siblings without gene silencing and disease symptoms, the authors showed that hypomethylation could be maintained in iPSCs after reprogramming (Brykczynska *et al.*, 2016). Importantly, a subset of iPSC clones with larger CGG expansions was found to carry a silenced *FMR1* gene, suggesting *de novo* methylation occurred at *FMR1* upon expansion of the CGG repeat size. Therefore the authors supported the existence of a repeat threshold required for methylation and silencing at *FMR1* that may vary between individuals. This suggested model from the Fragile X research field is similar to the hypothesis supported here for the existence of an HRE threshold for methylation at *C9orf72*.

Apart from *C9orf72* DNA methylation, studies have assessed additional epigenetic marks such as hydroxymethylation and histone modifications. In the study by Esanov *et al*., *C9orf72* promoter hypermethylation (5-methylcytosine, 5mC) was reduced upon reprogramming whilst hydroxymethylation (5-hydroxymethylcytosine, 5hmC) was increased (Esanov *et al.*, 2016). Moreover, trimethylation of lysine residues within histones H3 and H4 has also been detected in brain tissue and blood from C9orf72 HRE carriers and was associated with reduced *C9orf72* mRNA expression (Belzil *et al.*, 2013).

A frequent observation in *C9orf72* repeat expansion carriers is the reduction of *C9orf72* mRNA, mainly variant 2 transcript which contains the HRE in the promoter region (DeJesus-Hernandez *et al.*, 2011, Gijselinck *et al.*, 2012, van Blitterswijk *et al.*, 2015). This reduction in *C9orf72* expression has also been recapitulated in some

iPSC-derived neuronal models of FTD/ALS (Almeida *et al.*, 2013, Donnelly *et al.*, 2013, Shi *et al.*, 2018) whilst others have shown no differences in expression (Sareen *et al.*, 2013). In the present study, total *C9orf72* mRNA was significantly reduced in 150 DIV iPSC-CNs of *C9orf72* HRE carrier lines C9-2 and C9-3, whereas for C9-1, the expression was similar to controls. Interestingly, the *C9orf72* lines C9-2 and C9-3 were the ones that also exhibited *C9orf72* promoter hypermethylation, in contrast to C9-1 which had an unmethylated 5' CpG. These data reveal an inverse correlation between *C9orf72* total mRNA expression and promoter hypermethylation. Moreover, they support the hypothesis that hypermethylation acts as a disease modifier by reducing the production of RNA foci, as the percentage of RNA foci was reduced in C9-2 and C9-3 compared to C9-1 (see 3.3.6). *C9orf72* V2 transcript was not found to be significantly reduced due to the presence of some outlier values, however, *C9orf72* V3 was significantly upregulated in C9-1 compared to control and other *C9orf72* 150 DIV iPSC-CNs, in line with previous reports (Cohen-Hadad *et al.*, 2016).

Analysis of C9orf72 protein levels in *in vitro* models, and in particular iPSC-derived neurons, has been one of the least studied aspects of *C9orf72* haploinsufficiency. Moreover, considering the recently discovered role of *C9orf72* in FTD/ALS, little is known about normal C9orf72 protein levels during cortical development. This has been attributed to the lack of well-characterised and specific to C9orf72 protein commercial antibodies. For some anti-C9orf72 antibodies, cross-reactivity with glial fibrillary acidic protein (GFAP) has been reported due to a high homology epitope sequence between the two proteins (Satoh *et al.*, 2012, Waite *et al.*, 2014). Finally, according to a recent C9orf72 antibody validation study, highlighting the issues of C9orf72 antibodies in the field of FTD/ALS research, an anti-C9orf72 antibody (SC-138763) that failed to recognise C9orf72 protein in any human sample in any application has been used in 15 published studies which have been subsequently cited thousands of times (Laflamme *et al.*, 2019).

In the present study, a characterisation of several commercial anti-C9orf72 antibodies was carried out. Moreover, antibody characterisation for immunoblot was carried out on stable knockdown *C9orf72* HEK293 cell lines and CRISPR/Cas9 *C9orf72* knockout iPSC lines. The rabbit polyclonal anti-C9orf72 from Proteintech (22637-1-AP) was found to be specific against the long protein isoform of C9orf72 (C9-L). However, C9-S isoform could not be detected in iPSC-CNs using the antibody. C9-L levels were investigated across different timepoints throughout cortical differentiation in a control line (CTRL-1) and found to be stable during *in vitro* corticogenesis. This is the first description of C9orf72 protein levels characterisation during CN differentiation from iPSCs to extended differentiation timepoints of up to 365 DIV which are not conventionally assessed in published studies. C9orf72 total mRNA levels were reduced in two of the *C9orf72* HRE lines, pointing towards a potential reduction in C9orf72 protein too. However, analysis of C9-L levels in iPSC-CN of *C9orf72* HRE carriers compared to controls revealed no differences between the two groups. The levels of C9orf72 protein were analysed at both 150 DIV and 260 DIV mature CNs. Sareen *et al.* also found no differences in C9orf72 protein levels between *C9orf72* iPSC-MNs and controls, using a Genetex GTX119776 anti-C9orf72 antibody (Sareen *et al.*, 2013). However, Shi *et al*. reported haploinsufficiency of C9orf72 in induced motor neurons (iMNs) of *C9orf72* HRE carriers compared to controls (Shi *et al.*, 2018). Specifically they found a reduction in C9-L protein isoform, detected using anti-C9orf72 Proteintech 22637-1-AP, the same antibody used in the present study. Discrepancies in the detection of reduced protein between studies may be reflective of the phenotypic heterogeneity between *C9orf72* HRE patients in FTD/ALS, or it could be attributed to the maturity of neuronal cultures and the neuronal differentiation protocols employed. Indeed it is well demonstrated that iN models, such as the iMNs used in Shi *et al.* study, retain ageing signatures that offer better recapitulation of age-associated phenotypes. Therefore, despite the use of 150 and 260 DIV iPSC-CNs,

these cultures are still closer to fetal and changes in protein levels may be detected later on.

To summarise, *C9orf72* iPSC-CNs are useful tools for the investigation of haploinsufficiency phenotypes in FTD/ALS. Moreover, as shown here, iPSC-derived models are particular useful in the investigation of DNA methylation of *C9orf72* promoter. The next chapter will focus on loss of C9orf72 function, complementing the work presented in this chapter.

Chapter 5 Generation and characterisation of *C9orf72* knockout iPSCs

5.1. Introduction

Loss of C9orf72 protein function as a result of the HRE in *C9orf72* has been described as one potential disease mechanism in C9-FTD/ALS, based on the finding that *C9orf72* RNA and protein levels are reduced in brain tissue and cells from *C9orf72* HRE carriers (DeJesus-Hernandez *et al.*, 2011, Renton *et al.*, 2011, Gijselinck *et al.*, 2012, Almeida *et al.*, 2013, Belzil *et al.*, 2013, Donnelly *et al.*, 2013, Xi *et al.*, 2013, Liu *et al.*, 2014, Waite *et al.*, 2014). However, the description of a patient homozygous for the *C9orf72* HRE with clinicopathological presentation of similar severity to that of heterozygotes for the HRE, argued against a major role of *C9orf72* haploinsufficiency in the pathogenesis of C9-FTD/ALS (Fratta *et al.*, 2013). Nonetheless, it is still important to understand the function of C9orf72 and consequences of loss of function on neuronal viability.

Loss of *C9orf72* orthologs in zebrafish and *C. elegans* resulted in neurodegenerative phenotypes in these *in vivo* models (Ciura *et al.*, 2013, Therrien *et al.*, 2013) and *C9orf72* haploinsufficiency triggered neurodegenerative phenotypes in human *C9orf72* iPSC-MNs (Shi *et al.*, 2018). On the contrary, genetic ablation of *C9orf72* in mice did not lead to overt neurodegenerative phenotypes (Koppers *et al.*, 2015, Burberry *et al.*, 2016, Jiang *et al.*, 2016, O'Rourke *et al.*, 2016), instead it revealed an important role of C9orf72 in myeloid cell function and the regulation of inflammatory and autoimmunity responses to innate immune activation (Burberry *et al.*, 2016, Jiang *et al.*, 2016, O'Rourke *et al.*, 2016, Burberry *et al.*, 2020, McCauley *et al.*, 2020). More recently, the study of combined LOF and GOF iPSC-MNs and mouse models has shown that loss of C9orf72 function exacerbates the neurodegenerative phenotypes of the HRE-mediated toxicity, therefore supporting a synergistic model of GOF and LOF mechanisms in C9-FTD/ALS (Shi *et al.*, 2018, Shao *et al.*, 2019, Abo-Rady *et al.*, 2020, Zhu *et al.*, 2020).

Understanding the normal function of C9orf72 is crucial for defining its precise role in neurodegeneration. Many recent studies have shed light into a role of C9orf72 protein in regulation of autophagy and endosomal trafficking (Amick *et al.*, 2016, Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Webster *et al.*, 2016, Shi *et al.*, 2018). C9orf72 forms a complex with SMCR8 that regulates autophagy via interaction with the ULK1 complex which controls autophagy initiation (Sellier *et al.*, 2016, Sullivan *et al.*, 2016, Webster *et al.*, 2016, Yang *et al.*, 2016).Moreover, the C9orf72 protein was localised in early endosomes and required for normal vesicle trafficking and lysosomal biogenesis in iPSC-MNs (Shi *et al.*, 2018).

The study of *C9orf72* iPSCNs offers a unique tool to study early disease mechanisms. However, these models do not allow the dissection of loss versus gain of function mechanisms, as all three disease mechanisms co-exist in HRE carrier iPSCs. Therefore the generation of *C9orf72* knockout iPSC lines is essential for investigation of the loss of function disease mechanisms. This Chapter describes the generation and characterisation of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines and iPSC-CNs as useful tools in the study of C9orf72 protein function and its role in neurodegeneration. Importantly, no studies have investigated *C9orf72* knockout in iPSC-CNs. The study of these models can elucidate pathways affected in this specific neuronal type as opposed to described *C9orf72* knockout iPSC-MNs. Moreover, the expression of autophagy genes is assessed in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs to identify alterations in autophagy. Finally, due to the role of C9orf72 in autophagy and a recently discovered genetic interaction with mitochondrial fission 1 (FIS1) (Chai *et al.*, 2019), the selective autophagy of mitochondia (mitophagy) pathway is also investigated in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs.

5.2. Aims

The specific aims of this chapter were:

- To generate and characterise *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines.
- To differentiate and characterise *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ CNs.
- To determine whether autophagy and mitophagy are impaired in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ CNs.

5.3. Results

5.3.1. Generation and characterisation of TALEN and CRISPR-Cas9 nucleases targeting *C9orf72* exon 2

For the generation of *C9orf72* heterozygous and homozygous knockout iPSC lines, both TAL Effector Nuclease (TALEN) and CRISPR/Cas9 nickase technologies were employed. The gene knockout is mediated by NHEJ which is triggered by nuclease-induced DSBs at the targeted DNA region of an essential gene exon. Using online software tools (as described in 2.7.), two TALEN nuclease and two CRISPR/Cas9 nickase pairs targeting exon 2 of *C9orf72* gene were designed. Exon 2 was targeted as it is the most 5' exon present in all *C9orf72* splice variants (Figure 5.1 A). The genomic DNA sequences in *C9orf72* exon 2 targeted by all designer nucleases are shown in Figure 5.1 B.

Following assembly of the constructs, the nuclease pairs were functionally assessed via gold standard T7 endonuclease I mutation detection assay for their ability to cleave their gDNA targets in HEK293 cells (as described in 2.12). Briefly, gDNA cleavage is detected by incubation of the amplified, target site-containing DNA heteroduplexes with T7 endonuclease, which detects and cleaves the mismatches between wild-type and mutant strands, resulting in smaller cleavage products when resolved on an agarose gel. Figure 5.2 shows the results of the functional testing and the calculated modification efficiency of each nuclease pair. Overall, all generated nucleases were able to cleave their gDNA targets in *C9orf72* exon 2. In both gene editing technologies employed in this project, the nucleases are obligatory dimers and can only cleave gDNA in pairs.

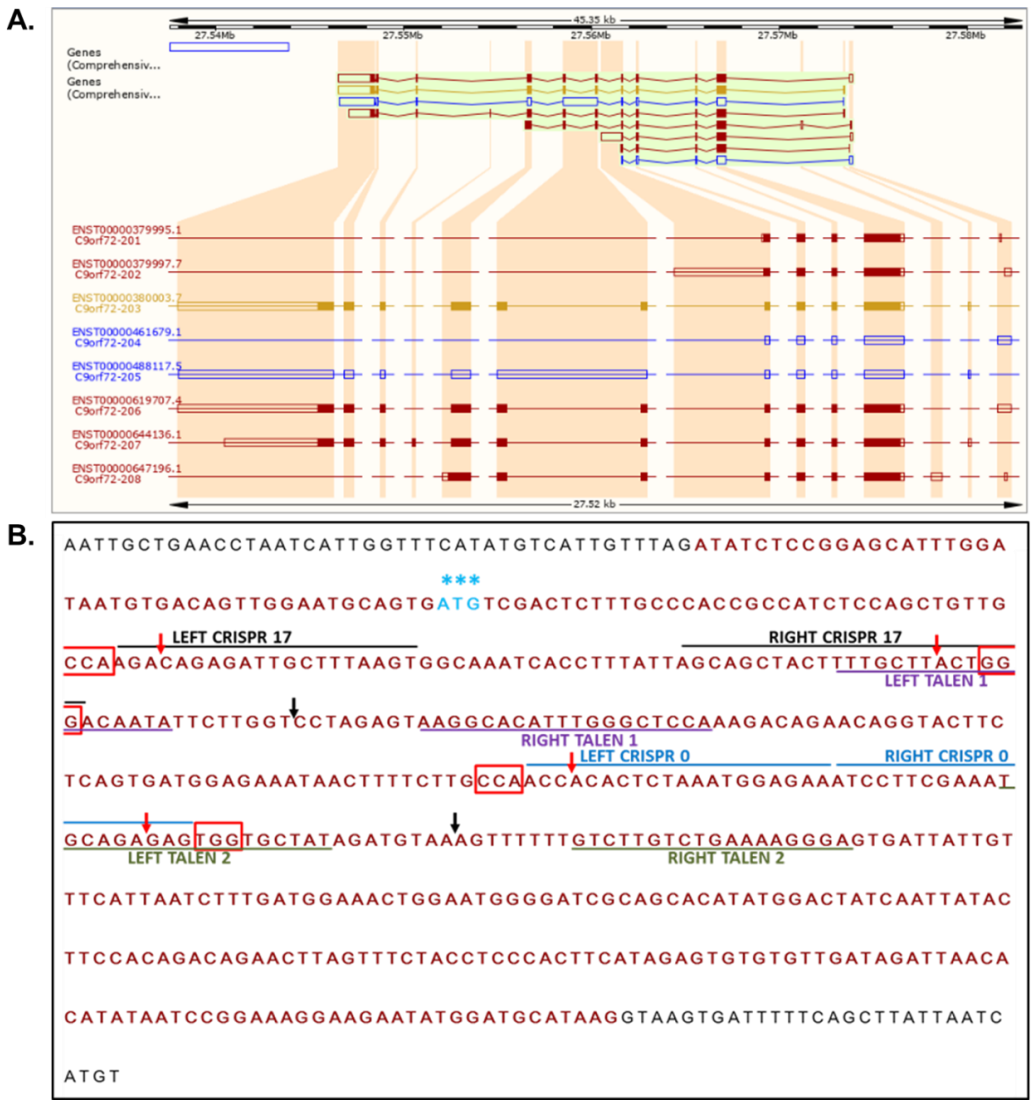


Figure 5.1 CRISPR/Cas9n and TALEN nucleases DNA binding sequences on *C9orf72* exon 2.

A. All *C9orf72* splice variant transcripts share a common exon 2 which also includes the transcription start codon (Ensembl GRCh38.p12, March 2019). B. Genomic DNA sequence of *C9orf72* exon 2 with annotated the nuclease recognition sites of CRISPR/Cas9n pair 0, CRISPR/Cas9n pair 17, TALEN pair 1 and TALEN pair 2. For CRISPR/Cas9n pairs, the PAM NGG motifs of the sgRNAs are shown in red squares and the DNA nicking sites are indicated by red arrows 3bp upstream of the PAM motif. The spacers for CRISPR/Cas9n pair 0 and CRISPR/Cas9n pair 17 are 0 bp and 17 bp respectively. For TALENs DSBs occur in the centre of the spacer sequences as indicated by black arrows. The transcription start codon ATG is indicated with blue asterisks. The true positions of each nuclease pair lie on opposite strands but here the genomic sequence of only one DNA strand is shown for simplicity.

Indeed, in the T7 endonuclease I assay results for both TALEN and CRISPR/Cas9 nucleases no gDNA cleavage was detected in HEK293 cells transfected with either left or right nuclease alone (Figure 5.2 A, C). Moreover, TALEN 1 and TALEN 2 pairs were able to induce DSBs with 12-17% and 6-7.6% efficiencies respectively (Figure 5.2 A, B). T7 endonuclease I assay results are assay-specific, therefore TALEN pairs from independent assays (Figure 5.2 A) were compared in same assay (Figure 5.2 B) and the TALEN pair 1 with highest efficiency was tested again in a third assay alongside the CRISPR/Cas9 nickase pairs (Figure 5.2 C). In the final assay, the calculated modification efficiencies for the CRISPR/Cas9 pair 0, CRISPR/Cas9 pair 17 and TALEN pair 1 nuclease pairs were 26.3-27.8%, 25.3-26.2% and 36% respectively (Figure 5.2 C). TALEN pair 1 showed the highest modification efficiency in this assay, however, GFP CRISPR/Cas9 pair 0 was used for downstream genome editing in iPSCs. The rationale for selecting the GFP CRISPR/Cas9 plasmid for the iPSC editing experiments instead of the TALEN pair was that: (i) it would allow monitoring of nucleofection efficiencies via GFP fluorescence and 96-well single iPSC GFP-FACS sorting and (ii) it had been otpimised for use in human stem cells.

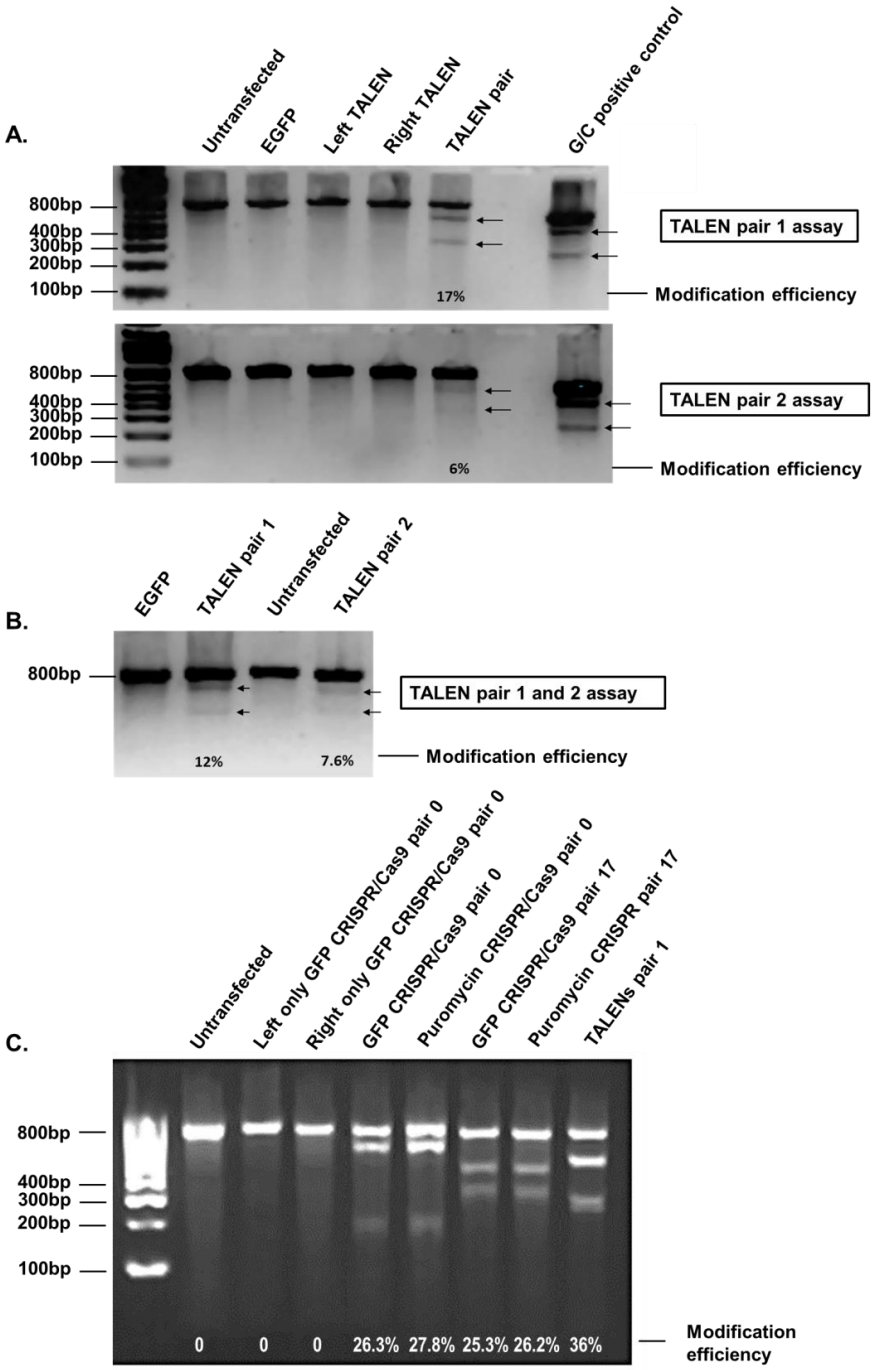


Figure 5.2 All generated nucleases successfully generate double strand breaks in *C9orf72* exon 2.

A. Functional testing of the generated TALEN pairs targeting *C9orf72* exon 2 via two independent T7 endonuclease I assays. The two TALEN pairs were both functional and able to induce DSB at the intended gDNA targets in exon 2 as indicated by the presence of the T7 endonuclease cleaved products below the main bands in the agarose gels. Transfection of HEK293 cells with either left or right TALEN only did not result in any DNA cleavage. The G/C positive control for the T7 endonuclease I assay was an amplicon amplified from a known sequence plasmid containing a G/C

(Continued Figure legend) mismatch. B. Direct comparison of TALEN pair 1 and pair 2 within the same T7 endonuclease I assay. The efficiencies were calculated as 12% for TALEN pair 1 and 7.6% for TALEN pair 2. C. Functional testing of CRISPR/Cas9n pairs targeting *C9orf72* exon 2 via T7 endonuclease I assay. CRISPR/Cas9n pairs were transfected to HEK293 cells and tested via T7 endonuclease I mutation detection assay. Transfections with either left or right CRISPR/Cas9n did not induce any detectable DSBs. TALEN pair 1 was included in the assay as a positive control.

5.3.2. Optimisation of the protocol for the generation of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines

5.3.2.1. Choice of Cas9 vector and sgRNAs

The Cas9 nickase was used for generating the *C9orf72* knockout iPSC cell lines as opposed to the wild-type Cas9 enzyme due to the reduced frequency of genomic off-target effects associated with the paired nickase strategy (Ran *et al.*, 2013a). Specifically the CRISPR/Cas9n GFP vector (pSpCas9n(BB)-2A-GFP) was chosen because, unlike pSpCas9n(BB)-2A-Puro or TALEN vectors which rely on clonal selection via resistance to either puromycin or hygromycin respectively, it allows for FACS sorting of nucleofected GFP+ iPSCs, monitoring of the nucleofection efficiency and enables single iPSC FACS sorting on 96 well plates for clonal expansion. The two generated sgRNA pairs, pair 0 and pair 17, were selected due to the low number of potential genomic off-targets and both pairs displayed similar modification efficiencies (Figure 5.2). Pair 0 was chosen for downstream experiments in iPSCs.

5.3.2.2. *Optimal conditions for single cell survival*

Initially, a direct comparison of three different defined pluripotent cell maintenance media for their ability to promote single cell survival was performed. CTRL-1 iPSCs were routinely maintained in Essential 8 and were also adapted to mTeSR1 and TeSR2 culture media, all under feeder-free conditions (2.13). The morphology of iPSCs was similar in all three defined media as shown in Figure 5.3 A. Following nucleofections of iPSCs with the right pair 0 pSpCas9n(BB)-2A-GFP vectors (2.14) and FACS sorting into 96-well plates containing the respective growth medium (Essential 8, mTeSR1 or TeSR2) supplemented by 10 μM Y-27632, plates were inspected daily for colony detection. Overall, as shown in Figure 5.3 B, survival of single iPSCs was generally low across all tested media conditions, however, the best survival was observed under mTeSR1 medium growth conditions (n=1). Therefore, CTLR-1 iPSCs were adapted to mTeSR1 for the generation of *C9orf72* edited lines.

Single cell survival was further enhanced by the addition of iPSC-conditioned mTeSR1 medium in a 1:1 ratio mTeSR1:conditioned mTeSR1 in single cells post FACS sorting in the presence of 10 μM Y-27632. However, overall the number of single iPSC-derived colonies on 96-well plates was very low to allow screening for *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines. Therefore iPSCs were plated post FACS-sorting at 5,000 cells per 10 cm dish seeding density. Emerging well-defined single iPSC-derived colonies were then manually picked and transferred to 96-well plates.

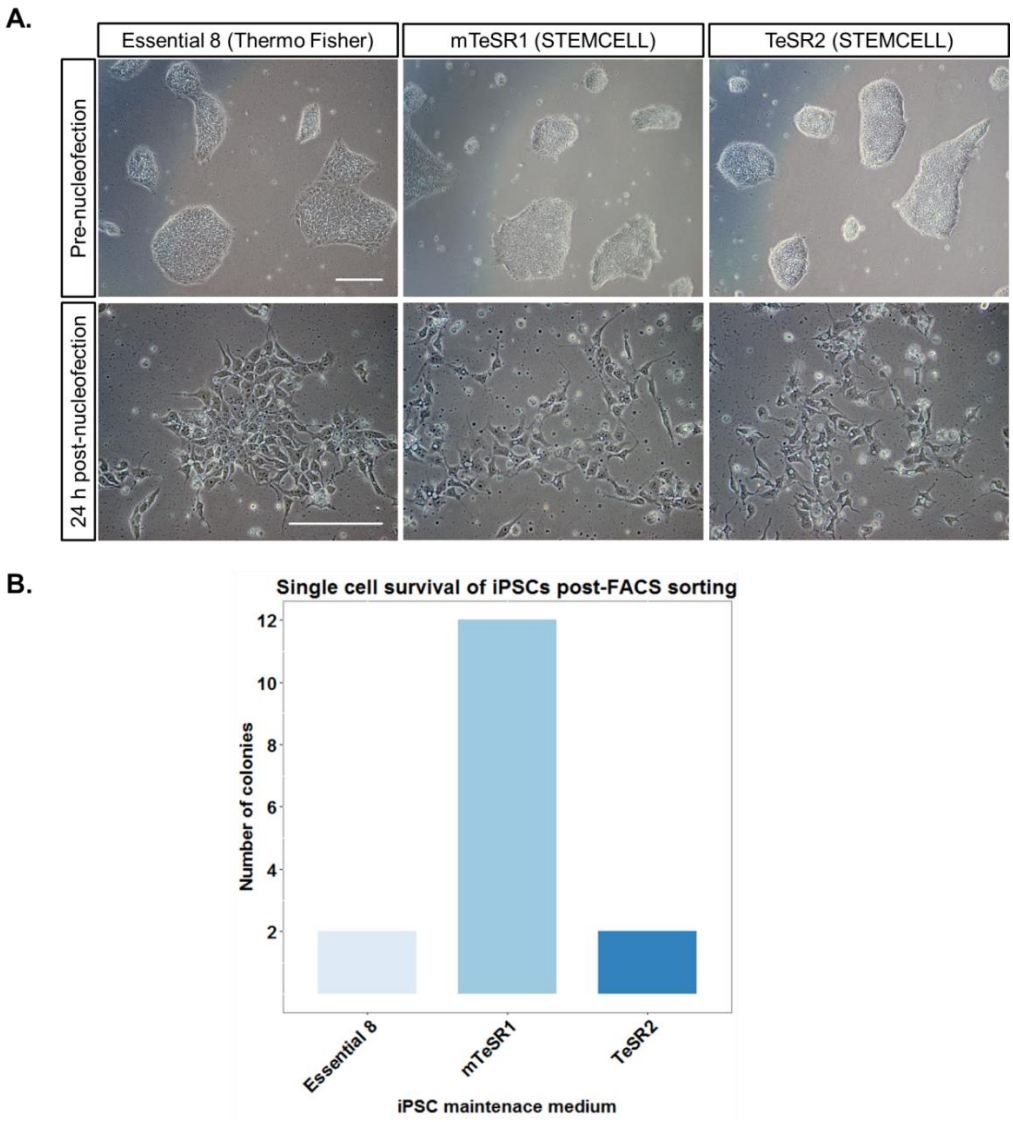


Figure 5.3 Preliminary data suggest mTeSR1 maintenance media offers better survival of single iPSCs post-sorting.

A. CTRL-1 iPSCs were maintained in Essential 8 media and for the purpose of this experiment they were adapted to mTeSR1 and TeSR2 maintenance media. The iPSC colony morphology was similar in all three pluripotent maintenance media. Moreover, 24 h post-nucleofection, the iPSC survival in all three media in the presence of 10μM Y-27632 was similar, with marginally higher cell death in the mTeSR1 and TeSR2 media, probably due to adaptation stress. Scale bars 200μm. B. GFP-positive single iPSCs from all three culture conditions were FACS-sorted into 96-well plates and grown in the respective media in the presence of 10μM Y-27632. The total number of surviving single-cell-derived colonies was low across all media conditions, but mTeSR1 resulted in best survival compared to Essential 8 and TeSR2 (n=1).

5.3.2.3. *Nucleofection*

Using the 4D Lonza nucleofection system there is limited flexibility in testing additional conditions for further protocol optimisation as all programs are pre-set and the buffers are provided by the manufacturer (all proprietary). Therefore, nucleofections were performed using the P3 Primary Cell Kit (Lonza) and the program for H9 Stem cells using CB-150 pulse (2.14). The use of the GFP CRISPR/Cas9n vector allowed the FACS sorting of the nucleofected cells and therefore the calculation of the nucleofection efficiencies. Figure 5.4 shows representative fluorescence scatter plots from three different nucleofection experiments in which the nucleofection efficiencies ranged from 11.84% to 15.5%. The low nucleofection efficiencies were a limiting factor in the generation of *C9orf72*$^{-/-}$ iPSC lines.

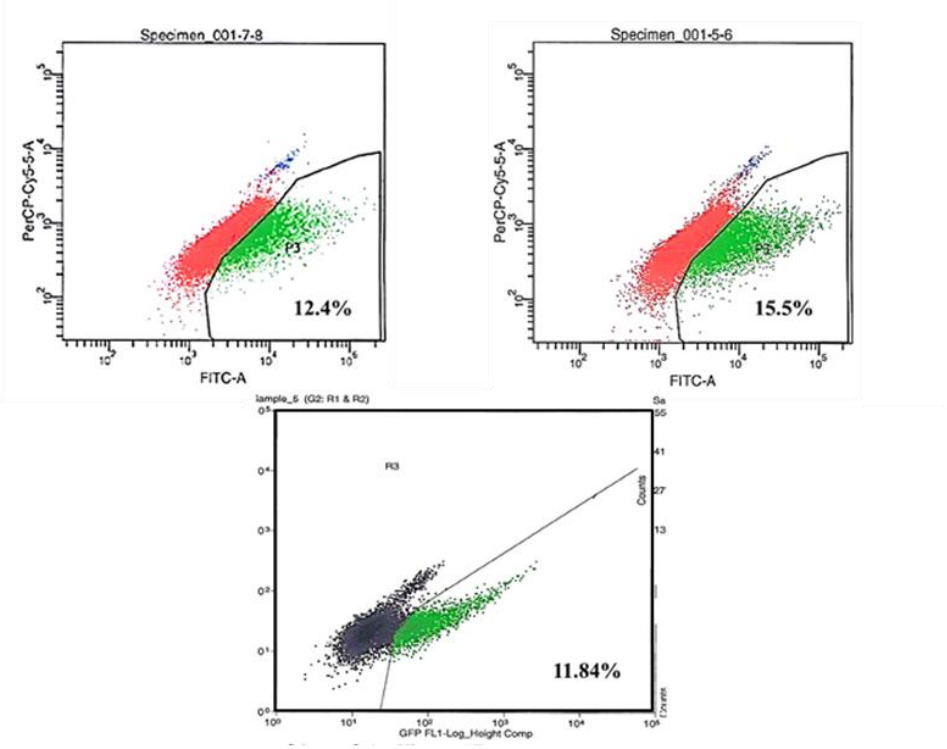


Figure 5.4 CRISPR/Cas9n nucleofection efficiencies.
FACS sorting of pSpCas9n(BB)-2A-GFP nucleofected, GFP positive, iPSCs using FACSAria III (top) and MoFlo XDP (bottom) cell sorters revealed nucleofection efficiencies ranging from 11.84% to 15.5%. Green indicated GFP positive iPSCs, whereas red on FACSAria and black on MoFlo XDP indicate wild-type, non-GFP cells. Gating was set using non-nucleofected, wild-type iPSCs as negative control.

5.3.2.4. *Low biallelic editing efficiency*

The choice of Cas9n instead of Cas9 has been an important parameter in the reduction of potential genomic off targets. However, the requirement for co-nucleofection of both right and left Cas9n plasmids in single iPSCs for successful induction of DSBs combined with the low nucleofection efficiency in iPSCs (see 5.3.2.3) and the overall low frequency of biallelic editing (0.4%), were the major limiting factors in the generation of *C9orf72*$^{-/-}$ iPSC lines. To overcome this problem and to obtain several *C9orf72*$^{-/-}$ clones to screen for biallelic freameshift indels, a *C9orf72*$^{+/-}$ iPSC line was edited again using CRISPR/Cas9n and pair 0 sgRNAs. This allowed the generation of several *C9orf72*$^{-/-}$ isogenic iPSC lines in one single round at a frequency similar to the monoallelic targeting frequency (see 5.3.3.1).

5.3.3. Characterisation of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines

5.3.3.1. *Sequencing of CRISPR-Cas9n edited iPSC lines*

Clonally-derived engineered iPSC lines were screened for NHEJ-generated indels by PCR amplification of 185 bp flanking the *C9orf72* exon 2 targeted region followed by high resolution agarose electrophoresis (as described in 2.17). Representative high resolution agarose electrophoresis (3% agarose) indel-screening gels are shown in Figure 5.5 A. Specifically, all non-modified, wild-type iPSC lines displayed a single PCR band at 185kb, whereas the engineered lines resolved as double/triple bands corresponding to different sizes of indels present at the nuclease cut site. Interestingly, the pattern of triple bands is explained by the presence of heterohybrids formed by annealing of different sequence DNA strands. The resulting wild-type/indel or indel/indel heteroduplexes display strand mismatches that lead to DNA secondary structures that alter the mobility of the product in the agarose gel. Indeed, the presence of DNA heteroduplexes has been confirmed by gel extraction of all PCR

bands followed by Sanger sequencing, with representative electropherograms shown in Figure 5.5 B. Sanger sequencing revealed modified iPSC clones that comprised of several indels generated by NHEJ, mostly monoallelic (8%) but also biallelic at a low frequency (0.4%). The mutation frequency in iPSCs is lower than the one calculated by the T7 endonuclease I mutation detection assay in HEK293 cells for CRISPR/Cas9n pair 0 (26.3-27.8%) and this difference is attributed to the different efficiencies of nucleofection and transfection respectively in these two differenct cell types. Figure 5.5 C shows the DNA sequences of some representative frameshift indels detected in modified iPSC clones.

The low frequency of biallelic frameshift indels in iPSCs and the poor survival of single iPSC clones were the main obstacles in acquiring enough *C9orf72*$^{-/-}$ iPSC lines for downstream use. This was overcome by a second CRISPR/Cas9n targeting round, whereby several *C9orf72*$^{-/-}$ iPSC lines were generated after editing a *C9orf72*$^{+/-}$ iPSC line (HET2) using the CRISPR/Cas9n and pair 0 sgRNAs. This resulted in several modified iPSC clones with biallelic frameshift indels. Three *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and three *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines were selected for further characterisation, differentiation to CNs and downstream experiments. The DNA sequences of all monoallelic and biallelic frameshift indels in the *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines respectively are shown in Figure 5.6 B.

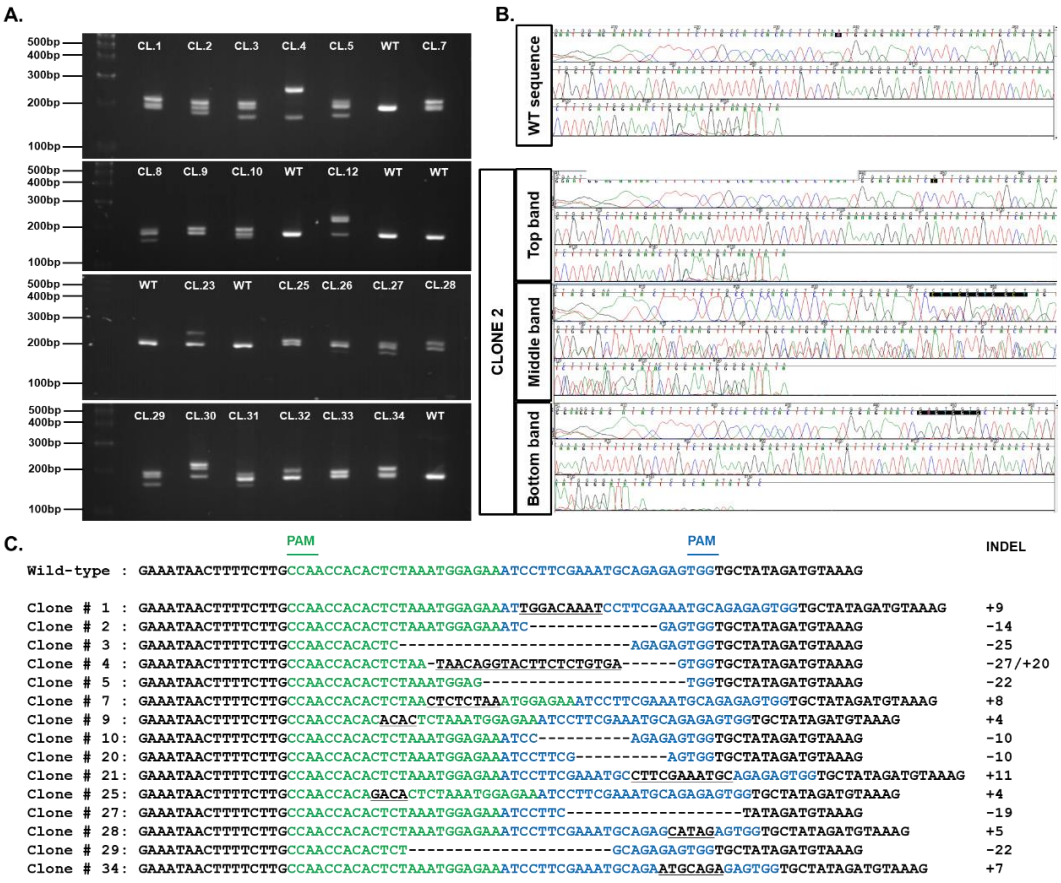
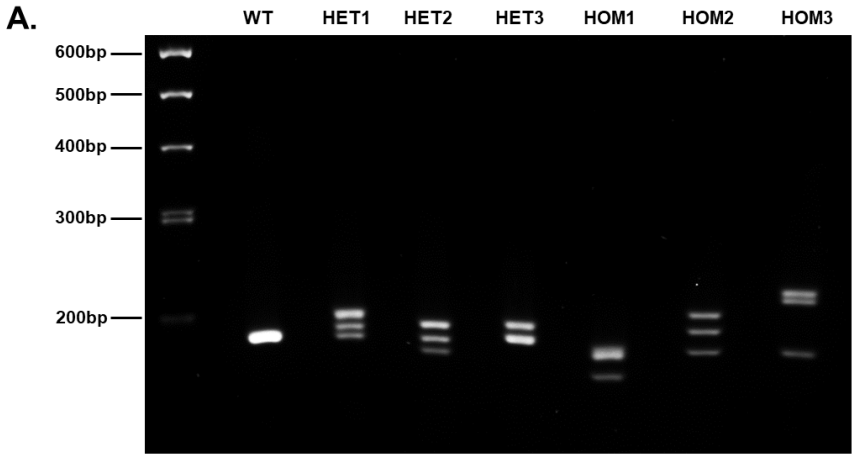


Figure 5.5 High resolution electrophoresis screening of edited iPSCs for NHEJ-induced indels.

A. For the detection of CRISPR/Cas9-modified iPSC lines, the DSB region was amplified by PCR and the PCR products were resolved on a high percentage (3%) agarose gel. During high resolution electrophoresis, a single band indicating unmodified, wild-type iPSCs or double/triple bands indicating modified iPSCs, were detected. The additional band in the triple band products was due to the formation of heterohybrid PCR products. PCR products resolving in bands of equal intensity are suggestive of single iPSC-derived colonies, whereas PCR products resolving in bands of unequal intensities are indicative of mixed clones, such as CL.23, CL.26, CL.31 and CL.32. B. Electropherograms showing the DNA sequencing results for the WT sequence and iPSC clone 2 which exhibited a three band PCR product on the 3% agarose gel (shown in A). The top band consisted of wild-type DNA sequence, the bottom was a modified product with a 14 bp deletion and the middle was a heterohybrid mixture of modified and wild-type DNA. C. Examples of sequenced iPSC clones comprising of NHEJ-induced frameshift indels in the region targeted by the sgRNAs in comparison to the wild-type sequence. Left sgRNA target sequence and PAM sequence are highlighted in green and right sgRNA target sequence and PAM sequence are highlighted in blue.



B.

	PAM	PAM	INDEL
Wild-type :	GAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAG		
HET # 1:	GAAATAACTTTTCTTGCCAACCACACTCTAACTCTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAG		+8
HET # 2:	GAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCC----------AGAGAGTGGTGCTATAGATGTAAAG		-10
HET # 3:	GAAATAACTTTTCTTGCCAACCACACATCATCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAG		+4
HOM # 1:	GAAATAACTTTTCTTGCCAACCACACTC----------AGAGAGTGGTGCTATAGATGTAAAG		-25/-10
HOM # 2:	GAAATAACTTTTCTTGCCAACCACACTCTACACTCTAAATGGAGAAATCC----------AGAGAGTGGTGCTATAGATGTAAAG		+7/-10
HOM # 3:	GAAATAACTTTTCTTGCCAACCACACTCTAAATGGAGAAATCC----------AGAACACTCTAAATGGAGAAATCCTTCGAAATGCAGAGAGTGGTGCTATAGATGTAAAG		-10/+34

Figure 5.6 Monoallelic and biallelic frameshift indels in iPSCs.

A. High resolution electrophoresis of three *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and three *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines. B. DNA sequences of monoallelic or biallelic NHEJ-induced frameshift indels in the three *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and three *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines respectively, in comparison to the wild-type sequence. All three *C9orf72*$^{-/-}$ iPSC lines carry the same -10 bp indel in one allele which is the same as in HET2 line. Left sgRNA target sequence and PAM sequence are highlighted in green and right sgRNA target sequence and PAM sequence are highlighted in blue.

5.3.3.2. Genomic off-targets

The use of Cas9n for genome editing has been shown to result in reduced genomic off-targets compared to Cas9, as Cas9n can only nick a single DNA strand rather than generating a mutagenic DSB (Ran *et al.*, 2013a, Cho *et al.*, 2014). Apart from the use of Cas9n, the selection of highly specific sgRNAs is another important factor for the reduction of potential off-targets. For pair 0, both left and right sgRNAs had no potential off-targets with either 1 or 2 nucleotide sequence mismatches, and the potential genomic off-targets were starting when the sgRNA sequence had 3 or 4 nucleotide mismatches.

Therefore, in order to identify all potential off-targets, an analysis was carried out interrogating all combinations of potential pair 0 sgRNAs with 3 or 4 mismatches within the same chromosome and their genomic distance (offset). Briefly, using a Python script, the lists of off-targets for left and right sgRNAs were merged when they were both present in the same chromosome. The start and end chromosomal locations for all sgRNAs were annotated and the genomic distance between all potential combinations of right and left sgRNAs within the same chromosome was calculated. Finally, the results were sorted by smallest genomic distance as shown in Table 5.1.

As demonstrated by Ran *et al.,* robust NHEJ was observed for sgRNA pairs with offsets from −4 to 20 bp, with modest indels forming in pairs offset by up to 100 bp (Ran *et al.*, 2013a). Therefore any combination of sgRNAs with an offset of up to 100bp would suggest a potential DSB and off-target in the region which should be subsequently sequenced via Sanger Sequencing.

The results of the analysis showed that, apart from the known combination of pair 0 sgRNAs targeting *C9orf72*, no other sgRNA pair exists at the same chromosome with a 100bp offset. The top 10 potential genomic off-targets and their genomic locations

are shown in Table 5.1. The shortest offset distance identified between a sgRNA pair on chromosome 3 was 14,141bp, well above the 100bp offset cutoff. Therefore, there are no reported potential off-targets with the use of pair 0 sgRNAs and Cas9n.

Table 5.1. Top 10 potential genomic off-targets for CRISPR pair 0 sgRNAs.

	Location Left sgRNA (Chromosome:genomic location)	Location Right sgRNA (Chromosome:genomic location)	Offset (bp)	Region Left sgRNA	Region Right sgRNA	Mismatches Left sgRNA	Mismatches Right sgRNA
Intended target	9:27566886-27566908	9:27566909-27566931	1	Exonic	Exonic	0	0
Off-target 1	3:24025788-24025810	3:24011625-24011647	14141	Intergenic	Intronic	4	4
Off-target 2	3:82048722-82048744	3:81940583-81940605	108117	Intergenic	Intergenic	3	4
Off-target 3	4:5656244-5656266	4:5506029-5506051	150193	Intronic	Intergenic	4	4
Off-target 4	21:37972058-37972080	21:38138242-38138264	166162	Intergenic	Intronic	3	3
Off-target 5	20:30740330-30740352	20:30525433-30525455	214875	Intronic	Intronic	4	4
Off-target 6	4:5204359-5204381	4:5506029-5506051	301648	Intronic	Intergenic	4	4
Off-target 7	11:45881809-45881831	11:45327323-45327345	554464	Intronic	Intergenic	4	4
Off-target 8	1:101355667-101355689	1:101973347-101973369	617658	Intronic	Intergenic	4	4
Off-target 9	2:112100949-112100971	2:111451083-111451105	649844	Intergenic	Intergenic	4	4
Off-target 10	5:147747409-147747431	5:148407612-148407634	660181	Intergenic	Exonic	4	4

5.3.3.3. *CRISPR-Cas9n edited iPSC lines maintain pluripotency and morphology*

All edited *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines were further assessed for typical stem cell morphology as well as maintenance of human stem cell identity markers. As shown in

Figure 5.7, all engineered iPSC lines displayed similar morphology to the parental CTRL-1 line, which is typical of ESCs grown in feeder-free conditions. Furthermore, all lines were further characterised for the maintenance of their stem cell identity via immunofluorescence staining. Figure 5.8 shows all *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines were positive for pluripotency markers OCT4 and TRA-1-81.

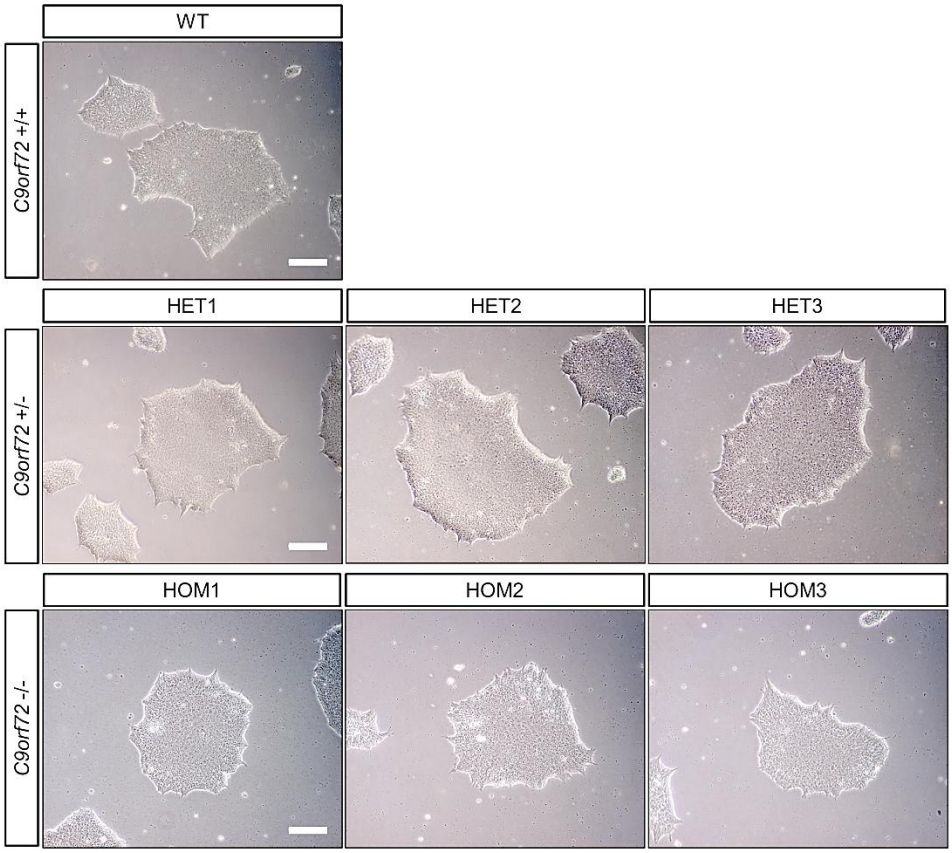


Figure 5.7 *C9orf72* knockout iPSC lines display typical ESC morphology.
All *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines display typical ESC morphology when grown *in vitro* under feeder-free conditions, similar to the parental wild-type iPSCs (CTRL-1). Scale bars 200µm.

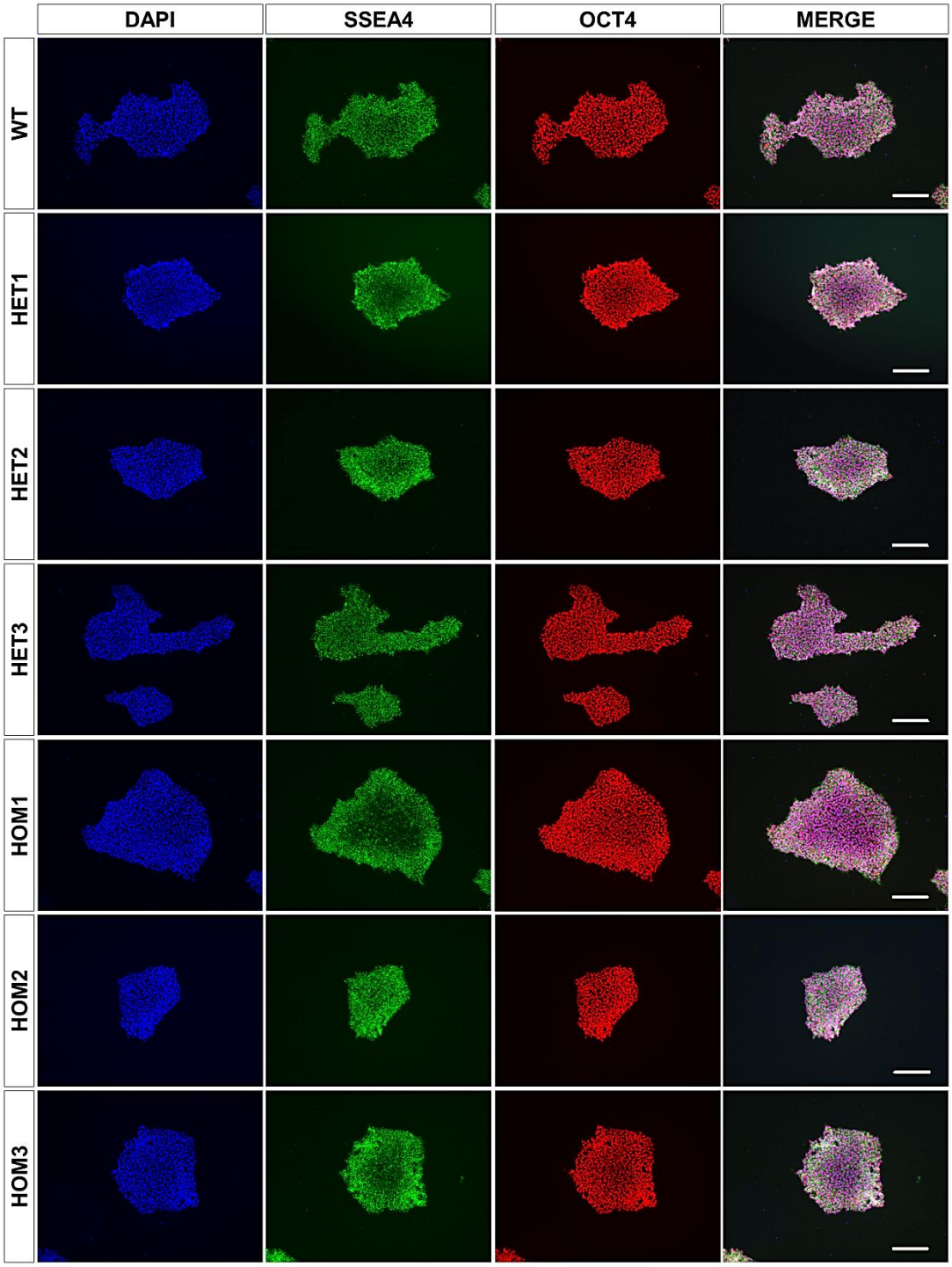


Figure 5.8 *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines maintain pluripotency.
C9orf72$^{+/-}$ (HET1, HET2 and HET3) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines were characterised by positive immunofluorescence for core pluripotency transcription factor OCT4 (red) and cell surface stage-specific embryonic antigen 4 (SSEA4) (green). Cell nuclei were counterstained with DAPI. Scale bars 200 μm.

Finally, karyotype analysis of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines was carried out using a qPCR-based genomic analysis kit, detecting the majority of karyotypic abnormalities reported in human ES and iPS cells (StemCell Technologies), as described in 2.21. The assays cover regions of genomic-instability on chromosomes 1q, 8q, 10p, 12p, 17q, 18q and 20q. According to the results of the analysis, there were abnormalities detected for chromosomes 17q and 20q across all edited lines, including the parental wild-type line. These abnormalities are consistent with a chromosomal gain, but due to the nature of the assay, it cannot be determined whether it is a chromosomal region or the whole chromosome. Therefore to identify the specific karyotypic abnormalities further karyotypic analysis by Giemsa banding will be required. This finding suggests that abnormalities were present in the parental line prior to genome editing. Of note, no karyotyping data is available for *C9orf72*$^{+/-}$ line HET3 which was not included in the qPCR analysis.

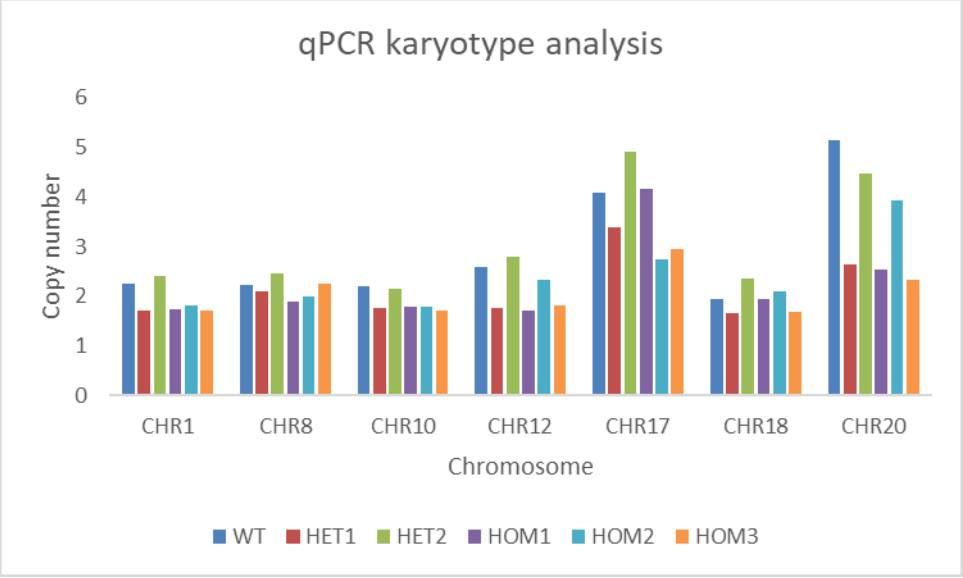


Figure 5.9 Karyotype analysis of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines.
C9orf72$^{+/-}$ (HET1 and HET2) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines were characterised using a qPCR-based genetic analysis kit (StemCell Technologies). Major karyotyping abnormalities were detected on chromosomes 17q and 20q across all edited iPSC lines and including the parental wild-type iPSCs, suggesting these abnormalities existed and are not caused by genome editing.

5.3.3.4. *Confirmation of C9orf72 downregulation in C9orf72$^{+/-}$ and C9orf72$^{-/-}$ iPSC lines*

The reduction of C9orf72 protein levels in the *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines was confirmed by western blot as described in 2.28 (Figure 5.10). The C9-L isoform was detected at 50 kDa in parental WT iPSCs whereas it was significantly reduced to haploinsufficiency levels in *C9orf72*$^{+/-}$ iPSC lines (HET1, HET2 and HET3) and undetected in *C9orf72*$^{-/-}$ iPSC lines (HOM1, HOM2 and HOM3) (n=3 biological replicates) (All *p* values are provided in Appendix B) (Figure 5.10). The C9-S isoform could not be detected specifically in wild-type iPSC lysates.

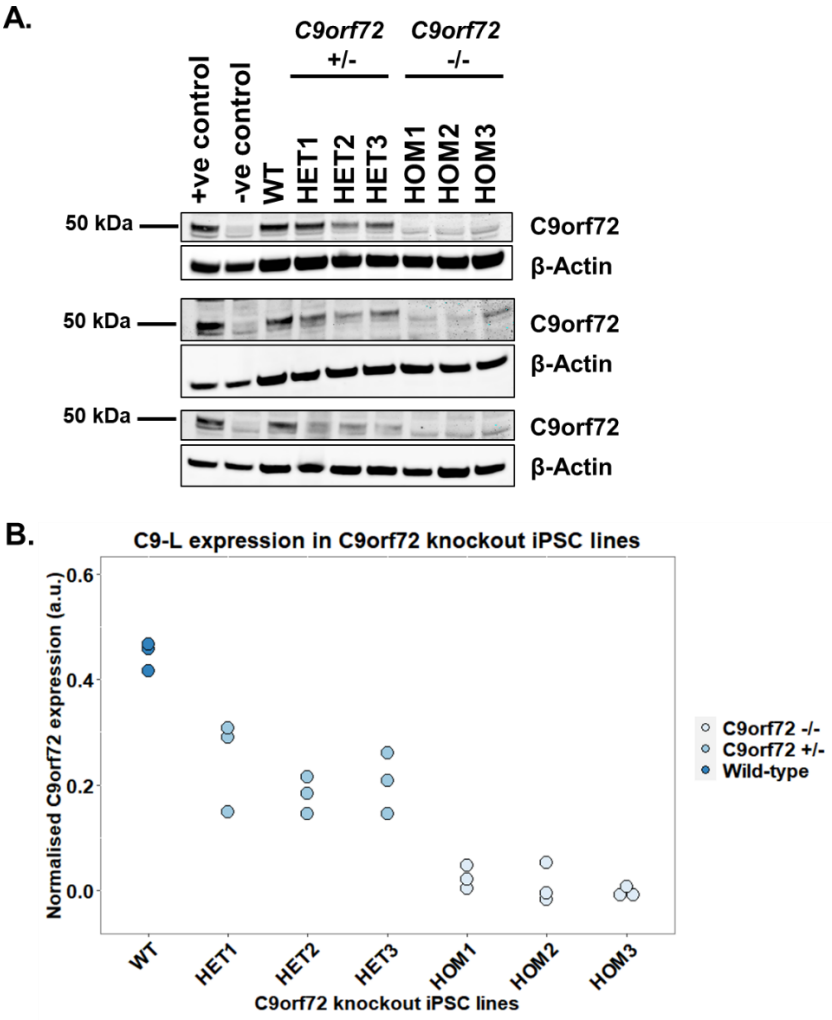


Figure 5.10 Confirmation of C9-L loss in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines.
Protein levels of C9orf72 were assessed by western blot in lysed iPSCs from the parental CTRL-1 (WT), *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) lines. Wild-type and *C9orf72* knock down HEK lysates were

(Continued Figure legend) included in each gel as positive and negative controls respectively, for the detection of C9orf72 protein. C9-L was detected at 50 kDa using anti-C9orf72 rabbit polyclonal antibody 22637-1-AP from Proteintech. However, no C9-S isoform could be detected using this antibody. β-Actin was used as a loading control. Each blot represents one biological replicate of different passage iPSCs (n=3). B. Densitometric quantification of western blot data shown in A. The levels of C9-L were reduced in a dose-dependent manner with the loss of each allele in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC lines.

5.3.4. Characterisation of differentiated *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs

5.3.4.1. C9orf72$^{+/-}$ *and C9orf72*$^{-/-}$ *iPSCs can generate deep layer CNs in vitro*

Wild-type, *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ isogenic iPSC lines were differentiated to CNs (as described in 2.5). All differentiated lines generated cortical rosettes at ~ 12-14 DIV. As shown in Figure 2.11, the 25 DIV cortical rosettes from all edited lines were positive for the early forebrain marker FOXG1, with proliferating NPCs also being positive for Ki67. Moreover, neuronal cultures at 40 DIV were positive for deep cortical layer marker CTIP2 as well as pan-neuronal marker TUJ1, confirming the cortical identity of differentiated neurons (Figure 5.12).

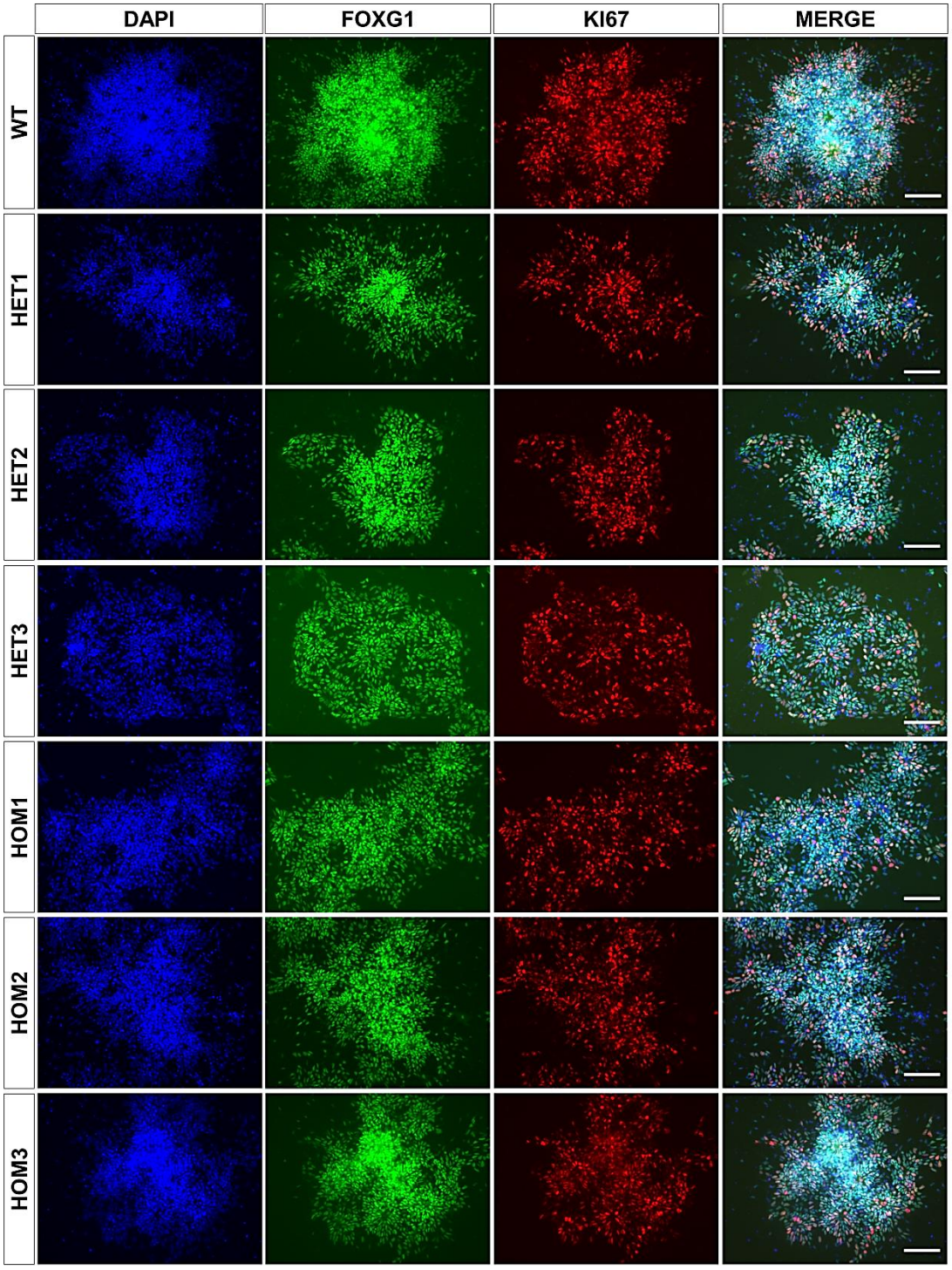


Figure 5.11 Cortical rosettes at 25 DIV express early forebrain marker FOXG1. Positive immunofluorescence for the early forebrain marker FOXG1 confirmed the cortical identity of the induced neural tissue in all *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) differentiated lines. The cycling cortical progenitors forming the polarised rosettes were also positive for the proliferation marker Ki-67 (red). Nuclei were counterstained with DAPI. Scale bars 100 µm.

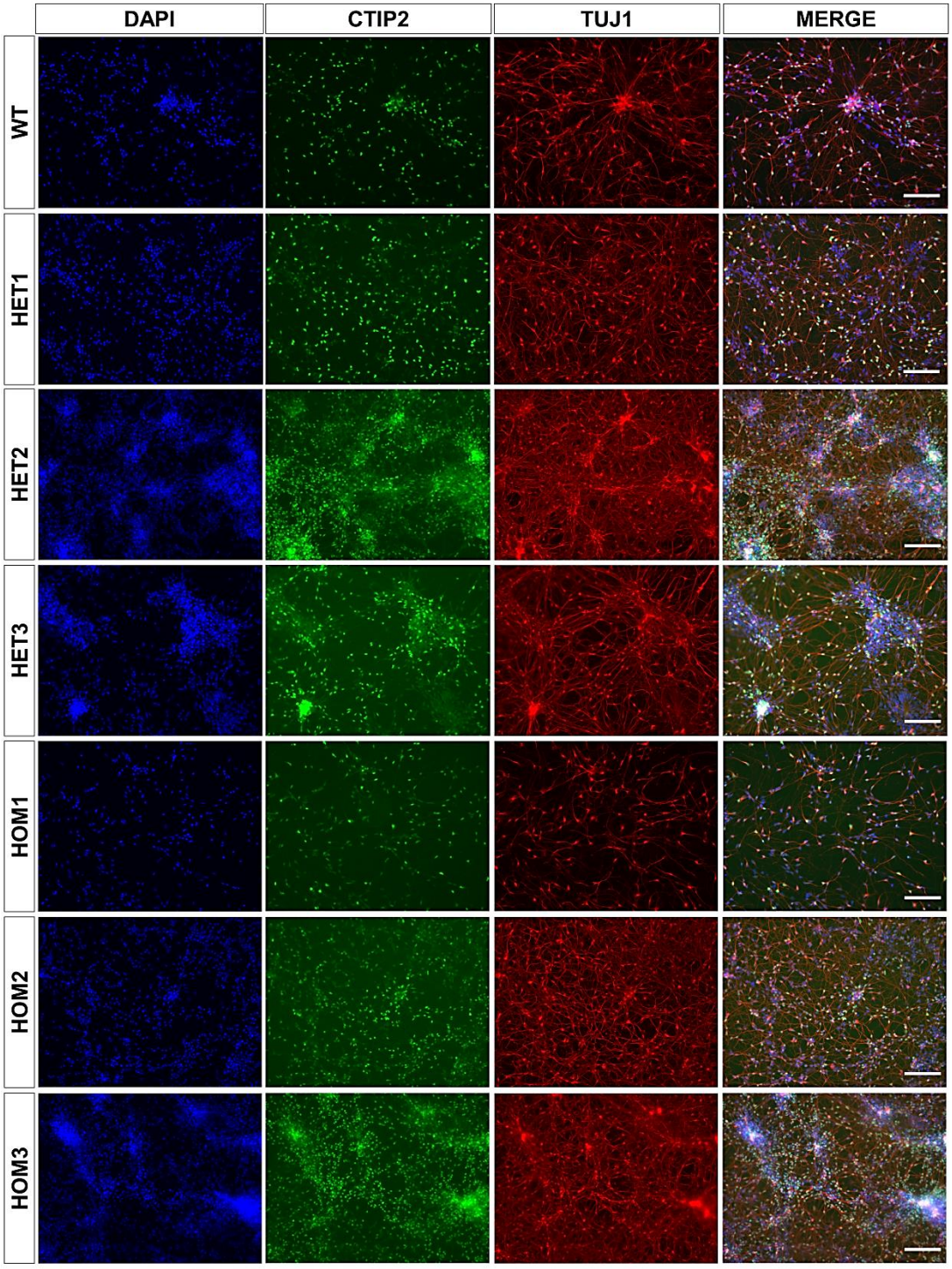


Figure 5.12 Cortical neurons at 40 DIV express deep-layer CN marker CTIP2. Cortical identity of the differentiated neurons from *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC lines was confirmed by positive immunofluorescence for deep cortical layer marker CTIP2. All neuronal axons were stained with TUJ1 (red). Nuclei were counterstained with DAPI. Scale bars 100 µm.

5.3.4.2. Confirmation of loss of C9orf72 protein in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs

The loss of C9orf72 protein was also confirmed in 80 DIV *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs by western blot. As shown in Figure 5.13, the loss of every *C9orf72* allele resulted in a respective dose-dependent significant reduction in C9-L protein isoform levels in the *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs (All *p* values are provided in Appendix B). In line with previous findings described in 4.2.7, 4.2.9 and 5.3.3.4, the C9-S isoform could not be specifically detected using this antibody in wild-type iPSC-CNs versus knockout iPSC-CNs.

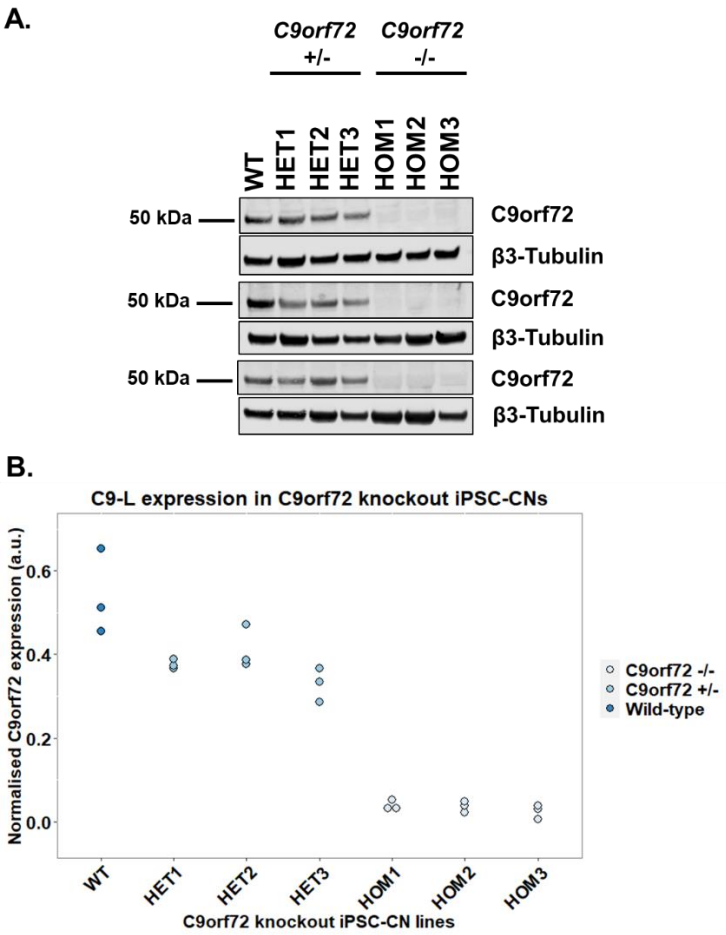


Figure 5.13 Confirmation of C9-L loss in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs.
Protein levels of C9orf72 were assessed via western blot in lysed day 80 *in vitro* iPSC-CNs from the parental CTRL-1 (WT), *C9orf72*$^{+/-}$ (HET1, HET2 and HET3) and *C9orf72*$^{-/-}$ (HOM1, HOM2 and HOM3) lines. C9-L was detected at 50 kDa using anti-C9orf72 rabbit polyclonal antibody 22637-1-AP from Proteintech, but no C9-S could be detected. B3-Tubulin was used as a loading control. Each blot represents one independent induction (n=3). B. Densitometric quantification of western blot data

(Continued Figure legend) shown in A. The levels of C9-L were reduced in a dose-dependent manner with the loss of each allele in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs.

5.3.5. Analysis of expression of autophagy-associated genes in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs

A wealth of evidence link C9orf72 function to autophagy regulation and lysosomal function (see 1.2.9). Therefore, the effect of C9orf72 loss on the expression of autophagy-related genes *ULK1*, *SQSTM1*, *LC3B* as well as *LAMP1* was investigated in *C9orf72*$^{+/-}$, *C9orf72*$^{-/-}$ and parental WT day 80 *in vitro* iPSC-CNs under basal conditions (n=3 independent inductions). Specifically, *ULK1* mRNA levels were significantly downregulated in *C9orf72*$^{-/-}$ iPSC-CNs compared to *C9orf72*$^{+/-}$ (p_{adj} = 0.019) and parental WT (p_{adj} = 0.049) iPSC-CNs, whereas *LC3B* mRNA levels were significantly reduced in *C9orf72*$^{-/-}$ iPSC-CNs compared to parental WT (p_{adj} = 0.049) but not *C9orf72*$^{+/-}$ (p_{adj} = 0.086) iPSC-CNs. Similarly, mRNA levels of the lysosomal gene *LAMP1* were significantly downregulated in *C9orf72*$^{-/-}$ iPSC-CNs compared to *C9orf72*$^{+/-}$ iPSC-CNs (p_{adj} = 0.007). No significant difference in the expression levels of *SQSTM1* between WT, *C9orf72*$^{-/-}$ and *C9orf72*$^{+/-}$ iPSC-CNs was observed. Further experiments measuring the autophagic flux in the *C9orf72*$^{-/-}$ and *C9orf72*$^{+/-}$ iPSC-CNs are necessary to assess the autophagy phenotypes caused by loss of C9orf72 function in iPSC-CNs.

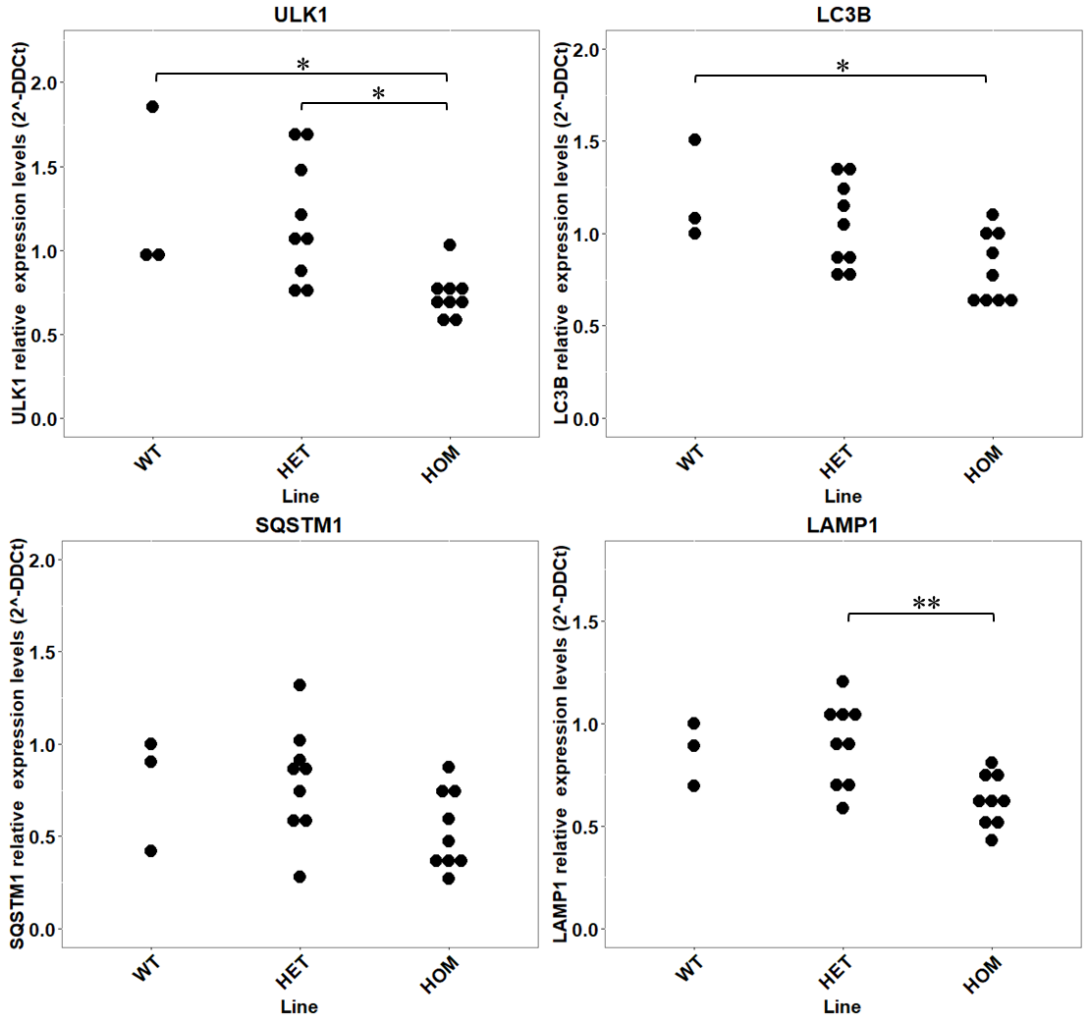


Figure 5.14 Altered expression of autophagy-related genes in *C9orf72*$^{-/-}$ iPSC-CNs. The relative mRNA levels of autophagy-related genes *ULK1*, *LC3B*, *SQSTM1* and *LAMP1* were assessed via RT-qPCR in day 80 *in vitro* parental CTRL-1 (WT), C9orf72$^{+/-}$ (HET1, HET2 and HET3) and C9orf72$^{-/-}$ (HOM1, HOM2 and HOM3) iPSC-CNs. *ULK1* mRNA levels were significantly reduced in *C9orf72*$^{-/-}$ compared to *C9orf72*$^{+/-}$ (p_{adj} = 0.019) and parental WT (p_{adj} = 0.049) iPSC-CNs. *LC3B* mRNA levels were significantly reduced in *C9orf72*$^{-/-}$ compared to parental WT (p_{adj} = 0.049) iPSC-CNs. Similarly, mRNA levels of the lysosomal gene *LAMP1* were significantly downregulated in *C9orf72*$^{-/-}$ compared to *C9orf72*$^{+/-}$ (p_{adj} = 0.007) iPSC-CNs. No significant difference in the expression levels of *SQSTM1* between WT, *C9orf72*$^{-/-}$ and *C9orf72*$^{+/-}$ iPSC-CNs was observed. Relative expression was calculated using the $2^{(-\Delta\Delta Ct)}$ method. Ct values were normalised to the housekeeping gene *ACTB* Ct values. RT-qPCR assays were performed in technical duplicates using iPSC-CNs from three independent inductions (n=3).

5.3.6. Investigation of mitophagy in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs

The PINK1/Parkin-dependent mitophagy is a type of selective macroautophagy for the degradation of damaged mitochondria (see 1.4.1). Upon mitochondrial damage, PINK1 accumulates in the OMM and phorphorylates ubiquitin (Ub) chains bound to OMM protein substrates leading to the recruitment of Parkin which then ubiquitinates several OMM protein substrates, including MFN2. Ubiquitin chains are further phosphorylated by PINK1, leading to amplification of phospho-ubiquitin in the OMM that is bound by autophagy receptor proteins connecting the damaged mitochondria to LC3 on autophagosomes for degradation by macroautophagy.

The integrity of the PINK1/Parkin-dependent mitophagy pathway was assessed in day 90 *in vitro* *C9orf72*$^{+/-}$ (HET1 and HET3), *C9orf72*$^{-/-}$ (HOM1 and HOM2) and parental WT iPSC-CNs via western blotting, as described in 2.29 and 2.30. Briefly, mitochondrial damage was induced by combined treatment with Oligomycin, an inhibitor of ATP synthase, and Antimycin A, an inhibitor of cellular respiration, causing increased superoxide generation coupled with loss of mitochondrial membrane potential (ΔΨm). All iPSC-CNs were treated with 1 μM Oligomycin/Antimycin A for 12, 24 and 36 hours. The levels of major components of the mitophagy pathway including PINK1, MFN2 and phosphorylated Ubiquitin, were assessed via western blot. The preliminary data shown in Figure 5.15 indicate that PINK1 accumulation on OMM in response to Oligomycin/Antimycin A treatment was severely impaired in HOM1 and HOM2 *C9orf72*$^{-/-}$ iPSC-CNs, compared to *C9orf72*$^{+/-}$ and parental WT iPSC-CNs (n=1 experiment). This was accompanied by downstream effects in *C9orf72*$^{-/-}$ iPSC-CNs consistent with lack of MFN2 ubiquitination and no increase in phospho-ubiquitin levels, in contrast to the increase in MFN2 ubiquitination and phospho-ubiquitin levels observed in *C9orf72*$^{+/-}$ and parental WT iPSC-CNs. Collectively, the above preliminary results suggest an impairment in the induction of mitophagy in *C9orf72*$^{-/-}$ iPSC-CNs and require further replication and investigation.

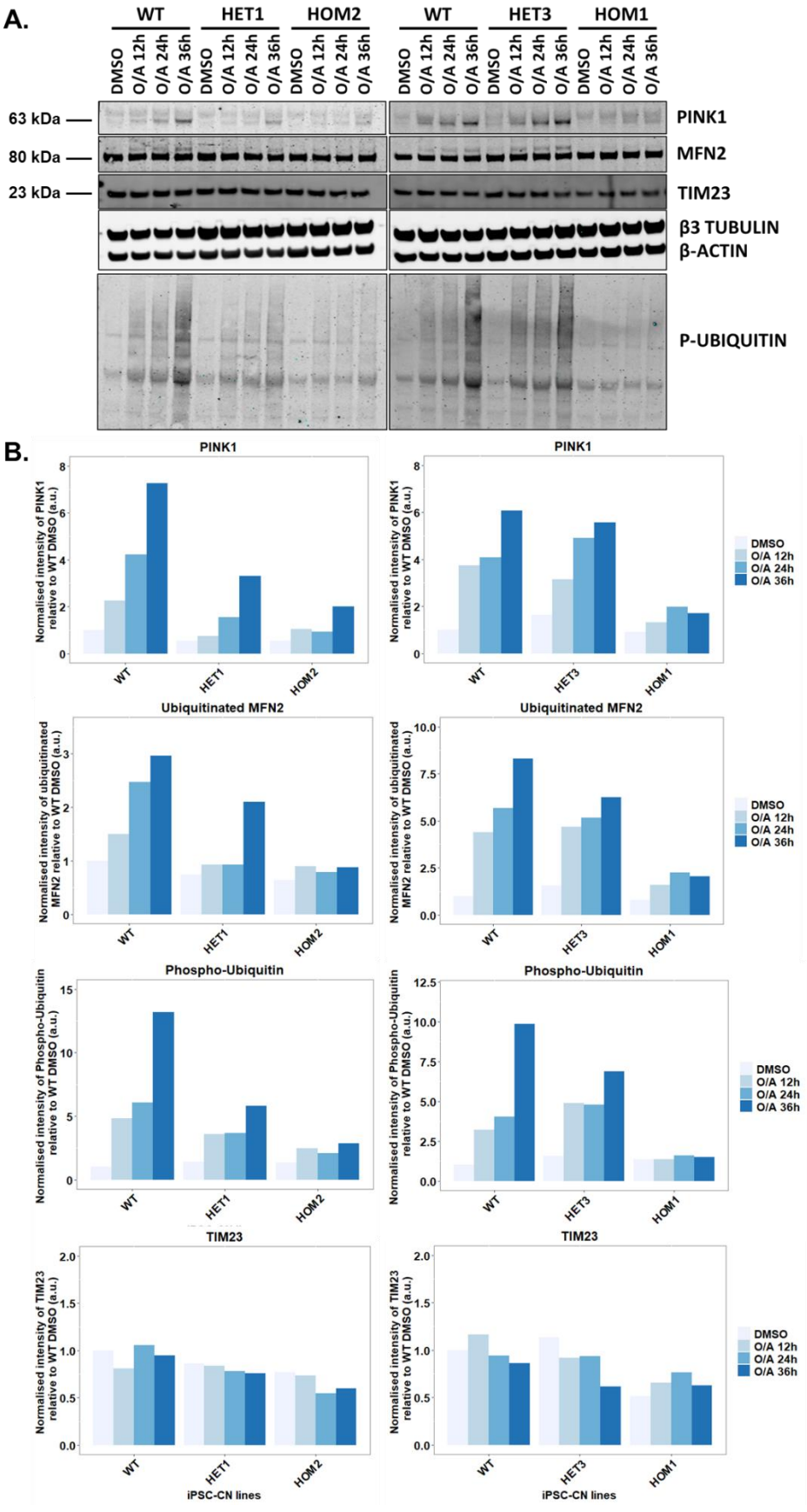


Figure 5.15 Impaired mitophagy in *C9orf72*$^{-/-}$ iPSC-CNs.

(Continued Figure 5.13 legend)

A. Western blot analysis of major components of the PINK1/Parkin-dependent mitophagy pathway PINK1, MFN2 and phospho-ubiquitin in day 90 *in vitro C9orf72*$^{+/-}$ (HET1 and HET3), *C9orf72*$^{-/-}$ (HOM1 and HOM2) and parental WT iPSC-CNs. Mitochondrial damage was induced by treatment of iPSC-CNs with 1 μM Oligomycin/Antimycin A for 12, 24 and 36 hours. Preliminary data showed impaired PINK1 accumulation on OMM in response to Oligomycin/Antimycin A treatment in HOM1 and HOM2 *C9orf72*$^{-/-}$ iPSC-CNs, compared to *C9orf72*$^{+/-}$ and parental WT iPSC-CNs (n=1 experiment). This was accompanied by downstream effects in *C9orf72*$^{-/-}$ iPSC-CNs consistent with lack of MFN2 ubiquitination and no increase in phospho-ubiquitin levels, in contrast to the increase in MFN2 ubiquitination and phospho-ubiquitin levels observed in *C9orf72*$^{+/-}$ and parental WT iPSC-CNs. β-Actin and β3-Tubilin were used as loading controls. B. Densitometric quantification of western blot data shown in A.

5.4. Discussion

This Chapter describes the generation, characterisation and study of isogenic *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs, generated via CRISPR/Cas9n editing. Some of the issues encountered during the generation of the CRIPSR/Cas9n edited iPSCs included the low nucleofection efficiency, the poor survival of single-cell iPSCs post FACS sorting and the low frequency of biallelic knockout iPSC clones. Despite optimisation efforts that significantly improved the survival of edited iPSCs, the low nucleofection efficiency combined with the low biallelic CRISPR/Cas9n targeting efficiency, resulted in very low percentage of generated *C9orf72*$^{-/-}$ iPSC lines (2%). In order to obtain several *C9orf72*$^{-/-}$ clones to screen for biallelic freameshift indels in a cost-effective, time-efficient and not as labour-intensive way, a second round of CRISPR/Cas9n editing of a *C9orf72*$^{+/-}$ iPSC line (HET2) was performed, resulting in the generation of several *C9orf72*$^{-/-}$ isogenic iPSC lines at a frequency similar to the monoallelic targeting frequency.

Two parameters considered for reducing genome editing off target effects were the use of CRISPR/Cas9 nickase rather than wild-type CRISPR/Cas9 and sgRNA pairs with low number of potential off-targets. Indeed, there were no potential genomic off-targets detected for the sgRNA pair 0 used for editing *C9orf72* gene. Specifically, the minimal offset between the two sgRNAs for off-target effects was ~14 kb, well above the maximum 100 bp offset cutoff required for induction of a DSB by the CRISPR/Cas9n. Despite the lack of CRISPR/Cas9n off-targets, genetic analysis for identification of karyotypic abnormalities revealed abnormalities on chromosomes 17 and 20 in all tested *C9orf72* knockout iPSC lines, including the parental wild-type iPSCs. Therefore this suggests that karyotypic abnormalities have not emerged during genome editing process since they were present in the parental line.

C9orf72$^{+/-}$ and *C9orf72*$^{-/-}$ iPSCs were differentiated to CNs for the study of the effect of C9orf72 loss in FTD. Of note, there are no published studies of *C9orf72* knockout iPSC-CNs to date, instead, there have been reports of *C9orf72* knockout iPSC-MN models (see 1.3.4). Parental, *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSCs did not differ in their ability to generate deep-layer CNs. However, due to time constraints, further characterisation of the *C9orf72* knockout CN cultures for the presence of upper layer CN markers could not carried out and will be completed in future work on this project.

The choice of three *C9orf72*$^{+/-}$ and three *C9orf72*$^{-/-}$ isogenic iPSC lines carrying different frameshift indels increases the power of the study and robustness in the detection of phenotypes. However, the use of one parental WT iPSC line resulted in reduced power for detection of statistically significant differences between WT and *C9orf72*$^{+/-}$/*C9orf72*$^{-/-}$ study groups.

Gene expression analysis of key autophagy genes *ULK1*, *LC3B* and lysosomal gene *LAMP1* revealed significant downregulation of these genes in day 80 *C9orf72*$^{-/-}$ iPSC-CNs compared to *C9orf72*$^{+/-}$ and/or parental WT iPSC-CNs. The dysregulated expression of these autophagy-related genes in *C9orf72*$^{-/-}$ iPSC-CNs is in line with the described role of C9orf72 in the induction of ULK1-mediated autophagy and suggest altered basal autophagy that requires further investigation by assessing the autophagic flux in *C9orf72* knockout iPSC-CNs.

Furthermore, the PINK1/Parkin-dependent mitophagy pathway was found to be compromised in day 90 *in vitro C9orf72*$^{-/-}$ (HOM1 and HOM2) iPSC-CNs compared to *C9orf72*$^{+/-}$ (HET1 and HET3) and parental WT iPSC-CNs. Specifically, upon Oligomycin/Antimycin A-induced mitochondrial damage, PINK1 accumulation on OMM was severely impaired in HOM1 and HOM2 *C9orf72*$^{-/-}$ iPSC-CNs, compared to *C9orf72*$^{+/-}$ and parental WT iPSC-CNs. This was also accompanied by a lack of MFN2 ubiquitination and no increase in phospho-ubiquitin levels in *C9orf72*$^{-/-}$ iPSC-CNs

sompared to isogenic controls WT and *C9orf72*$^{+/-}$. These are preliminary findings (n=1) that require further replication in future work. However, they suggest that the induction of mitophagy is impaired in *C9orf72*$^{-/-}$ iPSC-CNs possibly leading to the accumulation of compromised mitochondria in neurons lacking C9orf72.

No published studies on PINK1/Parkin-dependent mitophagy pathway in *C9orf72*$^{-/-}$ iPSC-CNs currently exist. However, this pathway is a good candidate for further investigation in *C9orf72* knockout models due to the great overlap with autophagy pathway in which C9orf72 protein has a well described role. Moreover, a genetic interaction between *C9orf72* and mitochondrial fission protein gene *FIS1* suggest a potential role of C9orf72 in mitophagy that requires futher investigation (Chai *et al.*, 2019). Mitochondrial defects have also been reported in *C9orf72* HRE iPSC-MNs, such as swollen mitochondria (Dafinca *et al.*, 2016), reduced (Dafinca *et al.*, 2016) or increased (Lopez-Gonzalez *et al.*, 2016) mitochondrial membrane potential and poly-GR toxicity has been shown to compromise mitochondrial function.

Interestingly, overexpression of another ALS gene *SOD1* or its G85R or G93A mutants was also found to interfere with PINK1 accumulation or mitophagy induced by carbonyl cyanide-4-(trifluoromethoxy)phenylhydrazone (FCCP), a mitochondrial oxidative phosphorylation uncoupler, in HEK293 cells (Gao *et al.*, 2020). This effect was mediated by inhibition of the CHCHD4/GFER system in the mitochondrial intermembrane space (IMS). Heterozygous mutations in the lysosomal enzyme β-Glucocerebrosidase (*GBA*) cause Gaucher disease (GD), the most common lysosomal disorder, whereas heterozygous *GBA* mutations are the most common genetic risk factor for Parkinson disease (PD). Heterozygous *GBA* L444P mutation was also found to inhibit mitophagy *in vivo* and *in vitro* through a dual mechanism: the loss of GBA activity impeded lysosomal clearance of autophagosomes, whereas the mutant GBA protein inhibited autophagy induction and mitochondrial priming (Li *et al.*, 2019).

In conclusion, the work presented in this chapter has confirmed a role for C9orf72 in autophagy and revealed mitophagy as a novel pathway potentially affected in C9orf72 knockout iPSC-CNs, but future work will be necessary to replicate these findings.

Chapter 6 General discussion

6.1. Main conclusions

The purpose of this study was to investigate FTD-related pathologies and *C9orf72* haploinsufficiency in *C9orf72* HRE iPSC-CNs, and secondly, the generation and study of *C9orf72* knockout iPSC-CNs.

The use of iPSC-CNs derived from *C9orf72* HRE patients was a useful tool in the study of C9-FTD disease mechanisms. The majority of published iPSC studies to date have used iPSC-MNs (Lines *et al.*, 2020), therefore in this thesis, the use of iPSC-CNs offers an insight into the relevant pathologies in CNs. *C9orf72* iPSC-CNs exhibited pathological hallmarks of C9-FTD/ALS, including sense RNA foci and the DRPs poly-GP and poly-GR. Moreover, *C9orf72* iPSC-CNs showed early signs of haploinsufficiency, evident by hypermethylation of the *C9orf72* promoter region as well as significant reduction in *C9orf72* total mRNA levels compared to control iPSC-CNs, though with considerable variability between *C9orf72* HRE patient lines. *C9orf72* haploinsufficiency in protein levels could not be detected in *C9orf72* iPSC-CNs, despite assessing C9orf72 protein levels in iPSC-CNs at extended timepoints of 150 DIV and 260 DIV. This may be associated with the immaturity of iPSC-derived neuronal cultures (Stein *et al.*, 2014, Burke *et al.*, 2020), as described previously.

Another important feature of this study was the temporal assessment of the *C9orf72* promoter methylation and C9orf72 protein expression in *C9orf72* iPSC-CNs and control iPSC-CNs respectively. Temporal analysis is particularly important in such developmental models as it can reveal dynamic changes that are usually missed when focussing on a single timepoint (Burke *et al.*, 2020). One such example was the temporal analysis of *C9orf72* iPSC-MN excitability (Devlin *et al.*, 2015) that revealed differences in the *C9orf72* iPSC-MN electrical properties at different developmental stages *in vitro*, consistent with hyperexcitability in 3-4 weeks-old *C9orf72* iPSC-MNs

versus a loss of excitability in 9-10 weeks-old *C9orf72* iPSC-MNs. Similarly, the temporal analysis of *C9orf72* promoter methylation revealed a temporal pattern of 5' CpG hypermethylation establishment in one *C9orf72* HRE patient line as well as HRE methylation instability in another *C9orf72* HRE patient line that would have otherwise not been detected.

The comprehensive analysis of HRE size, RNA foci, *C9orf72* mRNA expression and promoter methylation in the three *C9orf72* iPSC-CN lines revealed strong correlations between these variables. The 5' CpG island methylation levels were negatively correlated with the percentage of RNA foci and the *C9orf72* total mRNA levels, in line with previous reports showing that 5' CpG hypermethylation is associated with reduced RNA foci pathology and *C9orf72* expression (Liu *et al.*, 2014, Russ *et al.*, 2014, Bauer, 2016). Moreover, the 5' CpG island methylation levels were positively correlated with HRE size. Studies on a larger number of patient lines will be required to better understand the relationships between these variables.

The generation of CRISPR/Cas9n *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ isogenic iPSC lines provided useful *in vitro* models for the dissection and study of *C9orf72* loss of function mechanisms. *C9orf72*$^{-/-}$ iPSC-CNs at 80 DIV exhibited significant downregulation of autophagy-related genes *ULK1*, *LC3B* and *LAMP1* compared to *C9orf72*$^{+/-}$ and/or WT iPSC-CNs, indicative of altered basal autophagy. This finding is in agreement with previous descriptions of a role of C9orf72 in the induction of ULK1-mediated autophagy (Webster *et al.*, 2016, Yang *et al.*, 2016), and requires further investigation by functional assessment of the autophagy pathway in *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-CNs.

Furthermore, the work presented here has identified the PINK1/Parkin-dependent mitophagy, a type of selective macroautophagy, as a novel pathway potentially affected by C9orf72 loss of function. Preliminary findings suggest that induction of

mitophagy is impaired in *C9orf72*$^{-/-}$ iPSC-CNs, but further replication of these findings is required. Considering the role of C9orf72 in the induction of ULK1-mediated autophagy, a genetic interaction reported between *C9orf72* and mitochondrial fission protein gene *FIS1* (Chai *et al.*, 2019), as well as the wider role of mitophagy impairements in neurodegenerative diseases, the mitophagy pathway constitutes a good candidate for further investigation in *C9orf72* knockout iPSC-CNs.

6.2. iPSC models in C9-FTD/ALS; challenges and limitations

The use of *in vivo* and *in vitro* models has elucidated several novel pathways affected in C9-FTD/ALS and significantly accelerated research. The use of these models has been particularly useful at identifying the earliest pathogenic events, before the manifestation of end-stage pathologies detected at *post-mortem*. In many studies, these models have been used interchangeably, complementing each other. For instance, screens for HRE toxicity modifiers have been performed in *in vivo* models such as Drosophila or yeast, allowing an easy phenotypic readout, whereby the validity of the top candidates and their relevance to human disease has been subsequently confirmed in iPSCNs. Indeed, iPSC models offer the unparalleled advantage of allowing the study of patient-derived neurons, a more physiologically relevant model system not relying on overexpression of transgenes. *C9orf72* HRE iPSCNs, carrying the patient's precise HRE, can be used to assess the endogenous levels of toxic species of RNA foci and DPRs, as well as the *C9orf72* promoter methylation and *C9orf72* haploinsufficiency. Exploring the variability of these phenotypes between different patients that may help identify additional modifiers of disease in C9-FTD/ALS. Moreover, *C9orf72* HRE iPSCNs can also be used as drug screening platforms for the identification of candidate compounds that reduce toxicity and improve neuronal viability.

Despite the many advantages, modelling adult-onset neurodegenerative diseases such as C9-FTD/ALS with a developmental model such as iPSCNs is challenging and has limitations. In *C9orf72* HRE carriers, the production of toxic species is tolerated for many years by neurons and glia before the onset of clinical symptoms, and during these years additional stressors and other epigenetic factors may act to induce or accelerate the manifestation of symptoms. Indeed, iPSCNs more closely resemble embryonic rather than adult neurons and the toxic species detected are not sufficient to induce neuronal death. Hence, in several studies neuronal death was observed following application of additional stressors in the cultures. Apart from stressors, induced ageing (Miller *et al.*, 2013) or transdifferentiation of fibroblasts to neurons (iNs) (Mertens *et al.*, 2015) offer alternate ways to accelerate ageing *in vitro*. Poly-GP and poly-GR were detected in this study in iPSC-CNs by MSD immunoassay, however, further investigation of other DPR species was not possible due to unavailability of optimised MSD assays.

The use of iPSCs in the study of C9-FTD/ALS is tied to additional challenges due to the unique *C9orf72* HRE features associated with genomic instability of the HRE and hypermethylation of the *C9orf72* promoter which are also sources of significant variability. The use of southern blot has become an important characterisation step in *C9orf72* HRE iPSC studies by addressing the presence of mosaicism in cultures or the stability of the HRE during different differentiation stages. Moreover, tissue mosaicism in *C9orf72* HRE carriers has implications in the faithfulness of the disease models as the HRE size in the peripheral tissues used for reprogramming can differ from the actual size in the brain. Therefore this can result in iPSCNs consisting of different size HREs compared to patients' neurons. In most cases, post-mortem brain tissue is not available to confirm the *C9orf72* HRE size.

One study comparing the *C9orf72* promoter methylation in *C9orf72* ESCs and haploidentical *C9orf72* iPSCs suggested that C9orf72 iPSCs may be prone to

hypermethylation of the *C9orf72* promoter (Cohen-Hadad *et al.*, 2016), which has further implications in the use of *C9orf72* iPSCs as models in C9-FTD/ALS. Currently, there is not enough evidence to support this, as there have been very few iPSC studies investigating *C9orf72* promoter methylation, comprising of a very limited amount of study lines. One example contradicting this hypothesis is the *C9orf72* iPSC line of the present study that does not have a hypermethylated promoter. More studies on a larger number of *C9orf72* iPSC lines would be necessary to investigate *C9orf72* promoter methylation. Furthermore, *C9orf72* promoter methylation assessment could also be carried out in iPSC studies, similar to southern blot, as *C9orf72* promoter hypermethylation is an important disease modifier in C9-FTD/ALS. Other important factors also need consideration when studying disease-associated hypermethylation in iPSCs, as reprogramming leads to global DNA hypomethylation (Leitch *et al.*, 2013). DNA demethylation is associated with naïve pluripotency state, whereas partial demethylation has been associated with epigenetic memory, which is the partial retention of epigenetic patterns of the originally reprogrammed cell type (Kim *et al.*, 2010b). Reprogramming-induced global demethylation is followed by epigenetic reconfiguration during neuronal development that allows expression of neuron-specific genes, with the DNA methylation landscape dynamically reconfigured throughout human brain development (Lister *et al.*, 2013). Inherent variation in several master regulators responsible for shaping the epigenetic landscapes in iPSCs or in differentiating neurons may lead to more variability in the model system.

Apart from the HRE and hypermethylation-induced variability, additional source of variability in iPSC cultures can be due to the presence of additional genetic modifiers, such as *TMEM106B* or epigenetic modifiers, influencing the phenotypes. Therefore, characterisation of iPSC lines for known modifiers could help explain some of the observed phenotypic variability.

In the present study, the use of three *C9orf72* HRE lines is a limiting factor in drawing general conclusions about the observations made. Ideally, more patient lines as well as more than one clones per line are important to increase the power for detection of phenotypes and true associations. Nevertheless, similar to other cellular studies, sample availability, labour- and cost-effectiveness are important factors influencing the study design.

An important step in the characterisation of iPSC cultures is the detection of a normal karyotype. Here, despite an initial detection of a karyotypically normal CTRL-1 line, karyotypic abnormalities were detected in this line as well as the CRISPR/Cas9 *C9orf72* knockout isogenic lines derived from it. Indeed, iPSCs comprise of normal diploid karyotypes at early passages, and karyotypic abnormalities are associated with increased passage number. The most commonly detected abnormalities in iPSCs are gains of whole or parts of chromosomes 1, 12, 17 and 20 (Amps *et al.*, 2011, Taapken *et al.*, 2011, Assou *et al.*, 2020). Such changes are favoured as they confer a growth advantage. One such example is the gain of chromosome 20q11.21 which leads to upregulation of the anti-apoptotic protein BCL-XL. It is not possible to identify the exact karyotypic abnormalities present in the CRISPR/Cas9 iPSCs using the genetic analysis qPCR method, therefore further cytogenetic karyotypic analysis should be carried out to fully characterise these lines.

The differentiation of iPSCs to CNs offers the best way to explore pathologies relevant to FTD. The protocol used in the current study is closely mimicking embryonic cortical development by a stepwise generation of both deep and upper layer CNs as well as glia. However, one limitation of the current differentiation protocol is the lack of additional cell types, such as the GABAergic inhibitory interneurons and microglia, from the CN cultures. Microglia are derived from erythro-myeloid progenitors of extraembryonic yolk sac origin and enter the CNS early in embryonic development whereas GABA interneurons originate from the ventral telencephalon and migrate

during embryonic development to the different layers of the cortex (Zhu *et al.*, 1999, Gomez Perdiguero *et al.*, 2015).

Finally, the work presented in this thesis focused mostly on cell-autonomous phenotypes of iPSC-CNs, with potential influences of glia also present in mature cultures. However, there is a wealth of evidence suggesting that contribution of non-cell-autonomous mechanisms is important in the pathogenesis of C9-FTD/ALS. Furthermore, an additional limitation is the preliminary nature in some major outcomes from the present study, due to COVID-19-related work disruption, which are pending as future work (see 6.3).

6.3. Future directions

The work presented here provides another example of the use of iPSCs in C9-FTD/ALS modelling and the dissection of C9orf72 loss of function mechanisms using CRISPR/Cas9n genome editing. Whilst this work has provided insight into the C9-FTD/ALS pathogenic mechanisms and suggested novel pathways potentially affected by C9orf72 loss in iPSC-CNs, further experiments need to be carried out as part of future work.

Several experiments currently interrupted due to the COVD-19 pandemic will be completed in future work. Specifically, a thorough investigation of autophagy by assessment of the autophagic flux using autophagy inhibitors and inducers will be performed in *C9orf72* knockout iPSC-CNs and *C9orf72* iPSC-CNs. Live *in vitro* measuring of the autophagic flux can also be performed using mCherry-GFP-LC3 construct. Moreover, PINK1/Parkin mitophagy western blots will be replicated and additional mitochondrial assays such as measurement of mitochondrial membrane potential with the fluorescent dye Tetramethylrhodamine Methyl Ester (TMRM) will be performed on these iPSC-CNs in a 96-well format.

It would also be of interest to understand the consequences of *C9orf72* HRE and loss of function on neuronal function, which could be achieved by the temporal analysis of the electrical properties of the *C9orf72* patient iPSC-CNs as well as the C9orf72 knockout iPSC-CNs at different stages throughout CN differentiation. For this purpose, a multi-electrode array (MEA) protocol has been optimised for use in iPSC-CNs.

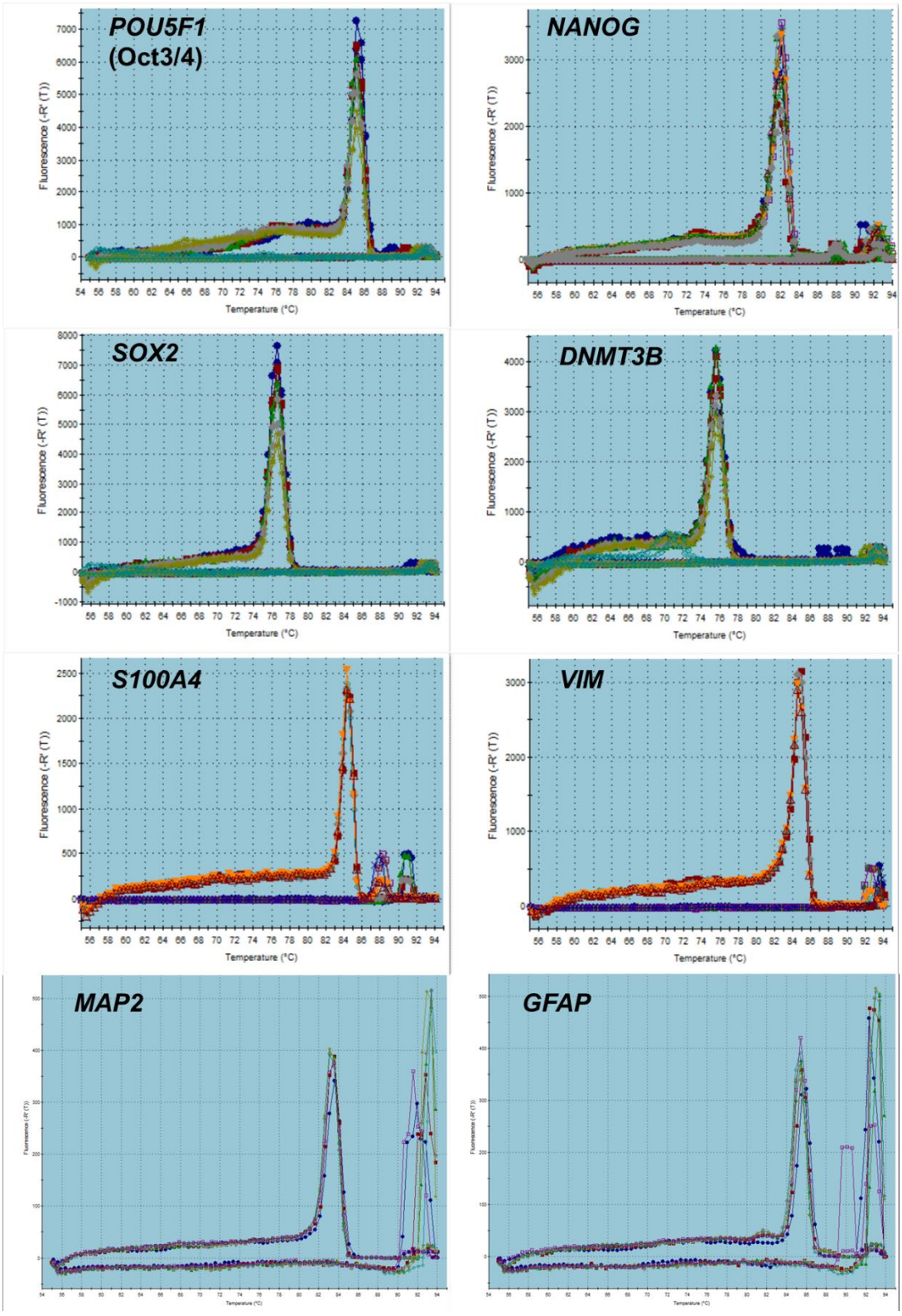
C9orf72 expression is highest in myeloid cells (O'Rourke *et al.*, 2016, Rizzu *et al.*, 2016) and recent evidence support a major function in regulation of innate immune response and inflammation (Burberry *et al.*, 2020, McCauley *et al.*, 2020). Therefore, the study of *C9orf72* iPSC-microglia or other myeloid lineage cells will be very important to address cell-autonomous as well as non-cell-autonomous disease mechanisms. Moreover, study of *C9orf72*$^{+/-}$ and *C9orf72*$^{-/-}$ iPSC-microglia will offer an insight on the effects of C9orf72 loss in these cells.

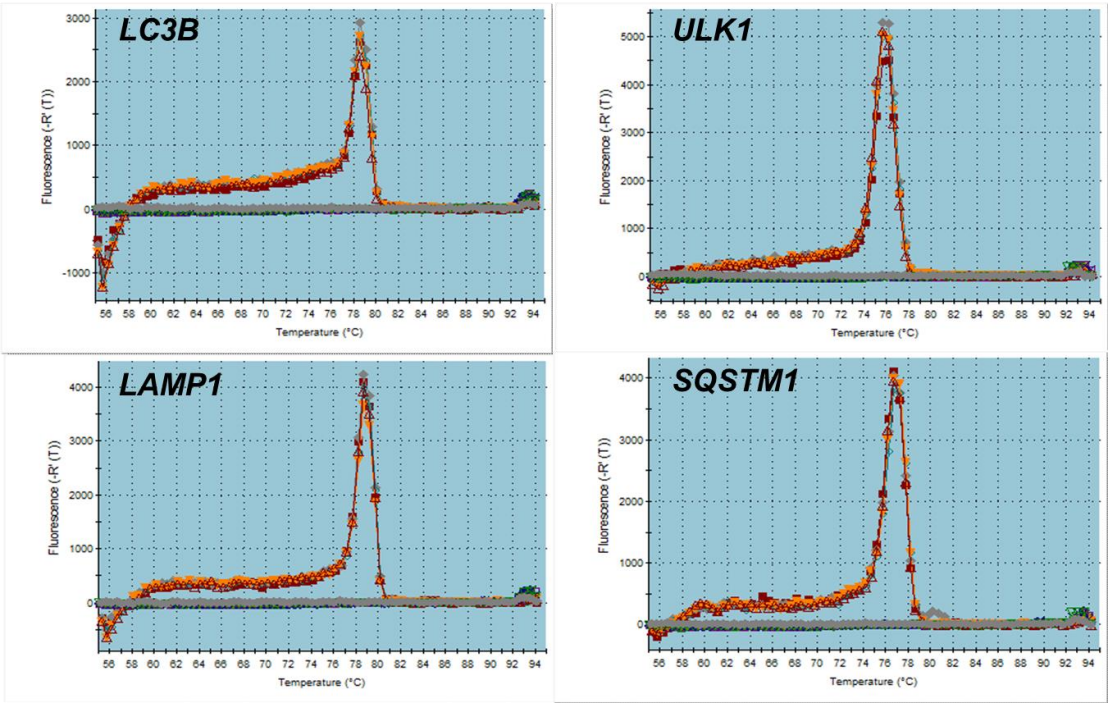
In conlusion, the work presented in this thesis has provided some further insight into the C9-FTD/ALS pathogenic mechanisms in *C9orf72* iPSC-CNs and has identified some novel pathways that are potentially affected by C9orf72 loss in *C9orf72* knockout iPSC-CNs. Future work will be important in confirming these findings.

Appendix A

Primers used for RT-qPCR; sequences, efficiencies and dissociation curves.

Gene	Sequence (5' to 3')	Efficiency (%)
POU5F1 (Oct3/4)	F: TTCTGGCGCCGGTTACAGAACCA R: GACAACAATGAAAATCTTCAGGAGA	90.3
NANOG	F: GCTTGCCTTGCTTTGAAGCA R: TTCTTGACTGGGACCTTGTC	91.8
SOX2	F: CATGGCAATCAAAATGTCCA R: TTTCACGTTTGCAACTGTCC	99.9
DNMT3B	F: TTTAGGGAGAACGGGAAT R: AGCACCAGTAAGAAGAGT	93.4
S100A4	F: TTCTTTCTTGGTTTGATCC R: TTAGTTCTGACTTGTTGAGC	109.7
VIM	F: GTACGTCAGCAATATGAAAG R: AGTGTCTTGGTAGTTAGCAG	103.8
GFAP	F: GTACCAGGACCTGCTCAAT R: CAACTATCCTGCTTCTGCTC	93.5
MAP2	F: CCACCTGAGATTAAGGATCA R: GGCTTACTTTGCTTCTCTGA	106
LC3B	F: AGTAGTTATCACTCTTAGG R: CTCTTGACATTAGTATCTG	95.5
ULK1	F: GTCACACGCCACATAACAG R: TCTTCTAAGTCCAAGCACAGT	98.7
LAMP1	F: CATACACTCACTCTCAAT R: ATGTCAGTTATAGATTCCA	99.6
SQSTM1	F: CCTCCTAACAAGTGTATCTC R: ACGACTATGTGACCTCTT	100.5
TBP	F: TGCACAGGAGCCAAGAGTGAA R: CACATCACAGCTCCCCACCA	99.9
ACTB	F: GAGCACAGAGCCTCGCCTTT R: TCATCATCCATGGTGAGCTGG	99.7


POU5F1
(Oct3/4)
NANOG
SOX2
DNMT3B
S100A4
VIM
MAP2
GFAP
Fluorescence (-R' (T))
Temperature (°C)


LC3B
ULK1
LAMP1
SQSTM1
Fluorescence (-R' (T))
Temperature (°C)

Appendix B

ANOVA and Tukey's post hoc *p* values for all experiments in the thesis.

Analysis (Chapter)	ANOVA *p* value	Tukey's p_{adj} value for multiple comparisons
VIM (3.3.4.)	0.0195	CTRL-2-C9-1 0.0495332
GFAP (3.3.4.)	6.17e-09	C9-2-C9-1 0.0000000 C9-3-C9-1 0.0000001 CTRL-1-C9-1 0.0000000 CTRL-2-C9-1 0.0000000 CTRL-3-C9-1 0.0000000 CTRL-2-C9-3 0.0239185
C9orf72 total mRNA (4.2.5.1.)	1.1e-05	C9-2-C9-1 0.0006728 C9-3-C9-1 0.0111321 CTRL-1-C9-2 0.0000675 CTRL-2-C9-2 0.0001627 CTRL-3-C9-2 0.0001372 CTRL-1-C9-3 0.0009901 CTRL-2-C9-3 0.0025312 CTRL-3-C9-3 0.0021122
C9orf72 V1 semi-quantitative RT-PCR (4.2.5.2.)	7.04e-05	C9-2-C9-1 0.0002344 C9-3-C9-1 0.0026361 CTRL-1-C9-1 0.0001388 CTRL-2-C9-1 0.0000860 CTRL-3-C9-1 0.0005303
C9orf72 V3 semi-quantitative RT-PCR (4.2.5.2.)	8.09e-05	C9-2-C9-1 0.0001542 C9-3-C9-1 0.0008419 CTRL-1-C9-1 0.0002370 CTRL-2-C9-1 0.0000998 CTRL-3-C9-1 0.0009725
C9orf72 V3 LNA RT-qPCR (4.2.5.2.)		C9-2-C9-1 0.0005357 C9-3-C9-1 0.0072516 CTRL-1-C9-1 0.0004847 CTRL-2-C9-1 0.0001816 CTRL-3-C9-1 0.0007249
C9-L protein in CRISPR-iPSCs (5.3.3.4.)	7.35e-08	HOM1-HET1 0.0005088 HOM2-HET1 0.0002780 HOM3-HET1 0.0001592 WT-HET1 0.0018163 HOM1-HET2 0.0124378 HOM2-HET2 0.0064015 HOM3-HET2 0.0034349 WT-HET2 0.0000906 HOM1-HET3 0.0041289 HOM2-HET3 0.0021535 HOM3-HET3 0.0011763 WT-HET3 0.0002375 WT-HOM1 0.0000004 WT-HOM2 0.0000002 WT-HOM3 0.0000002

Analysis (Chapter)	ANOVA *p* value	Tukey's p_{adj} value for multiple comparisons
C9-L protein in CRISPR-iPSC-CNs (5.3.4.2.)	2.8e-09	HOM1-HET1 0.0000073 HOM2-HET1 0.0000069 HOM3-HET1 0.0000046 WT-HET1 0.0105514 HOM1-HET2 0.0000022 HOM2-HET2 0.0000021 HOM3-HET2 0.0000014 WT-HET2 0.0569028 HOM1-HET3 0.0000404 HOM2-HET3 0.0000375 HOM3-HET3 0.0000241 WT-HET3 0.0012268 WT-HOM1 0.0000001 WT-HOM2 0.0000001 WT-HOM3 0.0000000
ULK1 (5.3.5.)	0.0113	HOM-HET 0.0193970 WT-HOM 0.0498060
LC3B (5.3.5)	0.0267	WT-HOM 0.0435562
LAMP1 (5.3.5.)	0.00805	HOM-HET 0.0076868

References

Ababneh, N. A., Scaber, J., Flynn, R., Douglas, A., Barbagallo, P.*, et al.* 2020. Correction of amyotrophic lateral sclerosis related phenotypes in induced pluripotent stem cell-derived motor neurons carrying a hexanucleotide expansion mutation in C9orf72 by CRISPR/Cas9 genome editing using homology-directed repair. *Human Molecular Genetics,* 29**,** 2200-2217.

Abo-Rady, M., Kalmbach, N., Pal, A., Schludi, C., Janosch, A.*, et al.* 2020. Knocking out C9ORF72 Exacerbates Axonal Trafficking Defects Associated with Hexanucleotide Repeat Expansion and Reduces Levels of Heat Shock Proteins. *Stem Cell Reports,* 14**,** 390-405.

Abu Diab, M., Mor-Shaked, H., Cohen, E., Cohen-Hadad, Y., Ram, O.*, et al.* 2018. The G-rich Repeats in FMR1 and C9orf72 Loci Are Hotspots for Local Unpairing of DNA. *Genetics,* 210**,** 1239-1252.

Almeida, S., Gascon, E., Tran, H., Chou, H. J., Gendron, T. F.*, et al.* 2013. Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons. *Acta Neuropathol,* 126**,** 385-99.

Almeida, S., Krishnan, G., Rushe, M., Gu, Y., Kankel, M. W.*, et al.* 2019. Production of poly(GA) in C9ORF72 patient motor neurons derived from induced pluripotent stem cells. *Acta Neuropathol,* 138**,** 1099-1101.

Amick, J., Roczniak-Ferguson, A. & Ferguson, S. M. 2016. C9orf72 binds SMCR8, localizes to lysosomes, and regulates mTORC1 signaling. *Mol Biol Cell,* 27**,** 3040-3051.

Amps, K., Andrews, P. W., Anyfantis, G., Armstrong, L., Avery, S.*, et al.* 2011. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nature biotechnology,* 29**,** 1132.

Andrade, N. S., Ramic, M., Esanov, R., Liu, W., Rybin, M. J.*, et al.* 2020. Dipeptide repeat proteins inhibit homology-directed DNA double strand break repair in C9ORF72 ALS/FTD. *Molecular Neurodegeneration,* 15**,** 13.

Aoki, Y., Manzano, R., Lee, Y., Dafinca, R., Aoki, M.*, et al.* 2017. C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia. *Brain,* 140**,** 887-897.

Ash, P. E., Bieniek, K. F., Gendron, T. F., Caulfield, T., Lin, W. L.*, et al.* 2013. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron,* 77**,** 639-46.

Assou, S., Girault, N., Plinet, M., Bouckenheimer, J., Sansac, C.*, et al.* 2020. Recurrent Genetic Abnormalities in Human Pluripotent Stem Cells: Definition and Routine Detection in Culture Supernatant by Targeted Droplet Digital PCR. *Stem cell reports,* 14**,** 1-8.

Atanasio, A., Decman, V., White, D., Ramos, M., Ikiz, B.*, et al.* 2016. C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci Rep,* 6**,** 23204.

Baborie, A., Griffiths, T. D., Jaros, E., Perry, R., Mckeith, I. G.*, et al.* 2014. Accumulation of dipeptide repeat proteins predates that of TDP-43 in frontotemporal lobar degeneration associated with hexanucleotide repeat expansions in C9ORF72 gene. *Neuropathol Appl Neurobiol,* 41**,** 601-12.

Baker, J. C., Beddington, R. S. & Harland, R. M. 1999. Wnt signaling in Xenopus embryos inhibits bmp4 expression and activates neural development. *Genes Dev,* 13**,** 3149-59.

Baker, M., Mackenzie, I. R., Pickering-Brown, S. M., Gass, J., Rademakers, R.*, et al.* 2006. Mutations in progranulin cause tau-negative frontotemporal dementia linked to chromosome 17. *Nature,* 442**,** 916-9.

Bang, J., Spina, S. & Miller, B. L. 2015. Frontotemporal dementia. *The Lancet,* 386**,** 1672-1682.

Bardelli, D., Sassone, F., Colombrita, C., Volpe, C., Gumina, V.*, et al.* 2020. Reprogramming fibroblasts and peripheral blood cells from a C9ORF72 patient: A proof-of-principle study. *Journal of Cellular and Molecular Medicine,* 24**,** 4051-4060.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P.*, et al.* 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science,* 315**,** 1709-1712.

Bauer, P. O. 2016. Methylation of C9orf72 expansion reduces RNA foci formation and dipeptide-repeat proteins expression in cells. *Neuroscience Letters,* 612**,** 204-209.

Bauer, P. O., Dunmore, J. H., Sasaguri, H. & Matoska, V. 2019. Neurons Induced From Fibroblasts of c9ALS/FTD Patients Reproduce the Pathology Seen in the Central Nervous System. *Frontiers in Neuroscience,* 13.

Beaulieu, C. 1993. Numerical data on neocortical neurons in adult rat, with special reference to the GABA population. *Brain research,* 609**,** 284-292.

Beck, J., Poulter, M., Hensman, D., Rohrer, J. D., Mahoney, C. J.*, et al.* 2013. Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am J Hum Genet,* 92**,** 345-53.

Bell, M. V., Hirst, M. C., Nakahori, Y., Mackinnon, R. N., Roche, A.*, et al.* 1991. Physical mapping across the fragile X: hypermethylation and clinical expression of the fragile X syndrome. *Cell,* 64**,** 861-6.

Belzil, V. V., Bauer, P. O., Gendron, T. F., Murray, M. E., Dickson, D.*, et al.* 2014. Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Res,* 1584**,** 15-21.

Belzil, V. V., Bauer, P. O., Prudencio, M., Gendron, T. F., Stetler, C. T.*, et al.* 2013. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol,* 126**,** 895-905.

Benussi, A., Alberici, A., Buratti, E., Ghidoni, R., Gardoni, F.*, et al.* 2019. Toward a Glutamate Hypothesis of Frontotemporal Dementia. *Frontiers in neuroscience,* 13**,** 304-304.

Bibikova, M., Beumer, K., Trautman, J. K. & Carroll, D. 2003. Enhancing gene targeting with designed zinc finger nucleases. *Science,* 300**,** 764-764.

Bird, T., Knopman, D., Vanswieten, J., Rosso, S., Feldman, H.*, et al.* 2003. Epidemiology and genetics of frontotemporal dementia/Pick's disease. *Ann Neurol,* 54 Suppl 5**,** S29-31.

Birger, A., Ben-Dor, I., Ottolenghi, M., Turetsky, T., Gil, Y.*, et al.* 2019. Human iPSC-derived astrocytes from ALS patients with mutated C9ORF72 show increased oxidative stress and neurotoxicity. *EBioMedicine,* 50**,** 274-289.

Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. 1998. FokI dimerization is required for DNA cleavage. *Proceedings of the national academy of sciences,* 95**,** 10570-10575.

Blokhuis, A. M., Koppers, M., Groen, E. J. N., Van Den Heuvel, D. M. A., Dini Modigliani, S.*, et al.* 2016. Comparative interactomics analysis of different ALS-associated proteins identifies converging molecular pathways. *Acta Neuropathologica,* 132**,** 175-196.

Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S.*, et al.* 2009. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science,* 326**,** 1509-1512.

Boeve, B. F., Boylan, K. B., Graff-Radford, N. R., Dejesus-Hernandez, M., Knopman, D. S.*, et al.* 2012. Characterization of frontotemporal dementia and/or

amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. *Brain,* 135**,** 765-83.
Boeynaems, S., Bogaert, E., Kovacs, D., Konijnenberg, A., Timmerman, E.*, et al.* 2017. Phase separation of C9orf72 dipeptide repeats perturbs stress granule dynamics. *Molecular cell,* 65**,** 1044-1055. e5.
Bonnard, M., Mirtsos, C., Suzuki, S., Graham, K., Huang, J.*, et al.* 2000. Deficiency of T2K leads to apoptotic liver degeneration and impaired NF-kappaB-dependent gene transcription. *EMBO J,* 19**,** 4976-85.
Boxer, A. L., Mackenzie, I. R., Boeve, B. F., Baker, M., Seeley, W. W.*, et al.* 2011. Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *J Neurol Neurosurg Psychiatry,* 82**,** 196-203.
Brady, O. A., Zheng, Y., Murphy, K., Huang, M. & Hu, F. 2013. The frontotemporal lobar degeneration risk factor, TMEM106B, regulates lysosomal morphology and function. *Human molecular genetics,* 22**,** 685-695.
Brykczynska, U., Pecho-Vrieseling, E., Thiemeyer, A., Klein, J., Fruh, I.*, et al.* 2016. CGG Repeat-Induced FMR1 Silencing Depends on the Expansion Size in Human iPSCs and Neurons Carrying Unmethylated Full Mutations. *Stem Cell Reports,* 7**,** 1059-1071.
Burberry, A., Suzuki, N., Wang, J. Y., Moccia, R., Mordes, D. A.*, et al.* 2016. Loss-of-function mutations in the C9ORF72 mouse ortholog cause fatal autoimmune disease. *Sci Transl Med,* 8**,** 347ra93.
Burberry, A., Wells, M. F., Limone, F., Couto, A., Smith, K. S.*, et al.* 2020. C9orf72 suppresses systemic and neural inflammation induced by gut bacteria. *Nature,* 582**,** 89-94.
Burke, E. E., Chenoweth, J. G., Shin, J. H., Collado-Torres, L., Kim, S.-K.*, et al.* 2020. Dissecting transcriptomic signatures of neuronal differentiation and maturation using iPSCs. *Nature Communications,* 11**,** 462.
Butti, Z. & Patten, S. A. 2019. RNA Dysregulation in Amyotrophic Lateral Sclerosis. *Frontiers in Genetics,* 9.
Byrne, S., Elamin, M., Bede, P., Shatunov, A., Walsh, C.*, et al.* 2012. Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: a population-based cohort study. *Lancet Neurol,* 11**,** 232-40.
Cacace, R., Van Cauwenberghe, C., Bettens, K., Gijselinck, I., Van Der Zee, J.*, et al.* 2013. C9orf72 G4C2 repeat expansions in Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging,* 34**,** 1712.e1-7.
Cairns, N. J., Bigio, E. H., Mackenzie, I. R., Neumann, M., Lee, V. M.*, et al.* 2007. Neuropathologic diagnostic and nosologic criteria for frontotemporal lobar degeneration: consensus of the Consortium for Frontotemporal Lobar Degeneration. *Acta Neuropathol,* 114**,** 5-22.
Castaldo, I., Pinelli, M., Monticelli, A., Acquaviva, F., Giacchetti, M.*, et al.* 2008. DNA methylation in intron 1 of the frataxin gene is related to GAA repeat length and age of onset in Friedreich ataxia patients. *J Med Genet,* 45**,** 808-12.
Chai, N., Haney, M. S., Couthouis, J., Morgens, D. W., Benjamin, A.*, et al.* 2019. Genome-wide synthetic lethal CRISPR screen identifies FIS1 as a genetic interactor of ALS-linked C9ORF72. *Brain Res,* 1728**,** 146601.
Chambers, S. M., Fasano, C. A., Papapetrou, E. P., Tomishima, M., Sadelain, M.*, et al.* 2009. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol,* 27**,** 275-80.
Chen, Y., Lin, Z., Chen, X., Cao, B., Wei, Q.*, et al.* 2016. Large C9orf72 repeat expansions are seen in Chinese patients with sporadic amyotrophic lateral sclerosis. *Neurobiol Aging,* 38**,** 217 e15-217 e22.
Cheng, W., Wang, S., Zhang, Z., Morgens, D. W., Hayes, L. R.*, et al.* 2019. CRISPR-Cas9 Screens Identify the RNA Helicase DDX3X as a Repressor of C9ORF72

(GGGGCC)n Repeat-Associated Non-AUG Translation. *Neuron,* 104**,** 885-898 e8.
Chew, J., Gendron, T. F., Prudencio, M., Sasaguri, H., Zhang, Y. J.*, et al.* 2015. Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science,* 348**,** 1151-4.
Chiang, H. H., Forsell, C., Lindstrom, A. K., Lilius, L., Thonberg, H.*, et al.* 2017. No common founder for C9orf72 expansion mutation in Sweden. *J Hum Genet,* 62**,** 321-324.
Chio, A., Borghero, G., Restagno, G., Mora, G., Drepper, C.*, et al.* 2012. Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. *Brain,* 135**,** 784-93.
Chitiprolu, M., Jagow, C., Tremblay, V., Bondy-Chorney, E., Paris, G.*, et al.* 2018. A complex of C9ORF72 and p62 uses arginine methylation to eliminate stress granules by autophagy. *Nature Communications,* 9**,** 2794.
Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S.*, et al.* 2014. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome research,* 24**,** 132-141.
Chow, C. Y., Landers, J. E., Bergren, S. K., Sapp, P. C., Grant, A. E.*, et al.* 2009. Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. *The American Journal of Human Genetics,* 84**,** 85-88.
Chow, T. W., Miller, B. L., Hayashi, V. N. & Geschwind, D. H. 1999. Inheritance of frontotemporal dementia. *Arch Neurol,* 56**,** 817-22.
Cirulli, E. T., Lasseigne, B. N., Petrovski, S., Sapp, P. C., Dion, P. A.*, et al.* 2015. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science,* 347**,** 1436-1441.
Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H.*, et al.* 2013. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann Neurol,* 74**,** 180-7.
Cohen-Hadad, Y., Altarescu, G., Eldar-Geva, T., Levi-Lahad, E., Zhang, M.*, et al.* 2016. Marked Differences in C9orf72 Methylation Status and Isoform Expression between C9/ALS Human Embryonic and Induced Pluripotent Stem Cells. *Stem Cell Reports,* 7**,** 927-940.
Colak, D., Zaninovic, N., Cohen, M. S., Rosenwaks, Z., Yang, W.-Y.*, et al.* 2014. Promoter-Bound Trinucleotide Repeat mRNA Drives Epigenetic Silencing in Fragile X Syndrome. *Science,* 343**,** 1002-1005.
Cooper-Knock, J., Walsh, M. J., Higginbottom, A., Robin Highley, J., Dickman, M. J.*, et al.* 2014. Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain,* 137**,** 2040-51.
Corbier, C. & Sellier, C. 2016. C9ORF72 is a GDP/GTP exchange factor for Rab8 and Rab39 and regulates autophagy. *Small GTPases***,** 1-6.
Coyne, A. N., Lorenzini, I., Chou, C. C., Torvund, M., Rogers, R. S.*, et al.* 2017. Post-transcriptional Inhibition of Hsc70-4/HSPA8 Expression Leads to Synaptic Vesicle Cycling Defects in Multiple Models of ALS. *Cell Rep,* 21**,** 110-125.
Cradick, T. J., Fine, E. J., Antico, C. J. & Bao, G. 2013. CRISPR/Cas9 systems targeting β-globin and CCR5 genes have substantial off-target activity. *Nucleic Acids Res,* 41**,** 9584-92.
Cruchaga, C., Graff, C., Chiang, H. H., Wang, J., Hinrichs, A. L.*, et al.* 2011. Association of TMEM106B gene polymorphism with age at onset in granulin mutation carriers and plasma granulin protein levels. *Arch Neurol,* 68**,** 581-6.
Cruts, M., Gijselinck, I., Van Der Zee, J., Engelborghs, S., Wils, H.*, et al.* 2006. Null mutations in progranulin cause ubiquitin-positive frontotemporal dementia linked to chromosome 17q21. *Nature,* 442**,** 920-4.
Dafinca, R., Scaber, J., Ababneh, N., Lalic, T., Weir, G.*, et al.* 2016. 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 Sclerosis and Frontotemporal Dementia. *Stem Cells,* 34**,** 2063-78.

Davidson, Y., Robinson, A. C., Liu, X., Wu, D., Troakes, C.*, et al.* 2016. Neurodegeneration in frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9orf72 is linked to TDP-43 pathology and not associated with aggregated forms of dipeptide repeat proteins. *Neuropathol Appl Neurobiol,* 42**,** 242-54.

De Esch, C. E., Ghazvini, M., Loos, F., Schelling-Kazaryan, N., Widagdo, W.*, et al.* 2014. Epigenetic characterization of the FMR1 promoter in induced pluripotent stem cells from human fibroblasts carrying an unmethylated full mutation. *Stem Cell Reports,* 3**,** 548-55.

De Vos, K. J., Mórotz, G. M., Stoica, R., Tudor, E. L., Lau, K.-F.*, et al.* 2011. VAPB interacts with the mitochondrial protein PTPIP51 to regulate calcium homeostasis. *Human Molecular Genetics,* 21**,** 1299-1311.

Deas, E., Plun-Favreau, H., Gandhi, S., Desmond, H., Kjaer, S.*, et al.* 2010. PINK1 cleavage at position A103 by the mitochondrial protease PARL. *Human Molecular Genetics,* 20**,** 867-879.

Dejesus-Hernandez, M., Finch, N. A., Wang, X., Gendron, T. F., Bieniek, K. F.*, et al.* 2017. In-depth clinico-pathological examination of RNA foci in a large cohort of C9ORF72 expansion carriers. *Acta Neuropathol,* 134**,** 255-269.

Dejesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M.*, et al.* 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron,* 72**,** 245-56.

Deltcheva, E., Chylinski, K., Sharma, C. M., Gonzales, K., Chao, Y.*, et al.* 2011. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature,* 471**,** 602-607.

Deng, H. X., Chen, W., Hong, S. T., Boycott, K. M., Gorrie, G. H.*, et al.* 2011. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature,* 477**,** 211-5.

Devenney, E. M., Ahmed, R. M., Halliday, G., Piguet, O., Kiernan, M. C.*, et al.* 2018. Psychiatric disorders in C9orf72 kindreds. *Study of 1,414 family members,* 91**,** e1498-e1507.

Devlin, A. C., Burr, K., Borooah, S., Foster, J. D., Cleary, E. M.*, et al.* 2015. Human iPSC-derived motoneurons harbouring TARDBP or C9ORF72 ALS mutations are dysfunctional despite maintaining viability. *Nat Commun,* 6**,** 5999.

Devon, R. S., Orban, P. C., Gerrow, K., Barbieri, M. A., Schwab, C.*, et al.* 2006. Als2-deficient mice exhibit disturbances in endosome trafficking associated with motor behavioral abnormalities. *Proc Natl Acad Sci U S A,* 103**,** 9595-600.

Dobson-Stone, C., Hallupp, M., Bartley, L., Shepherd, C. E., Halliday, G. M.*, et al.* 2012. C9ORF72 repeat expansion in clinical and neuropathologic frontotemporal dementia cohorts. *Neurology,* 79**,** 995-1001.

Donnelly, C. J., Zhang, P. W., Pham, J. T., Haeusler, A. R., Mistry, N. A.*, et al.* 2013. RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron,* 80**,** 415-28.

Eiraku, M., Watanabe, K., Matsuo-Takasaki, M., Kawada, M., Yonemura, S.*, et al.* 2008. Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell,* 3**,** 519-32.

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N. D., Tabar, V.*, et al.* 2008. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes & development,* 22**,** 152-165.

Esanov, R., Belle, K. C., Van Blitterswijk, M., Belzil, V. V., Rademakers, R.*, et al.* 2016. C9orf72 promoter hypermethylation is reduced while hydroxymethylation is acquired during reprogramming of ALS patient cells. *Exp Neurol,* 277**,** 171-177.

Farg, M. A., Sundaramoorthy, V., Sultana, J. M., Yang, S., Atkinson, R. A*., et al.* 2014. C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum Mol Genet,* 23**,** 3579-95.

Fecto, F., Yan, J., Vemula, S. P., Liu, E., Yang, Y*., et al.* 2011. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Archives of neurology,* 68**,** 1440-1446.

Ferguson, C. J., Lenk, G. M. & Meisler, M. H. 2009. Defective autophagy in neurons and astrocytes from mice deficient in PI (3, 5) P2. *Human molecular genetics,* 18**,** 4868-4878.

Ferraiuolo, L., Meyer, K., Sherwood, T. W., Vick, J., Likhite, S*., et al.* 2016. Oligodendrocytes contribute to motor neuron death in ALS via SOD1-dependent mechanism. *Proceedings of the National Academy of Sciences,* 113**,** E6496-E6505.

Fratta, P., Mizielinska, S., Nicoll, A. J., Zloh, M., Fisher, E. M. C*., et al.* 2012. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Scientific Reports,* 2**,** 1016.

Fratta, P., Polke, J. M., Newcombe, J., Mizielinska, S., Lashley, T*., et al.* 2015. Screening a UK amyotrophic lateral sclerosis cohort provides evidence of multiple origins of the C9orf72 expansion. *Neurobiol Aging,* 36**,** 546 e1-7.

Fratta, P., Poulter, M., Lashley, T., Rohrer, J. D., Polke, J. M*., et al.* 2013. Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. *Acta Neuropathol,* 126**,** 401-9.

Fredi, M., Cavazzana, I., Biasiotto, G., Filosto, M., Padovani, A*., et al.* 2019. C9orf72 Intermediate Alleles in Patients with Amyotrophic Lateral Sclerosis, Systemic Lupus Erythematosus, and Rheumatoid Arthritis. *Neuromolecular Med,* 21**,** 150-159.

Freibaum, B. D., Lu, Y., Lopez-Gonzalez, R., Kim, N. C., Almeida, S*., et al.* 2015. GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature,* 525**,** 129-33.

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V*., et al.* 2015. Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nature Neuroscience,* 18**,** 631-636.

Frick, P., Sellier, C., Mackenzie, I. R. A., Cheng, C.-Y., Tahraoui-Bories, J*., et al.* 2018. Novel antibodies reveal presynaptic localization of C9orf72 protein and reduced protein levels in C9orf72 mutation carriers. *Acta Neuropathologica Communications,* 6**,** 72.

Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D*., et al.* 2013. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nature biotechnology,* 31**,** 822-826.

Fu, Y., Sander, J. D., Reyon, D., Cascio, V. M. & Joung, J. K. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nature biotechnology,* 32**,** 279-284.

Gallagher, M. D., Suh, E., Grossman, M., Elman, L., Mccluskey, L*., et al.* 2014. TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. *Acta Neuropathol,* 127**,** 407-18.

Gami, P., Murray, C., Schottlaender, L., Bettencourt, C., De Pablo Fernandez, E*., et al.* 2015. A 30-unit hexanucleotide repeat expansion in C9orf72 induces pathological lesions with dipeptide-repeat proteins and RNA foci, but not TDP-43 inclusions and clinical disease. *Acta Neuropathol,* 130**,** 599-601.

Gao, F., Zhang, Y., Hou, X., Tao, Z., Ren, H*., et al.* 2020. Dependence of PINK1 accumulation on mitochondrial redox system. *Aging Cell,* 19**,** e13211.

Gaspard, N., Bouschet, T., Hourez, R., Dimidschstein, J., Naeije, G*., et al.* 2008. An intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature,* 455**,** 351-357.

Gecz, J., Gedeon, A. K., Sutherland, G. R. & Mulley, J. C. 1996. Identification of the gene FMR2, associated with FRAXE mental retardation. *Nature genetics,* 13**,** 105-108.

Gegg, M. E., Cooper, J. M., Chau, K. Y., Rojo, M., Schapira, A. H.*, et al.* 2010. Mitofusin 1 and mitofusin 2 are ubiquitinated in a PINK1/parkin-dependent manner upon induction of mitophagy. *Hum Mol Genet,* 19**,** 4861-70.

Gendron, T. F., Bieniek, K. F., Zhang, Y. J., Jansen-West, K., Ash, P. E.*, et al.* 2013. Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol,* 126**,** 829-44.

Gijselinck, I., Van Langenhove, T., Van Der Zee, J., Sleegers, K., Philtjens, S.*, et al.* 2012. A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol,* 11**,** 54-65.

Gijselinck, I., Van Mossevelde, S., Van Der Zee, J., Sieben, A., Engelborghs, S.*, et al.* 2016. The C9orf72 repeat size correlates with onset age of disease, DNA methylation and transcriptional downregulation of the promoter. *Mol Psychiatry,* 21**,** 1112-24.

Gomez-Deza, J., Lee, Y. B., Troakes, C., Nolan, M., Al-Sarraj, S.*, et al.* 2015. Dipeptide repeat protein inclusions are rare in the spinal cord and almost absent from motor neurons in C9ORF72 mutant amyotrophic lateral sclerosis and are unlikely to cause their degeneration. *Acta Neuropathol Commun,* 3**,** 38.

Gomez-Suaga, P., Paillusson, S., Stoica, R., Noble, W., Hanger, D. P.*, et al.* 2017. The ER-Mitochondria Tethering Complex VAPB-PTPIP51 Regulates Autophagy. *Current Biology,* 27**,** 371-385.

Gomez-Tortosa, E., Gallego, J., Guerrero-Lopez, R., Marcos, A., Gil-Neciga, E.*, et al.* 2013. C9ORF72 hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. *Neurology,* 80**,** 366-70.

Gomez Perdiguero, E., Klapproth, K., Schulz, C., Busch, K., Azzoni, E.*, et al.* 2015. Tissue-resident macrophages originate from yolk-sac-derived erythro-myeloid progenitors. *Nature,* 518**,** 547-551.

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L.*, et al.* 2004. Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol,* 55**,** 335-46.

Gu, K., Yang, B., Tian, D., Wu, L., Wang, D.*, et al.* 2005. R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature,* 435**,** 1122-1125.

Gu, Y., Shen, Y., Gibbs, R. A. & Nelson, D. L. 1996. Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. *Nature genetics,* 13**,** 109-113.

Guilinger, J. P., Thompson, D. B. & Liu, D. R. 2014. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nature biotechnology,* 32**,** 577.

Hadano, S., Hand, C. K., Osuga, H., Yanagisawa, Y., Otomo, A.*, et al.* 2001. A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat Genet,* 29**,** 166-73.

Haeusler, A. R., Donnelly, C. J., Periz, G., Simko, E. a. J., Shaw, P. G.*, et al.* 2014. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature,* 507**,** 195-200.

Hammaker, D., Boyle, D. L. & Firestein, G. S. 2011. Synoviocyte innate immune responses: TANK-binding kinase-1 as a potential therapeutic target in rheumatoid arthritis. *Rheumatology,* 51**,** 610-618.

Hansen, D. V., Lui, J. H., Flandin, P., Yoshikawa, K., Rubenstein, J. L.*, et al.* 2013. Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. *Nature Neuroscience,* 16**,** 1576-1587.

Hao, Z., Liu, L., Tao, Z., Wang, R., Ren, H.*, et al.* 2019. Motor dysfunction and neurodegeneration in a C9orf72 mouse line expressing poly-PR. *Nature Communications,* 10**,** 2906.

Harms, M., Benitez, B. A., Cairns, N., Cooper, B., Cooper, P.*, et al.* 2013. C9orf72 hexanucleotide repeat expansions in clinical Alzheimer disease. *JAMA Neurol,* 70**,** 736-41.

Hendry, S. H., Schwark, H. D., Jones, E. G. & Yan, J. 1987. Numbers and proportions of GABA-immunoreactive neurons in different areas of monkey cerebral cortex. *Journal of Neuroscience,* 7**,** 1503-1519.

Heo, J.-M., Ordureau, A., Paulo, Joao a., Rinehart, J. & Harper, J. W. 2015. The PINK1-PARKIN Mitochondrial Ubiquitylation Pathway Drives a Program of OPTN/NDP52 Recruitment and TBK1 Activation to Promote Mitophagy. *Molecular Cell,* 60**,** 7-20.

Herman, M., Ciancanelli, M., Ou, Y. H., Lorenzo, L., Klaudel-Dreszler, M.*, et al.* 2012. Heterozygous TBK1 mutations impair TLR3 immunity and underlie herpes simplex encephalitis of childhood. *J Exp Med,* 209**,** 1567-82.

Herranz-Martin, S., Chandran, J., Lewis, K., Mulcahy, P., Higginbottom, A.*, et al.* 2017. Viral delivery of C9orf72 hexanucleotide repeat expansions in mice leads to repeat-length-dependent neuropathology and behavioural deficits. *Dis Model Mech,* 10**,** 859-868.

Hjerpe, R., Bett, J. S., Keuss, M. J., Solovyova, A., Mcwilliams, T. G.*, et al.* 2016. UBQLN2 mediates autophagy-independent protein aggregate clearance by the proteasome. *Cell,* 166**,** 935-949.

Ho, W. Y., Tai, Y. K., Chang, J. C., Liang, J., Tyan, S. H.*, et al.* 2019. The ALS-FTD-linked gene product, C9orf72, regulates neuronal morphogenesis via autophagy. *Autophagy***,** 1-16.

Holm, I. E., Englund, E., Mackenzie, I. R., Johannsen, P. & Isaacs, A. M. 2007. A reassessment of the neuropathology of frontotemporal dementia linked to chromosome 3. *J Neuropathol Exp Neurol,* 66**,** 884-91.

Holm, I. E., Isaacs, A. M. & Mackenzie, I. R. A. 2009. Absence of FUS-immunoreactive pathology in frontotemporal dementia linked to chromosome 3 (FTD-3) caused by mutation in the CHMP2B gene. *Acta Neuropathologica,* 118**,** 719.

Houck, S. A., Ren, H. Y., Madden, V. J., Bonner, J. N., Conlin, M. P.*, et al.* 2014. Quality control autophagy degrades soluble ERAD-resistant conformers of the misfolded membrane protein GnRHR. *Molecular cell,* 54**,** 166-179.

Hsiung, G. Y., Dejesus-Hernandez, M., Feldman, H. H., Sengdy, P., Bouchard-Kerr, P.*, et al.* 2012. Clinical and pathological features of familial frontotemporal dementia caused by C9ORF72 mutation on chromosome 9p. *Brain,* 135**,** 709-22.

Hutton, M., Lendon, C. L., Rizzu, P., Baker, M., Froelich, S.*, et al.* 1998. Association of missense and 5'-splice-site mutations in tau with the inherited dementia FTDP-17. *Nature,* 393**,** 702-5.

Iacoangeli, A., Al Khleifat, A., Jones, A. R., Sproviero, W., Shatunov, A.*, et al.* 2019. C9orf72 intermediate expansions of 24-30 repeats are associated with ALS. *Acta Neuropathol Commun,* 7**,** 115.

Imamura, K., Izumi, Y., Watanabe, A., Tsukita, K., Woltjen, K.*, et al.* 2017. The Src/c-Abl pathway is a potential therapeutic target in amyotrophic lateral sclerosis. *Sci Transl Med,* 9.

Inoue, H., Nojima, H. & Okayama, H. 1990. High efficiency transformation of Escherichia coli with plasmids. *Gene,* 96**,** 23-8.

Ishikawa, K.-I., Saiki, S., Furuya, N., Imamichi, Y., Tsuboi, Y.*, et al.* 2019. p150glued deficiency impairs effective fusion between autophagosomes and lysosomes due to their redistribution to the cell periphery. *Neuroscience Letters,* 690**,** 181-187.

Jackson, J. L., Finch, N. A., Baker, M. C., Kachergus, J. M., Dejesus-Hernandez, M.*, et al.* 2020. Elevated methylation levels, reduced expression levels, and frequent contractions in a clinical cohort of C9orf72 expansion carriers. *Mol Neurodegener,* 15**,** 7.

Jiang, H., Mankodi, A., Swanson, M. S., Moxley, R. T. & Thornton, C. A. 2004. Myotonic dystrophy type 1 is associated with nuclear foci of mutant RNA, sequestration of muscleblind proteins and deregulated alternative splicing in neurons. *Human Molecular Genetics,* 13**,** 3079-3088.

Jiang, J., Zhu, Q., Gendron, T. F., Saberi, S., Mcalonis-Downes, M.*, et al.* 2016. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron,* 90**,** 535-50.

Jin, S. M., Lazarou, M., Wang, C., Kane, L. A., Narendra, D. P.*, et al.* 2010. Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *Journal of Cell Biology,* 191**,** 933-942.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A.*, et al.* 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science,* 337**,** 816-821.

Jovicic, A., Mertens, J., Boeynaems, S., Bogaert, E., Chai, N.*, et al.* 2015. Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat Neurosci,* 18**,** 1226-9.

Kaivorinne, A. L., Bode, M. K., Paavola, L., Tuominen, H., Kallio, M.*, et al.* 2013. Clinical Characteristics of C9ORF72-Linked Frontotemporal Lobar Degeneration. *Dementia and Geriatric Cognitive Disorders Extra,* 3**,** 251-262.

Kanekura, K., Yagi, T., Cammack, A. J., Mahadevan, J., Kuroda, M.*, et al.* 2016. Poly-dipeptides encoded by the C9ORF72 repeats block global protein translation. *Hum Mol Genet,* 25**,** 1803-13.

Kay, S., Hahn, S., Marois, E., Hause, G. & Bonas, U. 2007. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. *Science,* 318**,** 648-651.

Kiernan, M. C., Vucic, S., Cheah, B. C., Turner, M. R., Eisen, A.*, et al.* 2011. Amyotrophic lateral sclerosis. *The lancet,* 377**,** 942-955.

Kim, D. S., Lee, J. S., Leem, J. W., Huh, Y. J., Kim, J. Y.*, et al.* 2010a. Robust enhancement of neural differentiation from human ES and iPS cells regardless of their innate difference in differentiation propensity. *Stem Cell Rev,* 6**,** 270-81.

Kim, H. & Kim, J.-S. 2014. A guide to genome engineering with programmable nucleases. *Nature Reviews Genetics,* 15**,** 321-334.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R.*, et al.* 2010b. Epigenetic memory in induced pluripotent stem cells. *Nature,* 467**,** 285-290.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R.*, et al.* 2010c. Epigenetic memory in induced pluripotent stem cells. *Nature,* 467**,** 285-90.

Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T.*, et al.* 2016. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature,* 529**,** 490-495.

Kleinstiver, B. P., Prew, M. S., Tsai, S. Q., Topkar, V. V., Nguyen, N. T.*, et al.* 2015. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature,* 523**,** 481-485.

Knight, S. J. L., Flannery, A. V., Hirst, M. C., Campbell, L., Christodoulou, Z.*, et al.* 1993. Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. *Cell,* 74**,** 127-134.

Kohli, M. A., John-Williams, K., Rajbhandary, R., Naj, A., Whitehead, P.*, et al.* 2013. Repeat expansions in the C9ORF72 gene contribute to Alzheimer's disease in Caucasians. *Neurobiol Aging,* 34**,** 1519 e5-12.

Kondapalli, C., Kazlauskaite, A., Zhang, N., Woodroof, H. I., Campbell, D. G.*, et al.* 2012. PINK1 is activated by mitochondrial membrane potential depolarization and stimulates Parkin E3 ligase activity by phosphorylating Serine 65. *Open Biology,* 2**,** 120080.

Koppers, M., Blokhuis, A. M., Westeneng, H. J., Terpstra, M. L., Zundel, C. A.*, et al.* 2015. C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann Neurol,* 78**,** 426-38.

Koressaar, T. & Remm, M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics,* 23**,** 1289-91.

Kovacs, G. G., Murrell, J. R., Horvath, S., Haraszti, L., Majtenyi, K.*, et al.* 2009. TARDBP variation associated with frontotemporal dementia, supranuclear gaze palsy, and chorea. *Movement Disorders,* 24**,** 1842-1847.

Kovacs, G. G., Van Der Zee, J., Hort, J., Kristoferitsch, W., Leitha, T.*, et al.* 2015. Clinicopathological description of two cases with SQSTM1 gene mutation associated with frontotemporal dementia. *Neuropathology,* 36**,** 27-38.

Koyano, F., Okatsu, K., Kosako, H., Tamura, Y., Go, E.*, et al.* 2014. Ubiquitin is phosphorylated by PINK1 to activate parkin. *Nature,* 510**,** 162-166.

Kwiatkowski, T. J., Jr., Bosco, D. A., Leclerc, A. L., Tamrazian, E., Vanderburg, C. R.*, et al.* 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science,* 323**,** 1205-8.

Kwon, I., Xiang, S., Kato, M., Wu, L., Theodoropoulos, P.*, et al.* 2014. Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science,* 345**,** 1139-45.

Kyotani, A., Azuma, Y., Yamamoto, I., Yoshida, H., Mizuta, I.*, et al.* 2015. Knockdown of the Drosophila FIG4 induces deficient locomotive behavior, shortening of motor neuron, axonal targeting aberration, reduction of life span and defects in eye development. *Exp Neurol,* 277**,** 86-95.

Laaksovirta, H., Peuralinna, T., Schymick, J. C., Scholz, S. W., Lai, S. L.*, et al.* 2010. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol,* 9**,** 978-85.

Laflamme, C., Mckeever, P. M., Kumar, R., Schwartz, J., Kolahdouzan, M.*, et al.* 2019. Implementation of an antibody characterization procedure and application to the major ALS/FTD disease gene C9ORF72. *eLife,* 8**,** e48363.

Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P.*, et al.* 2013. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci U S A,* 110**,** E4530-9.

Laird, F. M., Farah, M. H., Ackerley, S., Hoke, A., Maragakis, N.*, et al.* 2008. Motor neuron disease occurring in a mutant dynactin mouse model is characterized by defects in vesicular trafficking. *Journal of Neuroscience,* 28**,** 1997-2005.

Lashley, T., Rohrer, J. D., Mead, S. & Revesz, T. 2015. Review: an update on clinical, genetic and pathological aspects of frontotemporal lobar degenerations. *Neuropathol Appl Neurobiol,* 41**,** 858-81.

Lattante, S., Rouleau, G. A. & Kabashi, E. 2013. TARDBP and FUS mutations associated with amyotrophic lateral sclerosis: summary and update. *Hum Mutat,* 34**,** 812-26.

Lazarou, M., Sliter, D. A., Kane, L. A., Sarraf, S. A., Wang, C.*, et al.* 2015. The ubiquitin kinase PINK1 recruits autophagy receptors to induce mitophagy. *Nature,* 524**,** 309-314.

Le Ber, I. 2013. Genetics of frontotemporal lobar degeneration: an up-date and diagnosis algorithm. *Rev Neurol (Paris),* 169**,** 811-9.

Le Ber, I., Camuzat, A., Guerreiro, R., Bouya-Ahmed, K., Bras, J.*, et al.* 2013. SQSTM1 mutations in French patients with frontotemporal dementia or

frontotemporal dementia with amyotrophic lateral sclerosis. *JAMA neurology,* 70**,** 1403-1410.
Lee, K.-H., Zhang, P., Kim, H. J., Mitrea, D. M., Sarkar, M.*, et al.* 2016. C9orf72 dipeptide repeats impair the assembly, dynamics, and function of membrane-less organelles. *Cell,* 167**,** 774-788. e17.
Lee, V. M., Goedert, M. & Trojanowski, J. Q. 2001. Neurodegenerative tauopathies. *Annu Rev Neurosci,* 24**,** 1121-59.
Lee, Y. B., Chen, H. J., Peres, J. N., Gomez-Deza, J., Attig, J.*, et al.* 2013. Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep,* 5**,** 1178-86.
Leitch, H. G., Mcewen, K. R., Turp, A., Encheva, V., Carroll, T.*, et al.* 2013. Naive pluripotency is associated with global DNA hypomethylation. *Nat Struct Mol Biol,* 20**,** 311-6.
Lesage, S., Le Ber, I., Condroyer, C., Broussolle, E., Gabelle, A.*, et al.* 2013. C9orf72 repeat expansions are a rare genetic cause of parkinsonism. *Brain,* 136**,** 385-91.
Levine, T. P., Daniels, R. D., Gatta, A. T., Wong, L. H. & Hayes, M. J. 2013. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics,* 29**,** 499-503.
Li, H., Ham, A., Ma, T. C., Kuo, S.-H., Kanter, E.*, et al.* 2019. Mitochondrial dysfunction and mitophagy defect triggered by heterozygous GBA mutations. *Autophagy,* 15**,** 113-130.
Liang, F., Han, M., Romanienko, P. J. & Jasin, M. 1998. Homology-directed repair is a major double-strand break repair pathway in mammalian cells. *Proceedings of the National Academy of Sciences,* 95**,** 5172-5177.
Lindquist, S. G., Duno, M., Batbayli, M., Puschmann, A., Braendgaard, H.*, et al.* 2013. Corticobasal and ataxia syndromes widen the spectrum of C9ORF72 hexanucleotide expansion disease. *Clin Genet,* 83**,** 279-83.
Lines, G., Casey, J. M., Preza, E. & Wray, S. 2020. Modelling Frontotemporal Dementia using patient-derived induced pluripotent stem cells. *Mol Cell Neurosci***,** 103553.
Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A.*, et al.* 2013. Global Epigenomic Reconfiguration During Mammalian Brain Development. *Science,* 341**,** 1237905.
Liu, E. Y., Russ, J., Wu, K., Neal, D., Suh, E.*, et al.* 2014. C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. *Acta Neuropathol,* 128**,** 525-41.
Liu, Y., Pattamatta, A., Zu, T., Reid, T., Bardhi, O.*, et al.* 2016. C9orf72 BAC Mouse Model with Motor Deficits and Neurodegenerative Features of ALS/FTD. *Neuron,* 90**,** 521-34.
Liu, Y., Wang, T., Ji, Y. J., Johnson, K., Liu, H.*, et al.* 2018. A C9orf72–CARM1 axis regulates lipid metabolism under glucose starvation-induced nutrient stress. *Genes & Development,* 32**,** 1380-1397.
Livak, K. J. & Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.
Livesey, M. R., Magnani, D., Cleary, E. M., Vasistha, N. A., James, O. T.*, et al.* 2016. Maturation and electrophysiological properties of human pluripotent stem cell-derived oligodendrocytes. *Stem Cells,* 34**,** 1040-53.
Lomen-Hoerth, C., Anderson, T. & Miller, B. 2002. The overlap of amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology,* 59**,** 1077-9.
Lopez-Gonzalez, R., Lu, Y., Gendron, Tania f., Karydas, A., Tran, H.*, et al.* 2016. Poly(GR) in C9ORF72-Related ALS/FTD Compromises Mitochondrial Function and Increases Oxidative Stress and DNA Damage in iPSC-Derived Motor Neurons. *Neuron,* 92**,** 383-391.

Mackenzie, I. R. 2007. The neuropathology and clinical phenotype of FTD with progranulin mutations. *Acta Neuropathol,* 114**,** 49-54.
Mackenzie, I. R., Arzberger, T., Kremmer, E., Troost, D., Lorenzl, S.*, et al.* 2013a. Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathol,* 126**,** 859-79.
Mackenzie, I. R., Frick, P., Grässer, F. A., Gendron, T. F., Petrucelli, L.*, et al.* 2015. Quantitative analysis and clinico-pathological correlations of different dipeptide repeat protein pathologies in C9ORF72 mutation carriers. *Acta Neuropathol,* 130**,** 845-61.
Mackenzie, I. R., Frick, P. & Neumann, M. 2013b. The neuropathology associated with repeat expansions in the C9ORF72 gene. *Acta Neuropathol,* 127**,** 347-57.
Mackenzie, I. R. & Neumann, M. 2016. Molecular neuropathology of frontotemporal dementia: insights into disease mechanisms from postmortem studies. *J Neurochem,* 138 Suppl 1**,** 54-70.
Mackenzie, I. R., Rademakers, R. & Neumann, M. 2010. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol,* 9**,** 995-1007.
Madill, M., Mcdonagh, K., Ma, J., Vajda, A., Mcloughlin, P.*, et al.* 2017. Amyotrophic lateral sclerosis patient iPSC-derived astrocytes impair autophagy via non-cell autonomous mechanisms. *Molecular Brain,* 10**,** 22.
Maharjan, N., Künzli, C., Buthey, K. & Saxena, S. 2017. C9ORF72 Regulates Stress Granule Formation and Its Deficiency Impairs Stress Granule Assembly, Hypersensitizing Cells to Stress. *Molecular Neurobiology,* 54**,** 3062-3077.
Mahoney, C. J., Beck, J., Rohrer, J. D., Lashley, T., Mok, K.*, et al.* 2012. Frontotemporal dementia with the C9ORF72 hexanucleotide repeat expansion: clinical, neuroanatomical and neuropathological features. *Brain,* 135**,** 736-50.
Majounie, E., Renton, A. E., Mok, K., Dopper, E. G., Waite, A.*, et al.* 2012. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol,* 11**,** 323-30.
Mali, P., Aach, J., Stranges, P. B., Esvelt, K. M., Moosburner, M.*, et al.* 2013. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature biotechnology,* 31**,** 833-838.
Mann, D. M., Rollinson, S., Robinson, A., Bennion Callister, J., Thompson, J. C.*, et al.* 2013. Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun,* 1**,** 68.
Marchlik, E., Thakker, P., Carlson, T., Jiang, Z., Ryan, M.*, et al.* 2010. Mice lacking Tbk1 activity exhibit immune cell infiltrates in multiple tissues and increased susceptibility to LPS-induced lethality. *Journal of Leukocyte Biology,* 88**,** 1171-1180.
Marogianni, C., Rikos, D., Provatas, A., Dadouli, K., Ntellas, P.*, et al.* 2019. The role of C9orf72 in neurodegenerative disorders: a systematic review, an updated meta-analysis, and the creation of an online database. *Neurobiol Aging,* 84**,** 238.e25-238.e34.
Martins-Taylor, K., Nisler, B. S., Taapken, S. M., Compton, T., Crandall, L.*, et al.* 2011. Recurrent copy number variations in human induced pluripotent stem cells. *Nat Biotechnol,* 29**,** 488-91.
Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H.*, et al.* 2010. Mutations of optineurin in amyotrophic lateral sclerosis. *Nature,* 465**,** 223-226.
Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J. C., Yakir, B.*, et al.* 2010. Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell,* 7**,** 521-31.

Mccauley, M. E., O'rourke, J. G., Yáñez, A., Markman, J. L., Ho, R.*, et al.* 2020. C9orf72 in myeloid cells suppresses STING-induced inflammation. *Nature,* 585**,** 96-101.

Mckhann, G. M., Albert, M. S., Grossman, M., Miller, B., Dickson, D.*, et al.* 2001. Clinical and pathological diagnosis of frontotemporal dementia: report of the Work Group on Frontotemporal Dementia and Pick's Disease. *Arch Neurol,* 58**,** 1803-9.

Meissner, C., Lorenz, H., Weihofen, A., Selkoe, D. J. & Lemberg, M. K. 2011. The mitochondrial intramembrane protease PARL cleaves human Pink1 to regulate Pink1 trafficking. *Journal of Neurochemistry,* 117**,** 856-867.

Mertens, J., Paquola, A. C. M., Ku, M., Hatch, E., Bohnke, L.*, et al.* 2015. Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects. *Cell Stem Cell,* 17**,** 705-718.

Meyer, K., Ferraiuolo, L., Miranda, C. J., Likhite, S., Mcelroy, S.*, et al.* 2014. Direct conversion of patient fibroblasts demonstrates non-cell autonomous toxicity of astrocytes to motor neurons in familial and sporadic ALS. *Proc Natl Acad Sci U S A,* 111**,** 829-32.

Millecamps, S., Boillee, S., Le Ber, I., Seilhean, D., Teyssou, E.*, et al.* 2012. Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. *J Med Genet,* 49**,** 258-63.

Miller, F. D. & Gauthier, A. S. 2007. Timing Is Everything: Making Neurons versus Glia in the Developing Cortex. *Neuron,* 54**,** 357-369.

Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J.*, et al.* 2011. A TALE nuclease architecture for efficient genome editing. *Nature biotechnology,* 29**,** 143-148.

Miller, J. D., Ganat, Y. M., Kishinevsky, S., Bowman, R. L., Liu, B.*, et al.* 2013. Human iPSC-based modeling of late-onset disease via progerin-induced aging. *Cell Stem Cell,* 13**,** 691-705.

Miller, Z. A., Sturm, V. E., Camsari, G. B., Karydas, A., Yokoyama, J. S.*, et al.* 2016. Increased prevalence of autoimmune disease within C9 and FTD/MND cohorts: Completing the picture. *Neurol Neuroimmunol Neuroinflamm,* 3**,** e301.

Mizielinska, S., Gronke, S., Niccoli, T., Ridler, C. E., Clayton, E. L.*, et al.* 2014. C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science,* 345**,** 1192-4.

Mizielinska, S., Lashley, T., Norona, F. E., Clayton, E. L., Ridler, C. E.*, et al.* 2013. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol,* 126**,** 845-57.

Moens, T. G., Niccoli, T., Wilson, K. M., Atilano, M. L., Birsa, N.*, et al.* 2019. C9orf72 arginine-rich dipeptide proteins interact with ribosomal proteins in vivo to induce a toxic translational arrest that is rescued by eIF1A. *Acta Neuropathol,* 137**,** 487-500.

Mok, K., Traynor, B. J., Schymick, J., Tienari, P. J., Laaksovirta, H.*, et al.* 2012. Chromosome 9 ALS and FTD locus is probably derived from a single founder. *Neurobiol Aging,* 33**,** 209 e3-8.

Moore, S., Alsop, E., Lorenzini, I., Starr, A., Rabichow, B. E.*, et al.* 2019. ADAR2 mislocalization and widespread RNA editing aberrations in C9orf72-mediated ALS/FTD. *Acta Neuropathol,* 138**,** 49-65.

Mori, K., Arzberger, T., Grasser, F. A., Gijselinck, I., May, S.*, et al.* 2013a. Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol,* 126**,** 881-93.

Mori, K., Lammich, S., Mackenzie, I. R., Forne, I., Zilow, S.*, et al.* 2013b. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-

negative inclusions in the hippocampus of patients with C9orf72 mutations. *Acta Neuropathol,* 125**,** 413-23.

Mori, K., Weng, S. M., Arzberger, T., May, S., Rentzsch, K.*, et al.* 2013c. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science,* 339**,** 1335-8.

Morita, M., Al-Chalabi, A., Andersen, P. M., Hosler, B., Sapp, P.*, et al.* 2006. A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology,* 66**,** 839-44.

Moscou, M. J. & Bogdanove, A. J. 2009. A simple cipher governs DNA recognition by TAL effectors. *Science,* 326**,** 1501-1501.

Münch, C., Sedlmeier, R., Meyer, T., Homberg, V., Sperfeld, A. D.*, et al.* 2004. Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. *Neurology,* 63**,** 724-726.

Murphy, N. A., Arthur, K. C., Tienari, P. J., Houlden, H., Chio, A.*, et al.* 2017. Age-related penetrance of the C9orf72 repeat expansion. *Sci Rep,* 7**,** 2116.

Murray, M. E., Dejesus-Hernandez, M., Rutherford, N. J., Baker, M., Duara, R.*, et al.* 2011. Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72. *Acta Neuropathol,* 122**,** 673-90.

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D.*, et al.* 1998. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology,* 51**,** 1546-54.

Neumann, M., Bentmann, E., Dormann, D., Jawaid, A., Dejesus-Hernandez, M.*, et al.* 2011. FET proteins TAF15 and EWS are selective markers that distinguish FTLD with FUS pathology from amyotrophic lateral sclerosis with FUS mutations. *Brain,* 134**,** 2595-609.

Neumann, M., Mackenzie, I. R., Cairns, N. J., Boyer, P. J., Markesbery, W. R.*, et al.* 2007. TDP-43 in the ubiquitin pathology of frontotemporal dementia with VCP gene mutations. *J Neuropathol Exp Neurol,* 66**,** 152-7.

Neumann, M., Rademakers, R., Roeber, S., Baker, M., Kretzschmar, H. A.*, et al.* 2009. A new subtype of frontotemporal lobar degeneration with FUS pathology. *Brain : a journal of neurology,* 132**,** 2922-2931.

Niblock, M., Smith, B. N., Lee, Y.-B., Sardone, V., Topp, S.*, et al.* 2016. Retention of hexanucleotide repeat-containing intron in C9orf72 mRNA: implications for the pathogenesis of ALS/FTD. *Acta Neuropathologica Communications,* 4**,** 18.

Nordin, A., Akimoto, C., Wuolikainen, A., Alstermark, H., Jonsson, P.*, et al.* 2015. Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Hum Mol Genet,* 24**,** 3133-42.

O'rourke, J. G., Bogdanik, L., Muhammad, A., Gendron, T. F., Kim, K. J.*, et al.* 2015. C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron,* 88**,** 892-901.

O'rourke, J. G., Bogdanik, L., Yanez, A., Lall, D., Wolf, A. J.*, et al.* 2016. C9orf72 is required for proper macrophage and microglial function in mice. *Science,* 351**,** 1324-9.

Oberle, I., Rousseau, F., Heitz, D., Kretz, C., Devys, D.*, et al.* 1991. Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science***,** 1097-1102.

Onyike, C. U. & Diehl-Schmid, J. 2013. The epidemiology of frontotemporal dementia. *Int Rev Psychiatry,* 25**,** 130-7.

Ordureau, A., Heo, J.-M., Duda, D. M., Paulo, J. A., Olszewski, J. L.*, et al.* 2015. Defining roles of PARKIN and ubiquitin phosphorylation by PINK1 in mitochondrial quality control using a ubiquitin replacement strategy. *Proceedings of the National Academy of Sciences,* 112**,** 6637-6642.

Otomo, A., Hadano, S., Okada, T., Mizumura, H., Kunita, R.*, et al.* 2003. ALS2, a novel guanine nucleotide exchange factor for the small GTPase Rab5, is implicated in endosomal dynamics. *Hum Mol Genet,* 12**,** 1671-87.

Papadopoulos, C., Kirchner, P., Bug, M., Grum, D., Koerver, L.*, et al.* 2017. VCP/p97 cooperates with YOD1, UBXD1 and PLAA to drive clearance of ruptured lysosomes by autophagy. *The EMBO Journal,* 36**,** 135-150.

Pearson, J. P., Williams, N. M., Majounie, E., Waite, A., Stott, J.*, et al.* 2011. Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *J Neurol,* 258**,** 647-55.

Pera, E. M., Ikeda, A., Eivers, E. & De Robertis, E. M. 2004. Integration of IGF, FGF, and anti-BMP signals via Smad1 phosphorylation in neural induction. *Genes Dev,* 17**,** 3023-8.

Peters, O. M., Cabrera, G. T., Tran, H., Gendron, T. F., Mckeon, J. E.*, et al.* 2015. Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron,* 88**,** 902-909.

Philippidou, P. & Dasen, J. S. 2013. Hox genes: choreographers in neural development, architects of circuit organization. *Neuron,* 80**,** 12-34.

Pickles, S., Vigié, P. & Youle, R. J. 2018. Mitophagy and Quality Control Mechanisms in Mitochondrial Maintenance. *Current Biology,* 28**,** R170-R185.

Pieretti, M., Zhang, F. P., Fu, Y. H., Warren, S. T., Oostra, B. A.*, et al.* 1991. Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell,* 66**,** 817-22.

Pilz, G.-A., Shitamukai, A., Reillo, I., Pacary, E., Schwausch, J.*, et al.* 2013. Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type. *Nature Communications,* 4**,** 2125.

Poole, A. C., Thomas, R. E., Yu, S., Vincow, E. S. & Pallanck, L. 2010. The mitochondrial fusion-promoting factor mitofusin is a substrate of the PINK1/parkin pathway. *PloS one,* 5**,** e10054.

Porterfield, V., Khan, S. S., Foff, E. P., Koseoglu, M. M., Blanco, I. K.*, et al.* 2020. A three-dimensional dementia model reveals spontaneous cell cycle re-entry and a senescence-associated secretory phenotype. *Neurobiology of Aging,* 90**,** 125-134.

Pottier, C., Bieniek, K. F., Finch, N., Van De Vorst, M., Baker, M.*, et al.* 2015. Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta Neuropathol,* 130**,** 77-92.

Prince, M. J., Wimo, A., Guerchet, M. M., Ali, G. C., Wu, Y.-T.*, et al.* 2015. *World Alzheimer Report 2015 - The Global Impact of Dementia*, Alzheimer's Disease International.

Ran, F. A., Hsu, P. D., Lin, C.-Y., Gootenberg, J. S., Konermann, S.*, et al.* 2013a. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell,* 154**,** 1380-1389.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A.*, et al.* 2013b. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc,* 8**,** 2281-308.

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H.*, et al.* 2011. Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain,* 134**,** 2456-77.

Ratti, A., Corrado, L., Castellotti, B., Del Bo, R., Fogh, I.*, et al.* 2012. C9ORF72 repeat expansion in a large Italian ALS cohort: evidence of a founder effect. *Neurobiol Aging,* 33**,** 2528 e7-14.

Ratti, A., Gumina, V., Lenzi, P., Bossolasco, P., Fulceri, F.*, et al.* 2020. Chronic stress induces formation of stress granules and pathological TDP-43 aggregates in human ALS fibroblasts and iPSC-motoneurons. *Neurobiol Dis***,** 105051.

Reddy, K., Schmidt, M. H., Geist, J. M., Thakkar, N. P., Panigrahi, G. B.*, et al.* 2014. Processing of double-R-loops in (CAG).(CTG) and C9orf72

(GGGGCC).(GGCCCC) repeats causes instability. *Nucleic Acids Res,* 42**,** 10473-87.

Reddy, K., Zamiri, B., Stanley, S. Y., Macgregor, R. B., Jr. & Pearson, C. E. 2013. The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J Biol Chem,* 288**,** 9860-6.

Renton, A. E., Majounie, E., Waite, A., Simon-Sanchez, J., Rollinson, S.*, et al.* 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron,* 72**,** 257-68.

Richter, B., Sliter, D. A., Herhaus, L., Stolz, A., Wang, C.*, et al.* 2016. Phosphorylation of OPTN by TBK1 enhances its binding to Ub chains and promotes selective autophagy of damaged mitochondria. *Proceedings of the National Academy of Sciences,* 113**,** 4039-4044.

Rizzu, P., Blauwendraat, C., Heetveld, S., Lynes, E. M., Castillo-Lizardo, M.*, et al.* 2016. C9orf72 is differentially expressed in the central nervous system and myeloid cells and consistently reduced in C9orf72, MAPT and GRN mutation carriers. *Acta Neuropathol Commun,* 4**,** 37.

Rohrer, J. D., Guerreiro, R., Vandrovcova, J., Uphill, J., Reiman, D.*, et al.* 2009. The heritability and genetics of frontotemporal lobar degeneration. *Neurology,* 73**,** 1451-6.

Rohrer, J. D., Isaacs, A. M., Mizielinska, S., Mead, S., Lashley, T.*, et al.* 2015. C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *Lancet Neurol,* 14**,** 291-301.

Römer, P., Hahn, S., Jordan, T., Strauss, T., Bonas, U.*, et al.* 2007. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science,* 318**,** 645-648.

Rowland, L. P. & Shneider, N. A. 2001. Amyotrophic lateral sclerosis. *N Engl J Med,* 344**,** 1688-700.

Rubino, E., Rainero, I., Chiò, A., Rogaeva, E., Galimberti, D.*, et al.* 2012. SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Neurology,* 79**,** 1556-62.

Russ, J., Liu, E. Y., Wu, K., Neal, D., Suh, E.*, et al.* 2014. Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta Neuropathol,* 129**,** 39-52.

Rutherford, N. J., Heckman, M. G., Dejesus-Hernandez, M., Baker, M. C., Soto-Ortolaza, A. I.*, et al.* 2012. Length of normal alleles of C9ORF72 GGGGCC repeat do not influence disease phenotype. *Neurobiol Aging,* 33**,** 2950 e5-7.

Sabatelli, M., Conforti, F. L., Zollino, M., Mora, G., Monsurro, M. R.*, et al.* 2012. C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. *Neurobiol Aging,* 33**,** 1848 e15-20.

Saberi, S., Stauffer, J. E., Jiang, J., Garcia, S. D., Taylor, A. E.*, et al.* 2017. Sense-encoded poly-GR dipeptide repeat proteins correlate to neurodegeneration and uniquely co-localize with TDP-43 in dendrites of repeat-expanded C9orf72 amyotrophic lateral sclerosis. *Acta Neuropathol,* 135**,** 459-474.

Sakae, N., Bieniek, K. F., Zhang, Y. J., Ross, K., Gendron, T. F.*, et al.* 2018. Poly-GR dipeptide repeat polymers correlate with neurodegeneration and Clinicopathological subtypes in C9ORF72-related brain disease. *Acta Neuropathol Commun,* 6**,** 63.

Sampson, T. R., Debelius, J. W., Thron, T., Janssen, S., Shastri, G. G.*, et al.* 2016. Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. *Cell,* 167**,** 1469-1480.e12.

Sanjana, N. E., Cong, L., Zhou, Y., Cunniff, M. M., Feng, G.*, et al.* 2012. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc,* 7**,** 171-92.

Sareen, D., O'rourke, J. G., Meera, P., Muhammad, A. K., Grant, S.*, et al.* 2013. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med,* 5**,** 208ra149.

Sasai, Y. 2013. Cytosystems dynamics in self-organization of tissue architecture. *Nature,* 493**,** 318-26.

Satoh, J., Tabunoki, H., Ishida, T., Saito, Y. & Arima, K. 2012. Dystrophic neurites express C9orf72 in Alzheimer's disease brains. *Alzheimers Res Ther,* 4**,** 33.

Satoh, J., Yamamoto, Y., Kitano, S., Takitani, M., Asahina, N.*, et al.* 2014. Molecular network analysis suggests a logical hypothesis for the pathological role of c9orf72 in amyotrophic lateral sclerosis/frontotemporal dementia. *J Cent Nerv Syst Dis,* 6**,** 69-78.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M.*, et al.* 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods,* 9**,** 676-82.

Schludi, M. H., Becker, L., Garrett, L., Gendron, T. F., Zhou, Q.*, et al.* 2017. Spinal poly-GA inclusions in a C9orf72 mouse model trigger motor deficits and inflammation without neuron loss. *Acta Neuropathol,* 134**,** 241-254.

Seelaar, H., Kamphorst, W., Rosso, S. M., Azmani, A., Masdjedi, R.*, et al.* 2008. Distinct genetic forms of frontotemporal dementia. *Neurology,* 71**,** 1220-6.

Sellier, C., Campanari, M. L., Julie Corbier, C., Gaucherot, A., Kolb-Cheynel, I.*, et al.* 2016. Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death. *EMBO J,* 35**,** 1276-97.

Selvaraj, B. T., Livesey, M. R., Zhao, C., Gregory, J. M., James, O. T.*, et al.* 2018. C9ORF72 repeat expansion causes vulnerability of motor neurons to Ca(2+)-permeable AMPA receptor-mediated excitotoxicity. *Nat Commun,* 9**,** 347.

Shao, Q., Liang, C., Chang, Q., Zhang, W., Yang, M.*, et al.* 2019. C9orf72 deficiency promotes motor deficits of a C9ALS/FTD mouse model in a dose-dependent manner. *Acta Neuropathologica Communications,* 7**,** 32.

Shatunov, A., Mok, K., Newhouse, S., Weale, M. E., Smith, B.*, et al.* 2010. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol,* 9**,** 986-94.

Shi, Y., Kirwan, P. & Livesey, F. J. 2012a. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. *Nat Protoc,* 7**,** 1836-46.

Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. C. & Livesey, F. J. 2012b. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nature Neuroscience,* 15**,** 477-486.

Shi, Y., Lin, S., Staats, K. A., Li, Y., Chang, W. H.*, et al.* 2018. Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat Med,* 24**,** 313-325.

Shiba-Fukushima, K., Arano, T., Matsumoto, G., Inoshita, T., Yoshida, S.*, et al.* 2014. Phosphorylation of mitochondrial polyubiquitin by PINK1 promotes Parkin mitochondrial tethering. *PLoS Genet,* 10**,** e1004861.

Shiba-Fukushima, K., Imai, Y., Yoshida, S., Ishihama, Y., Kanao, T.*, et al.* 2012. PINK1-mediated phosphorylation of the Parkin ubiquitin-like domain primes mitochondrial translocation of Parkin and regulates mitophagy. *Scientific Reports,* 2**,** 1002.

Silverman, H. E., Goldman, J. S. & Huey, E. D. 2019. Links Between the C9orf72 Repeat Expansion and Psychiatric Symptoms. *Current Neurology and Neuroscience Reports,* 19**,** 93.

Simon-Sanchez, J., Dopper, E. G., Cohn-Hokke, P. E., Hukema, R. K., Nicolaou, N.*, et al.* 2012. The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. *Brain,* 135**,** 723-35.

Simone, R., Balendra, R., Moens, T. G., Preza, E., Wilson, K. M.*, et al.* 2018. G-quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *EMBO Molecular Medicine,* 10**,** 22-31.

Sivadasan, R., Hornburg, D., Drepper, C., Frank, N., Jablonka, S.*, et al.* 2016. C9ORF72 interaction with cofilin modulates actin dynamics in motor neurons. *Nat Neurosci,* 19**,** 1610-1618.

Skibinski, G., Parkinson, N. J., Brown, J. M., Chakrabarti, L., Lloyd, S. L.*, et al.* 2005. Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat Genet,* 37**,** 806-8.

Smith, B. N., Newhouse, S., Shatunov, A., Vance, C., Topp, S.*, et al.* 2013. The C9ORF72 expansion mutation is a common cause of ALS+/-FTD in Europe and has a single founder. *Eur J Hum Genet,* 21**,** 102-8.

Smith, K. R., Damiano, J., Franceschetti, S., Carpenter, S., Canafoglia, L.*, et al.* 2012. Strikingly different clinicopathological phenotypes determined by progranulin-mutation dosage. *Am J Hum Genet,* 90**,** 1102-7.

Snowden, J. S., Rollinson, S., Thompson, J. C., Harris, J. M., Stopford, C. L.*, et al.* 2012. Distinct clinical and pathological characteristics of frontotemporal dementia associated with C9ORF72 mutations. *Brain,* 135**,** 693-708.

Spillantini, M. G., Murrell, J. R., Goedert, M., Farlow, M. R., Klug, A.*, et al.* 1998. Mutation in the tau gene in familial multiple system tauopathy with presenile dementia. *Proc Natl Acad Sci U S A,* 95**,** 7737-41.

Sreedharan, J., Blair, I. P., Tripathi, V. B., Hu, X., Vance, C.*, et al.* 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science,* 319**,** 1668-72.

Stagi, M., Klein, Z. A., Gould, T. J., Bewersdorf, J. & Strittmatter, S. M. 2014. Lysosome size, motility and stress response regulated by fronto-temporal dementia modifier TMEM106B. *Molecular and Cellular Neuroscience,* 61**,** 226-240.

Starr, A. & Sattler, R. 2018. Synaptic dysfunction and altered excitability in C9ORF72 ALS/FTD. *Brain research,* 1693**,** 98-108.

Stein, Jason l., De la torre-Ubieta, L., Tian, Y., Parikshak, Neelroop n., Hernández, Israel a.*, et al.* 2014. A Quantitative Framework to Evaluate Modeling of Cortical Development by Neural Stem Cells. *Neuron,* 83**,** 69-86.

Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. 2014. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature,* 507**,** 62-67.

Stewart, H., Rutherford, N. J., Briemberg, H., Krieger, C., Cashman, N.*, et al.* 2012. Clinical and pathological features of amyotrophic lateral sclerosis caused by mutation in the C9ORF72 gene on chromosome 9p. *Acta Neuropathol,* 123**,** 409-17.

Stoica, R., De Vos, K. J., Paillusson, S., Mueller, S., Sancho, R. M.*, et al.* 2014. ER-mitochondria associations are regulated by the VAPB-PTPIP51 interaction and are disrupted by ALS/FTD-associated TDP-43. *Nat Commun,* 5**,** 3996.

Sudria-Lopez, E., Koppers, M., De Wit, M., Van Der Meer, C., Westeneng, H. J.*, et al.* 2016. Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta Neuropathol,* 132**,** 145-7.

Suh, E., Lee, E. B., Neal, D., Wood, E. M., Toledo, J. B.*, et al.* 2015. Semi-automated quantification of C9orf72 expansion size reveals inverse correlation between hexanucleotide repeat number and disease duration in frontotemporal degeneration. *Acta Neuropathol,* 130**,** 363-72.

Sullivan, P. M., Zhou, X., Robins, A. M., Paushter, D. H., Kim, D.*, et al.* 2016. The ALS/FTLD associated protein C9orf72 associates with SMCR8 and WDR41 to regulate the autophagy-lysosome pathway. *Acta Neuropathol Commun,* 4**,** 51.

Sutcliffe, J. S., Nelson, D. L., Zhang, F., Pieretti, M., Caskey, C. T.*, et al.* 1992. DNA methylation represses FMR-1 transcription in fragile X syndrome. *Hum Mol Genet,* 1**,** 397-400.

Suzuki, N., Maroof, A. M., Merkle, F. T., Koszka, K., Intoh, A.*, et al.* 2013. The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nature Neuroscience,* 16**,** 1725-1727.

Symington, L. S. & Gautier, J. 2011. Double-strand break end resection and repair pathway choice. *Annual review of genetics,* 45**,** 247-271.

Taapken, S. M., Nisler, B. S., Newton, M. A., Sampsell-Barron, T. L., Leonhard, K. A.*, et al.* 2011. Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nat Biotechnol,* 29**,** 313-4.

Takahashi, K., Okita, K., Nakagawa, M. & Yamanaka, S. 2007. Induction of pluripotent stem cells from fibroblast cultures. *Nat Protoc,* 2**,** 3081-9.

Tanaka, A., Cleland, M. M., Xu, S., Narendra, D. P., Suen, D.-F.*, et al.* 2010. Proteasome and p97 mediate mitophagy and degradation of mitofusins induced by Parkin. *Journal of Cell Biology,* 191**,** 1367-1380.

Tanaka, Y. & Chen, Z. J. 2012. STING specifies IRF3 phosphorylation by TBK1 in the cytosolic DNA signaling pathway. *Science signaling,* 5**,** ra20-ra20.

Tang, D., Sheng, J., Xu, L., Zhan, X., Liu, J.*, et al.* 2020. Cryo-EM structure of C9ORF72-SMCR8-WDR41 reveals the role as a GAP for Rab8a and Rab11a. *bioRxiv***,** 2020.04.16.045708.

Tang, Y., Liu, M. L., Zang, T. & Zhang, C. L. 2017. Direct Reprogramming Rather than iPSC-Based Reprogramming Maintains Aging Hallmarks in Human Motor Neurons. *Front Mol Neurosci,* 10**,** 359.

Therrien, M., Rouleau, G. A., Dion, P. A. & Parker, J. A. 2013. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in C. elegans. *PLoS One,* 8**,** e83450.

Tran, H., Almeida, S., Moore, J., Gendron, T. F., Chalasani, U.*, et al.* 2015. Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron,* 87**,** 1207-1214.

Tresse, E., Salomons, F. A., Vesa, J., Bott, L. C., Kimonis, V.*, et al.* 2010. VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD. *Autophagy,* 6**,** 217-27.

Tsai, S. Q., Wyvekens, N., Khayter, C., Foden, J. A., Thapar, V.*, et al.* 2014. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nature biotechnology,* 32**,** 569-576.

Ugolino, J., Ji, Y. J., Conchina, K., Chu, J., Nirujogi, R. S.*, et al.* 2016. Loss of C9orf72 Enhances Autophagic Activity via Deregulated mTOR and TFEB Signaling. *PLoS Genet,* 12**,** e1006443.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C.*, et al.* 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Res,* 40**,** e115.

Urbach, A., Bar-Nur, O., Daley, G. Q. & Benvenisty, N. 2010. Differential modeling of fragile X syndrome by human embryonic stem cells and induced pluripotent stem cells. *Cell Stem Cell,* 6**,** 407-11.

Urnov, F. D., Miller, J. C., Lee, Y.-L., Beausejour, C. M., Rock, J. M.*, et al.* 2005. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature,* 435**,** 646-651.

Van Blitterswijk, M., Baker, M. C., Dejesus-Hernandez, M., Ghidoni, R., Benussi, L.*, et al.* 2013a. C9ORF72 repeat expansions in cases with previously identified pathogenic mutations. *Neurology,* 81**,** 1332-41.

Van Blitterswijk, M., Dejesus-Hernandez, M., Niemantsverdriet, E., Murray, M. E., Heckman, M. G.*, et al.* 2013b. Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol,* 12**,** 978-88.

Van Blitterswijk, M., Gendron, T. F., Baker, M. C., Dejesus-Hernandez, M., Finch, N. A.*, et al.* 2015. Novel clinical associations with specific C9ORF72 transcripts in patients with repeat expansions in C9ORF72. *Acta Neuropathol,* 130**,** 863-76.

Van Blitterswijk, M., Mullen, B., Heckman, M. G., Baker, M. C., Dejesus-Hernandez, M.*, et al.* 2014a. Ataxin-2 as potential disease modifier in C9ORF72 expansion carriers. *Neurobiol Aging,* 35**,** 2421 e13-7.

Van Blitterswijk, M., Mullen, B., Nicholson, A. M., Bieniek, K. F., Heckman, M. G.*, et al.* 2014b. TMEM106B protects C9ORF72 expansion carriers against frontotemporal dementia. *Acta Neuropathol,* 127**,** 397-406.

Van Deerlin, V. M., Sleiman, P. M., Martinez-Lage, M., Chen-Plotkin, A., Wang, L. S.*, et al.* 2010. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet,* 42**,** 234-9.

Van Der Zee, J., Gijselinck, I., Dillen, L., Van Langenhove, T., Theuns, J.*, et al.* 2013. A pan-European study of the C9orf72 repeat associated with FTLD: geographic prevalence, genomic instability, and intermediate repeats. *Hum Mutat,* 34**,** 363-73.

Van Es, M. A., Veldink, J. H., Saris, C. G., Blauw, H. M., Van Vught, P. W.*, et al.* 2009. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet,* 41**,** 1083-7.

Van Langenhove, T., Van Der Zee, J., Gijselinck, I., Engelborghs, S., Vandenberghe, R.*, et al.* 2013. Distinct clinical characteristics of C9orf72 expansion carriers compared with GRN, MAPT, and nonmutation carriers in a Flanders-Belgian FTLD cohort. *JAMA Neurol,* 70**,** 365-73.

Van Mossevelde, S., Van Der Zee, J., Cruts, M. & Van Broeckhoven, C. 2017a. Relationship between C9orf72 repeat size and clinical phenotype. *Current Opinion in Genetics & Development,* 44**,** 117-124.

Van Mossevelde, S., Van Der Zee, J., Gijselinck, I., Sleegers, K., De Bleecker, J.*, et al.* 2017b. Clinical Evidence of Disease Anticipation in Families Segregating a C9orf72 Repeat Expansion. *JAMA Neurol,* 74**,** 445-452.

Vance, C., Al-Chalabi, A., Ruddy, D., Smith, B. N., Hu, X.*, et al.* 2006. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain,* 129**,** 868-76.

Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K. J., Nishimura, A. L.*, et al.* 2009. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science,* 323**,** 1208-1211.

Varcianna, A., Myszczynska, M. A., Castelli, L. M., O'neill, B., Kim, Y.*, et al.* 2019. Micro-RNAs secreted through astrocyte-derived extracellular vesicles cause neuronal network degeneration in C9orf72 ALS. *EBioMedicine,* 40**,** 626-635.

Vera, E., Bosco, N. & Studer, L. 2016. Generating Late-Onset Human iPSC-Based Disease Models by Inducing Neuronal Age-Related Phenotypes through Telomerase Manipulation. *Cell Reports,* 17**,** 1184-1192.

Verkerk, A. J. M. H., Pieretti, M., Sutcliffe, J. S., Fu, Y.-H., Kuhl, D. P. A.*, et al.* 1991. Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell,* 65**,** 905-914.

Vieira, R. T., Caixeta, L., Machado, S., Silva, A. C., Nardi, A. E.*, et al.* 2013. Epidemiology of early-onset dementia: a review of the literature. *Clin Pract Epidemiol Ment Health,* 9**,** 88-95.

Wainger, B. J., Kiskinis, E., Mellin, C., Wiskow, O., Han, S. S.*, et al.* 2014. Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep,* 7**,** 1-11.

Waite, A. J., Baumer, D., East, S., Neal, J., Morris, H. R.*, et al.* 2014. Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and

frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiol Aging,* 35**,** 1779.e5-1779.e13.
Watanabe, K., Kamiya, D., Nishiyama, A., Katayama, T., Nozaki, S.*, et al.* 2005. Directed differentiation of telencephalic precursors from embryonic stem cells. *Nat Neurosci,* 8**,** 288-96.
Watts, G. D., Wymer, J., Kovach, M. J., Mehta, S. G., Mumm, S.*, et al.* 2004. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat Genet,* 36**,** 377-81.
Webster, C. P., Smith, E. F., Bauer, C. S., Moller, A., Hautbergue, G. M.*, et al.* 2016. The C9orf72 protein interacts with Rab1a and the ULK1 complex to regulate initiation of autophagy. *EMBO J,* 35**,** 1656-76.
Wen, X., Tan, W., Westergard, T., Krishnamurthy, K., Markandaiah, S. S.*, et al.* 2014. Antisense proline-arginine RAN dipeptides linked to C9ORF72-ALS/FTD form toxic nuclear aggregates that initiate in vitro and in vivo neuronal death. *Neuron,* 84**,** 1213-25.
Westergard, T., Jensen, B. K., Wen, X., Cai, J., Kropf, E.*, et al.* 2016. Cell-to-Cell Transmission of Dipeptide Repeat Proteins Linked to C9orf72-ALS/FTD. *Cell Rep,* 17**,** 645-652.
Williams, K. L., Fifita, J. A., Vucic, S., Durnall, J. C., Kiernan, M. C.*, et al.* 2013. Pathophysiological insights into ALS with C9ORF72 expansions. *J Neurol Neurosurg Psychiatry,* 84**,** 931-5.
Woollacott, I. O. & Mead, S. 2014. The C9ORF72 expansion mutation: gene structure, phenotypic and diagnostic issues. *Acta Neuropathol,* 127**,** 319-32.
Wu, J. J., Cai, A., Greenslade, J. E., Higgins, N. R., Fan, C.*, et al.* 2020. ALS/FTD mutations in UBQLN2 impede autophagy by reducing autophagosome acidification through loss of function. *Proc Natl Acad Sci U S A,* 117**,** 15230-15241.
Wyvekens, N., Topkar, V. V., Khayter, C., Joung, J. K. & Tsai, S. Q. 2015. Dimeric CRISPR RNA-guided FokI-dCas9 nucleases directed by truncated gRNAs for highly specific genome editing. *Human gene therapy,* 26**,** 425-431.
Xi, Z., Rainero, I., Rubino, E., Pinessi, L., Bruni, A. C.*, et al.* 2014a. Hypermethylation of the CpG-island near the C9orf72 G(4)C(2)-repeat expansion in FTLD patients. *Hum Mol Genet,* 23**,** 5630-7.
Xi, Z., Van Blitterswijk, M., Zhang, M., Mcgoldrick, P., Mclean, J. R.*, et al.* 2015a. Jump from pre-mutation to pathologic expansion in C9orf72. *Am J Hum Genet,* 96**,** 962-70.
Xi, Z., Yunusova, Y., Van Blitterswijk, M., Dib, S., Ghani, M.*, et al.* 2014b. Identical twins with the C9orf72 repeat expansion are discordant for ALS. *Neurology,* 83**,** 1476-1478.
Xi, Z., Zhang, M., Bruni, A. C., Maletta, R. G., Colao, R.*, et al.* 2015b. The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathol,* 129**,** 715-27.
Xi, Z., Zinman, L., Grinberg, Y., Moreno, D., Sato, C.*, et al.* 2012. Investigation of c9orf72 in 4 neurodegenerative disorders. *Arch Neurol,* 69**,** 1583-90.
Xi, Z., Zinman, L., Moreno, D., Schymick, J., Liang, Y.*, et al.* 2013. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am J Hum Genet,* 92**,** 981-9.
Xiao, S., Macnair, L., Mcgoldrick, P., Mckeever, P. M., Mclean, J. R.*, et al.* 2015. Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Ann Neurol,* 78**,** 568-83.
Xiao, S., Macnair, L., Mclean, J., Mcgoldrick, P., Mckeever, P.*, et al.* 2016. C9orf72 isoforms in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration. *Brain Res,* 1647**,** 43-9.

Xiao, Y., Zou, Q., Xie, X., Liu, T., Li, H. S.*, et al.* 2017. The kinase TBK1 functions in dendritic cells to regulate T cell homeostasis, autoimmunity, and antitumor immunity. *Journal of Experimental Medicine,* 214**,** 1493-1507.

Yamano, K. & Youle, R. J. 2013. PINK1 is degraded through the N-end rule pathway. *Autophagy,* 9**,** 1758-1769.

Yang, M., Liang, C., Swaminathan, K., Herrlinger, S., Lai, F.*, et al.* 2016. A C9ORF72/SMCR8-containing complex regulates ULK1 and plays a dual role in autophagy. *Sci Adv,* 2**,** e1601167.

Yang, Y., Hentati, A., Deng, H. X., Dabbagh, O., Sasaki, T.*, et al.* 2001. The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. *Nat Genet,* 29**,** 160-5.

Yu, J., Zhou, X., Chang, M., Nakaya, M., Chang, J. H.*, et al.* 2015. Regulation of T-cell activation and migration by the kinase TBK1 during neuroinflammation. *Nat Commun,* 6**,** 6074.

Yu, T., Yi, Y.-S., Yang, Y., Oh, J., Jeong, D.*, et al.* 2012. The Pivotal Role of TBK1 in Inflammatory Responses Mediated by Macrophages. *Mediators of Inflammation,* 2012**,** 979105.

Yuva-Aydemir, Y., Almeida, S., Krishnan, G., Gendron, T. F. & Gao, F.-B. 2019. Transcription elongation factor AFF2/FMR2 regulates expression of expanded GGGGCC repeat-containing C9ORF72 allele in ALS/FTD. *Nature Communications,* 10**,** 5466.

Zetsche, B., Gootenberg, Jonathan s., Abudayyeh, Omar o., Slaymaker, Ian m., Makarova, Kira s.*, et al.* 2015. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell,* 163**,** 759-771.

Zhang, D., Iyer, L. M., He, F. & Aravind, L. 2012. Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Front Genet,* 3**,** 283.

Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G. M.*, et al.* 2011. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nature biotechnology,* 29**,** 149-153.

Zhang, K., Daigle, J. G., Cunningham, K. M., Coyne, A. N., Ruan, K.*, et al.* 2018a. Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. *Cell,* 173**,** 958-971.e17.

Zhang, K., Donnelly, C. J., Haeusler, A. R., Grima, J. C., Machamer, J. B.*, et al.* 2015. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature,* 525**,** 56-61.

Zhang, M., Ferrari, R., Tartaglia, M. C., Keith, J., Surace, E. I.*, et al.* 2018b. A C6orf10/LOC101929163 locus is associated with age of onset in C9orf72 carriers. *Brain,* 141**,** 2895-2907.

Zhang, M., Tartaglia, M. C., Moreno, D., Sato, C., Mckeever, P.*, et al.* 2017. DNA methylation age-acceleration is associated with disease duration and age at onset in C9orf72 patients. *Acta Neuropathologica,* 134**,** 271-279.

Zhang, S.-C., Wernig, M., Duncan, I. D., Brüstle, O. & Thomson, J. A. 2001. In vitro differentiation of transplantable neural precursors from human embryonic stem cells. *Nature Biotechnology,* 19**,** 1129-1133.

Zhang, Y.-J., Gendron, T. F., Ebbert, M. T. W., O'raw, A. D., Yue, M.*, et al.* 2018c. Poly(GR) impairs protein translation and stress granule dynamics in C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *Nature Medicine,* 24**,** 1136-1142.

Zhang, Y. J., Gendron, T. F., Grima, J. C., Sasaguri, H., Jansen-West, K.*, et al.* 2016. C9ORF72 poly(GA) aggregates sequester and impair HR23 and nucleocytoplasmic transport proteins. *Nat Neurosci,* 19**,** 668-677.

Zhang, Y. J., Jansen-West, K., Xu, Y. F., Gendron, T. F., Bieniek, K. F.*, et al.* 2014. Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol,* 128**,** 505-24.

Zhao, B., Du, F., Xu, P., Shu, C., Sankaran, B.*, et al.* 2019a. A conserved PLPLRT/SD motif of STING mediates the recruitment and activation of TBK1. *Nature,* 569**,** 718-722.

Zhao, C., Devlin, A. C., Chouhan, A. K., Selvaraj, B. T., Stavrou, M.*, et al.* 2019b. Mutant C9orf72 human iPSC-derived astrocytes cause non-cell autonomous motor neuron pathophysiology. *Glia,* 68**,** 1046-1064.

Zhou, Y., Kumari, D., Sciascia, N. & Usdin, K. 2016. CGG-repeat dynamics and FMR1 gene silencing in fragile X syndrome stem cells and stem cell-derived neurons. *Mol Autism,* 7**,** 42.

Zhu, Q., Jiang, J., Gendron, T. F., Mcalonis-Downes, M., Jiang, L.*, et al.* 2020. Reduced C9ORF72 function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in C9orf72. *Nat Neurosci,* 2**3,** 615-624.

Zhu, Y., Li, H., Zhou, L., Wu, J. Y. & Rao, Y. 1999. Cellular and molecular guidance of GABAergic neuronal migration from an extracortical origin to the neocortex. *Neuron,* 23**,** 473-85.

Zirra, A., Wiethoff, S. & Patani, R. 2016. Neural Conversion and Patterning of Human Pluripotent Stem Cells: A Developmental Perspective. *Stem Cells Int,* 2016**,** 8291260.

Zu, T., Liu, Y., Bañez-Coronel, M., Reid, T., Pletnikova, O.*, et al.* 2013. RAN proteins and RNA foci from antisense transcripts in C9ORF72 ALS and frontotemporal dementia. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** E4968-E4977.